A unified model of rule-set learning and selection

Pierson Fleischer *, Sébastien Hélie

Purdue University, United States of America

A R T I C L E   I N F O

A B S T R A C T

The ability to focus on relevant information and ignore irrelevant information is a fundamental part of intelligent behavior. It not only allows faster acquisition of new tasks by reducing the size of the problem space but also allows for generalizations to novel stimuli. Task-switching, task-sets, and rule-set learning are all intertwined with this ability. There are many models that attempt to individually describe these cognitive abilities. However, there are few models that try to capture the breadth of these topics in a unified model and fewer still that do it while adhering to the biological constraints imposed by the findings from the field of neuroscience. Presented here is a comprehensive model of rule-set learning and selection that can capture the learning curve results, error-type data, and transfer effects found in rule-learning studies while also replicating the reaction time data and various related effects of task-set and task-switching experiments. The model also factors in many disparate neurological findings, several of which are often disregarded by similar models.

© 2020 Elsevier Ltd. All rights reserved.

1. Introduction

The world is not random. The way something looks, feels, or smells can provide cues about its possible uses. An example of this would be the knowledge that the hard, inedible things inside of fruits are seeds. Objects that share the same features can often be used in the same ways. To keep up with the previous example, that knowledge about seeds could help a person identify the seeds of fruits that they had never encountered before so they could grow more. An intelligent agent can learn these consistencies and use them to build associations between an object's observed properties and its behavior. These associations then allow the agent to predict the behavior of new stimuli that share relevant features with familiar stimuli. Relevant factors can also include previous actions taken and other broader contexts such as goals. These associations are called rules. They cover a broad spectrum of conditional relations that can be as specific as "Don't eat red mushrooms with white spots" to as broad and as multifaceted as "Be polite as possible (and all the specific behaviors that entails) when interacting with customers". The first example relates a specific stimulus feature to a specific action while the second specifies a whole range of different and otherwise unrelated actions to be associated with some internal classification of customer (versus, for instance, coworker). Within the actions that make up polite behavior additional rules specify when each action should be taken. A collection of rules that are mutually exclusive and thus only one rule in the collection can apply to any given stimulus is a rule-set.

This article proposes a new biological theory of rule learning, switching, and selection. The new theory proposes that rules are regular working memory items. Working memory refers to the cognitive system responsible for short-term storage and manipulation of information in real-time. Therefore, rule switching and selection should be similar to switching and selecting working memory items. The theory is implemented using a computational model that simulates a working memory circuit centered around the prefrontal cortex using spiking neurons. The model is validated by simulating several key phenomena in rule learning, switching, and selection such as transfer, priming, and representational effects as demonstrated by Collins and Frank (2013), van t Wout, Lavric, and Monsell (2015), and Badre, Kayser, and D'Esposito (2010) respectively. The remainder of this article is organized as follows. First, the relevant biological findings are described along with how the model incorporates them. Second, a selection of modern and classic effects is replicated and compared to human data from outside studies. The article concludes with a discussion of the implications of the proposed theory for future research in the neuroscience of rule learning, switching, and selection.

1.1. Rules

Rules can cover a wide variety of situations and circumstances. In the simplest case they directly indicate a response. These kinds of rules will be referred to as *concrete* rules and they directly select a response/category, usually through discrimination of one

---

or more features of a stimulus. Concrete rules discriminate based on a single feature or a subset of features that is shared across a group of stimuli. For instance, the rule "Red stimuli are part of category B" will apply to all red stimuli regardless of shape. Concrete rules can also discriminate using multiple features of a stimulus, an example of this being "When I encounter a red traffic light while driving, I should press the brake pedal". In this case, there are two relevant stimulus features and one state-based condition: 1. traffic signal, 2. red light, 3. while driving. These rules apply to multiple stimuli including stimuli that have never been encountered before, such as traffic lights in a new city. This means that when a novel stimulus is encountered, learning can progress very rapidly or even be skipped altogether, so long as the rule-set is still valid (Hélie, Ell, Filoteo, & Maddox, 2015b).

In contrast to concrete rules, the term *abstract* rule is used to define rules that do not assign specific responses to stimuli and instead are used to select a set of concrete rules from all of the possible sets of concrete rules that could be applied to the stimulus. If one were sorting coins, abstract rules would indicate whether to sort by size or by monetary value, while concrete rules would indicate which category each coin should belong to. Any stimulus features that are used by abstract rules rather than concrete rules are defined as cues.

1.2. Physiological basis for rules

The prefrontal cortex (PFC) is commonly understood as the location of high-order processes and cognitive control (Miller, Freedman, & Wallis, 2002). Rules, especially abstract rules, fall under this category and thus investigations into the location of rules should point towards the PFC. An fMRI study by Hélie, Roeder, and Ashby (2010) found that subjects that successfully learned the rule-sets of a rule-based task exhibited corresponding activity in the PFC. In another fMRI experiment, Badre and Wagner (2004) used a task that required subjects to use rules when there was a mismatch between the primed response and the requested response. In these cases, fMRI results showed increased activity in the dorsolateral PFC (DLPFC).

