From concrete to abstract rules : A computational sketch

Snigdha Dagar, Frédéric Alexandre, Nicolas P. Rougier

HAL Id: hal-03695814
https://hal.inria.fr/hal-03695814

Submitted on 15 Jun 2022

From concrete to abstract rules : A computational sketch

Snigdha Dagar, Frederic Alexandre, and Nicolas Rougier

Inria Bordeaux Sud Ouest, University of Bordeaux, CNRS

Abstract. A multi-dimensional stimulus can elicit a range of responses depending on which dimension or combination of dimensions is considered. Such selection can be implicit, providing a fast and automatic selection, or explicit, providing a slower but contextualized selection. Both forms are important but do not derive from the same processes. Implicit selection results generally from a slow and progressive learning that leads to a simple response (concrete / first-order) while explicit selection derives from a deliberative process that allows to have more complex and structured response (abstract / second-order). The prefrontal cortex (PFC) is believed to provide the ability to contextualize concrete rules that leads to the acquisition of abstract rules even though the exact mechanisms are still largely unknown. The question we address in this paper is precisely about the acquisition, the representation and the selection of such abstract rules. Using two models from the literature (PBWM and HER), we explain that they both provide a partial but differentiated answer such that their unification offers a complete picture.

Keywords: Cognitive control · Prefrontal cortex · Computational model · Abstract rules

1   Introduction

Two main strategies are generally reported for the selection of behavior [5,6]. On the one hand, implicit memory elaborated by slow learning processes can generate a rigid behavior (also called default behavior), robust in stable worlds, easy to generate but difficult to quickly adapt to changes. On the other hand, explicit memory manipulating models of the world can be used for the prospective and explicit exploration of possible behaviors, yielding a flexible and rapidly changing strategy, where behavioral rules can be associated to contexts and selected quickly as the environment changes. In the simplest case, this means learning rules defined as associations between an object's properties and a direct response. Such rules can be called concrete, while more complex or abstract rules may involve the learning of second order relations on top of the first-order rules. The prefrontal cortex (PFC) is believed to provide the ability to contextualize concrete rules that leads to the acquisition of abstract rules [6]. Considering the number of contexts we encounter every day and the ease with which we select appropriate strategies for each, some relevant questions arise : How do we represent these strategies or rules and how do we determine which one is appropriate ? An important way of understanding how the PFC supports contextual learning and implements cognitive control is thus to understand how its representations are organized and manipulated.

There is sufficient evidence to suggest that the PFC is organized hierarchically [3] with more caudal areas learning first-order associations and more rostral areas putting them in context to facilitate learning of abstract rules. This can be done by top-down modulation in the PFC, which underlies the ability to focus attention on task-relevant stimuli and ignore irrelevant distractors, in two ways : either as a result of weight changes in modulated pathways and predictions, or through activation-based biasing provided by a working memory system. These mechanisms have been explored in two prominent models of the PFC. One well established model for cognitive control function through the working memory system is the Prefrontal cortex and Basal ganglia Working Memory model (PBWM) [10] in which a flexible working memory system with an adaptive gating mechanism is implemented. At the biological level, the model proposes that the PFC facilitates active maintenance for sustaining task-relevant information, while the Basal Ganglia (BG) provides the selective gating mechanism. A hierarchical extension of this model [7] proposes that hierarchical control can arise from multiple such nested frontostriatal loops (loops between the PFC and the BG). The system adaptively learns to represent and maintain higher order information in rostral regions which conditionalize attentional selection in more caudal regions.

A second hierarchical model, Hierarchical Error Representation (HER) [1], explains cognitive control in terms of the interaction between the dlPFC (dorsolateral prefrontal cortex) and the mPFC

