

HIERARCHICAL STRUCTURE IN PREFRONTAL CORTEX IMPROVES PERFORMANCE AT ABSTRACT TASKS

R. TUKKER* and A. C. VAN ROSSUM

*Almende BV, Westerstraat 50,
3016 DJ, Rotterdam, The Netherlands
E-mail: remco@almende.org

S. L. FRANK

*Centre for Language Studies,
Radboud University Nijmegen, The Netherlands*

W. F. G. HASELAGER

*Donders Institute for Brain, Cognition and Behaviour
Radboud University Nijmegen, The Netherlands*

We used an Echo State Network (ESN) to investigate cognitive control tasks that involve temporal and policy abstraction, namely an n -back task and a Wisconsin Card Sorting task. In the task conditions where those abstractions were most relevant, separating the abstraction levels increased ESN performance, while the performance was reduced by such a separation in the conditions where the abstractions were less important. This result gives a possible explanation for the topological mapping of abstraction levels over the prefrontal cortex, because of wiring economy in the brain.

1. Introduction

Over the past years, multiple fMRI and other neuroscience results led to the proposal that the Prefrontal Cortex (PFC) is organized hierarchically, with a topological mapping of abstraction levels over the PFC.¹⁻⁵ Anterior regions of the PFC are thought to facilitate higher abstraction levels, while posterior regions would be responsible for lower abstraction levels. Badre and D'Esposito⁶ distinguished four abstraction types in those studies: domain generality, relational integration, temporal abstraction and policy abstraction. We will investigate tasks involving the two most studied abstraction types, policy and temporal abstraction, suggesting a reason why they may be mapped topologically on the PFC.

Temporal abstraction simply means that certain information is relevant for a longer duration than other information. A prime example is the n -back test. *Policy abstraction* (also known as context abstraction) occurs when a goal or context representation does not deal with physical reality, but instead only concerns other representations. This is for example the case with Wisconsin Card Sorting (WCS). A particular color, shape or number of items determines the motor response (which is not abstract) in the WCS task; however, policy abstraction is introduced because we need the current context to determine which property of the card is currently relevant for the sorting.

We hypothesize that control tasks that involve policy or temporal abstraction share the requirement of information *preservation between* the abstraction levels and information *integration within* the abstraction levels. This simple hypothesis can explain the topological mapping of abstraction levels, because it is likely that more neural connections are required for integration of information than for keeping information separated. Therefore, the total axon length that is required in the PFC is smallest when information is integrated within each area, and preservation of information is obtained by processing it over different areas. That total axon length is in fact minimized by evolution is known as the wiring economy principle.⁷⁻⁹ Thus, a confirmation of our hypothesis would provide evidence and a possible explanation for the topological mapping of abstractions in the PFC.

So far, the imaging results led to several theories of abstraction levels in the PFC, of which we will discuss the most significant. One very influential theory is presented in the fMRI study by Koechlin *et al.*¹⁰ Three abstraction levels were investigated by varying the number of possible motor responses, the amount of information given by a context cue and the amount of information given by an instruction cue in the beginning of a block of trials. This resulted in differential activation in dorsal premotor cortex (PMd), caudal lateral PFC (IPFC) and rostral IPFC, respectively. This is evidence for a topological mapping of policy and temporal abstraction. Structural equation modelling was used to support the idea of a “cascade of control”, from rostral IPFC that supplies temporal information, via caudal IPFC that supplies context information, to the less abstract representations in PMd.