Non-fMRI studies also show that the PFC has a strong association with rules. Damage to the PFC is known to cause deficits in performance in the Wisconsin Card Sorting Test (WCST) (Bunge, 2004; Fuster, 2008), which relies on rule-sets and rule-set switching. In addition, Wallis and Miller (2003) found that rule-selective firing patterns were more prevalent in the DLPFC than stimulus-selective or response-selective firing patterns. Another study by O'Reilly, Noelle, Braver, and Cohen (2002) has shown impairments in the ability to switch between rule-sets in subjects with PFC lesions.

In addition to rule representation, a large amount of research suggests that the lateral prefrontal cortex (LPFC) is associated with tasks that require abstract rule-sets and frequent rule-set switching. A few examples of such research include an fMRI study by Koechlin, Ody, and Kouneiher (2003) which found that anterior areas of the PFC were associated with greater levels of abstraction. Crone, Wendelken, Donohue, and Bunge (2005) found more activity in the LPFC when abstract rules were used, when a new rule-set is selected, and when the next rule-set that would be used was not predictable. Similarly, Zanolie, Van Leijenhorst, Rombouts, and Crone (2008) found increased activity in the dorsolateral PFC relative to medial areas during various rule-set related processes. The greatest increase in activity was found when the subject made an error due to interference from irrelevant rule-sets — meaning the association with the relevant rule-set needs to be reinforced. A smaller increase in LPFC activation occurred when a new rule-set was selected through trial and error from a finite set of possible rule-sets. The smallest

increase in activity occurred when the environment changed and the currently active rule-set was found to no longer be applicable.

Similar results are found when the anterior PFC is not fully developed. In a study by Bunge and Zelazo (2006), structural MRI scans were used to track PFC development in young children. Over the same period of time their ability to use rules was evaluated through a set of behavioral tasks. Anterior areas of the PFC develop later than posterior areas and as they develop, children become more capable of using rules in complex ways. Children go from being able to use only a single rule to multiple rules within a single rule set, until finally being able to switch between multiple rule-sets. Badre, Hoffman, Cooney, and D'Esposito (2009) found that patients with lesions in the PFC that perform categorization tasks become impaired as contingencies are added to the task. The more anterior the lesion the more abstract the contingencies have to be before impairment. Other supporting evidence can be found in review articles by Badre (2008) and Buckner (2003).

Together these findings suggest a hierarchical organization where concrete rule-sets are represented in posterior PFC regions and abstract rules are more anterior (Badre & Nee, 2018; Hélie et al., 2010; Nee & D'Esposito, 2016; Zarr & Brown, 2016). This creates a tiered system where physical responses are coded in motor and premotor areas. Anterior to these motor-related areas is the posterior PFC where the concrete rule-sets that govern the selection of motor responses are represented. Anterior to that is the anterior PFC, location of the abstract rule-sets, which selects from among the concrete rule-sets.

1.3. Physiological constraints in the PFC

The results of physiological studies of the PFC have provided several constraints and qualities that will be used in constructing the new theory.

When it comes to learning, dopamine is typically used as the reward signal in models, differentiating reward and punishment at the cellular level. Dopamine is known to modulate synaptic strength and its release is correlated with reward, making it a prime candidate for the reward signal. The problem with using this approach in the PFC is that the PFC lacks significant quantities of the protein DAT, which facilitates the rapid reuptake of dopamine (Seamans & Robbins, 2010). Without this protein dopamine lingers in the synaptic space and thus cannot provide a signal with the temporal resolution that LTP and LTD would require for the synaptic modifications to be meaningful and improve task performance (Cass & Gerhardt, 1995; Hélie, Ell, & Ashby, 2015a). Izhikevich (2007a) suggested that even a temporally imprecise dopamine signal can facilitate learning. However, Izhikivich's experiments still used a dopamine reuptake rate 5 times faster than that of the PFC and, in his instrumental conditioning experiments, there was a 10 s delay between trials which did not solve the problem of credit assignment but instead avoided it. In addition, the instrumental conditioning task only presented a single repeated stimulus over blocks of 400 trials, which enormously streamlined learning. This is problematic, not just for using dopamine as a reward signal in the PFC, but for any theory that relies on dopamine signals in the PFC having trial-level temporal resolution. Part of the motivation for the proposed new theory is the incorporation of alternative methods of reward signaling and guiding synaptic learning.

Space is also a limiting factor in the PFC. Selecting a response means considering both rule-sets and stimulus features. This creates a combinatorial space that might be too vast to be represented at the neuron level. Most representational methods require unreasonably large neuronal populations (Crawford, Gingerich, & Eliasmith, 2016). Few brain areas are as large as the PFC, thus making this objection doubly problematic for theories

that localize representational associations in other regions. While there are some neuron-level representational systems that could theoretically encode real-world quantities of information, these approaches have not yet attempted to tackle the problem of learning (Crawford et al., 2016). This being the case, alternatives to neuron-level representation were sought for use in the proposed new theory.

