

**Remember as we empathize. Do brain mechanisms engaged in autobiographical memory retrieval causally affect empathy awareness? A combined TMS and EEG registered report**  
**Meconi, Federica; Hodsoll, John; Goranova, Zheni; Degano, Giulio; Lello, Nicolò Di; Miniussi, Carlo; Avenanti, Alessio; Mevorach, Carmel**

---

Review timeline:

Submission date: 21 October 2020

Editorial Decision: Major Modification (29 December 2020)

Revision Received: 12 April 2021

Accepted: 17 May 2021

Editor 1: Eric Prager

Editor 2: David McArthur

Editor 3: Junie Warrington

Reviewer 1: Brendan Gaesser

Reviewer 2: Elia Valentini

1st Editorial Decision

**Decision letter**

Dear Dr Meconi:

Thank you for submitting your manuscript to the Journal of Neuroscience Research. We have now received the reviewer feedback and have appended those reviews below. As you will see, the reviewers find the question addressed to be of potential interest. Yet, they do not find the manuscript suitable for publication in its current form.

If you feel that you can adequately address the concerns of the reviewers, you may revise and resubmit your paper within 90 days. It will require further review. Please explain in your cover letter how you have changed the present version and submit a point by point response to the editors' and reviewers' comments. If you require longer than 90 days to make the revisions, please contact Dr Junie Warrington (jpwarrington@umc.edu). To submit your revised manuscript: Log in by clicking on the link below <https://rex-prod.resxchange.com/submissionBoard/1/8695eb4d-b5d3-430c-814e-e7efae01e5a4/current>

(If the above link space is blank, it is because you submitted your original manuscript through our old submission site. Therefore, to return your revision, please go to our new submission site here ([submission.wiley.com/jnr](https://www.submission.wiley.com/jnr)) and submit your revision as a new manuscript; answer yes to the question "Are you returning a revision for a manuscript originally submitted to our former submission site (ScholarOne Manuscripts)? If you indicate yes, please enter your original manuscript's Manuscript ID number in the space below" and including your original submission's Manuscript ID number (jnr-2020-Oct-9112) where indicated. This will help us to link your revision to your original submission.)

Thank you again for your submission to the Journal of Neuroscience Research; we look forward to reading your revised manuscript.

Best Wishes,

Dr Eric Prager

Associate Editor, Journal of Neuroscience Research

Dr Junie Warrington

Editor-in-Chief, Journal of Neuroscience Research

Editorial Comments to the Author:

1. Please ensure that the transparent science questionnaire includes references to page numbers and lines where possible instead of "y."

Associate Editor: Prager, Eric

Comments to the Author:

Thank you for your submission and again I apologize for the delay. The reviewers found the proposal to be of potential interest, but they did note some major concerns. I believe, based on my review, that many of these comments can be discussed and detailed in the text. One of the biggest concerns raised by the reviewers and statistical editor was the power analysis and statistical approaches. I'd recommend that you have a biostatistician consult with you on this project. If you are able to revise the manuscript to the satisfaction of the reviewers and especially the statistical editor, we would be happy to consider the manuscript again.

Statistics Editor: McArthur, David

Comments to the Author:

Every one of the comments made by the other reviewers will need close attention. In particular, from a statistical perspective, computing statistical power on basis of t-test results for a complex 4-way repeated-measures design can hardly be endorsed. Using pilot data results on the basis of merely 5 subjects who underwent a different protocol is very risky, not to mention that the largest interaction observed there was not statistically significant, and effect size calculations based on very small sample sizes is hardly stable. The issue of how many factors are in play in any given analysis must be fully clarified before proceeding, along with a clear rationale for dropping factors if indeed that is the plan.

Presently the four numbered hypotheses are not independent but inherently chained together. Overall this study must be rethought and constructed in sequential stages in order for the findings to make sense: first is to separate participants into a meaningful empathic spectrum, as there is no question but that different people even with very similar backgrounds have widely differing empathic capabilities; second is to get consistent and repeatable results from what is undoubtedly an exhausting and challenging behavioral task with all the fine details such as time pressures and intertrial intervals carefully optimized in advance; third is to obtain findings that precuneous stimulation is truly different from control and to ensure that the choice of control is the correct one. At present it reads like every consideration is in flux simultaneously and no amount of statistical wizardry will be sufficient to untangle the findings satisfactorily.

Reviewer: 1

Comments to the Author

The proposed study aims investigate whether disrupting activation of a core brain area of AM retrieval, the precuneus, interferes with an effect of memory recall on empathy. Overall, this study is positioned to provide a novel contribution to work on the role of autobiographical memory in empathy. Although there some existing work that hints at a neural relationship, much of this work is correlational or, as is the case with the amnesic patient studies, mostly underpowered with mixed results. As such, this work could make a valuable contribution with the potential to provide causal evidence for the role neural regions supporting autobiographical memory influencing how people feel for and connect with others.

While there is much to like about the proposed study, the manuscript/study may benefit from (i) further unpacking the nuances of a proposed relationship between the components of autobiographical memory the precuneus supports and how these specific components relate to empathy, (ii) better explain and justify behavioral design choices, and (iii) connecting with related work—though this last point is more minor than the first two.

COMMENTS

(1) Explanation for Including the Face Manipulation of Pain vs. No Pain. The rationale for including painful vs. neutral faces in the study could be made a little more clear. I could imagine different reasons for needing to include painful and neutral faces (e.g., it might be that photos are needed to make the materials more realistic and avoid floor effects in empathy ratings, or you might have the prediction that congruency in the emotion of the stories and faces is needed to see other effects of interest, etc.) but you should make the reasoning for the inclusion of the photos in your design explicit as well as state any predictions you have with regard to this factor (e.g., if you do indeed predict some sort of role of congruency between stories and faces and how this interacts with memory). This added manipulation seems to offer greater complexity in analysis without necessarily adding greater substance to the claims that the authors make in the paper. That said, a clearer motivation could mollify this concern.

(2) Explanation for Including the Polygon Cues

Similar to faces, the inclusion of polygon cues is also a little obscure. I have seen these used in other work in place of participants re-reading stories or cue words (e.g., having participants memorize associations between the stories and random polygons, and then testing memory retrieval in response just to the polygons), but being that you're already presenting cue words, what purpose do these polygons serve to improve the design and claims the study aims to evaluate?

(3) Sampling plan and analysis.

The authors nicely articulate and justify their sampling plan, including powering analyses for specific effects. In your power analysis, you power for the focal Memory x Stimulation site interaction, but your design includes four IVs. What do the author's plan on doing regarding higher-order interactions? With so many variables in play, the study has the potential to examine three-way and four-way interactions. However, if the study finds nonsignificant higher-order (e.g., three-way or four-way) interactions it would be difficult to tell whether this was because of the lack of a true interaction effect or because the study was underpowered to detect the higher-order interaction. It is thus possible that you would find the two-way interaction of interest but that this interaction would be partially qualified by higher-order interactions. Could you speak to this possibility and state why (or why not) this is a concern for the proposed study?

(4) Relevant Neural and Behavioral Findings on Memory and Empathy

There are a few recent papers that are particularly relevant to this work and would bolster the background in the Introduction. Wagner and colleagues (2020, *Social Cognitive and Affective Neuroscience*) found similar patterns of activation for first-hand pain and empathy for another's pain in the hippocampus, temporoparietal junction, retrosplenial cortex, and anterior insula. Notably, connectivity between the hippocampus and the fusiform gyrus during the pain empathy task scaled with increased self-reported perspective-taking skills, often considered a key component in empathy. There is also relevant work from Vollberg and colleagues (2019, preprint: [10.31234/osf.io/r6wmx](https://doi.org/10.31234/osf.io/r6wmx)) which found that participants who completed an episodic specificity induction later reported greater empathy for others experiencing negative events that were unrelated to the ESI, suggesting that increasing spontaneously generated episodic detail while processing other's suffering influences feelings of empathy.

