Representations of Sensory Signals and Abstract Categories in Brain Networks

Dimitris A. Pinotsis (pinotsis@mit.edu)
Centre for Mathematical Neuroscience and Psychology and Department of Psychology
City —University of London, London EC1V 0HB, United Kingdom
The Picower Institute for Learning & Memory and Department of Brain and Cognitive Sciences
Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Markus Siegel (markus.siegel@uni-tuebingen.de)
Center for Integrative Neuroscience and MEG Center,
University of Tubingen, 72076 Tübingen, Germany

Earl K. Miller (ekmiller@mit.edu)
The Picower Institute for Learning & Memory and Department of Brain and Cognitive Sciences
Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Abstract:

Many recent advances in artificial intelligence (AI) are rooted in visual neuroscience. However, ideas from more complicated paradigms like decision-making are less used. Although automated decision-making systems are ubiquitous (driverless cars, pilot support systems, medical diagnosis algorithms etc.), achieving human-level performance in decision making tasks is still a challenge. At the same time, these tasks that are hard for AI are easy for humans. Thus, understanding human brain dynamics during these tasks and modeling them using deep neural networks could improve AI performance. Here we modelled some of the complex neural interactions during a sensorimotor decision making task. We investigated how brain dynamics flexibly represented and distinguished between sensory processing and categorization in two sensory domains: motion direction and color. We found that neural representations changed depending on context. We also trained deep recurrent neural networks to perform the same tasks as the animals. By comparing brain dynamics with network predictions, we found that computations in different brain areas also changed flexibly depending on context. Color computations appeared to rely more on sensory processing, while motion computations more on abstract categories. Our results shed light to the biological basis of categorization and differences in selectivity and computations in brain areas.

Keywords: decision making; LSTM; cortical hierarchies; categories; deep neural networks.

A Flexible Decision Making Task

We reanalyzed data previously published in (Siegel et al., 2015) using Representation Similarity Analysis (RSA; Kriegeskorte et al., 2008) and deep recurrent neural networks (RNNs). Monkeys performed the task shown in Fig. 1. They categorized motion direction and color of centrally presented, colored random dot stimuli (Fig. 1A). Before stimulus onset, a central cue indicated which feature to categorize. Monkeys indicated their choice with a leftward or rightward saccade and held central fixation throughout each trial until their response. Monkeys were free to respond any time up to 3 s past stimulus onset. We analyzed data from the epoch after stimulus onset until average response latency (1s to 1.27s; t=0 corresponds to cue onset). Stimuli systematically covered motion direction, and color space between opposite motion directions (up and down) and opposite colors (red and green; Fig. 1B). There were 7 possible stimulus motion directions and 7 possible colors. In total, there were 42 stimulus conditions. Depending on the task cued at the beginning of each trial, the animals categorized either the motion direction (up vs. down) or color (red vs. green) of the stimulus. We recorded LFP data from 6 cortical areas shown in Fig. 1C.
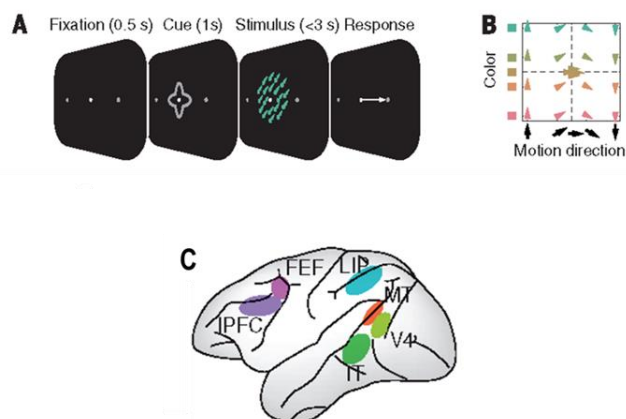

Figure 1: (A) Monkeys categorized the motion direction, or color, of centrally presented, colored random dot stimuli. Before stimulus onset, a central cue indicated which feature to categorize. Monkeys

indicated their choice with a leftward or rightward saccade and held central fixation throughout each trial until their response. Monkeys were required to respond within 3 seconds after the stimulus onset. For each trial, we analyzed the data from the stimulus onset to the average response latency (1s to 1.270s) (B) Stimuli systematically covered motion, direction, and color space between opposite motion directions (up and down) and opposite colors (red and green). All stimuli were 100% coherent, iso-speed, iso-luminant, and iso-saturated. (C) Schematic display of the recorded brain regions. See also (Siegel et al., 2015) for more details.