Several research groups replicated these findings, but all in a slightly different way.^{2,11,12} For us, one of the studies by Badre and D’Esposito² is most important. They varied the amount of competition on three successive policy abstraction levels and additionally one temporal abstraction

level. The fMRI result was that PMd, anterior PMd, the inferior frontal sulcus and the frontopolar cortex (FPC) were respectively activated. This led Badre and D'Esposito to propose that each additional level of abstraction might activate an increasingly anterior region in PFC. While both the abstraction levels and the results are similar to those from Koechlin *et al.*,¹⁰ the experimental conditions and the activated regions are not exactly the same.²

The most recent addition to the collection of theories and fMRI results comes from a study by Reynolds *et al.*³ The different abstraction levels were compared over five conditions: baseline, low temporal and policy abstraction, high temporal and low policy abstraction, low temporal and high policy abstraction and finally high temporal and high policy abstraction. Because the activation levels in a particular region of interest in dorsolateral PFC (dlPFC) were not completely in line with either Ref. 10 or Ref. 2, Reynolds *et al.*³ suggest there may not be a topological mapping in the PFC at all. Instead, they propose the “adaptive context maintenance” hypothesis, which states that task-relevant information is maintained in the PFC as long as necessary. We will show that while the adaptive context maintenance hypothesis might be correct, we have a reason not to reject the idea of a topological mapping: it can be beneficial for the execution of abstract tasks, assuming wiring economy.

2. Methods

2.1. Model

To test our hypothesis, we will employ an Echo State Network (ESN), somewhat similar to the model by Dominey.¹³ ESNs are a relatively new type of recurrent artificial neural network, that addresses problems with training procedures of traditional recurrent neural networks.¹⁴ The reservoir of our ESN is formed by a continuous time recurrent neural network. The activity of node i in the reservoir is updated according to a discrete time approximation of the following equation:

$$\tau_i \dot{a}_i = -a_i + f \left(\sum_j W_{ji} a_j - b_i \right) + u_i \quad (1)$$

With τ the time constants vector, a the activation vector, f the activation function (which is the tanh function in our case), W the sparse connection weight matrix, b the bias vector, and finally, u the external input vector.

The reservoir is initialized in such a way that two clusters can be discerned: a ‘top’ cluster and a ‘bottom’ cluster, representing anterior and posterior PFC, respectively. The top cluster gives output to the bottom cluster, but does not receive input from it. The bottom cluster is in turn used to train the output layer (see Fig. 1). Because of this hierarchical structure and the random connectivity in the reservoir, information is better integrated within a cluster and better preserved between the clusters. This makes our model well suited for the investigation of the effect of information separation versus information preservation: either less abstract information is offered to the bottom cluster and more abstract information to the top cluster, or all information is offered to both clusters.

The input to the network is given by setting a pattern of high and low activation values in a number of input nodes. These input nodes are connected to the reservoir through a sparse input connectivity matrix for the duration of the first timestep of each trial. Of course, the input pattern over the input nodes changes from trial to trial. To offer specific information to only part of the network, the input connectivity matrix is initialized to allow connections from this specific input node to one of the two clusters only. This is done without changing the overall connectivity of the input weight matrix, thus without changing the expected total input activity (only the location of this input activity is changed). Apart from enforcing those two clusters in the weight matrices, the connectivity is random.

To obtain output from the network, output nodes are connected to the bottom cluster of the reservoir. Each output node corresponds to one possible response and the output node that shows most activation at the end of the trial defines the response of the model. The fully connected output weight matrix connecting the bottom cluster with the output nodes is trained on the desired response with a simple regression procedure:

$$\vec{w}_i = S^\dagger \vec{d}_i \quad (2)$$

This equation gives \vec{w}_i , the weight vector for output node i , from S^\dagger , the pseudoinverse of the state collection matrix which contains the activation values of all nodes at the end of each training trial, and \vec{d}_i the vector containing the desired outputs for the node per trial. The desired output is set to an arbitrary positive value if the node should be selected at this trial, and to zero otherwise. The parameter settings for the model are summarized in Table 1. Further details of the implementation are described in Ref. 15, which is available at <https://github.com/RemcoTukker/PFC-ESN/> together with the Matlab implementation of the model.

