Human brain networks encode creativity state and level

Rikki Rabinovich (rikki.rabinovich@utah.edu)
Department of Neurosurgery, University of Utah, Salt Lake City, UT, USA

Mark Libowitz (mark.libowitz@hsc.utah.edu)
Department of Neurosurgery, University of Utah, Salt Lake City, UT, USA

Tyler S. Davis (tyler.davis@hsc.utah.edu)
Department of Neurosurgery, University of Utah, Salt Lake City, UT, USA

Rhiannon Cowan (rhiannon.cowan@utah.edu)
Department of Neurosurgery, University of Utah, Salt Lake City, UT, USA

Roger Beaty (rebeaty@psu.edu)
Department of Psychology, Pennsylvania State University; University Park, PA, USA

Shervin Rahimpour (shervin.rahimpour@hsc.utah.edu)
Department of Neurosurgery, University of Utah, Salt Lake City, UT, USA

Elliot H. Smith (e.h.smith@utah.edu)
Department of Neurosurgery, University of Utah, Salt Lake City, UT, USA

Ben Shofty (ben.shofty@hsc.utah.edu)
Department of Neurosurgery, University of Utah, Salt Lake City, UT, USA

Abstract

Creativity is an essential human cognitive process with an unclear neural basis. Here, we combine invasive intracranial recordings in the human brain with machine learning tools to explore the network-level representations underlying creative thinking. Our investigations unearthed widespread encoding of cognitive state, revealing unique brain states that underlie creative thinking vs. mathematical reasoning. Further, we identified nonlinear, high-dimensional representations of moment-by-moment creativity level. Finally, our findings hint at distinct roles of different cortical networks in promoting creativity, with default mode network "gating" creativity and dorsal attention network "regulating" the quality of creative output.

Keywords: creativity; default mode network; dorsal attention network; intracranial electrophysiology; sEEG; brain decoding; dimensionality reduction

As a vital component of human experience, creativity has received considerable attention from the scientific community, yet the neural underpinnings of creative thinking remain unexplained. Prior research has led to a consensus that no single brain region is responsible for creativity; rather, interactions within and between cortical networks likely underlie creative ideation (Shofty et al., 2023; Beaty et al., 2016; Marron et al., 2018; Girn et al., 2020). One such network, the default mode network (DMN), has been implicated in a range of cognitive processes—episodic memory retrieval, dream recall, and creativity (Das, de Los Angeles, & Menon, 2022; Vallat et al., 2022).

While creativity is a multifaceted set of processes, here, we specifically refer to divergent thinking, or the mode of thought that spontaneously generates multiple diverse ideas. Hereafter, we use the two terms (creativity and divergent thinking) interchangeably.

To probe the role of brain networks in divergent thinking, we recorded neural signals from stereo-EEG electrodes implanted in the human brain of 15 epilepsy patients, while they engaged in either a creativity-evoking Alternate Uses Task (AUT), which prompts participants to invent creative uses for everyday objects, or a control basic arithmetic task (BAT; Fig. 1a). AUT responses were automatically scored for originality using natural language processing tools (semantic distance). For the prompt "cup", the response "drink tea" scores low on originality, while "build a sandcastle" scores high. Networks sampled were default mode network (DMN), frontoparietal network (FPN), somatomotor network (SMN), limbic network, visual network, dorsal attention network (DAN), and ventral attention network (VAN), based on projecting a predefined network atlas and sEEG electrode locations into MNI space.

DMN dynamics reflect creative vs. mathematical cognitive mode.

We found cognitive state (arithmetic vs. creativity) to be linearly encoded in DMN responses, as a linear support vector machine (SVM) was sufficient to decode task state with high accuracy. Decoding was performed separately for each patient, holding out data from entire trials for cross-validation, thereby preserving temporal structure. While DMN was not the only network from which we were able to decode cognitive state, it was the most predictive (Fig. 1b).