The authors note inconsistent findings from patient populations with deficits in autobiographical memory (e.g., Rosenbaum et al., 2007; Beadle et al., 2013). I think it is worth drawing attention to readers that these studies are not only mixed but that they are severely underpowered with only a few patients in each study (2 in Rosenbaum, 3 in Beadle). Relatedly, a more recently published, and higher-powered study by Sawczak et al. (*Neuropsychologia*, 2019) showed that patients with medial temporal lobe epilepsy do not show an increase in empathy for others suffering (and willingness to help) after being prompted to construct specific episodes compared to healthy controls. This study did not examine spontaneous reactivation nor did they localize the anatomy underlying this difference with any level of precision, however, the results align with the current manuscript and should be discussed. It may also be worth citing Dermody et al. (2016, *Journal of Alzheimer's disease*) in your discussion of existing work on patients with memory impairments and empathy. Dermody and colleagues studied patients with Alzheimer's Disease and fronto-temporal dementia and find that both patient groups have deficits in cognitive empathy, while only the fronto-temporal impairments in affective empathy.

(5) Elaborating on the specific components of autobiographical memory supported by precuneus and how those components relate to empathy

There are many aspects of the proposed studies design and interpretation that seem astutely thought out and articulated (e.g., distinguishing the selectivity of stimulating the precuneus on the empathy vs. perceptual task; the author's openness to the possibility that because the precuneus is involved in explicit mentalizing, disrupting it could lead to an overall reduction in empathy ratings). And, the author's do a nice job of motivating the overall importance of studying the relationship between memory and empathy on the one hand, and linking the precuneus to autobiographical memory on the other hand.

"Compelling evidence supports the role of precuneus in both AM and specific empathic processes, i.e., explicit reasoning about others. Hebscher and colleagues (Hebscher et al., 2020) established in a transcranial magnetic stimulation (TMS) and magnetoencephalography study that the precuneus plays a causal role in AM retrieval. The authors nicely showed in an explicit recollection task that decreasing precuneus activation interferes with neural dynamics of early stages of AM retrieval and the vividness of recollection, see also (St. Jacques et al., 2017)." (p. 4)

However, I want to slightly push for a more detailed theoretical account of the mechanisms involved. I

think it would be beneficial to further unpack the details and specific processes/components of autobiographical memory associated with precuneus in St. Jacques et al. (2017), and particularly Hebscher et al. (2020), and how these components relate to or are distinct from previous working connecting components of memory with empathy such as in Vollberg et al., 2019, Sawczak et al., 2019, or Beadle et al., 2013.

(7) Empathy and related constructs “Mechanisms of inner simulation that are implemented by the human mirror neuron system (Molenberghs et al., 2012) and anterior and mid-cingulate cortex (Danziger et al., 2009) underpin the empathic feelings. Perspective shifting and mindreading that are implemented by a circuit of fronto-parietal and temporal areas, which includes the precuneus (Lamm et al., 2011; Schurz et al., 2014; Molenberghs et al., 2016), underlie the ability to explicitly reason about others.” (p. 3)

“Compelling evidence supports the role of precuneus in both AM and specific empathic processes i.e., explicit reasoning about others.” (p. 4)

The way empathy and empathic processes are sometimes discussed seems to blur the distinction between the emotional experience of empathy, empathic/prosocial concern, and the explicit mentalizing (perspective-taking or theory of mind), which are often associated with feelings of empathy, but not always. For the most part, I feel like the authors are clear in how they will measure empathy in the proposed study, but I think the distinction between constructs and how they are, or are not, related to one another could be a little more crisp in the Introduction.

Reviewer: 2

Comments to the Author  
Please see the attached file

**General comment:**

The authors propose an interesting study pertaining the investigation of a causal relationship between empathy and autobiographic memory (AM from now on). Importantly, the project proposal heavily draws on a recent unpublished study where the same task and procedure was applied and evidence supporting the involvement of AM in empathy. Thus, some of my specific comments do apply to the previous study as well for they are very much related to the theoretical and methodological foundations. Overall, I feel there is lack of clarity and detail in the justification of rationale and design.

**Background and rationale**

1. The authors define empathy as “the ability to mirror and to explicitly infer others’ inner states [...]”. However, they do not provide any reference nor draw distinctions between different ideas about empathy or related constructs. So where is this definition coming from? While I appreciate the advantage of keeping it concise and essential, I find this limited information as a cue of naivety? Maybe not, as I am confident the authors know well the theoretical and methodological landscape. For example, the introduction leaves several questions open: is empathy an automatism? Is it effortful? Is it a monolithic construct or a constellation of subfunctions? Is there agreement on its definition and across the empirical findings? In sum, I think a lengthier elaboration on the definition of empathy is perhaps worth as well as a definition of AM for the readers to understand better the rationale for their interaction.

2. It is noteworthy that the findings mentioned at line 61 highlight empathy as an explicit process based on metacognition and self-reflection rather than an automatic implicit process. The same argument stands for AM. This is not a trivial aspect of the topic as it may undermine rationale and limit interpretation (depending on the specific questions).

3. I don't think there is enough rationale to legitimate the relevance of the experimental questions. For example, if the authors posit a bidirectional link between empathy and AM (or more in general between empathy and memory) what do we know about the impact of empathy deficit on memory? The authors argued that there is conflicting evidence as amnesic patients do not show relevant empathy impairment. Likewise, do psychopaths holding clear affective empathy deficits show an AM

(or even just memory) deficit? This question follows up on my previous observation on the author's simplistic definition of empathy. Psychopaths seem to have remarkable theory of mind but lack "affect sharing" or otherwise named empathic concern. Some authors believe the latter concepts being the actual empathy core construct (which does not require a mirror neuron system to be conceived and be neuronally defined).

4. My feeling is that the Hypotheses 3 and 4 are propaedeutic and necessary for testing 1 and 2. With this in mind, there is a series of assumptions made by the authors and implicitly unaddressed in the intro. I will summarize them:

☒ There is an assumption (apparently supported by the previous unpublished study) that participants may "feel" empathy for stranger's faces associated with a sensory event previously experienced by the participant. However, how can we tell that this artifactual association really triggers affect sharing in the participants instead of a mere response bias based on logical compliance/social desirability? (this criticism applies to the study in preprint too).

☒ We have not enough evidence to support a specific role of precuneus, compared with another structure involved in empathy, and the rationale alone does not seem to be strong enough. So, support evidence is required (and the pilot data do not help with that; see later comments).

☒ There is assumption that theta burst TMS (and offline stimulation) is the most adequate stimulation pattern for this investigation but there is no supporting information nor literature reference.

☒ The same reasoning applies to one of the DVs, the P3 wave. This is a classical and widely investigated cross-modal deflection that has little or no sensory specificity with large multisource contribution and task-dependency. In other words, is a very gross unspecific brain index which provides a cortical arousal index that (at best) can be harnessed in specific designs. This choice begs the question why this wave should be preferred over time-frequency indexes, for example.

#### **Methods and statistical design**

5. The design presents an impressive number of factors. The authors seem to calculate power based on a 3-way interaction vs. a 4-way here. Also, the design is not fully spelled out in the cited preprint nor the power analysis. Thus, they for example, the authors say nothing here about the assumptions (e.g. non-independence). Clearly within subject correlation can dramatically affect power calculation.

6. Again, at line 193, are the authors investigating a main effect? My understanding is that the different hypotheses would test interactions besides main effect. Should we then accept the power assessment as valid for the interactions too?

☒ Again, at line 199, it is not clear whether the calculation is including 1, 2 or 3 factors. Ideally a simplified

version of the effect size calculation could entail only

mean AMpainful&painful - mean non-AMneutral&neutral / SD non-AMneutral&neutral

7. At line 206 I finally encounter one of the important details. And the calculation is made for a t-test. However, the authors present a **2\*2\*2\*2 design** which does not entail a t-test as first analysis but rather either an anova or a mixed model (because the interaction is crucial).