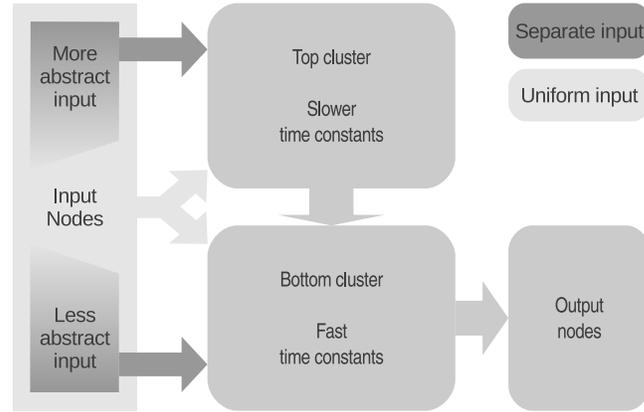


Fig. 1. The architecture of the Echo State Network. In dark gray and light gray the two input methods as determined by the connectivity in the input weight matrix, separated and uniform respectively. In the middle the hierarchical topology of the reservoir fashioned after the PFC. Apart from the constraint on the topology, the connectivity is random. At the right the output nodes for which the weights are trained with a simple regression method. Each output node corresponds to a possible response for a trial in the tasks.

Table 1. Parameter settings of the ESNs.

Property	Value (temporal / policy abstraction)
Number of reservoir nodes	200 / 100
Number of nodes top : bottom	1 : 3
Reservoir connectivity	0.05
Reservoir weight distribution	Uniform, centered at 0 ^a
Top down scaling	1.0 / 0.6
Spectral radius	0.9
Input connectivity	0.1
Input weight distribution	$\mathcal{U}(-2.5, 2.5)$
Bias distribution	$\mathcal{N}(0, 0.3)$
Timesteps per trial	15
Number of training trials	800
Number of test trials	800
Time constants	10 to 10^3 / 10 to 12.5

Note: ^a Top down weights are scaled and bottom-up connections removed. The range of the distribution is determined by the spectral radius and the random connectivity; no fixed values can therefore be given.

2.2. Tasks

The first task was designed to investigate policy abstraction and corresponds to the sorting in a WCS task. At each trial, one input node gives abstract information and is set to either low or high activity. This abstract information corresponds to the context in a simple version of WCS: either the cards have to be sorted according to color or according to shape. Depending on this context (or current policy), either the first half of the rest of the input nodes is relevant to determine the desired output, or the second half. These input nodes thus represent the color and shape in WCS, respectively. The input nodes that are not relevant in the current trial are distractors. The task is illustrated in Fig. 2.

The experimental condition in this task is the number of low-level input nodes. If our hypothesis is correct, preserving the information within the abstraction levels is most important with a large number of low-level input nodes, as the abstraction is most important in this situation. On the other hand, integrating information between the abstraction levels is most important in the conditions with a small number of low-level input

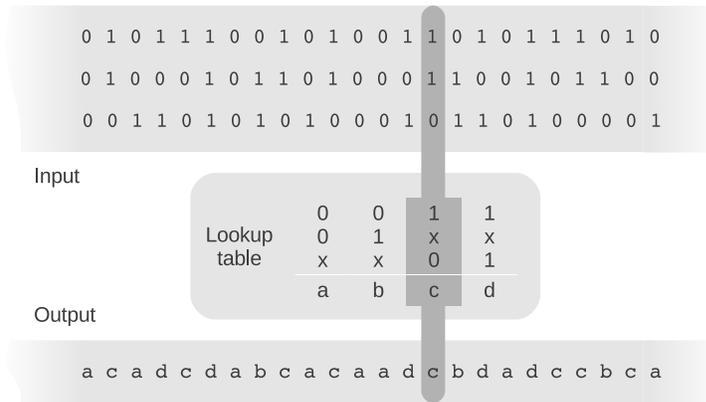


Fig. 2. The version of the policy abstraction task with three input nodes (and thus two low abstraction input nodes). Each column represents one trial. The input pattern at the top is used to set high and low activities in the input nodes at the beginning of each trial. In the lookup table the top row represents the abstract input that is used to determine whether the middle row or bottom row is currently important to find the desired output. The x in the lookup table is a wildcard: the input node can be either 1 or 0, it is irrelevant for the task and thus a distractor. The letters in the output row correspond to the four output nodes in this task, with the highlighted letter corresponding to the output node that should show the highest activation at the end of the current trial.

nodes. Thus, the ESN with topological input should perform better in the conditions with many low-level input nodes, while the ESN with uniform input is expected to perform better in the conditions with a small number of low-level input nodes.