Differences in representations between brain areas

To understand neural representations in different brain areas during this flexible decision task, we used two approaches. We computed 1) the similarity of neural representation in a brain area with the geometry of the sensory or category domain represented (which we call *domain selectivity*; motion vs color vs motion categories vs color categories). 2) The similarity of neural computation performed by a brain area with predictions from 2 deep RNNs: one trained to distinguish categories (like the behavioural task) and the other to process visual information (this we call, *computation selectivity*). The assumption here was that to perform the behavioural task both kinds of computations should take place in different brain areas, i.e. categorization required also sensory processing.

The two approaches we used are distinct. Being selective to a sensory domain (domain selectivity) is not the same as performing computations like sensory processing and abstract categorization (computation selectivity). Domain selectivity refers to representation content only, while computation selectivity characterizes how these representations are manipulated and compared to each other to find their similarities and differences. Also, sensory processing requires integrating sensory inputs while abstract categorization requires combining these integrated inputs with prior knowledge about learned categories. All these computations take time. Thus, understanding which computations each area performs requires analyzing temporal information in brain dynamics. Although distinct, domain and computation selectivity should give *similar results*. We found this below.

Domain Selectivity

We first considered the selectivity of each brain area to motion direction and color (Fig. 1B). To understand what kind of representations (motion direction vs color, sensory processing vs categorization) were encoded in each brain area, we computed the dissimilarity between

brain Representation Dissimilarity Matrices (RDMs) and sensory/category DMs (SDMs/CDMs) on the other. Brain RDMs were obtained using LFP recordings from each brain area. We followed (Kriegeskorte et al., 2008) and used the dissimilarity between dissimilarity matrices (called deviation) as metric to compare brain RDMs and SDMs/CDMs. These deviations are shown in panels of Figure 2. There are 6 panels (for the 6 brain areas). Each panel has 4 bars (deviation of each brain RDM from color SDM; motion SDM; color CDM; motion CDM).

Interestingly, for most brain areas domain selectivity depended not only on the stimulus but also on the domain of categorization. It switched between the two domains depending on task (motion direction or color categorization). This is a surprising result, not previously shown to the best of our knowledge. Also, related work in the literature usually focuses on sensory perception only, and does not normally involve flexible switching between sensory domains, contrary to the paradigm considered here.
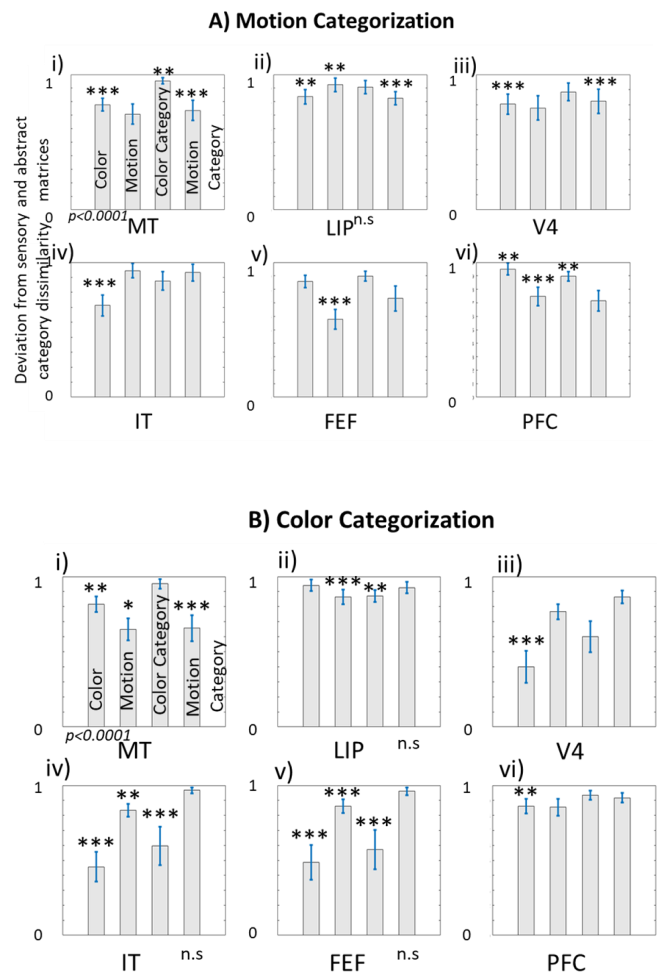