8. Further, I am afraid I would tend to disagree that the main hypothesis is Mem\*stim site. In my understanding the main Hyp is **mem\*emo\*face\*stim site** and even if emo and face are collapsed in one dummy variable then the authors would end up with **mem\*cue\*stim site**. This is a core aspect that needs to be addressed too. The authors may calculate indexes to reduce the number of variables to be submitted to statistical analyses, but this should be clearly specified upfront.

9. Because of the decision to change the stimulation protocol, I think the pilot data results are little

informative. Even if that was the case I am not convinced that a  $2 \times 2 \times 2 \times 2$  design with ? trials per cell (this has not been specified yet either) can provide robust statistical information. Please note that the pilot experiment not only targets a different area but also exploits a different stimulation protocol...

10. A minor consideration: because the tests are run in two different days the authors may want to consider a normalization of the data (e.g. z scores).

11. I take the 3 memory scenarios will be repeated across the experiment thus reducing the salience of this information across the experiment. Do the authors have rationale/evidence that this amount of AM is enough to trigger the desired effects or is a further assumption? Same reasoning applies to the number of faces chosen and their selection method.

12. I see a potential problem with the empathy task. If participants are asked to rate empathy and after that judge the face, a logical link is established thereby if they rate high in empathy then they won't be rating a face as non-painful. Hence, I wonder whether this approach may not actually be investigating theory of mind/perspective taking rather than affective sharing. This comment taps on the simplistic interpretation of empathy that permeate many studies despite several warnings and calls for theoretical/methodological precision in the previous years. In addition, the same sequitur applies to the reverse scenario (the face first and then the sentence), which is not happening by the way. Hence there must be another assumption about no order effects as well.

13. There is no indication of pauses between blocks. Considering how taxing the experiment can be for the participant I would expect at least some pauses.

14. It is not clear if and what hypotheses have been developed for the additional self-report measures, particularly the empathy questionnaires. It is very important that a distinction is drawn from exploratory and a priori investigation, and above all that the rationale and hypotheses cover the entire study.

#### **Data analysis**

15. I would provide detail on how this blinding will be achieved.

16. In this section there is no indication about if and what effect size will be reported.

17. Also, there is no indication of outliers handling as per checklist. Equally, no mention of distribution assessment nor correction of family-wise error (if deemed necessary).

18. No hypothesis has been established for RTs so perhaps no worth mentioning unless these data are thought to be used for exploratory analyses?

19. Th ANOVA approach is exactly reflecting my concerns. I would recommend using mixed models in conjunction with the already planned Bayesian approach, together with a simplification of the design and/or a reduction of factors in the analyses.

20. ERPs analysis: there is not enough detail about this analysis (e.g. number of permutations, type of clusterization approach, p val., etc.).

21. I frankly do not understand the need for source analysis for a study based on time-domain measures only and even more only focused on the P3 (which has got multiple sources too). In my opinion this is unnecessary especially because it does not seem to be a priori driven but another exploratory component of the study. I would change my mind if the source analysis was tightly linked to the hypotheses. But even in that case a Laplacian transform would be already appropriate to and be added to avg-ref analyses.

22. I am glad the authors spent a word of caution on the fact that the offline TMS will have non-specific effects. This taps on my previous comment about the rationale of using theta burst and offline TMS in the first place.

#### **Authors' Response**

---

Birmingham, April 12<sup>th</sup>, 2021

Dear Prof. Prager,

I am presently submitting a revision of the registered report entitled: “Remember as we empathize. Do brain mechanisms engaged in autobiographical memory retrieval causally affect empathy awareness? A combined TMS and EEG registered report.” (jnr-2020-Oct-9112), authored by myself, John Hodsoll, Zheni Goranova, Giulio Degano, Nicolò Di Lello, Carlo Miniussi, Alessio Avenanti, and Carmel Mevorach.

We are indebted to you, the statistical editor, and the reviewers for the attention devoted to our work. We thank you for giving us the chance to respond to the comments and invitation to submit a revision

of our paper. We took each comment seriously and addressed them carefully by responding to them in the response letter and changing the manuscript's respective sections. As a result, our manuscript is much improved and now makes an even more compelling case for investigating the causal role of autobiographical memory retrieval in empathy.

We would like to highlight the addition of Prof. John Hodsoll as an author of our manuscript. Prof. Hodsoll provided the statistical consultancy we were recommended to have. He critically contributed to the

statistical approach adopted in this study. The research will significantly benefit from his collaboration.

In closing, let me thank you again for your interest in our work, and for your patience in going through the letter and revised manuscript. I hope you, the statistical editor, and the reviewers will now find

the paper much improved and acceptable for publication in Journal of Neuroscience Research.

Sincerely,

Federica Meconi

**Associate Editor: Prager, Eric**

Comments to the Author:

Thank you for your submission and again I apologize for the delay. The reviewers found the proposal to be

of potential interest, but they did note some major concerns. I believe, based on my review, that many of these comments can be discussed and detailed in the text. One of the biggest concerns raised by the reviewers and statistical editor was the power analysis and statistical approaches. I'd recommend that you have a biostatistician consult with you on this project. If you are able to revise the manuscript to the satisfaction of the reviewers and especially the statistical editor, we would be happy to consider the manuscript again.

**Answer.** We thank the Editor for giving us the opportunity to resubmit our work. We replied to every comment raised by the reviewers and the statistical editor and added to our team work a biostatistician, Prof. John Hodsoll, as suggested. Prof. Hodsoll offered his consultancy and advised us how to improve our statistical approach. He is now listed in our author list.

**Statistics Editor: McArthur, David**

Comments to the Author:

Every one of the comments made by the other reviewers will need close attention. In particular, from a statistical perspective, computing statistical power on basis of t-test results for a complex 4-way repeatedmeasures

design can hardly be endorsed. Using pilot data results on the basis of merely 5 subjects who underwent a different protocol is very risky, not to mention that the largest interaction observed there was not statistically significant, and effect size calculations based on very small sample sizes is hardly stable. The issue of how many factors are in play in any given analysis must be fully clarified before proceeding, along with a clear rationale for dropping factors if indeed that is the plan. Presently the four numbered hypotheses are not independent but inherently chained together. Overall this study must be rethought and constructed in sequential stages in order for the findings to make sense:

**Answer.** We have taken on board these important comments especially with respect to the simplification of the design and related power analysis. We modified the Hypotheses section. We

provide replies for the three specific comments below.

**Statistics Editor:** first is to separate participants into a meaningful empathic spectrum, as there is no question but that different people even with very similar backgrounds have widely differing empathic capabilities;

**Answer.** While individual difference certainly exist in empathetic capacity, the question we are addressing here is regarding a general mechanism underlying empathy/memory interaction. We do not have a reason to predict that such a mechanism only applies to high empathizers. Indeed, in our previous studies we found a robust within-subject behavioural effect across all participants. Consequently, our intention here is to establish whether a causal relationship between autobiographical memory retrieval and empathy judgements exists in a sample of participants with normal empathy abilities. We will therefore assess participants' trait empathy using the Empathy Quotient questionnaire (Baron-Cohen & Wheelwright, 2004), which is the gold standard for measuring dispositional empathy abilities. In addition, instead of separating participants based on their empathy scores, we have also adjusted our plans slightly and simplified the task in order to reduce the variability of the empathic processes that participants would rely on for performing the task. We clarify this in points 1 and 6 raised by reviewer 1 and point 1 raised by reviewer 2.

**Statistics Editor:** second is to get consistent and repeatable results from what is undoubtedly an exhausting and challenging behavioral task with all the fine details such as time pressures and intertrial intervals carefully optimized in advance;

**Answer.** It is important to note that we already have extensive experience with the task and its delivery using a variety of methodologies (e.g., EEG, fMRI), including the use of time-pressured responses. We have also piloted it recently (pre-pandemic) with older adults and clinical populations. Across all these implementations and cohorts we recorded consistent and robust behavioural and EEG/fMRI signals. In addition, we had no evidence for stress or fatigue induced by the tasks across all these previous studies. In the present context, the session will be split into two parts. The first experimental session will be no longer than 30 minutes (split into blocks of trials interleaved by self-paced breaks) and the second session will be no longer than 90 minutes (again interleaved by several short self-paced breaks between blocks and one long break between tasks). This represents quite a standard length for EEG or TMS studies reported in the literature. Thus, although the specific experimental setting is new, it has been carefully designed and optimized based on extensive previous experiences and piloting with similar tasks and methodologies. We can confidently say that this is a setting that can be handled well by healthy young adults.