The second task is a generalized version of the n -back test to investigate temporal abstraction. The information presented at one specific, abstract input node has to be retained for n trials, after which it has to be integrated with the current information at the other input nodes to find the desired output. This integration is again dictated by a lookup table, this time with each possible pattern corresponding to one output node. A graphical representation of the task can be found in Fig. 3. To facilitate the retention of information in the reservoir, some of the time constants in the top cluster are randomly increased for this task.

According to our hypothesis, when n and thus the temporal abstraction is small, the most decisive factor in the performance of the network should be the ability to integrate information. Reversely, when n is large, the preservation of the more abstract information should be most important. Thus, we expect the ESN with segregated, topological input to perform better in n -back test when n is large, while the ESN with uniform input

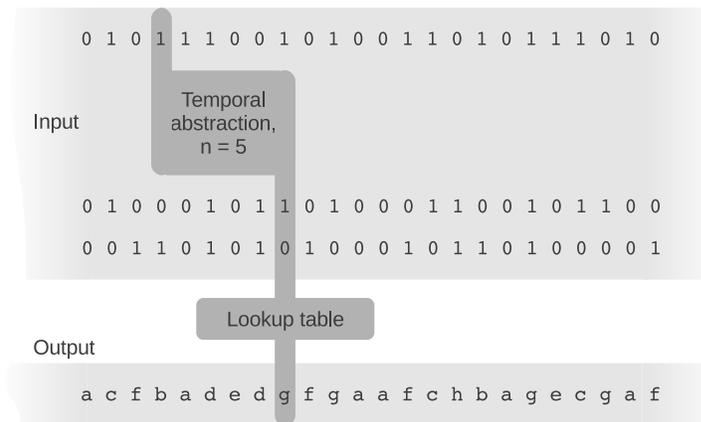


Fig. 3. The implementation of the temporal abstraction task used in the simulations. The basis is a lookup task with three binary input nodes, with each of the eight possible input patterns corresponding to an output node, which is represented by a letter in the output row. Each column in this figure represents a trial and temporal abstraction is thus introduced in the task by presenting the relevant bit from one of the input nodes n trials earlier than the data at the other input nodes, so that this information has to be retained for n trials to find the desired output.

should perform better when n is small.

2.3. Performance Evaluation

To evaluate the performance of the model on the different tasks, each reservoir was trained on 800 trials and directly afterwards tested on 800 trials, without resetting the network activity at any time. The final performance of each reservoir is simply calculated by dividing the number of correct responses to the test trials by the total number of test trials for the reservoir, yielding a number between zero and one. As each reservoir is random, 200 different reservoirs are initialized and evaluated in this way per condition and per input type, to achieve statistical significance.

3. Results

In Fig. 4 the resulting performance distributions for the temporal abstraction task are shown. In the easiest condition, with memory length $n = 0$, the task is purely the lookup of the current input pattern and both the ESN with uniform input and the ESN with topological input perform perfectly; hence, there is no difference between the two input patterns. Likewise, in the hardest conditions both networks perform weakly and the difference between the ESNs diminishes. However, between those two extremes, we can see that the performance distributions of topological ESNs and uniform ESNs show differences. The same pattern can be observed in Fig. 6, which is a plot of the performance distributions for the policy abstraction task. In both tasks, the difference between the uniform and topological input conditions is highly significant in most cases thanks to the large number of evaluated reservoirs, with $p < 0.001$ in a Mann Whitney U test, as specified in Fig. 4 and Fig. 6.