1.4. Behavioral investigations of rule-set switching

Rules have naturally garnered substantial interest amongst cognitive psychologists as well. A particularly common area of investigation involving rule-sets is the study of task-switching (Grant & Berg, 1948; Hübner, Futterer, & Steinhauser, 2001; Koch & Allport, 2006; Logan & Bundesen, 2003; Monsell & Mizon, 2006; Rogers & Monsell, 1995; Spector & Biederman, 1976; Stroop, 1935; Sudevan & Taylor, 1987). Since a rule-set is a collection of stimulus/response associations, switching tasks is akin to switching rule-sets. The rule-set for a color-naming task can have the same response assigned to a green frog and green tree while the plant or animal task set will have two different responses assigned to those stimuli. Subjects rarely make errors in these familiar tasks but the time it takes to make a response varies based on several factors such as: switching to a new task (Grant & Berg, 1948; Hübner et al., 2001; Jersild, 1927; Koch & Allport, 2006; Logan & Bundesen, 2003; Monsell & Mizon, 2006; Rogers & Monsell, 1995; Spector & Biederman, 1976; Stroop, 1935; Sudevan & Taylor, 1987), early knowledge of the next task (Hübner et al., 2001; Logan & Bundesen, 2003; Monsell & Mizon, 2006; Rogers & Monsell, 1995; Sudevan & Taylor, 1987), prior expectations about the next task (Alport, Styles, & Hsieh, 1994; Stroop, 1935; Sudevan & Taylor, 1987), and the number of rule-sets relevant to a given stimulus (Koch & Allport, 2006; Rogers & Monsell, 1995; Spector & Biederman, 1976; van t Wout et al., 2015). The general trends show that (1) switching to a new task, (2) using an uncommon task-set, or (3) using stimuli compatible with many task-sets, all increase reaction time. In contrast, (4) early knowledge of the next task, (5) using a common task-set, or (6) using stimuli compatible with only few task-sets, all decrease reaction time. Several of these findings, particularly switch costs and priming (1 and 4), suggest that there is a mechanism for maintaining rule-sets even when they are not currently in use.

All of the above experiments use tasks and cues that are already familiar to the subjects. Of course, this familiarity had to be developed somehow. Investigations into rule learning utilize tasks such as the WSCT (Grant & Berg, 1948), the intradimensional/extradimensional task (O'Reilly et al., 2002), and the Same/Different task (Bamber, 1969). Previous results suggest that subjects use hierarchical rule-sets to represent stimulus–response associations when there is no prior indication that rule-sets would be advantageous (Badre et al., 2010) and even when there is no advantage to using a hierarchical representation (Collins & Frank, 2013).

2. A new theory of rule-set representation and learning

2.1. Motivation for a new theory

The previous section identified four major goals for the new theory in addition to the central goal of describing rule-set learning and selection. (1) A means to solve or avoid combinatorial explosion when scaling up the feature space and number of rule-sets. (2) Use of dopamine in PFC areas limited to signals with coarse temporal resolution. (3) Some method for maintaining rule-sets across delay periods and between trials. (4) The

establishment of a hierarchical organization that explains the differences between the anterior and posterior PFC. Adherence to these constraints can be found in various models of rule learning and switching, but there is no model that simultaneously conforms to all of them.

For instance, O'Reilly et al. (2002) modeled the effects of PFC lesions on intra-dimensional vs. extra-dimensional target switching using a connectionist architecture. Subjects selected stimuli that contained a target feature and the feature could change intra-dimensionally (e.g., red to green in the 'color' dimension) or extra-dimensionally (e.g., red to circle — from the 'color' dimension to the 'shape' dimension). In the proposed terminology, an intra-dimensional shift would be a shift to a different stimulus–response pair within a rule-set while an extra-dimensional shift would be a change in rule-set. Rule-sets and stimulus features were integrated using an intermediate neural layer which would explode in size if the model were to be scaled up to handle more stimuli, responses, and stimulus–response associations. Rules were hard-coded in the model and the rule-sets were learned using cortical dopamine as the reward signal. There was no maintenance of rule-sets across trials meaning that cross-trial effects, like switch costs, could not be replicated. On a positive note, the hierarchical organization of the PFC was recognized with orbital PFC encoding features and lateral PFC encoding rule-sets.