284

Figure 2: Deviations between RDMs and SDMs/CDMs. (A) Motion categorization task. Each panel depicts deviations between RDM of a brain area and the SDM ("color", "motion" 1st and 2nd bars from left) or CDM ("color category", "motion category", 3rd and 4th bars) respectively. Error bars denote standard errors. All deviations were significant at the p<0.0001 level with the exception of those with "n.s" at the bottom(not significant; fixed-effects category-index randomization test, see Methods and (Kriegeskorte et al. 2008)). (B) Same results for the color categorization task. Note that deviation is based on correlation distance, thus smaller bars indicate better similarity between RDMs and SDMs/CDMs. Asterisks above each bar denote the significance level of the corresponding partial correlations. See (Pinotsis et al.,2019 for more details).

V4 showed preference towards the color domain in both tasks and motion categories in the motion task. MT was more selective for the motion category domain in both tasks. FEF exhibited selectivity for motion in the motion task and color in the color task. PFC selectivity was for the motion domain in the motion task and the color domain in the color task. Finally, IT seemed to prefer more color in both tasks and color categories in the color task. See (Pinotsis et al., 2019) for more details.[1] We then confirmed the above results using deep neural networks. We turned to computation selectivity. This was defined based on the similarity of brain activity with predictions from RNNs performing either sensory processing or categorization.

Computation Selectivity

To understand this, we built deep RNNs. Although they comprised six LSTM layers (the same number of layers as the cortical network from which we recorded LFP responses), we use them only for simulating brain computation, not as precise descriptions of anatomy. We considered 2 variants of the same RNN. One trained to perform sensory processing and the other abstract categorization (sensory and category RNN respectively). We assumed that sensory processing would be based on low level visual features, while categorization would be based on information that the animal had learned after being trained to perform the task. Then we compared the RNN predictions to neural activity. We concluded that the computation a brain area performed would be similar to that of the RNN whose predictions were more similar to (had smallest deviations) and significantly correlated with brain activity. We trained them using LFPs as inputs and labels corresponding to different sensory stimuli or

categories as outputs (depending on whether the RNN was processing sensory information or categorizing). We used RSA again and compared brain and network RDMs. Results are presented in Figures 3 and 4.
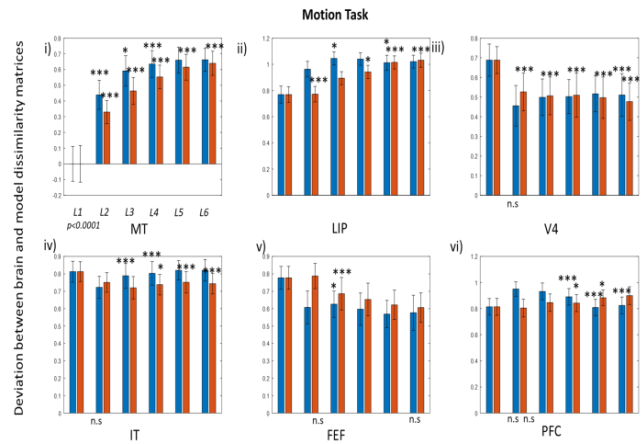

Figure 3: This figure follows the format of Figure 2. Bars in each panel depict deviations between RDM of a brain area and each layer in a deep RNN performing motion processing and categorization. There are six pairs of bars, equal to the number of layers. The left bar in each pair corresponds to deep RNN predictions when the network performs sensory processing, while the right bar corresponds to predictions during categorization. Error bars denote standard errors. All deviations were significant at the p<0.0001 level with the exception of those with "n.s" at the bottom (Kriegeskorte et al. 2008)). Asterisks above each bar denote the significance level of the corresponding partial correlations.