To simplify the interpretation of these graphs, we choose the difference between the medians of the ESNs with topological input and the ESN with uniform input as a measure of performance difference. In Fig. 5 we can find the differences for the temporal abstraction task. For short memory lengths, the uniformly processing ESNs have an advantage, while the topologically processing ESNs perform better for long memory lengths. The differences of the medians on the policy abstraction task are shown in Fig. 7. Over the conditions in which our performance measure gives a useful result, uniformly processing ESNs have an advantage when a small number of low-level features has to be integrated, while topologically processing ESNs have an advantage when a large number of features has to be integrated.

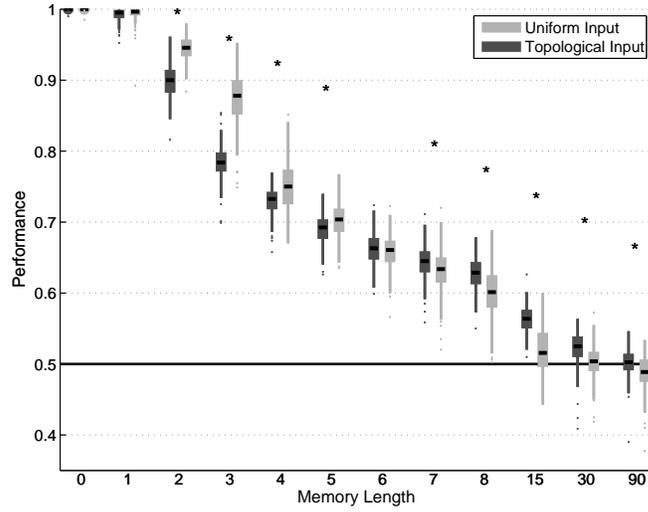


Fig. 4. The distributions of the performance of 200 random ESNs in each condition of the temporal abstraction task. Each box extends from the lower quartile to the upper quartile and the whiskers extend to 1.5 times the interquartile range. The gray line is the performance level that can be reached when the temporal abstract information is not taken into account. Asterisks denote a significant difference ($p < 0.001$) between the ESN with uniform input and the ESN with topological input.

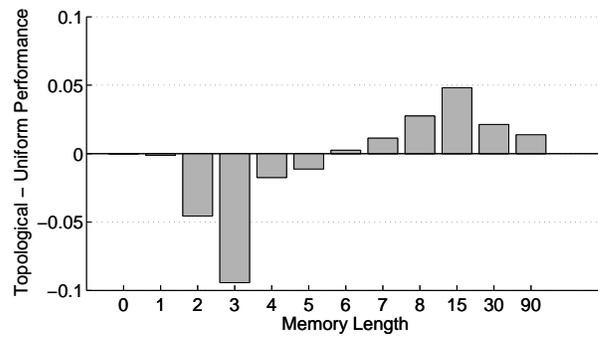


Fig. 5. The difference in median performance between a uniformly processing ESN and a topologically processing ESN as a function of the memory length in a n -back test. The number is positive if the topologically processing ESN performs better.

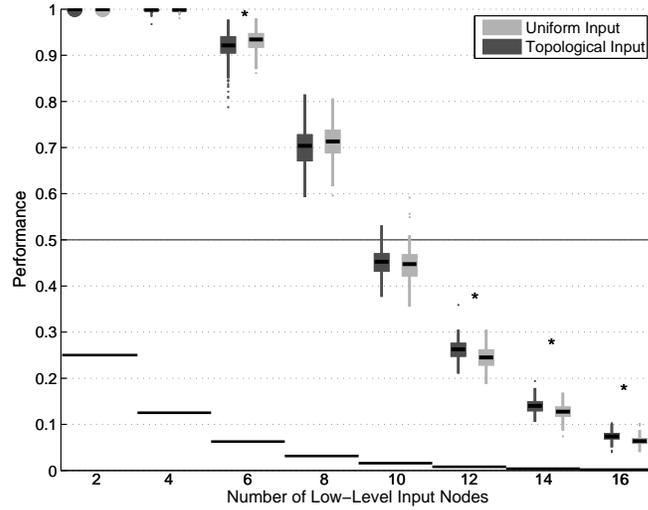


Fig. 6. The distributions of the performance of 200 random ESNs in each condition of the policy abstraction task. Each box extends from the lower quartile to the upper quartile and the whiskers extend to 1.5 times the interquartile range. The dark gray lines are at chance level for each of the conditions, while the light gray line is at the performance level that can be reached by integrating all low-level information but ignoring the high-level input node. Asterisks denote a significant difference ($p < 0.001$) between the ESN with uniform input and the ESN with topological input.