A connectionist model by Ardid and Wang (2013) was used to simulate a number of task-switch findings including switch cost, congruency effect, task-response interaction, and single-neuron activity. Rule-sets were maintained across trials through various feedback loops within the PFC. Associations were predetermined as connections were only present between stimuli/cues and their associated categories/rule-sets. Plasticity was used to model switch costs by affecting the speed of responses. The intracortical feedback loops allowed for maintaining rule-sets across trials. However, Ashby, Ell, Valentin, and Casale (2005) noted several findings suggesting that working memory maintenance is not solely supported by the PFC. Delayed response tasks show that several brain areas outside of the PFC exhibit delay related activity, and lesions and dysfunction in the basal ganglia and thalamus have been shown to impair working memory. Phenomena such as switch-costs, interference, and repetition effects are observed in both working memory and rule-based tasks. This suggests that rules and working memory share the same cerebral structures. Like many of the other models, the Ardid and Wang model is also subject to combinatorial explosion as the number of category cells required is equal to the product of the responses and the rule-sets.

Another approach to rule-set maintenance is used in a model by Reynolds, Braver, Brown, and Van der Stigchel (2006) that simulated the residual switch-cost after priming. In the model, the rule-set is selected by the cue and stored in the PFC. Dopamine probabilistically maintains the rule-set until the stimulus is presented. The stimulus and rule-set (if maintained) form a compound representation which then determines a response. While this model does provide a mechanism for rule-set maintenance, it uses trial-level dopamine signals, so some other method must be found that adheres to all of the goals set for this theory. The model also suffers combinatorial explosion in the intermediate layer when scaled up and there is no location-based hierarchical organization. To avoid this issue of dopamine in the PFC, other models incorporate the cortico-basal ganglia (BG) loop and have the entirety of the reward-mediated learning take place in the BG, where the presence of DAT allows dopamine to function as a temporally precise reward signal. In 2013 Collins and Frank conducted a two-phase experiment where subjects had to learn associations between two-dimensional stimuli and four

responses. In the first phase there were two possible values in each of the two dimensions (2 × 2) and each feature combination was associated with a different response. In phase two new values were added in one dimension and the new combinations were assigned responses so that the subjects' performance would differ based on the way the subject represented the environment. Collins and Frank simulated these two representations using a connectionist model in which rule-sets were associated with cues using a cortico-BG loop. Another cortico-BG loop associated responses with stimulus feature/rule-set pairs. The experimental results suggested that task representations that were not rule-based were uncommon to the point of being unidentifiable, suggesting that humans naturally organize information hierarchically. The incorporation of the BG does mean that dopamine can be used as the reward signal without issue, but the Collins and Frank model is still subject to combinatorial explosion. In fact, the issue is more severe because the intermediate layer is proposed to be in the striatum, a much smaller structure than the PFC. There is also no means of rule-set maintenance and the organization does not explain the dissociation between the anterior and posterior PFC. That said, the rule selection architecture mirrors the response selection architecture. This repeating design is complementary to the hierarchical nature of the PFC.

A solution to the combinatorial problem is provided by the Heterosynaptic Inhibitory Criterion Learning (HICL) model presented in Hélie et al. (2015b). The task used in that article asked subjects to categorize lines that varied continuously in two dimensions: length and orientation. The model used the computational cognitive neuroscience approach (Ashby & Helie, 2011) with rule-sets encoded in the PFC through synaptic gating and rule-set switching governed by random-walk. A reward signal is provided by feedback sensitive cortical cells. The model reproduced learning curves as well as the effects of the different saliencies of the relevant dimension compared to the irrelevant dimension. Feature and rule-set integration was implemented at the synapse level, rather than the neuron level using pre-synaptic inhibition (Shepherd, 2004). Pre-synaptic inhibition involves an inhibitory synapse in which the post-synaptic component is not the dendritic terminals of a post-synaptic cell as is common. Instead, the inhibition is received at the axon terminals of the pre-synaptic cell and affects only the specific synapse that the terminal is part of by blocking neurotransmitter release. Pre-synaptic inhibition acts like a gate, allowing cells that represent stimulus features to selectively activate only a subset of the cells that they connect to depending on the active rule-set. This model also avoids the issue with using dopamine in the PFC by using activity of reward-sensitive neurons as the reward signal. However, the focus of the Hélie et al. model was on (intra-dimensional) criterion learning within the PFC and no biological mechanism was proposed to account for rule-set switching, selection, or maintenance.

2.2. Assembling the features of the new theory

The reviewed material sets up a framework for integrating previous work and the proposal of a new biological theory of rule switching and selection. The theory states that the PFC is hierarchically organized with abstract rules located in anterior regions and concrete rules located closer to the premotor cortex which codes physical responses. Rule-sets are stored in working memory and maintained via a positive feedback loop with the thalamus. Inhibition of the thalamus removes the active rule-set from working memory. Rule-sets are implemented as synapse-level gating allowing selection of rewarded responses while blocking punished responses. The reward signal is provided by cortical cells sensitive to reward and punishment. As is standard in the