To further interrogate cortical networks' representations of cognitive state, we applied CEBRA, the recently developed nonlinear neural network encoder and dimensionality reduction tool (Schneider et al., 2023). If cognitive state is a fundamental property of brain network activity, it might be possible to predict creativity vs. arithmetic cognitive states *without* providing labels to the model. In other words, we asked whether a *self-supervised* model could still discriminate between creativity and arithmetic cognitive states. To this end, we trained self-supervised CEBRA models, using only time as a label and providing no information about cognitive state, to generate embeddings in a low dimensional latent space. To visualize the patterns of activity that separated creativity from mathematical thinking, we aligned across

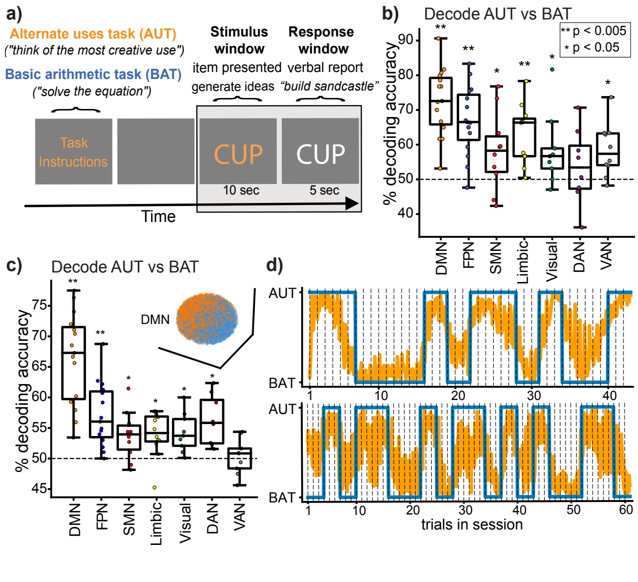

Fig. 1. a) Task design. b) Decoding trial type with linear SVM; p values based on two-sided 1-sample t-test; FDR corrected. c) Decoding trial type based on embeddings; p values calculated as in (b). Inset: embedding generated from self-supervised CEBRA model trained on DMN activity d) Projecting DMN embedding onto coding axis (2 example patients).

subjects and color-coded the latent space by trial type (Fig. 1c, inset), revealing a separation between creativity and arithmetic trials that was qualitatively strongest in DMN. Subsequently, we attempted to predict cognitive state from these embeddings, using a K-nearest neighbors classifier to classify AUT vs BAT trials, holding out one patient at a time. Decoding performance was significantly above chance for all networks except ventral attention network and was strongest for DMN (Fig. 1c). Reducing dimensionality to a single dimension along the coding axis (by projecting the data onto an axis defined by the weights of a linear regression model), we find that the moment-by-moment brain trajectory tracks the ground truth cognitive state (Fig. 1d).

Nonlinear network representations encode creativity level.

While a linear SVM was sufficient to decode cognitive state with high accuracy, we were not able to decode the degree of creativity linearly. Indeed, using a non-linear dimensionality reduction approach through CEBRA, we were able to identify creativity level-related patterns in latent embeddings of the neural activity. We trained a CEBRA model on subsets of data from each subject, using discretized creativity score as a supervisory signal, and then performed the same transformation on the training set and held-out data to generate low dimensional embeddings. We used the embeddings from the training set to decode creativity level and observed above-chance decoding performance for all brain networks; models trained on data with shuffled labels performed at chance level (Fig. 2a). However, unlike the cognitive state representations, for which DMN was the most predictive network, creativity level encoding was strongest in DAN (with DAN significantly more predictive than DMN; $p=0.02$; two-sided Mann Whitney U test). These results suggest that the DMN plays a role in "*gating*" creative thought, linearly, while the DAN is more involved in "*regulating*" levels of originality, nonlinearly, as assessed by the alternate uses task.

Besides the inter-network coding differences, we also noted the existence of a brain-wide creativity-coding axis (Fig. 2b). The projection of creativity level-related manifolds onto this axis (i.e., the moment-by-moment representation of creativity level) correlated well with experimentally measured originality scores (Fig 2c). Thus, despite the multi-dimensional geometry of the latent space, a discrete coding axis can be isolated.

When we projected all datapoints in each network's latent space onto its creativity-coding axis, the correlation of the resulting projection with true creativity values was highest for DMN (r=0.87) compared to the other networks, including DAN (r=0.79). Similarly, the goodness of fit of the linear regression model was higher for DMN (R^2=0.75) than for DAN (R^2=0.66), meaning the creativity coding axis in DMN better captured the variance in the neural data. Though these analyses both delve into similar questions about network-level representations of creativity level, there is a key difference: the decoding analysis was based on embeddings generated from binned creativity data, whereas the coding axis was inferred based on continuous originality scores. We speculate that the DAN's representation of creativity is higher-dimensional and more nonlinear. This conjecture is supported our observation of improved fit of curved coding axes generated by polynomial regression models, and that increasing the degree of the polynomial results in greater improvement of the DAN model's fit compared to the DMN model.