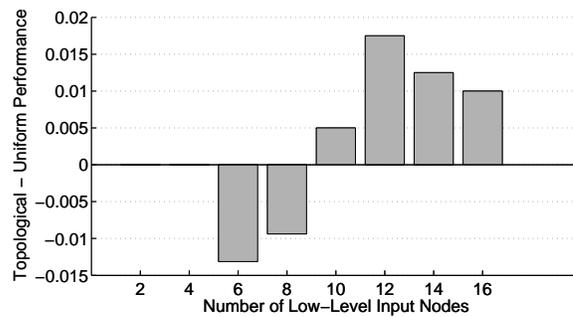


Fig. 7. The difference in median performances between topologically processing ESNs and uniformly processing ESNs at the policy abstraction task. The number is positive if the topologically processing ESN performs better.

4. Discussion

The results are exactly in line with what we would expect from the hypothesis that abstract cognitive control tasks require information to be integrated within an abstraction level and preserved between abstraction levels, as detailed in the methods section. This is in agreement with earlier results: Botvinick has already shown that a topological map of temporal abstraction can be the result of the backpropagation through time learning algorithm in a hierarchical network with an abstract task.⁴ Also, such a structure has been shown to partly self-organize using an evolutionary algorithm.¹⁶ The results for the mapping of policy abstraction are for example in line with research by Yamashita and Tani,¹⁷ who found self-organized policy abstraction over motor primitives. However, none of these investigations directly addressed the cause of this structure.

Having found requirements for tasks that can influence the way the PFC processes abstraction levels because of wiring economy, we can use these results for the interpretation of the imaging experiments that we described earlier. The work by Koechlin¹⁰ is in agreement with our results, because his experiments included a large difference in timescales: the context had to be remembered for mere seconds, while the episodic control signal had to be remembered for a much longer time, in the order of minutes. This seems reasonable for tasks in daily life as well. The work by Badre and D'Esposito² is partly in agreement: a hierarchical mapping of policy abstraction was shown, but the task that included policy abstraction had only a small number of low-level input nodes; a situation in which the uniform processing ESNs performed better in our research. This indicates that the task requirements are not the only factor of importance.

The most recent addition to the debate was made by Reynolds *et al.*,³ who argue that they did not find any evidence for topological mapping of abstraction and propose the “adaptive context maintenance” hypothesis. This is in conflict with our results: while Ref. 3 suggests an absence of topological mapping of abstractions, our results give a theoretical reason to expect such a mapping, regardless of how the information in the PFC is exactly maintained. Also note that the task used in Ref. 3 was slightly different from the other mentioned studies and that the analysis of the results was done using a rather restricted region of interest. Therefore, we suggest that while Reynolds *et al.* make a good point for the “adaptive context maintenance” hypothesis, their claim that abstraction levels are not mapped topologically over the PFC goes too far.

Some important limitations of our study have to be discussed as well.

First of all, our results may be dependent on our parameter settings. Thus, under different parameter settings the observed effect may not be present or may be unimportant compared to other mechanisms. To mitigate this problem we attempted to keep the number of critical parameters to a minimum and investigated a range of reasonable parameter settings. Unfortunately, an exhaustive gridsearch was not feasible. While we never observed the opposite effect, we could also not confirm the described relation under all parameter settings: especially changing the top-down connection strength led to results in which either the uniform processing ESN or the topological processing ESN always performed better.

