Perceptual decision making unfolds in a processing cascade within and across brain regions

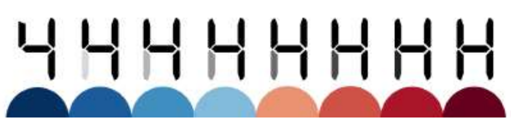
Laura Gwilliams[1,2], Jean-Rémi King[3,1]
New York University[1]; NYU Abu Dhabi[2]; Frankfurt Institute for Advance Studies[3]

1 Introduction

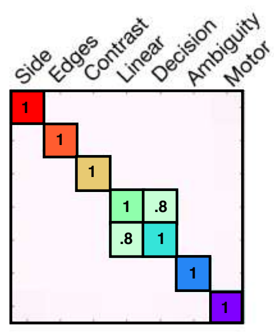
- Transforming environmental input (e.g. a pixelated screen) into a stable percept (e.g. recognising one's friend) is a process of **perceptual decision making**
- Neural activity ramps up in proportion to the evidence in favour of the ultimate selection, fed from the output of lower-level sensory regions[1,2]
- This previous evidence primarily comes from single-unit recordings in monkeys, which does not provide insight into the **macro-level architecture**
- Here we probe the **computations** and **representations** utilised across the whole human brain, and compare these to a pre-trained deep neural network optimised for image captioning

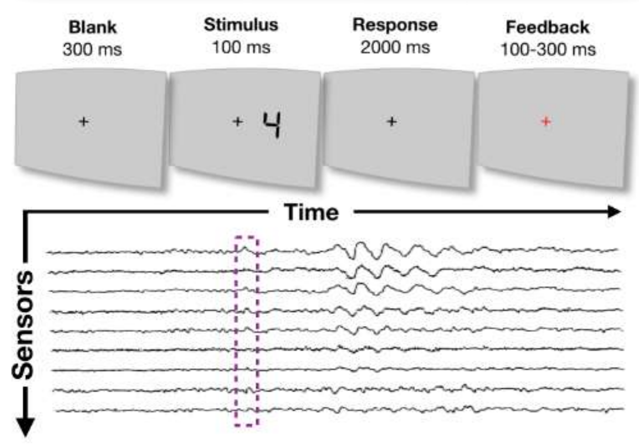
What *neural architecture* underlies perceptual decision making?

2a Experimental Design



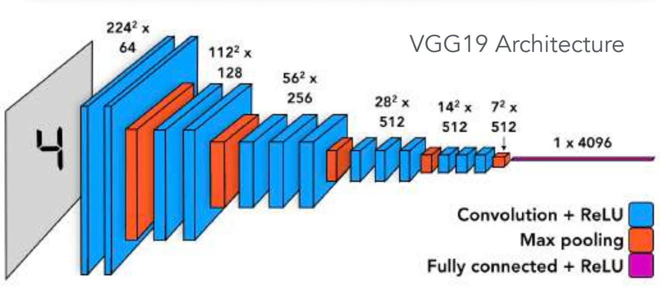
- 17 participants
- letter/digit discrimination
- 306-channel MEG
- 1960 trials
- Stimuli designed to be orthogonal on six variables of interest





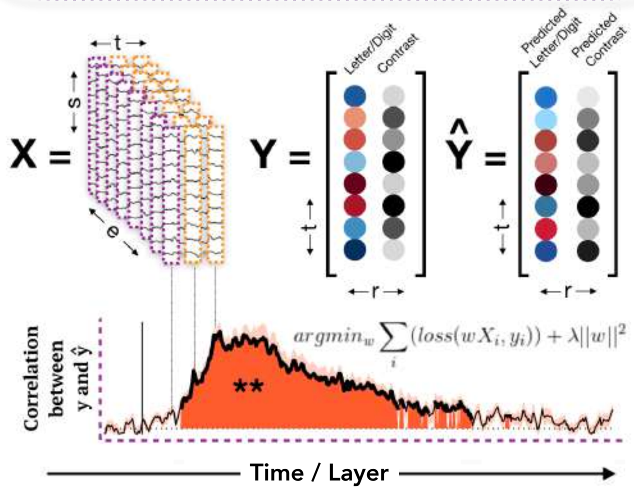
2b Neural Network

- Same stimuli input into pre-trained DNN **VGG19**[3]
- Activity in each layer projected into 306-dimensional space (equivalent to the MEG sensors)
- 17 random projections used, mirroring the distribution over subjects