**Statistics Editor:** third is to obtain findings that precuneus stimulation is truly different from control and to ensure that the choice of control is the correct one. At present it reads like every consideration is in flux simultaneously and no amount of statistical wizardry will be sufficient to untangle the findings satisfactorily.

**Answer.** For this precise reason, we decided to implement a stimulation protocol that has already been used successfully to investigate the role of the precuneus in autobiographical memory retrieval. In the present context we adopt the exact same stimulation protocol (including control area) as in this previous study.

#### **Reviewer: 1**

Comments to the Author

The proposed study aims investigate whether disrupting activation of a core brain area of AM retrieval, the

precuneus, interferes with an effect of memory recall on empathy. Overall, this study is positioned to provide a novel contribution to work on the role of autobiographical memory in empathy. Although there some existing work that hints at a neural relationship, much of this work is correlational or, as is the case with the amnesic patient studies, mostly underpowered with mixed results. As such, this work could make



a valuable contribution with the potential to provide causal evidence for the role neural regions supporting autobiographical memory influencing how people feel for and connect with others.

While there is much to like about the proposed study, the manuscript/study may benefit from (i) further unpacking the nuances of a proposed relationship between the components of autobiographical memory the

precuneus supports and how these specific components relate to empathy, (ii) better explain and justify behavioral design choices, and (iii) connecting with related work—though this last point is more minor than

the first two.

#### COMMENTS

1) **Reviewer Explanation for Including the Face Manipulation of Pain vs. No Pain.** The rationale for including painful vs. neutral faces in the study could be made a little more clear. I could imagine different reasons for needing to include painful and neutral faces (e.g., it might be that photos are needed to make the materials more realistic and avoid floor effects in empathy ratings, or you might have the prediction that congruency in the emotion of the stories and faces is needed to see other effects of interest, etc.) but you should make the reasoning for the inclusion of the photos in your design explicit as well as state any predictions you have with regard to this factor (e.g., if you do indeed predict some sort of role of congruency between stories and faces and how this interacts with memory). This added manipulation seems to offer greater complexity in analysis without necessarily adding greater substance to the claims that the authors make in the paper. That said, a clearer motivation could mollify this concern.

1) **Answer.** The main reason underlying the inclusion of the face manipulation is replication of previous studies that have used this particular variant of the pain decision task. Manipulation of facial expressions tends to be a good index of successful elicitation of an affective empathic reaction. Importantly, this empathy task has been used in several experimental settings and has successfully shown cognitive empathy effects beyond the affective empathy effect usually observed for the face manipulation. However, following the reviewer's point (as well as the statistical editor comments) we decided to reconsider this component of the task as in the present context we do not intend to show dissociation between affective and cognitive empathy and we have no prediction of interest regarding the effect (or interaction) for perturbation of the precuneus and the manipulation of the face. We will therefore propose to use only neutral faces.

2) **Explanation for Including the Polygon Cues Similar to faces,** the inclusion of polygon cues is also a little obscure. I have seen these used in other work in place of participants re-reading stories or cue words (e.g., having participants memorize associations between the stories and random polygons, and then testing memory retrieval in response just to the polygons), but being that you're already presenting cue words, what purpose do these polygons serve to improve the design and claims the study aims to evaluate?

2) **Answer:** The purpose of the retrieval task is for the participants to form a mental representation of each specific context in order to extract their neural fingerprints (which the classifier is trained to identify). The classifier is then tested on the main task to assess its ability to identify the same neuronal signal. As such, it is important that the retrieval task and the main task share the same memory retrieval aspect but not the same physical input. If a similar input is used across task we will not be able to ascertain the signal we classify represents the memory retrieval (and not the physical input).

3) **Sampling plan and analysis.** The authors nicely articulate and justify their sampling plan, including powering analyses for specific effects. In your power analysis, you power for the focal Memory x Stimulation site interaction, but your design includes four IVs. What do the author's plan on doing regarding higher-order interactions? With so many variables in play, the study has the potential to examine three-way and four-way interactions. However, if the study finds nonsignificant higher-order

(e.g., three-way or four-way) interactions it would be difficult to tell whether this was because of the lack of a true interaction effect or because the study was underpowered to detect the higher-order interaction. It is thus possible that you would find the two-way interaction of interest but that this interaction would be partially qualified by higher-order interactions. Could you speak to this possibility and state why (or why not) this is a concern for the proposed study?

**3) Answer.** The reviewer raises a critical point regarding the complexity of the task and the analysis we plan. Following the previous comment regarding the face manipulation parameter (and in accordance with other points mentioned by the editor and reviewers) we will simplify our design by removing the face manipulation factor. Furthermore, in all the hypotheses we stated, our main expected outcomes concern the comparison between AM and non-AM as the research question under investigation is the causal role of AM retrieval to empathy judgements, independently of the specific emotional content of the memory. Therefore, we decided to only compare painful AM with painful non-AM contexts, resulting in a 2 (Stimulation site: Precuneus vs. Cz) x 2 (Type of memory: AM vs. non-AM).

**4) a. Relevant Neural and Behavioral Findings on Memory and Empathy.** There are a few recent papers that are particularly relevant to this work and would bolster the background in the Introduction. Wagner and colleagues (2020, *Social Cognitive and Affective Neuroscience*) found similar patterns of activation for first-hand pain and empathy for another's pain in the hippocampus, temporoparietal junction, retrosplenial cortex, and anterior insula. Notably, connectivity between the hippocampus and the fusiform gyrus during the pain empathy task scaled with increased self-reported perspective-taking skills, often considered a key component in empathy. There is also relevant work from Vollberg and colleagues (2019, preprint: 10.31234/osf.io/r6wmx) which found that participants who completed an episodic specificity induction later reported greater empathy for others experiencing negative events that were unrelated to the ESI, suggesting that increasing spontaneously generated episodic detail while processing other's suffering influences feelings of empathy.

**4) Answer.** We thank the reviewer for highlighting these relevant recent papers. We now incorporated them in the introduction. We report below the revised paragraph from the introduction in green.

Lines 55–56:

“Similarly, empathy for others’ suffering can be influenced by spontaneous retrieval of memories details (Vollberg et al., 2021).”

[...]

Lines 91–94:

“Wagner and colleagues further supported this view by showing similar pattern of activation in these brain areas, in the hippocampus and in the anterior insula for pain perceived in first-hand and empathy for physical pain (Wagner et al., 2020). Furthermore, hippocampal-neocortical coupling during empathy for pain was larger for higher self-perceived cognitive empathy skills.”

**b.** The authors note inconsistent findings from patient populations with deficits in autobiographical memory (e.g., Rosenbaum et al., 2007; Beadle et al., 2013). I think it is worth drawing attention to readers that these studies are not only mixed but that they are severely underpowered with only a few patients in each study (2 in Rosenbaum, 3 in Beadle). Relatedly, a more recently published, and higherpowered

study by Sawczak et al. (*Neuropsychologia*, 2019) showed that patients with medial temporal lobe epilepsy do not show an increase in empathy for others suffering (and willingness to help) after being prompted to construct specific episodes compared to healthy controls. This study did not examine spontaneous reactivation nor did they localize the anatomy underlying this difference with any level of precision, however, the results align with the current manuscript and should be discussed. It may also be worth citing Dermody et al. (2016, *Journal of Alzheimer’s disease*) in your discussion of existing work on patients with memory impairments and empathy. Dermody and colleagues studied patients with

Alzheimer's Disease and fronto-temporal dementia and find that both patient groups have deficits in cognitive empathy, while only the fronto-temporal impairments in affective empathy.