Secondly, and arguably more importantly, the structure and organization of the PFC is of course not solely determined by the requirements of exerting control. The two most important factors that have been ignored in this study are the learning of new tasks and the evolutionary history of PFC. Learning new tasks is often considered to be updating or creation of task sets.¹⁸ This is one of the hard tasks the PFC has to solve, making an influence on the structure of the PFC likely. Reynolds and O'Reilly¹⁹ already started exploring this direction and suggested that a hierarchical PFC makes learning easier, by supporting generalization of knowledge or task sets to new domains. Unfortunately, learning is out of the scope of our study because ESNs probably learn very differently from PFC. Regarding the evolutionary history, the model by Hazy, Frank and O'Reilly^{20,21} seems to be on the right track already: they explicitly propose the PFC to be an evolutionary extension to the Basal Ganglia.

Finally, it is important to note that we are far from certain how the PFC exactly exerts cognitive control. We assumed that the tasks for the ESNs correspond with the task that the PFC has to solve and concluded that for these tasks, information integration and separation was relevant. However, an important difference is that while we used symbolic input nodes, the PFC might not. This may affect the results, especially because the number of relevant input nodes is one of the identified critical factors. Neural representation remains an active research topic and as yet there are few conclusive answers. Additionally, the natural tasks that have shaped the PFC through evolution are very different from the artificial tasks used in imaging studies and in this modelling work. Thus, we have to be careful when making claims about the causes of structure and organization in the PFC.

5. Conclusion and Future Research

In this article, we argued that the requirements of information integration and information preservation for cognitive control tasks can explain the topological mapping of policy and temporal abstraction levels over the prefrontal cortex. We supported this hypothesis with results from Echo State Networks: in abstract control tasks that require mainly preservation of information, topological processing was advantageous, while uniform processing benefited less abstract tasks that require mainly integration of information. This leads to a mapping over the PFC because of wiring economy: integration of information requires high connectivity while preservation of information requires less connectivity, leading to clustering at different abstraction levels. This result can help us to interpret neuroimaging studies: based on our results, we can expect a functional gradient over the PFC with temporal and policy abstraction levels taking similar roles. This contradicts the recent proposal in Ref. 3. However, we also mentioned some important limitations of our study. To address these limitations and to gain a better understanding of the PFC, we would like to suggest a number of avenues for future research.

First of all, it is important that the imaging results that have been obtained so far are replicated and extended. This could resolve much of the ambiguity of the different conclusions drawn in Refs. 2,10 and 3 and the present article. Special attention should be paid to the details of the tasks and paradigm: most imaging studies so far mixed different types of abstraction, making a direct comparison with other studies difficult.⁶ Moreover, the way in which participants solve tasks should be controlled, because in some cases it is possible to solve an abstract task without using abstraction or *vice versa*. Finally, it may be possible to quantify the amount of control that is necessary using the information-theoretic approaches of either Koehlin and Summerfield²² or Tononi and Sporns.²³

A second way forward is to integrate aforementioned imaging studies and our own work in a broader context. For example, recently it was found that healthy people show a hierarchical organization of the multimodal network (which includes DLPFC and other prefrontal regions), while this hierarchy was reduced in people with schizophrenia.²⁴ Interestingly, schizophrenia also compromises the performance on certain cognitive control tasks. Particularly, the masking paradigm (in which a very short stimulus is preceded or succeeded by a mask) shows that patients have a deficit in top-down control of visual information processing.²⁵ Other diseases, like Alzheimer's disease, may be investigated in a similar way.²⁶ Thus,

it may be possible to link control deficits with changes in PFC organization by using an ESN-based model. Recent developments in physiological studies and the publication of connectomes may help these efforts.²⁷

Finally, interesting and relevant connections can be made with developmental neuroscience²⁸ and evolutionary biology.²⁹ As the infant PFC takes quite some time to develop and achieve adult level of cognitive control, this gives good opportunities to investigate the construction of the hierarchy in the PFC. On a longer timescale, we can study the conditions in which a functional gradient in a hierarchical network will evolve using an evolutionary algorithm.^{16,30} Of particular interest in this line of research is the proposal that different types of abstraction may co-occur in nature.⁴

Acknowledgments

This work was supported by the European Project REPLICATOR within the work program “Cognitive Systems, Interaction, Robotics” under grant agreement number 216240.