(medial part of the PFC). The dlPFC learns to maintain representations of stimuli that reliably co-occur with outcome prediction error and these error representations are used by the mPFC to refine predictions about the likely outcomes of actions. The error is broadcasted through the PFC in a bottom-up manner, and modulated predictions from top-down facilitate selection of an appropriate response. Thanks to its recursive architecture, this model, presented in more details below, can elaborate hierarchical rules on the basis of learning by weight updating, both to select pertinent stimuli and to map a representation inspired with principles of predictive coding [2]. In addition to its elegant recursive mechanism, proposing an original computational mechanism to account for the hierarchical structure of the PFC, the HER model is also very interesting because its proposes to decompose the functioning of the PFC between, on the one hand, the prediction of the outcome and the monitoring of the error of prediction and, on the other hand, the elaboration of contextual (and possibly hierarchical) rules to compensate errors. This distribution of functions has also been reported between respectively the medial and lateral parts of the PFC [6], yielding more importance to the biological plausibility of the HER model. For these reasons, the HER model could be presented as a more elaborated and accurate model of the PFC, except for one point of discussion that we put forward here. All the adaptations of the HER model are made through learning by weight modifications, whereas the property of working memory of the PFC, as it is for example exploited in the PBWM model, is often presented as a key mechanisms for its adaptive capabilities. An important question is consequently to determine up to which point working memory and attentional modulations are necessary for the learning of hierarchical rules in cognitive control.

In the work presented here, we seek to answer specific questions about the nature of top-down modulation and selective attention, through the lens of hierarchical learning and representations. We start from the implementation of the hierarchical HER model and its study for a task in which individual first-order rules can be learned alone or associated within specific contexts to form second-order rules. We then evaluate the performances of the HER model in these different cases and compare them with a case where an attentional mechanism should be deployed to facilitate and orient its learning. As discussed in the concluding part, we observe that the attentional mechanism should be considered not only for the processing of information but also for the learning of rules, particularly in the hierarchical and contextual case.

2   Methods

This section first summarizes the HER model algorithm and equations, as described in the original paper [1] and subsequently presents the task that we have chosen for our study.

2.1   Model Details : HER

Working Memory Gating At each level of the hierarchy, external stimuli presented to the model may be stored in WM based on the learned value of storing that stimulus versus maintaining currently active WM representations.

External stimuli are represented as a vector $\mathbf{s}$, while internal representations of stimuli are denoted by $\mathbf{r}$. The value of storing the stimulus represented by $\mathbf{s}$ in WM versus maintaining current WM representation $\mathbf{r}$ is determined as :

$$\mathbf{v} = \mathbf{X^T s} \tag{1}$$

where $\mathbf{X}$ is a matrix of weights associating the external stimuli ($\mathbf{s}$) with corresponding WM representations ($\mathbf{r}$).

The value of storing stimulus $s_i(v_i)$ is compared to the value of maintaining the current contents r_j of WM (v_j) using a softmax function :

$$\text{probability of storing} s_i = \frac{(exp^{\beta v_i} + bias)}{(exp^{\beta v_i} + bias) + exp^{\beta v_j}} \tag{2}$$

Outcome Prediction Following the update of WM, predictions regarding possible responses and outcomes are computed at each hierarchical layer, using a simple feedforward network :

$$\mathbf{p} = \mathbf{W^T r} \tag{3}$$

where $\mathbf{p}$ is a vector of predictions of outcomes and $\mathbf{W}$ is a weight matrix associating $\mathbf{r}$ and $\mathbf{p}$

Top-Down Modulation Beginning at the top of the hierarchy, predictions are used to modulate weights at inferior layers and modulated predictions are computed, as shown with the red arrows in figure 1.

For a given layer, the prediction signal $\mathbf{p'}$ additively modulates stimulus-specific predictions $\mathbf{p}$ generated by the lower layer. In order to modulate predictive activity, $\mathbf{p'}$ is reshaped into a matrix $\mathbf{P'}$ and added to $\mathbf{W}$ in order to generate a modulated prediction of outcomes :

$$\mathbf{m} = (\mathbf{W} + \mathbf{P'})^{\mathbf{T}} \mathbf{r} \tag{4}$$

These modulated predictions are then used to modulate predictions of additional inferior layers (if any exist)

$$\mathbf{m} = (\mathbf{W} + \mathbf{M'})^{\mathbf{T}} \mathbf{r} \tag{5}$$

Response Selection Actions are learned as response-outcome conjunctions at the lowest layer of the hierarchy. To select a response, the model compares the modulated prediction of correct feedback to the prediction of error feedback, for each candidate response :

$$u_{response} = m_{response/correct} - m_{response/error} \tag{6}$$

This is then used in a softmax function to determine a response :

$$Prob(u_i) = \frac{\exp^{\gamma u_i}}{\sum \exp^{\gamma \mathbf{u}}} \tag{7}$$

Bottom-up Process Following the model's response, it is given feedback regarding its performance and two error signals are computed at the bottom most hierarchical layer, one comparing the unmodulated predictions to the outcome :

$$\mathbf{e} = \mathbf{a}(\mathbf{o} - \mathbf{p}) \tag{8}$$

and another comparing the modulated predictions to the outcome :

$$\mathbf{e} = \mathbf{a}(\mathbf{o} - \mathbf{m}) \tag{9}$$

where $\mathbf{o}$ is the vector of observed outcomes and $\mathbf{a}$ is a filter that is 0 for outcomes corresponding to unselected actions and 1 everywhere else.

