**Pseudo-approaches lead to pseudo-explanations: reply to Corlett**

Sepehr Razavi1,2, Michael Moutoussis3, Peter Dayan4,5, Nichola Raihani6,7, Vaughan Bell8,9, Joseph M. Barnby2,10 \*

1 University of Oxford, UK

2 Social Computation and Representation Lab, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, UK

3 Department of Imaging Neuroscience, University College London, London, UK

4 Max Planck Institute for Biological Cybernetics, Tübingen, Germany

5 University of Tübingen, Germany

6 School of Psychology, University of Auckland, New-Zealand

7 Department of Experimental Psychology, University College London, UK

8 Clinical, Educational and Health Psychology, University College London, UK

9 South London and Maudsley NHS Foundation Trust, London, UK

10 Centre for AI and Machine Learning, ECU, Perth, Australia

\*correspondence: joseph.barnby@kcl.ac.uk

**Keywords** : social cognition, computational models, paranoia, delusions

**Main**

Corlett et al. criticise a “social turn” in delusions research according to which paranoia is a result of a dysfunction in social cognition [1]. Instead, they propose that, despite appearances, paranoia is solely the result of alterations to domain-general responses to uncertainty. We appreciate the effort to find a parsimonious explanation and we agree that domain-general processes play an important role in understanding delusions. However, we reject the characterisation of previous work by us and others and question whether the dichotomies set up by Corlett et al. are helpful.

First, paranoia is a strikingly social phenomenon: it involves an exaggerated experience of other agents’ harmful intentions [2,3]. Corlett et al. understate this social content, declaring the inference of intentions an “ill-posed problem” (p2) and cast aside the role of mentalising more broadly (p4). They also argue that because threat misattribution involves domain-general processes, paranoia should be understood as “pseudo-social cognition”. This “pseudo-social” view misunderstands the scope of social cognition which is widely and consistently defined as the cognitive processes involved in social perception and behaviour. In one of many similar definitions, Frith and Frith define it as “everything that can, in principle, explain how animals (including humans) learn from each other about the world of objects and how they cooperate and compete with each other in the world of agents.”[4] From their misunderstanding, the authors then project puzzling commitments to those working on social cognition. Importantly, the contribution of domain-general processes and heuristics to social cognition is widely studied, uncontroversial, and no ‘pseudo’ framing has been needed to enable this [5].

Second, Corlett et al. object to the need for a “dedicated encapsulated module” (p2) responsible for social cognition which they claim is a central component in the theories they criticise. This is a grave misreading. At no point have any of us suggested this, and it forms no part of these theories. The work they cite has gone no further than arguing that social cognitive processes, likely shaped by selection and learning, should be considered an important component in cognitive models of paranoia and we have explicitly discussed the need for domain-general processes in these models. This account is motivated by the simple observation that humans are an overrepresented source of threat [6]. Since social cognition presents interesting computational challenges, individuals and groups likely use fine-tuned processes to distil complex social information at the algorithmic level. While identifying how statistical regularities might be harnessed in decision making is valuable, it still begs the question as to how humans transform these into structured, model-based representations of others. Corlett et al appear to suggest that because the accounts they criticise partly appeal to selection to explain paranoia[6], they must necessarily be invoking modularity, despite them never having done so. Assuming that functional specialisation necessarily implies an ‘encapsulated module’ is a widely recognised error even in evolutionary accounts [7].

Further, Corlett et al. argue that the alterations to social experience and behaviour seen in paranoia can be *entirely* explained without need to appeal to social cognition. This is a more controversial position than they admit. This approach would strip domain-general processes of context-specific inductive biases and strategies (such as stochastic or mixed policies) that prove crucial when dealing with intentional agents. We find unconvincing their strict dichotomy between domain-specific and general mechanisms, justified solely through a notional appeal to parsimony [8].

Third, Corlett et al. criticise the social interpretation of behaviour in game theory tasks used in paranoia research. They suggest that in such games “people are perhaps not regarding others and thinking recursively about their minds but rather thinking about their behaviour – about social norms” (p4). However, the paradigms we and collaborators have used to study social decisions in paranoia are deliberately ambiguous scenarios, lacking established norms; receivers can often benefit from inferring not just what happened, but must ask *why*. This demands more than simple heuristics because the intent that one *ought* to attribute when someone else behaves selfishly is under-determined. In such circumstances, model-based social reasoning (including the attribution of beliefs and intentions) tends to predominate over simpler, learned norms or habits [9]. It is more likely that different social cognitive strategies, including norms and recursive mentalising, may be flexibly deployed in different contexts, as evidenced by numerous studies [10]. Again, a dichotomy arises in the authors’ framing that erases important nuances: our formal examination of social cognition in delusions has included, but never been limited to, recursive mentalising [11].

We welcome adversarial collaborations but the diverging assumptions about the nature of social cognition make it difficult to envisage a common interpretation of the outcome. To this end, the authors need to clarify how they will capture the rich and varied range of intentional inference about the behaviour of others present in paranoia. As things stand, the pseudo-social approach only explains pseudo-paranoia.

**References**

1. Corlett, P., Rossi-Goldthorpe, R., Suthaharan, P., Sheffield, J. M., Obeso, S. C. de, & Heyes, C. (2025). Pseudosocial cognition and paranoia. *Trends in Cognitive Sciences*, *0*(0). https://doi.org/10.1016/j.tics.2025.05.019

2. Freeman, D., & Garety, P. A. (2000). Comments on the content of persecutory delusions: Does the definition need clarification? *British Journal of Clinical Psychology*, *39*(4), 407–414. https://doi.org/10.1348/014466500163400

3. Freeman, D. (2024). *Paranoia: A Psychologist’s Journey Into Extreme Mistrust and Anxiety*. William Collins.

4. Frith, C., & Frith, U. (2023). *What Makes Us Social?* The MIT Press. https://doi.org/10.7551/mitpress/10400.001.0001

5. Son, J.-Y., Bhandari, A., & FeldmanHall, O. (2021). Cognitive maps of social features enable flexible inference in social networks. *Proceedings of the National Academy of Sciences*, *118*(39), e2021699118. https://doi.org/10.1073/pnas.2021699118

6. Raihani, N. J., & Bell, V. (2019). An evolutionary perspective on paranoia. *Nature Human Behaviour*, *3*(2), 114–121. https://doi.org/10.1038/s41562-018-0495-0

7. Pietraszewski, D., & Wertz, A. E. (2022). Why Evolutionary Psychology Should Abandon Modularity. *Perspectives on Psychological Science*, *17*(2), 465–490. https://doi.org/10.1177/1745691621997113

8. Sober, E. (2015). *Ockham’s Razors: A User’s Manual* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781107705937

9. Hula, A., Montague, P. R., & Dayan, P. (2015). Monte Carlo Planning Method Estimates Planning Horizons during Interactive Social Exchange. *PLOS Computational Biology*, *11*(6), e1004254. https://doi.org/10.1371/journal.pcbi.1004254

10. Devaine, M., Hollard, G., & Daunizeau, J. (2014). Theory of Mind: Did Evolution Fool Us? *PLOS ONE*, *9*(2), e87619. https://doi.org/10.1371/journal.pone.0087619

11. Barnby, J. M., Dayan, P., & Bell, V. (2023). Formalising social representation to explain psychiatric symptoms. *Trends in Cognitive Sciences*, *27*(3), 317–332. https://doi.org/10.1016/j.tics.2022.12.004