**b. Answer.** We have now revised the manuscript mentioning the study by Sawczak and colleagues (2019), which is indeed of relevance for our argument.

Lines 100–101:

“Sawczak and colleagues showed that patients with medial temporal lobe damage fail to show empathy increases if prompted to build specific representations of others' suffering (Sawczak et al., 2019).”

However, clinical populations that show conditions that co-occur with the memory deficit contributed to offer the puzzled scenario concerning memory deficits and empathy impoverishment. We believe that in these cases, if empathy impoverishment is observed, this cannot be directly attributed to the memory deficit because of the co-occurring deficits that could further contribute to it. We did not mention any of those studies, as we wanted to highlight the pure role of memory deficit in empathy. Therefore, we chose not to mention the study by Demordy and colleagues (2016) within this particular step of our reasoning. We do expand on this study in our response to the reviewer's next point, which we believe is more suitable for the Discussion section.

**5) Elaborating on the specific components of autobiographical memory supported by precuneus and how those components relate to empathy.** There are many aspects of the proposed studies design and interpretation that seem astutely thought out and articulated (e.g., distinguishing the selectivity of stimulating the precuneus on the empathy vs. perceptual task; the author's openness to the possibility that because the precuneus is involved in explicit mentalizing, disrupting it could lead to an overall reduction in empathy ratings). And, the author's do a nice job of motivating the overall importance of a studying the relationship between memory and empathy on the one hand, and linking the precuneus to autobiographical memory on the other hand.

“Compelling evidence supports the role of precuneus in both AM and specific empathic processes, i.e., explicit reasoning about others. Hebscher and colleagues (Hebscher et al., 2020) established in a transcranial magnetic stimulation (TMS) and magnetoencephalography study that the precuneus plays a causal role in AM retrieval. The authors nicely showed in an explicit recollection task that decreasing precuneus activation interferes with neural dynamics of early stages of AM retrieval and the vividness of recollection, see also (St. Jacques et al., 2017).” (p. 4)

However, I want to slightly push for a more detailed theoretical account of the mechanisms involved. I think it would be beneficial to further unpack the details and specific processes/components of autobiographical memory associated with precuneus in St. Jacques et al. (2017), and particularly Hebscher et al. (2020), and how these components relate to or are distinct from previous working connecting components of memory with empathy such as in Vollberg et al., 2019, Sawczak et al., 2019, or Beadle et al., 2013.

**5) Answer.** The point that the reviewer raises here is very insightful and pushes the reasoning a bit forward, indeed. Targeting the precuneus would bring to at least two main advances in our knowledge of the contribution of memory to empathy. First, it would determine a cortical hub of the AM retrieval process involved in empathy. It would pose intriguing and unexplored research questions regarding the aspect of the self-referential processes, which are involved both in AM and in empathy. Such a result has the potential to contribute into explaining the mixed evidence regarding the empathy impoverishment after memory deficit. Precuneus is the parietal hub of the fronto-parietal network that supports the building of the sense of self. Precuneus underlies the retrieval of self-related information in long-term memories recall (Cavanna & Trimble, 2006; Mazzoni et al., 2019; Sajonz et al., 2010). It allows perspective shifting in autobiographical memory retrieval (Faul et al., 2020; St. Jacques et al., 2017), and is involved in making judgements on the self and the other (Lou et al., 2004; Murray et al., 2015; Northoff et al., 2006; Ochsner et al., 2004). Showing that interfering with the precuneus

activation when AM are retrieved in the service of empathy could suggest that the mixed results observed in clinical conditions with memory deficits depend on the extent to which precuneus can process self-related information. Although speculative at this stage, such a suggestion would find empirical justification on the differences between cognitive and affective empathy observed in patients with Alzheimer's disease (AD) and with fronto-temporal dementia (FTD) that the reviewer mentioned in their point #4b (Dermody et al., 2016). AD and FTD patients show a double dissociation regarding the onset of memory and social/empathy deficits (Bertoux et al., 2016; Zale & Dickerson, 2017). In patients with AD the first and most prominent symptom is the memory deficit whereas with prominent initial hippocampal atrophy; the first and most prominent symptom in FTD patients is a change of personality and severely impaired social skills with prominent frontal and temporal atrophy. On the one hand, cognitive empathy entails the ability to explicitly infer the others' reasoning, which is something that can be done or compensated by general semantic knowledge or episodic (non-autobiographical) memory. AD patients would show episodic memory deficits (Demichelis et al., 2020; Dermody et al., 2016; Fischer et al., 2019). On the other hand, affective empathy entails the ability to resonate the other's experiences in the inner self. To this aim, the empathizer requires the ability to keep a clear distinction between the self and the other. This is an ability that is impaired in FTD because of reduced interoception and general change of sense of self (Goodkind et al., 2015; Pasquini et al., 2020; Sturm et al., 2018).

We thank the reviewer for making this point - the above considerations will be incorporated in the discussion section.

**6) Empathy and related constructs** "Mechanisms of inner simulation that are implemented by the human mirror neuron system (Molenberghs et al., 2012) and anterior and mid-cingulate cortex (Danziger et al., 2009) underpin the empathic feelings. Perspective shifting and mindreading that are implemented by a circuit of fronto-parietal and temporal areas, which includes the precuneus (Lamm et al., 2011; Schurz et al., 2014; Molenberghs et al., 2016), underlie the ability to explicitly reason about others." (p. 3)

"Compelling evidence supports the role of precuneus in both AM and specific empathic processes i.e., explicit reasoning about others." (p. 4)

The way empathy and empathic processes are sometimes discussed seems to blur the distinction between the emotional experience of empathy, empathic/prosocial concern, and the explicit mentalizing (perspective-taking or theory of mind), which are often associated with feelings of empathy, but not always. For the most part, I feel like the authors are clear in how they will measure empathy in the proposed study, but I think the distinction between constructs and how they are, or are not, related to one another could be a little more crisp in the Introduction.

**6) Answer.** We agree with the reviewer's comment and in line with their and reviewer #2's opinion we revised the Introduction section clarifying the distinction between the two mechanisms and (as well as which one we refer to in the present study). We revised section appears below in green.

Lines 58 – 75:

"Empathy, which is the ability at the basis of social dynamics, has a multicomponent nature. It entails affective sharing as well as cognitive reasoning on others' inner states (cognitive empathy). These empathy components are dissociable in time-course and function, and a recent meta-analysis confirmed their anatomical dissociation (Lamm et al., 2011; Molenberghs et al., 2012, 2016).

Electroencephalography

(EEG) studies have revealed that affective sharing modulates event-related potentials (ERPs) in early timewindows:

within 250 ms from stimulus onset. Cognitive reasoning mainly modulates the P300 component, which is an electrophysiological index of motivated attention (Fan & Han, 2008; Hajcak et al., 2009; Hajcak

& Foti, 2020; Magliero et al., 1984; Meconi et al., 2018; Nieuwenhuis et al., 2005; Palmieri et al., 2021; Sessa, Meconi, Castelli, et al., 2014; Sessa, Meconi, & Han, 2014). Affective sharing (or affective empathy)

allows embodiment of others' inner states exploiting simulation mechanisms so that others' states are vicariously experienced in the self. Multiple functional magnetic resonance imaging (fMRI) studies investigating empathy for others' physical pain have repeatedly shown that this is grounded in the ability of the self to feel first-hand pain (Rütgen et al., 2015, 2020). Cognitive reasoning allows to build an accurate

representation of the others' inner states. Several studies have demonstrated that healthy adults infer others'

inner states by either relying on acquired general semantic knowledge (Pehrs et al., 2017) or on their own past experience (Mitchell et al., 2006; Gaesser et al., 2018; Gaesser, 2020)."

Throughout the manuscript, we disclosed whether we refer to affective sharing (or affective empathy) or to cognitive reasoning (or cognitive empathy). This is reported in green in the main text to facilitate inspection.