The outer product of the first error signal and the current contents of the WM at the bottom level is used as the feedback signal for the immediately superior layer where this process is repeated (figure 1).

$$\mathbf{O'} = \mathbf{re^T} \tag{10}$$

Effectively, at the second layer, the outcome matrix is a conjunction of stimuli, actions and outcomes. This matrix is reshaped into a vector $\mathbf{o'}$ and used to compute the prediction error at the superior layers :

$$\mathbf{e'} = \mathbf{a'}(\mathbf{o'} - \mathbf{p'}) \tag{11}$$

Weights Updating The second error signal is used to update weights within the bottom-most hierarchical layer, it updates the weights connecting the WM representation to prediction units ($\mathbf{W}$), as well as weights in the WM gating mechanism ($\mathbf{X}$):

$$\mathbf{X_{t+1}} = \mathbf{X_t} + (\mathbf{e_t^T W_t} \cdot \mathbf{r_t})\mathbf{d_t^T} \tag{12}$$

An eligibility vector $\mathbf{d}$ is used instead of the stimulus vector $\mathbf{s}$. When a stimulus i is presented, the value of d_i is set to 1, indicating a currently observed stimulus and at each iteration of the model, $\mathbf{d}$ is multiplied by a constant decay parameter indicating gradually decaying eligibility traces.

$$\mathbf{W_{t+1}} = \mathbf{W_t} + \alpha(\mathbf{e_t r_t^T}) \tag{13}$$

(a)   (b)
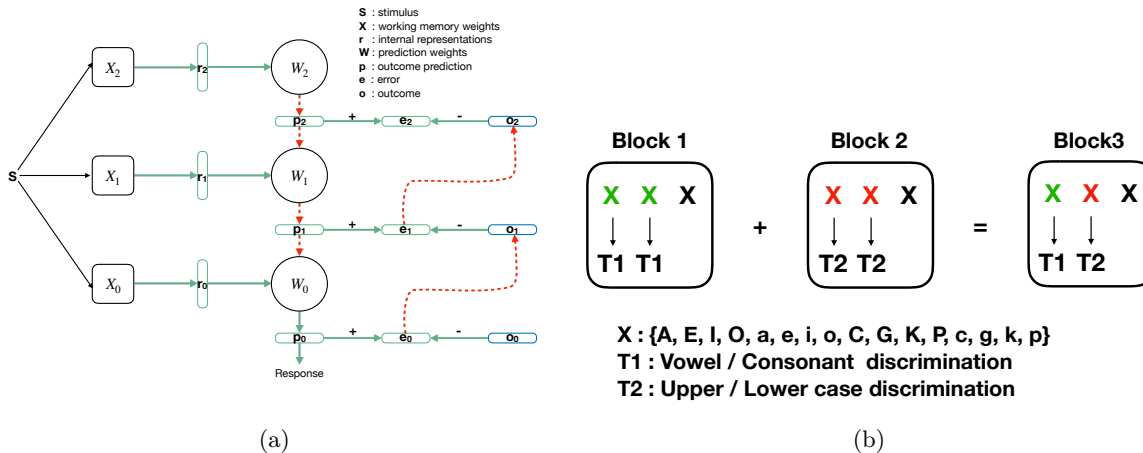
Fig. 1: **(a)** Model Schematics : Figure adapted from [1] **(b)** Task schematics : Figure adapted from [8]