---
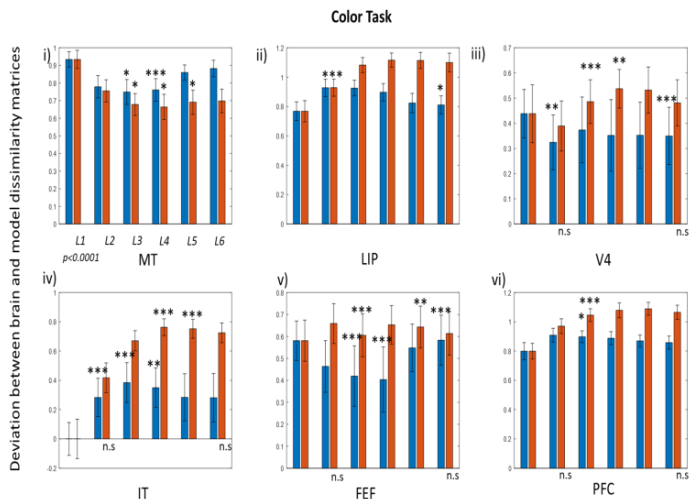
1 Due to space limitations, we did not include further details here.

Figure 4: Deviations between brain and network RDMs for color processing and categorization. This is similar to Figure 3 where the deep RNN has learned to process and categorize color as opposed to motion direction stimuli.

The results of Figures 3 and 4 confirmed those of Figure 2: V4 showed preference towards sensory processing in the color task and motion categorization in the motion task. MT was more selective for categorization during both the motion and color tasks. FEF showed clear preference for sensory processing during both tasks. PFC seemed to prefer more sensory processing in motion task and categorization in the color task. Finally, IT seemed to prefer sensory processing in both tasks which also coincided with its domain selectivity. See (Pinotsis et al., 2019) for more details.

Conclusions

Our results fit well with earlier results by (Mante et al., 2013). We found that sensory information reaches PFC. Gating of sensory input is absent and filtering out of irrelevant (sensory) information by earlier brain areas did not occur. Also, (Mante et al., 2013) found that PFC responses during the motion and colour tasks occupy different parts of state space, and the corresponding trajectories are well separated along the axis of context (task). This can explain the flexible domain selectivity switching between tasks we found here.

All in all, we found that representations changed flexibly depending on context (motion vs color task) and level of abstraction (sensory processing vs categorization). The motion task seemed to rely more on categorization, while the color task seemed to be driven by sensory computations. These results are in accord with earlier findings by (Brincatt et al., 2018). In that paper, coding in most areas was found to reflect a mixture of sensory and categorical effects. Similarly, we found significant similarities between brain RDMs and RDMs from neural networks that perform both sensory processing and abstract categorization. In the same work, categories arose gradually across the hierarchy. Our analysis, based on deep recurrent neural networks, revealed that gradual emergence is driven by sensory color and more abstract motion direction categorization.

All in all, our analysis sheds light to the biological basis of categorization and differences in selectivity and computations among different brain areas. It paves the way for constructing neural networks that can replicate brain dynamics underlying complex sensorimotor decision making tasks. Elucidating such differences can be important for building automated systems for intelligent decision making in multidimensional domains, like driverless cars, pilot support systems, medical diagnosis algorithms etc. We hope our work can help make progress in this direction.

References

Brincat, S.L., Siegel, M., von Nicolai, C., and Miller, E.K. (2018). Gradual progression from sensory to task-related processing in cerebral cortex. Proc. Natl. Acad. Sci. 115, E7202–E7211.

Kriegeskorte, N., Mur, M. & Bandettini, P.A. (2008). Representational similarity analysis-connecting the branches of systems neuroscience. Frontiers in Systems Neuroscience 2, 4.

Mante, V., Sussillo, D., Shenoy, K.V., & Newsome, W.T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature 503, 78.

Pinotsis, D.A., Siegel, M., and Miller, E.K. (2019). Sensory Processing and Categorization in Cortical and Deep Neural Networks. BioRxiv 647222.

Siegel, M., Buschman, T.J. & Miller, E.K. (2015). Cortical information flow during flexible sensorimotor decisions. Science 348, 1352–1355.