This result is consistent with our proposed framework in which the DAN regulates creativity through goal-directed attentional control of thought processes: unlike DMN's more simple and stable relationship with creativity, DAN generates a highly complex representation of creative states, which may fluctuate with task demands and cognitive engagement.

Together, our results imply that the default mode network maintains a stable representation of cognitive state and plays a key role in gating creative thought; meanwhile, dorsal attention network generates a complex, high-dimensional representation of creativity level that imposes context-dependent constraints and flexibly regulates the quality of creative output.

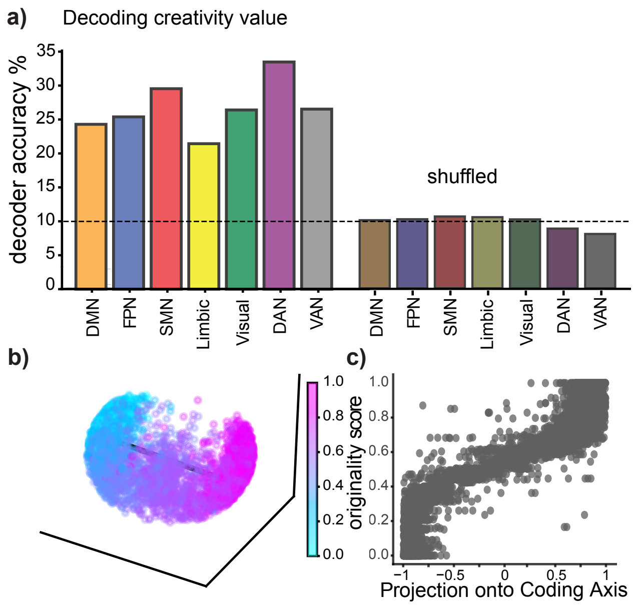

Fig. 2. a) Decoding creativity value based on latent embeddings generated by CEBRA models trained on real data (left) or data with shuffled labels (right). b) DMN embedding; black line is coding axis defined by linear regression. c) Projection of all datapoints of the latent embedding onto the coding axis compared to true normalized originality scores.

Acknowledgements

This work is supported by National Institutes of Health grants R01MH128187 (EHS) and U01NS121472 (EHS), University of Utah startup funds (EHS; BS), and the University of Utah Data Science Institute grant 4583 (EHS).

References

Beaty, R. E., Benedek, M., Silvia, P. J., & Schacter, D. L. (2016). Creative Cognition and Brain Network Dynamics. *Trends in Cognitive Sciences*, *20*(2), 87–95. https://doi.org/10.1016/j.tics.2015.10.004

Das, A., de Los Angeles, C., & Menon, V. (2022). Electrophysiological foundations of the human default-mode network revealed by intracranial-EEG recordings during resting-state and cognition. *NeuroImage*, *250*, 118927. https://doi.org/10.1016/j.neuroimage.2022.118927

Girn, M., Mills, C., Roseman, L., Carhart-Harris, R. L., & Christoff, K. (2020). Updating the dynamic framework of thought: Creativity and psychedelics. *NeuroImage*, *213*, 116726. https://doi.org/10.1016/j.neuroimage.2020.116726

Marron, T. R., Lerner, Y., Berant, E., Kinreich, S., Shapira-Lichter, I., Hendler, T., & Faust, M. (2018). Chain free association, creativity, and the default mode network. *Neuropsychologia*, *118*, 40–58. https://doi.org/10.1016/j.neuropsychologia.2018.03.018

Schneider, S., Lee, J. H., & Mathis, M. W. (2023). Learnable latent embeddings for joint behavioural and neural analysis. *Nature*, *617*(7960), 360–368. https://doi.org/10.1038/s41586-023-06031-6

Shofty et al., 2023 Shofty, B., Gonen, T., Bergmann, E., Mayseless, N., Korn, A., Shamay-Tsoory, S., Grossman, R., Jalon, I., Kahn, I., & Ram, Z. (2022). The default network is causally linked to creative thinking. *Molecular Psychiatry*, *27*(3), Article 3. https://doi.org/10.1038/s41380-021-01403-8

Vallat, R., Türker, B., Nicolas, A., & Ruby, P. (2022). High Dream Recall Frequency is Associated with Increased Creativity and Default Mode Network Connectivity. *Nature and Science of Sleep*, *14*, 265–275. https://doi.org/10.2147/NSS.S342137