models we have examined, responses are selected through the interaction of stimulus features and rule-sets in the posterior PFC. The rule-set selection mechanism in anterior PFC mirrors the architecture of the response selection, reflecting the observed hierarchy of the PFC and incorporating the iterative design philosophy seen in the Collins and Frank model. Integration of the rule-set and stimulus features is achieved using pre-synaptic inhibition. Each rule-set delivers a different pattern of inhibitory gating to signals from the features to the responses. This avoids the problem of combinatorial explosion by encoding the interaction of features and rule-sets in the pattern of synaptic gating rather than through an intermediate neural layer that grows proportionally to the product of the rule-sets and the responses. The inhibitory gates are plastic and can be strengthened or weakened based upon external feedback. This feedback is implemented using feedback sensitive cortical cells. These feedback cells are based on an fMRI experiment by O'Doherty, Kringelbach, Rolls, Hornak, and Andrews (2001) which found areas of the PFC that became more active with increasing reward and less active with increasing punishment. In other areas of the PFC the opposite behavior was observed. The activity of these cells is theorized to influence Hebbian learning so as to reinforce rewarded behavior and avoid punishing behavior. The term "pre-synaptic" is used extensively within the topic of neural connections and so, in an effort to avoid ambiguity, the term "non-somatic" shall be used from here on to refer to these connections that do not directly affect the spiking of the post-synaptic cell.

The Reynolds et al. model and the Ardid and Wang model used connectionist approaches to rule-set maintenance. However, The Reynolds et al. model's reliance on dopamine means that it does not meet all criteria for the proposed framework. Instead, inspiration is taken from FROST (Ashby et al., 2005), a model of working memory maintenance designed to reproduce the delayed-response task, which does not involve rule learning or task switching. In this task, the subject (typically a monkey) must remember where a reward is hidden until it is permitted to obtain it. FROST stores the reward location in the PFC and maintains that information through connection loops with the thalamus (Fig. 1). The maintenance is disabled by default and ceases when the model no longer has a need to maintain the information. In this regard, the FROST model behaves counter to what the task-switching experiments would suggest. Behavioral findings suggest that rule-sets must maintain some degree of activity after the removal of the cue and into the next trial. Since this interference is detrimental, the maintenance mechanism is likely to be active by default. To make the thalamic maintenance loop enabled by default the new theory incorporates the exterior segment of the globus pallidus (GPe). The GPe inhibits the internal segment of the globus pallidus (GPi) and, unlike the striatum, is naturally active and does not require external excitation to generate an inhibitory signal. With this the thalamus is freed and the maintenance loop is enabled by default. However, the old rule-sets are not maintained indefinitely so there must also be a way to halt the maintenance loop. Results published by Rushworth, Hadland, Paus, and Sipila (2002) and by Dove, Pollmann, Schubert, Wiggins, and von Cramon (2000) indicate that the pre-supplementary motor area (pre-SMA) is active when subjects engage in cue-mediated task switching. Rushworth et al. (2002) also show that temporary lesioning of the pre-SMA negatively affects performance only in trials involving a task-switch. The pre-SMA excites the subthalamic nucleus (Aron, Behrens, Smith, Frank, & Poldrack, 2007), which excites the GPi. The GPi overcomes the inhibition from the GPe and inhibits the thalamus, disrupting the cycle of excitatory input between the thalamus and the PFC cells that represent rule-sets.
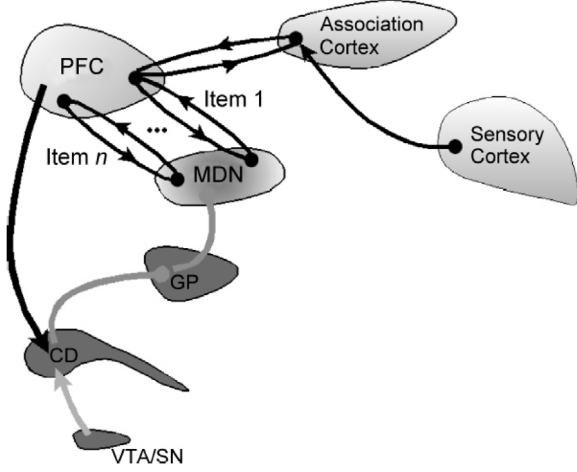
Fig. 1. The FROST model architecture showing the pathway through the basal ganglia and the parallel maintenance loops between the PFC and the thalamus. *Source:* From Ashby et al. (2005).
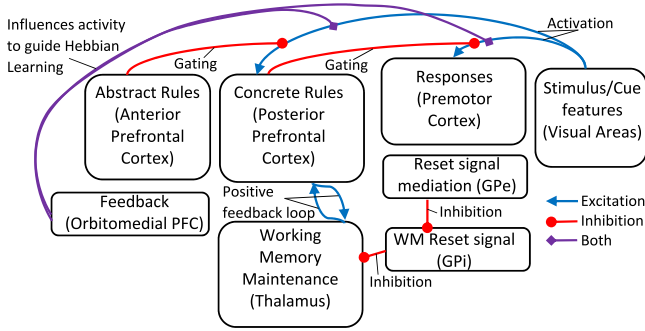

Fig. 2. Overview and connectivity diagram of the proposed theory.