**Reviewer: 2**

**General comment:**

The authors propose an interesting study pertaining the investigation of a causal relationship between empathy and autobiographic memory (AM from now on). Importantly, the project proposal heavily draws on a recent unpublished study where the same task and procedure was applied and evidence supporting the

involvement of AM in empathy. Thus, some of my specific comments do apply to the previous study as well for they are very much related to the theoretical and methodological foundations. Overall, I feel there is lack of clarity and detail in the justification of rationale and design.

**Background and rationale**

1) The authors define empathy as "the ability to mirror and to explicitly infer others' inner states [...]". However, they do not provide any reference nor draw distinctions between different ideas about empathy or related constructs. So where is this definition coming from? While I appreciate the advantage of keeping it concise and essential, I find this limited information as a cue of naivety? Maybe not, as I am confident the authors know well the theoretical and methodological landscape. For example, the introduction leaves several questions open: is empathy an automatism? Is it effortful? Is it a monolithic construct or a constellation of subfunctions? Is there agreement on its definition and across the empirical findings? In sum, I think a lengthier elaboration on the definition of empathy is perhaps worth as well as a definition of AM for the readers to understand better the rationale for their interaction.

1) **Answer.** Although we did present a distinction of affective sharing and cognitive empathy mechanisms (lines 107–114 in the resubmitted version of the manuscript), our aim was to present a straightforward line of reasoning as in whether interfering with memory retrieval would impoverish empathy responses. However, we take on board the reviewer's point that this choice was not helpful to our purposes and that finer-grained description of multicomponent aspect of empathy was not clear enough. We therefore thoroughly revised the presentation of the construct throughout the manuscript. We report the revised section below in green.

Lines 58 – 75:

"Empathy, which is the ability at the basis of social dynamics, has a multicomponent nature. It entails affective sharing as well as cognitive reasoning on others' inner states (cognitive empathy). These empathy components are dissociable in time-course and function, and a recent meta-analysis confirmed their anatomical dissociation (Lamm et al., 2011; Molenberghs et al., 2012, 2016).

Electroencephalography

(EEG) studies have revealed that affective sharing modulates event-related potentials (ERPs) in early timewindows:

within 250 ms from stimulus onset. Cognitive reasoning mainly modulates the P300 component, which is an electrophysiological index of motivated attention (Fan & Han, 2008; Hajcak et al., 2009; Hajcak

& Foti, 2020; Magliero et al., 1984; Meconi et al., 2018; Nieuwenhuis et al., 2005; Palmieri et al., 2021; Sessa, Meconi, Castelli, et al., 2014; Sessa, Meconi, & Han, 2014). Affective sharing (or affective empathy)

allows embodiment of others' inner states exploiting simulation mechanisms so that others' states are vicariously experienced in the self. Multiple functional magnetic resonance imaging (fMRI) studies investigating empathy for others' physical pain have repeatedly shown that this is grounded in the ability of the self to feel first-hand pain (Rütgen et al., 2015, 2020). Cognitive reasoning allows to build an accurate

representation of the others' inner states. Several studies have demonstrated that healthy adults infer others'

inner states by either relying on acquired general semantic knowledge (Pehrs et al., 2017) or on their own past experience (Mitchell et al., 2006; Gaesser et al., 2018; Gaesser, 2020)."

Throughout the manuscript, we disclosed whether we refer to affective sharing (or affective empathy) or to cognitive reasoning (or cognitive empathy). These changes appear in green in the main text.

2) It is noteworthy that the findings mentioned at line 61 highlight empathy as an explicit process based on metacognition and self-reflection rather than an automatic implicit process. The same argument stands for AM. This is not a trivial aspect of the topic as it may undermine rationale and limit interpretation (depending on the specific questions).

2) **Answer.** By refining the definitions of empathy and AM in the preceding point we believe this doubt is mitigated.

3) I don't think there is enough rationale to legitimate the relevance of the experimental questions. For example, if the authors posit a bidirectional link between empathy and AM (or more in general between empathy and memory) what do we know about the impact of empathy deficit on memory? The authors argued that there is conflicting evidence as amnesic patients do not show relevant empathy impairment. Likewise, do psychopaths holding clear affective empathy deficits show an AM (or even just memory) deficit? This question follows up on my previous observation on the author's simplistic definition of empathy. Psychopaths seem to have remarkable theory of mind but lack "affect sharing" or otherwise named empathic concern. Some authors believe the latter concepts being the actual empathy core construct (which does not require a mirror neuron system to be conceived and be neuronally defined).

3. **Answer.** The point the reviewer raises about the possible bidirectional link between empathy and memory is certainly intriguing but outside the scope of the present study. We have presented evidence from a series of studies on healthy and clinical populations supporting the logic step of reasoning regarding the role of memory in the service of empathy. This was highlighted by reviewer #1 in their general comment on the manuscript; we refer to this to further support our research question as timely and reasonable. With the proposed study, we are not excluding the possibility that the interplay between these processes could be in the opposite direction as the one we hypothesized, under different conditions. However, we are specifically testing one of these possible directions. We justified our research question by discussing studies on healthy subjects showing that memory is reactivated in empathy; that referring to autobiographical past experiences increases explicit reports of self-perceived empathy and prosocial attitude and behaviour in similar situations. We described the puzzled scenario offered by the studies on clinical conditions of memory loss. Building on this evidence, we address whether there is a causal role between memory retrieval (impaired in memory loss conditions) when participants are

requested to report explicit judgements of self-perceived empathy.

**4.** My feeling is that the Hypotheses 3 and 4 are propaedeutic and necessary for testing 1 and 2. With this in mind, there is a series of assumptions made by the authors and implicitly unaddressed in the intro. I will summarize them:

**a.** There is an assumption (apparently supported by the previous unpublished study) that participants may "feel" empathy for stranger's faces associated with a sensory event previously experienced by the participant. However, how can we tell that this artifactual association really triggers affect sharing in the participants instead of a mere response bias based on logical compliance/social desirability? (this criticism applies to the study in preprint too).

**a. Answer.** The reviewer raises a reasonable point. However, this would apply to every published work studying empathy. Requesting participants to judge their empathy, based on experimental instructions, is unavoidable in investigating this construct.

**b.** We have not enough evidence to support a specific role of precuneus, compared with another structure involved in empathy, and the rationale alone does not seem to be strong enough. So, support evidence is required (and the pilot data do not help with that; see later comments).

**b. Answer.** We are slightly puzzled by this comment - we are not targeting the precuneus for its role in empathy but rather for its role in autobiographical memory retrieval. Multiple studies, which we have mentioned in the Introduction section, support this function of the precuneus and we have specifically stated that no studies, to our knowledge, have shown that perturbation of the activation of the precuneus reduces empathy (see the section below from the introduction). The reviewer might want to expand on their point about the weakness of our rationale so that we can consider this argument in further detail.

Lines 123 – 130:

“Compelling evidence supports the role of precuneus in both AM and specific empathic processes, i.e., explicit reasoning about others. Hebscher and colleagues (Hebscher et al., 2020) established in a transcranial magnetic stimulation (TMS) and magnetoencephalography study that the precuneus plays a causal role in AM retrieval. The authors nicely showed in an explicit recollection task that decreasing precuneus activation interferes with neural dynamics of early stages of AM retrieval and the vividness of recollection, see also (St. Jacques et al., 2017). In contrast, to our knowledge, no study has shown that modulation of precuneus’ activation directly reduces any measure of empathic abilities. This therefore enables the proposition that interrupting the precuneus will affect AM retrieval, which in turn will affect cognitive empathy.”

**c.** There is assumption that theta burst TMS (and offline stimulation) is the most adequate stimulation pattern for this investigation but there is no supporting information nor literature reference.

**c. Answer.** We disagree with the reviewer’s point. We explained the reasoning underlying the choice of offline theta burst TMS. We are certainly not assuming that this is the most adequate stimulation pattern for the present investigation. The justification behind this experimental choice is grounded in the recent study by Hebscher and colleagues (2020) who have demonstrated reduced AM retrieval following precuneus stimulation using this exact same stimulation protocol. This is clearly stated in the Introduction section.