3 Analysis

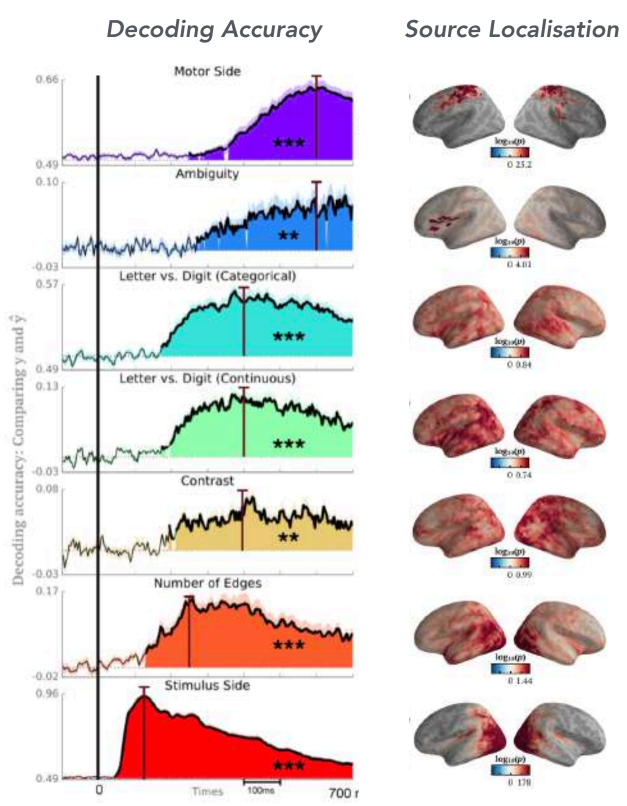
- MVPA applied at each millisecond (t) across the 306 sensors (s) for each epoch (e)
- Decoding scores computed within subjects/ projections, and then tested for significance across subjects with cluster-corrected statistics (p < .05)



4 Human Brain

Hierarchy of representations

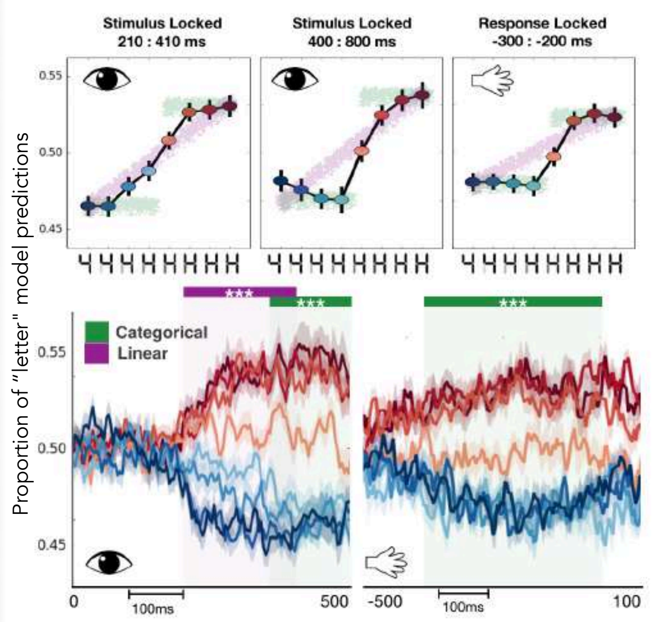
- Features of the stimulus are decodable in a cascade: sequentially activated and maintained in parallel.
- Each feature is supported by a different brain region



- **Suggests that:** the system continuously feeds the output of lower-level computations to higher stages

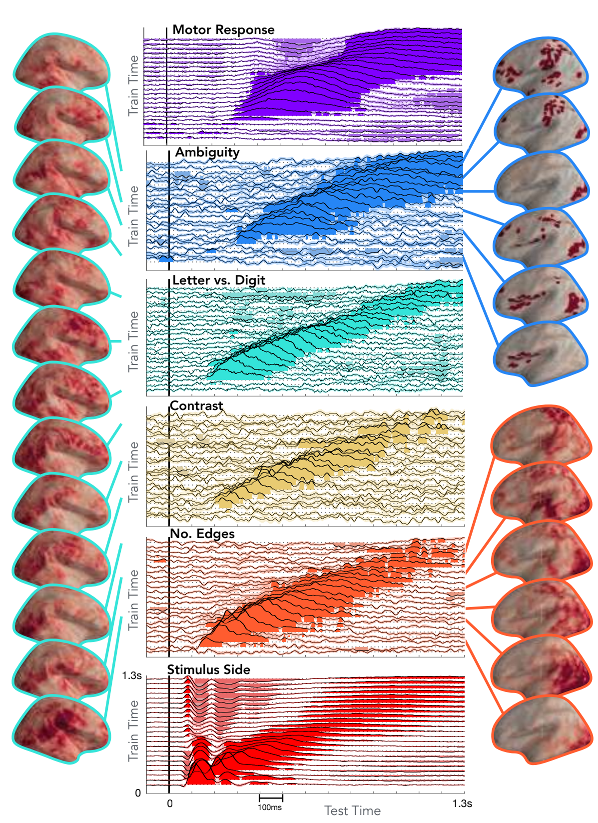
Sensory evidences accumulates to categorical decision