3. The model

A model of the proposed theory should represent varying numbers of stimulus dimensions, stimulus features within those dimensions, responses, cues, and rule-sets. Stimulus features are represented by PFC stimulus cells, cues by PFC cue cells, responses by PMC response cells, and rule-sets by PFC rule-set cells (Fig. 2). It should be pointed out that, while this model learns to associate features of the cue dimension with rule-sets, the distinction between the cue dimension and stimulus dimension(s) is defined beforehand. The model is already assumed to "know" which stimulus dimensions can be used in abstract rules and which can be used in concrete rules. There is no mechanism in this model for discovering a representational structure for a task. Stimulus dimensions are represented in the pattern of connections from the rule-set cells to the stimulus–response synapses. Rule-set cells strongly inhibit stimulus cells outside of their target dimension while weakly and plasticly inhibiting stimulus cells in their target dimension. Depending on the task there may be more rule-set cells than the number of rule-sets necessary for perfect performance and there may be multiple rule-set cells on the same dimension. Each rule-set cell has an associated thalamic cell that connects to it. In the case where a rule-set needs to use more than one stimulus dimension, each possible combination of the features in those dimensions is represented by a single stimulus cell.

3.1. Cortical pyramidal cells

The cells' membrane potential is modeled using Eqs. (1) and (2), parameterized to model a cortical pyramidal cell (Izhikevich, 2007b):

$$V_a(t+1) = V_a(t) + [0.7[V_a(t)+60][V_a(t)+40]$$
$$- U_a(t) + E_a(t) - I_a(t) + \varepsilon(t)]/100 \tag{1}$$

$$U_a(t+1) = U_a(t) + 0.03[-2[V_a(t)+60] - U_a(t)] \tag{2}$$

where $V_a(t)$ is the membrane potential (in mV) of cell a at time t, $U_a(t)$ is the value of the recovery variable (a dimensionless representation of the slow ionic current) of cell a at time t, $E_a(t)$ is the excitatory (glutamate) input to cell a at time t, and $I_a(t)$ is the inhibitory (GABA) input to cell a at time t. $\varepsilon_a(t)$ is the noise in cell a at time t and is drawn at each time-step from a Gaussian distribution with a mean of 0 and a standard deviation of 200. The noise and both inputs are represented as mA of injected current. If Eq. (1) would result in a $V_a(t)$ greater than 35 then a is said to have spiked and the following adjustments are made: $Va(t)$ is set to -50 mV and $Ua(t)$ is increased by 100. Additionally, the spike time t is concatenated onto the end of S_a, the list of the times at which cell a spiked. Each cell also generates output at each time step according to Eqs. (3) and (4).

$$O^{Glut}{}_a(t) = \sum_{s \in S_a} \left[\frac{t-s}{60} e^{1-\left[\frac{t-s}{60}\right]} \right]^+ \tag{3}$$

$$O^{GABA}{}_a(t) = \sum_{s \in S_a} \left[\frac{t-s}{30} e^{1-\left[\frac{t-s}{30}\right]} \right]^+ \tag{4}$$

where $O^{Glut}{}_a(t)$ is the glutamate output of cell a at time t and $O^{GABA}{}_a(t)$ is the GABA output of cell a at time t. This is an abstraction of how neurons work in the brain. Most biological cells release the same neurotransmitter at every one of their synapses. Within the model, some pyramidal cells have both excitatory and inhibitory outgoing connections. These cells are stand-ins for groups of cells that include excitatory pyramidal cells and inhibitory interneurons. $f(t)^+$ equals $f(t)$ when $f(t) \geq 0$, and 0 when $f(t) < 0$. Glutamate and GABA release are modeled with separate equations because the release and re-uptake of GABA can be faster than that of glutamate (Hélie et al., 2015b). The outputs of cells are transmitted to other cells through connections. Excitatory connections are denoted in equations and the text with a solid arrow and inhibitory connections with a hollow arrow, $\underset{s}{\rightarrow}$ and $\underset{s}{\Rightarrow}$ respectively. Standard cell-to-cell connections are indicated with the letter "s" below the arrow: $\underset{s}{\rightarrow}$ or $\underset{s}{\Rightarrow}$ and non-somatic connections with an "ns": $\underset{ns}{\rightarrow}$ or $\underset{ns}{\Rightarrow}$. The effect of connections on the E and I terms of the voltage equation is described by Eqs. (5), (6), and (7):

$$E_a(t) = \sum_{b \in C} w_{(b)\underset{s}{\rightarrow}(a)} \left[O^{Glut}{}_b(t) + P_{(b)\underset{s}{\rightarrow}(a)}(t) \right]^+ \tag{5}$$

$$I_a(t) = \sum_{b \in C} w_{(b)\underset{s}{\Rightarrow}(a)} \left[O^{GABA}{}_b(t) + P_{(b)\underset{s}{\Rightarrow}(a)}(t) \right]^+ \tag{6}$$

$$P_{connection}(t) = \sum_{c \in C} O^{Glut} \cdot w_{(c)\underset{ns}{\rightarrow}(connection)}$$
$$- \sum_{c \in C} O^{GABA} \cdot w_{(c)\underset{ns}{\Rightarrow}(connection)} \tag{7}$$

where $w_{(b)\rightarrow(a)}$ is the weight of an excitatory connection from cell a to cell b and $w_{(a)\underset{ns}{\rightarrow}\left[(b)\underset{s}{\rightarrow}(c)\right]}$ is the weight of an excitatory non-somatic connection from cell a to the excitatory synapse from cell b to cell c.