2.2 Task

To design our task, we consider the framework introduced by [8] which is composed of three subtasks where the stimuli are letters having three dimensions: color (red, green or black), case (upper or lower) and sound (vowel or consonant). In the first subtask (Block 1 in figure 1(b)), black color indicates to ignore the stimulus and green color indicates to discriminate the case (rule T1: left button for upper, right button for lower). In the second one (Block 2 in figure 1(b)), black color indicates to ignore the stimulus and red color indicates to discriminate the sound (rule T2: left button for vowel, right button for consonant). The third one (Block 3 in figure 1(b)) is a random mix of trials from the other two blocks. This framework is interesting because, whereas rules T1 and T2 in blocks 1 and 2 require the subject to attend to a single dimension of the stimulus, block 3 requires to pay attention to both and to decide which rule to apply based on the third (contextual) dimension. Let us also mention here that, while there is no apparent difficulty with such tasks, it is actually harder than it appears depending on the way a task is learnt. During block 1, one can either learn the rule : "green means case and black ignore" or the rule: "black ignore, else case". The same is true for block 2 with sound. If we now consider block 3 and depending on how a subject learnt the first two blocks, she may succeed or fail immediately. In this latter case, this means block 3 cannot exploit previous learning and has to be (re)learnt.

The original task was cued by instruction and corresponding performances were reported in the paper [8]. Here, we wish to explore the inherent capability of a model to learn an abstract and hierarchical rule task without instructional cues, as in the paradigm reported by [4] and also to consider how the hierarchy can be learnt, depending on how information is represented in the model. We used two types of learning paradigms for the simulations : the first paradigm in which rules T1 and T2 were learned one after the other, and the performance of the model was then tested on random trials interleaved from rule T1 and T2 (to say it differently, we apply successively block 1, 2 and 3). In

the second paradigm, an entire abstract rule that we call T3, corresponding to the selection on rules T1 and T2 depending on the contextual cue 'color' was directly learned (block 3 applied first) and performance of the model was subsequently tested on rule T1 and T2 (blocks 1 and 2). In the next section, we report performances observed with the HER model and with an adapted version that we propose subsequently.

3   Results

We have first studied how the HER model, as it has been designed (cf section 2.1), can address the tasks defined above, under the two mentioned paradigms (cf section 2.2). Due to the design of the HER model, each layer can only map or process one stimulus value, thus requiring as many layers as there are stimulus dimensions. The mapping in the model is also highly sensitive to the stimulus dimensions relative to one another, particularly higher-dimensional stimulus are preferentially mapped onto the lowest hierarchical layer. This rests on the assumption that stimulus dimensions better able to predict and reduce uncertainty about the response are mapped to lower layers.

This may not always be the case in real life situations though. We often have to adapt and generalize the same rules over several different contexts. In the task we consider as well, the context is determined by the color, which has 3 possible values - one of which always maps to the same response (to ignore) and the other 2 determine the response based on other stimulus dimensions.

3.1   Learning Curves

Performance observed for the first and second learning paradigms are reported in figures 2 (a) and (b) respectively. We see in the figure 2 (b) that due to its hierarchical structure, when there is an underlying abstract rule to learn (rule T3), the model is able to use the hierarchical information to acquire the rule while retaining performance in each of the sub-rules (Rule T1 and T2). It does so by monitoring an "error of errors" at each hierarchical layer, broadcasting this error to superior layers (bottom-up processing) that put it in context with the stimulus feature being attended to and finally sends this prediction information to the lower layers (top-down modulation) which are able to then select the appropriate response. In the figure 2 (a), we show that when the composite rules are first learnt sequentially, the model is not able to compose them into a single rule, but instead has to relearn its representations to reach optimal performance.

Next we show that due to the design of the model, a task which has only one level of hierarchy, such as the one considered here, can not be learnt with a model with 2 layers. In figure 2 (c) we see that with 2 layers, the model is able to learn the subparts of the rule (rules T1 and T2), but performance on the composite rule T3 saturates at 80%. By exploiting the gating mechanism, each sub-rule can be learnt individually by gating the 2 relevant feature dimensions at the 2 layers (color, vowel/consonant for rule T1 and color, lower/upper case for rule T2). However, in the third rule T3 when the 2 relevant features change from trial to trial to determine the correct response, the model fails to learn, since the contextual stimulus features don't provide top-down information about "which" other stimulus feature to attend to at the lower layer.

3.2   Gating weights

In the model, the gating weights determine both, when to update or maintain a stimulus feature, and also which of the stimulus features is to be gated. We observed the adjusted weights after each rule that is learned. In the first block, vowel, consonant and black have high values of getting updated at the lowest layer, while in rule T3 all the "lower level" cues have high values of getting updated. In such a case, there is again competition between which one of them to gate, and both can win with close probabilities, in the absence of any information from the superior layers. Depending on what is gated into the top two layers, any of those mappings could emerge.