- Neural responses first modulated relative to stimulus evidence, then match subjective experience



The architecture permits continuous feedback to/from all computations

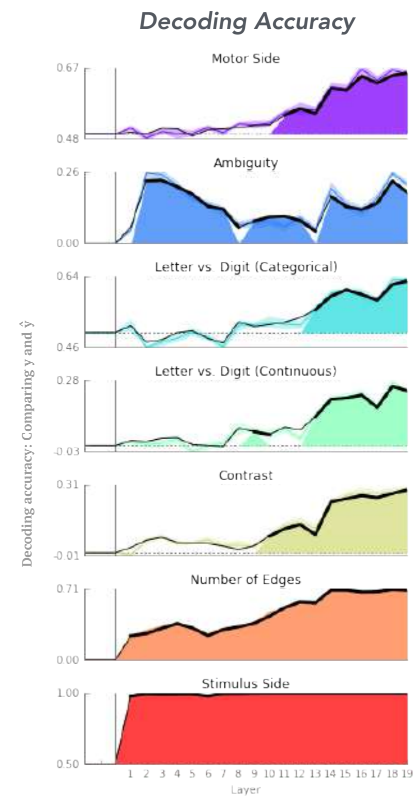
- The brain processes sensory features via multiple parallel cascades
- Neural activity moves across and within brain regions to support each stimulus feature



5 Neural Network

DNN *architecture* is similar

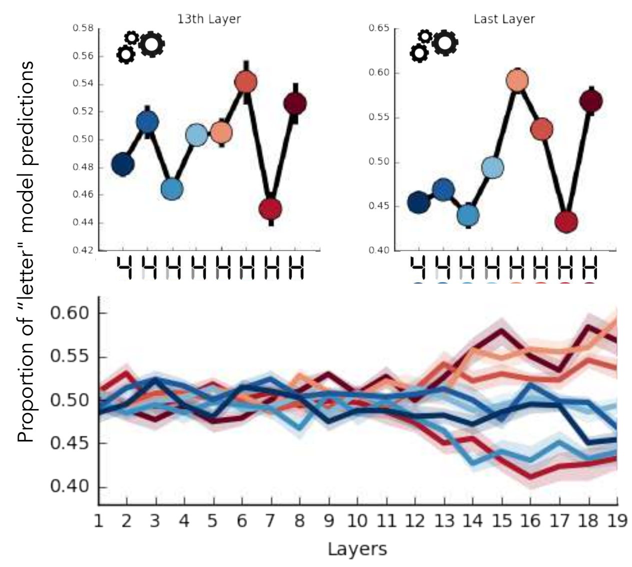
- Decoding stimulus features **DNN activity revealed a cascade** pattern, and similar decoding accuracy obtained on the human MEG data



- **Suggests that:** the cascade architecture is a robust computational strategy, common to both the biological and the artificial visual system

DNN *representation* is different

- Letters and digits are separable from activity in the last layer (89%); however, representations are not linearly *or* categorical, unlike the human observer



6 Conclusion

- **Both biological and artificial neural networks converge on a cascade architecture**
- **The DNN represents information differently from the human brain**
- **Activation patterns evolve within and across brain regions**
- **The brain processes stimulus features in multiple parallel hierarchical cascades**

Funding: G1001 NYUAD Institute (LG); European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No. 660086; Bettencourt-Schueller Foundation & the Philippe Foundation (JRK)

References

[1] Heekeren, H.R., S. Marrett, and L.G. Ungerleider, The neural systems that mediate human perceptual decision making. Nat Rev Neurosci, 2008. **9**(6): p. 467-79.

[2] Shadlen, M.N., et al., A computational analysis of the relationship between neuronal and behavioral responses to visual motion. Journal of Neuroscience, 1996. **16**(4): p. 1486-1510.

[3 Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556.*