Some of the non-somatic connections, like those from the feedback cells, are likely to be located on the axons of the stimulus cells because one weight applies to multiple outgoing synapses of the post-synaptic cell. Other non-somatic connections, like those from the rule-set or premotor cortex (PMC) cells in the model, have varying effects across the synapses of the post-synaptic cell. Biologically, the non-somatic connections that cause these varying effects could target either the axon near the terminals or individual synapses themselves. Within the model, those non-somatic connections targeting individual synapses correspond to pre-synaptic inhibition. Whatever the true biological implementation, they can be modeled in the same way because the only effect that the location of the synapse has is different travel times of the action potential. Unlike synapses, which take time in the new model, axon propagation is simplified as instantaneous, so synapse location is irrelevant in this new model.

If there is no connection of a given type between two model components, the corresponding w term can be treated as though it were 0. Within the model each connection has its own mean value for its weight term. When the model is initialized the starting value of the term is drawn from the uniform distribution of all numbers within $\pm 1\%$ of the corresponding mean unless the postsynaptic cell is an accumulator (as explained next).

3.2. Non-cortical cells

While most of the cells in the model are parameterized as Cortical Pyramidal cells, the model also makes use of thalamic and globus pallidus cells. All cells use Eqs. (3) through (7). However, each cell type has unique patterns of firing and responds differently to excitation. Therefore, in order to properly model these types of cells, different equations must be used for V and U and the effect of an action potential must be parameterized differently. Fortunately, Izhikevich (2007b) had already calculated parameters for all the neuron types used in the model. In cases where a parameter had a value that has no effect on the rest of the equation, such as a factor having a value of 1.0, it is intentionally included in the equations to make the relation to the general form clearer.

The voltage update equations for thalamic cells are:

$$V_a(t+1) = V_a(t) + [1.6\,[V_a(t) + 60]\,[V_a(t) + 50]$$
$$- U_a(t) + E_a(t) - I_a(t)]\,/200 \tag{8}$$

$$U_a(t+1)$$
$$= \begin{cases} U_a(t) + 0.01\,[15\,[V_a(t) + 65] - U_a(t)] & V(t) \leq -65 \\ U_a(t) + 0.01\,[-U_a(t)] & else \end{cases}$$
$$\tag{9}$$

where the parameters are the same as in the pyramidal cell (Izhikevich, 2007b). The two update equations for (9) model the two firing modes of thalamic neurons: bursting (top) and tonic (bottom). If $V_a(t)$ is greater than $[35 + 0.1 U_a(t)]$ then the cell has undergone an action potential. $V_a(t)$ is set to $[-60 - 0.1 U_a(t)]$ mV and $U_a(t)$ is increased by 10.

The voltage update equations for globus pallidus cells is:

$$V_a(t+1) = V_a(t) + [1\,[V_a(t) + 55]\,[V_a(t) + 40] + 140$$
$$- U_a(t) + E_a(t) - I_a(t)]\,/20 \tag{10}$$

$$U_a(t+1) = U_a(t) + 0.15\,[8\,[V_a(t) + 55] - U_a(t)] \tag{11}$$

where the parameters are the same as in the pyramidal cell. If $V_a(t)$ is greater than 25 then the cell has undergone an action potential, $V_a(t)$ is set to -50 mV, and $U_a(t)$ is increased by 200. These cells spontaneously fire unless they are inhibited by some other source. This is modeled by the addition of the constant factor of 140 in Eq. (10).
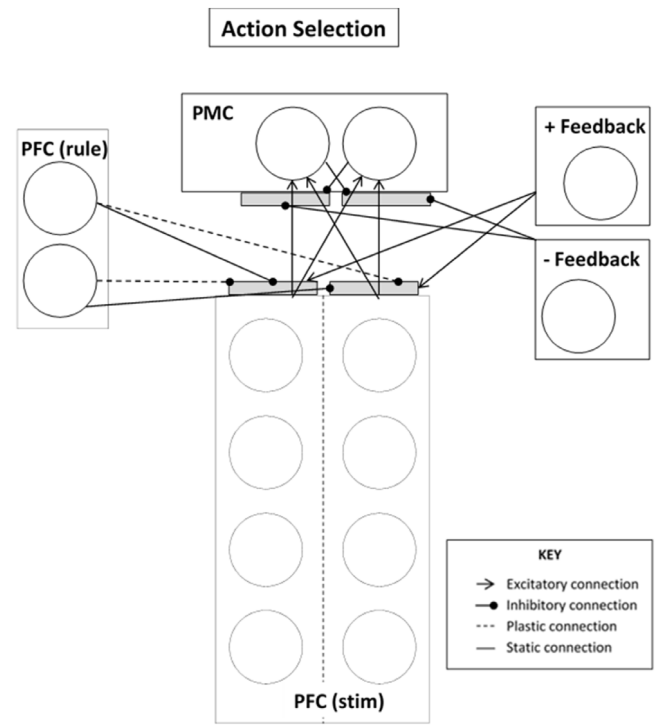