3.3   Prediction weights

The prediction weights at layer 0 are Stimulus-Action-Outcome conjugations and the gating mechanism determines which stimulus and in turn which action-outcome association is to be selected. The selected

Fig. 2: Performance of the model with 3 layers for the two paradigms **(a, b)**, plotted as an average over 100 runs, only for the runs that reached convergence criteria. The convergence criteria was defined as having a performance greater than 85% in the last 200 trials. **(c)** Performance for the model with 2 layers on the first learning paradigm.

associations are then modulated by superior layers and used the determine the response. At layer 1, the prediction errors of layer 0 are contextualized to make SxSxAxO conjugations and so on.

In the task considered for all our simulations, there are 5 concrete rules or S-A-O predictions to learn : Black - Action3, Vowel, Lower case - Action1 and Consonant, Upper case - Action2 (figure 1 (b)). In figure 3, we present examples of how a model with 3 layers selects a response by additive prediction modulation. We observed that elaborating a mapping between the stimulus and what is gated into the internal representation ($\mathbf{r}$) at different layers could be done in different ways, including randomly, as long as these mappings led to orthogonal and mutually exclusive activation of predictions (in $\mathbf{W}$). For example, in figure 3 (e), in Block 2, the color red was not gated into the internal representation, but the random gating of the other 2 dimensions still led to an appropriate modulated prediction that could initiate the correct response.

3.4   New Model

To explain the deficit of attentional mechanism in the HER model, and illustrate the advantage of our proposal, we performed some simple simulations. The model was trained individually on the two discrimination tasks ie, on the two concrete rules (T1 - vowel/consonant and T2 - lower/upper case), to obtain prediction weights or Stimulus-Action-Outcome associations as in figure 4 (b). We tested the ability of the HER model with 2 layers, to use this information and contextualize it to learn the abstract rule. The bottom layer of the model was initialized to the predictions previously learned and moreover, it was "frozen" such that no learning happened at this level, implying that these behaviors were rigid. At the upper layer, the gating weights were biased to update the internal representation with the context, which was the color in this case, implying saliency to previously unattended cues. As expected, the model failed to learn the abstract rule with these modifications. With the modified model, we used the same protocol ie the bottom layer was kept frozen, and there was a bias added

Prediction values for LCW in Block1

	R1/C	R1/E	R2/C	R2/E	R3/C	R3/E
L2 - C	0.00	0.00	0.00	0.00	-0.03	0.01
L1 - W	0.00	0.00	0.00	0.00	0.03	0.01
L0 - W	0.00	0.10	0.00	0.10	1.00	-0.02
L0 - modulated	0.00	0.10	0.00	0.10	1.00	0.00

(a)

Prediction values for LCW in Block2

	R1/C	R1/E	R2/C	R2/E	R3/C	R3/E
L2 - C	0.00	0.00	0.00	0.00	-0.00	0.13
L1 - W	0.00	0.00	0.00	0.00	0.00	0.10
L0 - W	0.00	0.10	0.00	0.10	1.00	-0.22
L0 - modulated	0.00	0.10	0.00	0.10	1.00	0.00

(b)

Prediction values for UVG in Block3

	R1/C	R1/E	R2/C	R2/E	R3/C	R3/E
L2 - V	0.08	0.33	-0.01	0.02	-0.04	0.05
L1 - G	0.18	-0.34	-0.01	0.04	-0.06	0.09
L0 - V	0.74	0.00	0.62	0.00	0.42	0.00
L0 - modulated	1.00	0.00	0.62	0.06	0.42	0.14

(c)

Prediction values for UVG in Block1

	R1/C	R1/E	R2/C	R2/E	R3/C	R3/E
L2 - V	0.23	-0.03	-0.17	0.19	0.00	0.00
L1 - G	0.04	0.03	-0.01	0.07	0.00	0.01
L0 - G	0.70	0.04	0.68	0.08	0.00	0.56
L0 - modulated	0.96	0.04	0.68	0.34	0.00	0.57

(d)