**d.** The same reasoning applies to one of the DVs, the P3 wave. This is a classical and widely investigated cross-modal deflection that has little or no sensory specificity with large multisource contribution and task-dependency. In other words, is a very gross unspecific brain index which provides a cortical arousal index that (at best) can be harnessed in specific designs. This choice begs the question why this wave should be preferred over time-frequency indexes, for example.

**d. Answer.** The P3 is an index of motivated attention that multiple previous works have robustly demonstrated is elicited in empathy tasks like the one proposed in the present study (Fan & Han, 2008; Meconi et al., 2018, 2019; Palmieri et al., 2021; Sessa, Meconi, Castelli, et al., 2014; Sessa, Meconi, & Han, 2014, 2014; Sheng & Han, 2012). It is therefore plausible to make predictions regarding this index in the context of our study. Timefrequency

analysis are not excluded and are part of the exploratory path of analysis. The reviewer might want to help us understanding better their point about the inadequacy of the P3 index. Since no support is provided for this strong claim, we cannot refute it, other than by providing the strong evidence for our claim.

### Methods and statistical design

**5.** The design presents an impressive number of factors. The authors seem to calculate power based on a 3-way interaction vs. a 4-way here. Also, the design is not fully spelled out in the cited preprint nor the power analysis. Thus, they for example, the authors say nothing here about the assumptions (e.g. non-independence). Clearly within subject correlation can dramatically affect power calculation.

**5. Answer.** We agree with the reviewer regarding the number of factors. After a careful reconsideration of this experimental choice based on both reviewers' comments we decided to reduce the number of factors to Stimulation and Type of Memory since our research question, as stated throughout all our hypothesis, is whether interfering with precuneus activity would reduce empathy rates for AM contexts. Correlations within measures was considered in the calculation of power and was reported in the manuscript in the power analysis paragraph.

**6.** Again, at line 193, are the authors investigating a main effect? My understanding is that the different hypotheses would test interactions besides main effect. Should we then accept the power assessment as valid for the interactions too? Again, at line 199, it is not clear whether the calculation is including 1, 2 or 3 factors. Ideally a simplified version of the effect size calculation could entail only mean AMpainful&painful - mean non-AMneutral&neutral/SD non-AMneutral&neutral

**6. Answer.** We were reporting the effect size of the main effect observed in previous studies. In the current study we expect to replicate that effect when the stimulation is delivered over the control site.

**7.** At line 206 I finally encounter one of the important details. And the calculation is made for a t-test. However, the authors present a **2\*2\*2\*2 design** which does not entail a t-test as first analysis but rather either an anova or a mixed model (because the interaction is crucial).

**7. Answer.** We took on board the reviewer's point about the complex design. We reported in point 5 that we simplified the design to a 2 x 2 design as our interest is about the role of AM in empathy regardless of the emotional content of the memory. Our power analysis is based on a t-test comparing the differential empathy rates for AM and non-AM conditions between stimulation sites.

**8.** Further, I am afraid I would tend to disagree that the main hypothesis is Mem\*stim site. In my understanding the main Hyp is **mem\*emo\*face\*stim site** and even if emo and face are collapsed in one dummy variable then the authors would end up with **mem\*cue\*stim site**. This is a core aspect that needs to be addressed too. The authors may calculate indexes to reduce the number of variables to be submitted to statistical analyses, but this should be clearly specified upfront.

**8. Answer.** We agreed with the reviewer and simplified our design, as specified in points 5 and 7. This is also specified in points 1 and 3 raised by reviewer #1.

**9.** Because of the decision to change the stimulation protocol, I think the pilot data results are little informative. Even if that was the case I am not convinced that a 2\*2\*2\*2 design with ? trials per cell (this has not been specified yet either) can provide robust statistical information. Please note that the pilot experiment not only targets a different area but also exploits a different stimulation protocol...

**9. Answer.** The reviewer raises a reasonable point. However, the experimental paradigm (apart from the stimulation protocol) was the same in the pilot. While we agree that the pilot study may not be significantly robust in its relevance, it does provide an estimation of effect sizes that can inform power analysis.

**10.** A minor consideration: because the tests are run in two different days the authors may want to consider a normalization of the data (e.g. z scores).



**10. Answer.** We agree with this point - we now mention normalization of the data in the manuscript.

Lines 453 – 454:

“Since participants will undergo two sessions in two different days, all the data will be normalized with z-scores.”

**11.** I take the 3 memory scenarios will be repeated across the experiment thus reducing the salience of this information across the experiment. Do the authors have rationale/evidence that this amount of AM is enough to trigger the desired effects or is a further assumption? Same reasoning applies to the number of faces chosen and their selection method.

**11. Answer.** In our previous study, under review, we used only one memory scenario per memory condition (1 painful AM, 1 painful non-AM, 1 neutral AM and 1 neutral non-AM) which was enough to observe relevant outcomes related to the AM processes. After removing the emotional content of the memory as a factor we will be able to have 6 different contexts per type of memory. The choice of dropping the face manipulation would allow us to double the number of identities we employ as neutral faces. By adopting these changes, the total number of trials we calculated for the task will not change.

**12.** I see a potential problem with the empathy task. If participants are asked to rate empathy and after that judge the face, a logical link is established thereby if they rate high in empathy then they won't be rating a face as non-painful. Hence, I wonder whether this approach may not actually be investigating theory of mind/perspective taking rather than affective sharing. This comment taps on the simplistic interpretation of

empathy that permeate many studies despite several warnings and calls for theoretical/methodological precision in the previous years. In addition, the same sequitur applies to the reverse scenario (the face first and then the sentence), which is not happening by the way. Hence there must be another assumption about

no order effects as well.

**12. Answer.** We have removed the face manipulation from the design and we revised the definition of the empathy construct we offer in the manuscript.

**13.** There is no indication of pauses between blocks. Considering how taxing the experiment can be for the participant I would expect at least some pauses.

**13. Answer.** We erroneously thought it would be clear that subdividing the trials in blocks would imply having breaks between the blocks. This is now explicitly mentioned in the relevant sections. The breaks between the blocks will be self-paced.

**14.** It is not clear if and what hypotheses have been developed for the additional self-report measures, particularly the empathy questionnaires. It is very important that a distinction is drawn from exploratory and a priori investigation, and above all that the rationale and hypotheses cover the entire study.

**14. Answer.** The empathy questionnaires are used to provide descriptive characteristics of the sample in terms of dispositional empathy and ability to describe emotions. The aim is to assess whether the sample is on average a sample of participants without abnormal empathy or ability to describe their own emotions. Participants scoring a clinical degree of alexithymia will be excluded from the sample. Correlational analysis with dispositional empathy is part of the exploratory analysis. This is now explicitly mentioned in the text.

Lines 366–371:

“Therefore, we will also assess the ability of participants to report their own emotions with the Toronto Alexithymia Scale, TAS-20 adults form (Bagby et al., 1994); participants that will report clinical degree of alexithymia will be excluded from the sample. Questionnaires will provide descriptive characteristics of the dispositional empathy and emotional resources of the sample. Correlational analysis between dispositional empathy and neural and behavioural data will be part of the exploratory path of analysis.”

## Data analysis

15. I would provide detail on how this blinding will be achieved.

15. **Answer.** This information is provided in lines 333 – 336:

“Blinding is involved in this study as participants will not be aware of which TMS session delivers pulses over the target or the control area and we will not disclose the reasons why we asked participants about their life episodes in the pre-screening phase until the whole study is completed.”

16. In this section there is no indication about if and what effect size will be reported.

16. **Answer.** This is now stated.

Line 440:

“Eta squared and standardised regression coefficients will be reported as effect sizes.”

17. Also, there is no indication of outliers handling as per checklist. Equally, no mention of distribution assessment nor correction of family-wise error (if deemed necessary).

17. **Answer.** We thank the reviewer for raising this point. We have now added this information in the Data Analysis section (please, see point 19).