References

1. E. Koechlin and A. Hyafil, *Science* **318**, p. 594 (2007).
2. D. Badre and M. D’Esposito, *Journal of Cognitive Neuroscience* **19**, 2082 (2007).
3. J. Reynolds, R. O’Reilly, J. Cohen and T. Braver, *PloS One* **7**, p. e30284 (2012).
4. M. Botvinick, *Philosophical Transactions of the Royal Society B: Biological Sciences* **362**, p. 1615 (2007).
5. R. O’Reilly, *Trends in Neurosciences* **33**, 355 (2010).
6. D. Badre and M. D’Esposito, *Nature Reviews Neuroscience* **10**, 659 (2009).
7. J. Allman, *Evolving Brains* (W. H. Freeman, New York, 2000).
8. C. Cherniak, *The Journal of Neuroscience* **14**, 2418 (1994).
9. O. Sporns, *Networks of the Brain* (The MIT Press, 2010).
10. E. Koechlin, C. Ody and F. Kouneiher, *Science* **302**, p. 1181 (2003).
11. K. Christoff, K. Keramatian, A. Gordon, R. Smith and B. Madler, *Brain Research* **1286**, 94 (2009).
12. D. Krawczyk, M. Michelle McClelland and C. Donovan, *Cortex* **47**, 588 (2011).
13. P. Dominey, *Biological Cybernetics* **73**, 265 (1995).
14. H. Jaeger, *The “Echo State” Approach to Analysing and Training Recurrent Neural Networks — with an Erratum Note*, tech. rep., German National Research Center for Information Technology (2001).
15. R. Tukker, Echo state networks for hierarchical cognitive control, Master’s thesis, Donders Institute, Radboud University Nijmegen (2012).
16. R. Paine and J. Tani, *Adaptive Behavior* **13**, p. 211 (2005).

17. Y. Yamashita and J. Tani, *PLoS Computational Biology* **4**, p. e1000220 (2008).
18. K. Sakai, *Annu. Rev. Neurosci.* **31**, 219 (2008).
19. J. Reynolds and R. O'Reilly, *Cognition* **113**, 281 (2009).
20. R. O'Reilly and M. Frank, *Neural Computation* **18**, 283 (2006).
21. T. Hazy, M. Frank and R. O'Reilly, *Philosophical Transactions of the Royal Society B: Biological Sciences* **362**, p. 1601 (2007).
22. E. Koechlin and C. Summerfield, *Trends in Cognitive Sciences* **11**, 229 (2007).
23. G. Tononi and O. Sporns, *BMC Neuroscience* **4**, p. 31 (2003).
24. D. Bassett, E. Bullmore, B. Verchinski, V. Mattay, D. Weinberger and A. Meyer-Lindenberg, *The Journal of Neuroscience* **28**, 9239 (2008).
25. C. Gilbert and M. Sigman, *Neuron* **54**, 677 (2007).
26. O. Sporns, *Frontiers in Computational Neuroscience* **5** (2011).
27. O. Sporns, *Annals of the New York Academy of Sciences* **1224**, 109 (2011).
28. K. Supekar, M. Musen and V. Menon, *PLoS Biology* **7**, p. e1000157 (2009).
29. R. Dunbar and S. Shultz, *Philosophical Transactions of the Royal Society B: Biological Sciences* **362**, 649 (2007).
30. J. Bullinaria, *The Importance of Neurophysiological Constraints for Modelling the Emergence of Modularity*, in *Computational Modelling in Behavioural Neuroscience: Closing the Gap Between Neurophysiology and Behaviour*, eds. D. Heinke and E. Mavritsaki (Psychology Press, 2009), pp. 187–208.