Prediction values for UVR in Block2

	R1/C	R1/E	R2/C	R2/E	R3/C	R3/E
L2 - V	-0.01	0.01	0.18	0.12	0.24	0.04
L1 - U	0.01	0.10	0.33	-0.19	-0.10	0.19
L0 - U	0.45	0.24	0.55	0.07	0.24	0.26
L0 - modulated	0.45	0.36	1.00	0.07	0.38	0.50

(e)

Prediction values for UVR in Block3

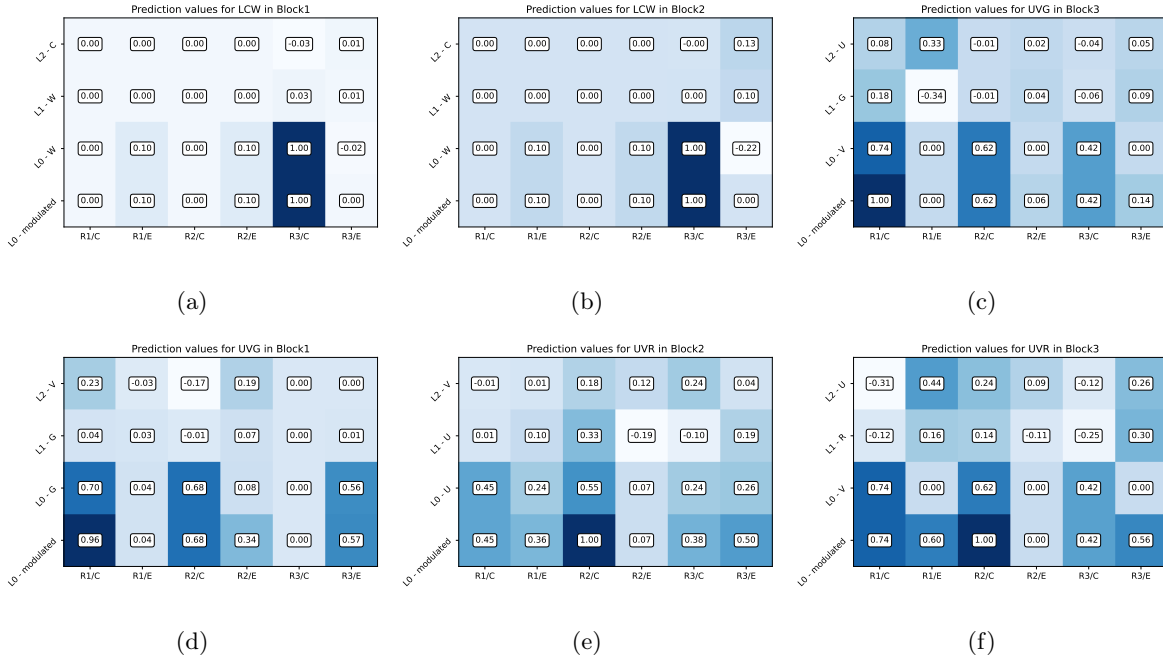|	R1/C	R1/E	R2/C	R2/E	R3/C	R3/E
L2 - U	-0.31	0.44	0.24	0.09	-0.12	0.26
L1 - R	-0.12	0.16	0.14	-0.11	-0.25	0.30
L0 - V	0.74	0.00	0.62	0.00	0.42	0.00
L0 - modulated	0.74	0.60	1.00	0.00	0.42	0.56

(f)

Fig. 3: Examples of how the model solves different cases of stimuli. The matrix shows the prediction values at different layers (first 3 rows), given the internal representation of the stimulus, and how they are modulated additively (row 4) to give the final Action-Outcome predictions that are used for response selection. **a, b** show the case when the stimulus is black, in rules 1 and 2 respectively. **d, e** show the case when the stimulus is Green, Vowel (rule T1) and Red, Upper case (rule T2). **c, f** show the case for Green, Vowel and Red, Upper case in rule T3

to the upper layer to encourage gating of the color. However, instead of an independent gating at the bottom layer, we included an output gating from the upper layer, which used the prediction errors at the upper layer to select which stimulus dimension was going to be gated into the bottom layer (figure 4 (a)). To put it more generally, the bottom layer was responsible for response selection while the upper layer was responsible for action-set selection through targeted attention (cf [6] for more details about the structuring concept of action-set and its role in PFC information processing). Our modified model achieved optimal performance fairly quickly, as shown in 4 (c).

