Mental imagery in the human brain: In space and time

ABSTRACT:

Background
Mental imagery is the ability to voluntarily generate internal representations via topdown modulation, without any concurrent bottom-up input.

Aims
While viewed and imagined objects from different categories can be reliably decoded from fMRI brain response patterns, it has proved more difficult to distinguish visually similar inputs, such as different instances of the same category. Here, we aimed to reconstruct the face a subject was looking at or imagining, and to investigate the time course of visual perception and mental imagery with intracranial EEG.

Method
We applied a recently developed deep learning system to the reconstruction of face images from human fMRI patterns. We trained a variational auto-encoder (VAE) neural network using a GAN (Generative Adversarial Network) unsupervised training procedure over a large dataset of celebrity faces. The auto-encoder latent space provides a meaningful, topologically organized 1024-dimensional description of each image. We then presented several thousand face images to human subjects, and learned a simple linear mapping between the multi-voxel fMRI activation patterns and the 1024 latent dimensions. Finally, we applied this mapping to novel test images, turning the obtained fMRI patterns into VAE latent codes, and ultimately the codes into face reconstructions.
The time course of mental imagery was examined with intra-cranial EEG. Participants were asked to view or imagine different object categories. Using MVPA techniques we decoded object information during visual perception and mental imagery.

Results
Qualitative and quantitative evaluation of the fMRI reconstructions revealed robust pairwise decoding (>95% correct), and a strong improvement relative to a baseline model (PCA decomposition). Furthermore, the technique allowed for accurate gender classification, and decoding which face was imagined, rather than seen by the subject. Decoding performance with the intracranial EEG signals revealed robust decoding between perception and imagery, and also robust item-level decoding in both conditions.

Conclusions

We hypothesize that the latent space of modern deep learning generative models could serve as a valid approximation for human brain representations. Furthermore, we show that the neural signatures of visual object information during perception and imagery are robust and allow for successful decoding. These representations share common information during the perception and imagery conditions.

Keywords

Mental imagery, Intracranial EEG, fMRI, LFPs in humans

Published Work:

VanRullen, R., & Reddy, L. (2019). Reconstructing faces from fMRI patterns using deep generative neural networks. *Communications Biology, 2*: 193. doi: 10.1038/s42003-019-0438-y

Researcher's Contacts:

Leila Reddy
Centre de Recherche Cerveau et Cognition (CerCo)
CNRS Pavillon Baudot
CHU Purpan Toulouse 31052
France
Tel: +33 5 62 74 61 53
Email: leila.reddy@cnrs.fr