18. No hypothesis has been established for RTs so perhaps no worth mentioning unless these data are thought to be used for exploratory analyses?

18. **Answer.** Our hypotheses are based on self-report measures. No specific hypotheses are set on RTs. We do not mention the RTs anymore in the main text.

19. The ANOVA approach is exactly reflecting my concerns. I would recommend using mixed models in conjunction with the already planned Bayesian approach, together with a simplification of the design and/or a reduction of factors in the analyses.

19. **Answer.** We took on board the reviewer’s suggestion and we now propose a linear mixed-effects model approach. The fixed effects will include Type of Memory (AM vs. non-AM) and Stimulation Site (Precuneus vs. Cz) and trials. The random structure will include participants and faces as random intercepts. This is now specified in the Data Analysis plan, which we report below for convenient inspection.

Lines 412 – 441:

*“Behaviour*

Empathy rating and perceptual scores for the empathy and perceptual tasks will be summarised with the overall mean and frequency of response per rating level according to experimental condition, along

with the number of missing responses per condition. The rating responses will be treated as continuous to facilitate using linear mixed models for the primary analysis.

For the primary analysis (Hypothesis 1) responses will be entered a linear mixed-effect model with fixed effects of type of memory (AM vs. non-AM), stimulation site (Precuneus vs. Cz) and a two-way interaction as a contrast between the size of the AM difference (AM vs non-AM) according to stimulation site. Trial number (as a continuous variable) will also be included as a fixed effect to allow for performance

to change over the course of the experiment (e.g., from practice effects). To account for the dependency between repeated measures, the model will include random effects for participants and faces. Random intercepts will be included for both participants and faces. Models will be estimated with REML and likelihood ratio tests will compare different random effects structures with random intercepts only being the minimal. Finally, to account for potential dependency across trials, the model will include marginal terms to model the auto-correlation between residuals, with responses to adjacent trials tending to be more similar than trials further away in the sequence. Model criticism will involve examining level 1 and 2 residuals for normality and the identification of influential observations using Cook’s distance.

In order to test the specificity of stimulation on empathy rates, a second model will extend the first to include perceptual ratings as an outcome and fixed effects of the type of Rate (Empathy vs. Perceptual) along with the type of Memory (AM vs. non-AM), stimulation site (Precuneus vs. Cz), and their

interactions. A 3-way interaction between rating type, memory type and stimulation site will not be sufficiently powered for a conclusive hypothesis test and so the focus of this assessment will be standardized

regression coefficients (ES) and 95% confidence intervals. For both analyses, these responses will be collected for each participant independently of the quality of their EEG data.

A third model will be formulated for the vividness rates collected in the retrieval task. The fixed effects will include the Type of Memory (AM vs. non-AM). The random effect structure will include participants and trials as random intercepts. Vividness rates and EEG data collected in the retrieval task will be part of an exploratory path of analysis.

Eta squared and standardised regression coefficients will be reported as effect sizes.”

Lines 455 – 459: *Univariate ERPs analysis*

“An alternative approach to analyse ERPs amplitudes will be the linear mixed-effects model approach. Mean amplitudes in the critical 0.3-0.6 s time-window will be calculated for each condition and be the dependent variable. The model will include the same fixed effects (Type of Memory, Stimulation Site and their interaction, trials) and the same random structure (participants, faces as random intercepts), as for the empathy rates.

**20.** ERPs analysis: there is not enough detail about this analysis (e.g. number of permutations, type of clusterization approach, p val., etc.).

**20. Answer.** We have added this information in the relevant section and we report it below for the reviewer’s convenience.

Lines 444 – 450:

“To this end, spatio-temporal cluster-based Montecarlo permutation tests as implemented in Fieldtrip will be performed over the whole scalp over a 1 sec time-window. Clusters will be defined with a minimum of 3 neighbour channels. The neighbourhood of channels will be determined with a triangulation method. The number of permutation will be set to 1000. ERPs will be time-locked to the onset of the face contrasting AM vs. non-AM contexts for both sites of stimulation. The accepted p-value will be set to .05.”

**21.** I frankly do not understand the need for source analysis for a study based on time-domain measures only and even more only focused on the P3 (which has got multiple sources too). In my opinion this is unnecessary especially because it does not seem to be a priori driven but another exploratory component of

the study. I would change my mind if the source analysis was tightly linked to the hypotheses. But even in that case a Laplacian transform would be already appropriate to and be added to avg-ref analyses.

**21. Answer.** We erroneously included the source analysis in the text as it is linked to an exploratory path of analysis. We intend to run a linear discriminant analysis (LDA) trained over a specific time-window of the retrieval task and test it over a specific time-window of the empathy task. Since we intend to run the experiment in two sessions, we will run the classifier in source space. We have removed the reference to the source analysis in the current version.

**22.** I am glad the authors spent a word of caution on the fact that the offline TMS will have non-specific effects. This taps on my previous comment about the rationale of using theta burst and offline TMS in the first place.

**22. Answer.** The rationale of using offline theta burst TMS has been clarified in reviewer’s point #4c.

## Decision Letter

Dear Dr Federica Meconi,

Thank you for submitting your paper to the Journal of Neuroscience Research for consideration as a Registered Report (Stage 1). The manuscript has now been reviewed by external expert reviewers.

We are pleased to offer in-principle acceptance of your manuscript. You may now progress to Stage 2 and complete the study as approved. We would be grateful if you could now update the Editorial Office as to the anticipated completion date of your study.

Following completion of your study, we invite you to resubmit your paper for peer review as a Stage 2 Registered Report. Please be reminded that your manuscript can still be rejected for publication at Stage 2 if the Editors consider any of the following conditions to be met:

1. Whether the data are able to test the authors' proposed hypotheses by satisfying the approved outcome-neutral conditions (such as quality checks, positive controls)
2. Whether the Introduction, rationale and stated hypotheses are the same as the approved Stage 1 submission (required)
3. Whether the authors adhered precisely to the registered experimental procedures
4. Whether any unregistered post hoc analyses added by the authors are justified, methodologically sound, and informative
5. Whether the authors' conclusions are justified given the data

We encourage you to read the complete guidelines for authors concerning Stage 2 submissions at <https://onlinelibrary.wiley.com/page/journal/10974547/homepage/forauthors.html>. Please especially note the requirements for data sharing. Submit your Stage 2 as a new manuscript selecting "Registered Report (Stage 2)" as the manuscript type upon submission.

As per the journal's policy, we will now be processing your Stage 1 for publication.

Your Stage 1 article cannot be published until you have signed the appropriate license agreement. Within the next few days you will receive an email from Wiley's Author Services system which will ask you to log in and will present you with the appropriate license for completion. After completion of the appropriate license your Stage 1 article will be sent for copyediting and typesetting.

If you wish to make your article Open Access, you can purchase it via our Author Services platform when you sign your license. Since Registered Reports are published in two stages, please use this discount code to split the APC between both stages of publication: 50RRP.

Permissions: If you wish to use all or part of any previously published illustration, please obtain written permission from the publisher/copyright holder and provide a copy of the written permission.

Kind regards,

Dr Eric Prager  
Associate Editor, Journal of Neuroscience Research  
[eric.prager@cohenbio.org](mailto:eric.prager@cohenbio.org)

Dr. Junie Warrington  
Editor-in-Chief, Journal of Neuroscience Research

### REVIEWS:

Reviewer: 1

#### Comments to the Author

The author's provide thoughtful and thorough responses to reviewer comments, clarifying key issues and streamlining the methodological design and analytical approach.

I am satisfied the revision registered report. I think the study is now well positioned to provide a novel and interesting contribution to the field and endorse moving ahead with data collection.

Reviewer: 2

#### Comments to the Author

I think the work is now greatly improved and I have no further comments that would significantly further improve the proposal

## Authors' Response

3<sup>rd</sup> Editorial Decision

Decision Letter

---

Authors' Response

---

4<sup>th</sup> editorial decision

Decision Letter

Author response