4 Discussion

The PFC plays a major role in cognitive control and particularly for learning, selecting and monitoring hierarchical rules. For example, in experimental paradigms, discrimination or categorization tasks can be considered as first-order rules which could be learned individually. However, when conflicting stimuli are presented simultaneously, a contextual cue is needed to identify which of the first order rules is to be applied, thus forming second-order rules.

The inner mechanisms of the PFC have been studied in computational models and among them, the property of working memory used for biasing by selective attention in the PBWM model and, more recently in the HER model, the separation between outcome prediction error monitoring, and hierarchical rule learning. Considering the indisputable progress brought by the design of the HER model, we questioned whether it was now a standalone model of the PFC to be used in any circumstances or if the contribution of certain mechanisms like selective attention was still to be considered in some cases and possibly added to the general framework of PFC modeling. More specifically, considering the deployment of cognitive control in realistic behavioral tasks and considering that most hierarchical representations arise from the intersection between agents and the problems they face, and are created over time in a learning process, in a rapid and flexible way, our question was to know if the HER model could account for this kind of process.
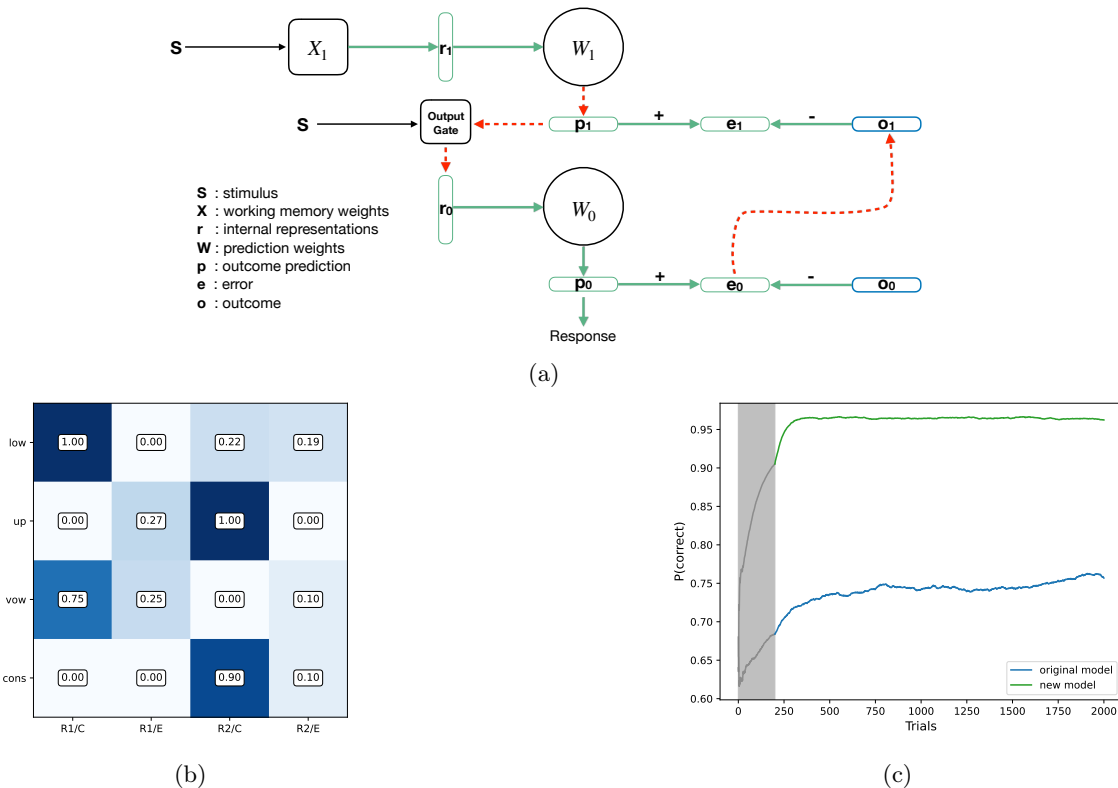
(a)

(b)

(c)

Fig. 4: **(a)** The modified model with output gating from layer 1. The gating weights in layer 1 (X_1) learn over time to gate the context into r_1. The selected prediction units from layer 1 (p_1) are then used to make a decision on which value of the stimulus s is gated into r_0 (the output gate). **(b)** Prediction weights (W_0) for the concrete rules at layer 0. These weights are pre-learned by training the model with rules T1 and T2, independently. **(c)** Performance of the original model compared to the modified model over a 100 runs, when layer 0 is fixed to the weights in figure (b) and only layer 1 prediction weights (W_1) and gating weights (X_1) are learned.

Using a task elaborated along two paradigms, we show that, when concrete rules are already learnt and need to be contextualized, the use of a biasing selective attention mechanism is more effective than modulated weights changes in displaying effective cognitive control. When concrete rules are acquired first, superior layers must learn to select the appropriate concrete rule by targeted attention, rather than by relearning representations. We observe that a subject can perform optimally on a given task even though she uses a different rule representation compared to the *official* one. On a single task, this has no consequence and there is actually no way to know which exact rule is used internally. However, when this rule needs to be composed with another rule such as to form a new rule, this may pose problem and lead to bad performance. This has been illustrated on the task: if a subject uses any of the alternative rules for tasks T1 or T2, she'll be unable to solve task T3 even though this task is merely made of a mix of T1 or T2 trials. The reason for the failure of the HER model in this case is to be found in the failure to attend the relevant dimension of the task, here, color, thus claiming for considering and incorporating this mechanism to a versatile PFC model. Analyzing these results in a more general view, we can remark that most experimental paradigms that study hierarchy break down the complexity of a task by providing instructional cues to the participant. Even in studies with rodents and non-human primates, shaping is used in learning paradigms to enable the learning of complex or abstract rules. In developmental learning, this kind of shaping is called curriculum learning. It is evident that such breaking down of complexity must facilitate the acquisition of abstract rules, and hence modeling approaches must demonstrate these behavioral results.

From a more conceptual point of view, the term *hierarchy* can be used in many different ways, two common ones being *processing hierarchies* and *representational hierarchies*. In the first, higher levels exert control over lower levels, for example by controlling the flow of information or by setting the

agenda for lower levels [9]. In the second one, higher levels form abstractions over lower levels, such that lower levels contain concrete, sensory and fine-grained information, whereas higher levels contain general, conceptual and integrated information [3], [11]. It is thus important that a model of the PFC to exploit both views, suggesting to incorporate an attentional mechanism for the flexible and controlled design of hierarchical rules from previously learned concrete rules, as we proposed in the new model sketched here.

References

1. Alexander, W.H., Brown, J.W.: Hierarchical error representation: a computational model of anterior cingulate and dorsolateral prefrontal cortex. Neural Computation **27**(11), 2354–2410 (2015)
2. Alexander, W.H., Brown, J.W.: Frontal cortex function as derived from hierarchical predictive coding. Scientific Reports **8**(1), 3843 (2018). https://doi.org/10.1038/s41598-018-21407-9
3. Badre, D.: Cognitive control, hierarchy, and the rostro–caudal organization of the frontal lobes. Trends in cognitive sciences **12**(5), 193–200 (2008)
4. Badre, D., Kayser, A.S., D'Esposito, M.: Frontal cortex and the discovery of abstract action rules. Neuron **66**(2), 315–326 (2010)
5. Daw, N.D., Niv, Y., Dayan, P.: Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nature neuroscience **8**(12), 1704–1711 (Dec 2005). https://doi.org/10.1038/nn1560, http://dx.doi.org/10.1038/nn1560
6. Domenech, P., Koechlin, E.: Executive control and decision-making in the prefrontal cortex. Current Opinion in Behavioral Sciences **1**, 101–106 (Feb 2015). https://doi.org/10.1016/j.cobeha.2014.10.007, http://dx.doi.org/10.1016/j.cobeha.2014.10.007
7. Frank, M.J., Badre, D.: Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 1: computational analysis. Cerebral cortex **22**(3), 509–526 (2012)
8. Koechlin, E., Ody, C., Kouneiher, F.: The architecture of cognitive control in the human prefrontal cortex. Science **302**(5648), 1181–1185 (2003)
9. Miller, E.K., Cohen, J.D.: An integrative theory of prefrontal cortex function. Annual review of neuroscience **24**(1), 167–202 (2001)
10. O'Reilly, R.C., Frank, M.J.: Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. Neural computation **18**(2), 283–328 (2006)
11. Tenenbaum, J.B., Kemp, C., Griffiths, T.L., Goodman, N.D.: How to grow a mind: Statistics, structure, and abstraction. science **331**(6022), 1279–1285 (2011)