Fig. 3. The connectivity diagram of the response selection module. Non-somatic connections are denoted using a gray field over the somatic connection that receives the non-somatic input. A dotted line separates stimulus neurons (circles) of different dimensions.

3.3. Response selection module

The model is composed of two major modules: a response selection module and a rule-set selection module. In the response selection module, an array of pyramidal cells represent the stimulus features (Fig. 3). If the features are discrete, each feature can be represented by a single cell. Stimulus features that vary continuously can be represented by cells with receptive fields that determine how much excitation each cell receives from a specific value in that continuum. None of the simulations included in this article used continuous stimulus values so no specific receptive field formulae is discussed (but see Hélie et al., 2015b). Cells representing discrete features receive a 500 mA input current when the feature they represent is present. Stimulus cells are part of the PFC and thus the features that they represent are highly preprocessed (Freedman, Riesenhuber, Poggio, & Miller, 2003). While they can represent something as simple as the color or orientation of a stimulus, complex ideas such as a compound 3D shape can be represented if the representation is informative to response selection.

To present a stimulus to the model, a 500 mA current is injected into each stimulus cell and cue cell that represent the features of the stimulus. The stimulus cells excite the pyramidal response cells in the PMC. The PMC cells represent different responses or stimulus categories. They are not meant to correspond to specific movements or action plans although in an experimental setting the two are often equivalent. Because responses are mutually exclusive, each PMC cell sends non-somatic inhibition to the incoming excitatory connections of every other PMC cell, creating winner-take-all behavior (Rumelhart & Zipser, 1985). This design should scale well as Rutishauser, Douglas, and Slotine (2011) showed that WTA components tend to stabilize large networks. This feature, when combined with the avoidance of
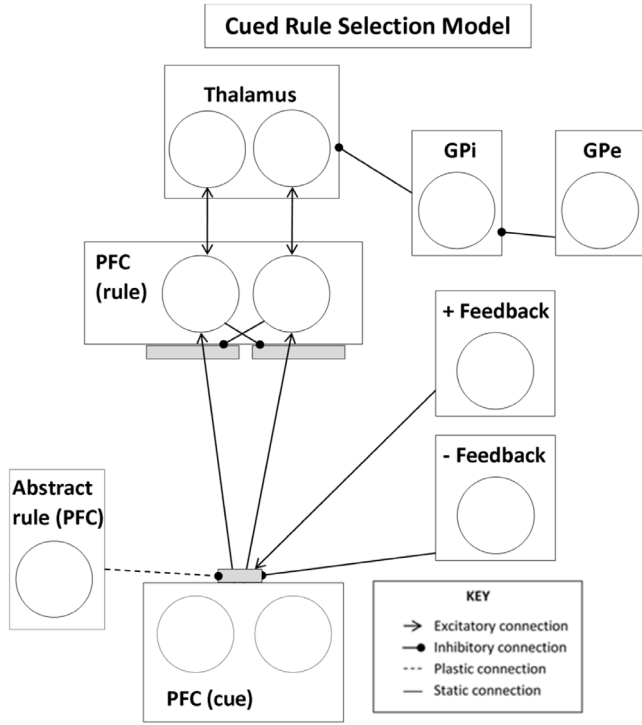
Fig. 4. The connectivity diagram of the response selection model segment. Non-somatic connections are denoted using a gray field over the somatic connection that receives the non-somatic input.

combinatorial explosion through pre-synaptic inhibition, should allow this model to scale up to real-world problem sizes.

Rules are represented by pyramidal cells in the PFC. Each of the rule-set cells non-somatically inhibit connections from stimulus cells to PMC cells. Each rule-set cell has a stimulus dimension associated with it. Connections from stimulus cells that represent features outside of this dimension are inhibited statically and severely. Within the associated dimension the inhibition is weaker and can change. This means that when a rule-set cell is active, the model is focusing on the associated stimulus dimension and ignoring the other dimensions. A running sum of each PMC cell's O_{Glut} term is tracked over the course of each trial. The response made corresponds to the first PMC cell whose sum reaches the response threshold (see Table A.1). Response cells are conceptualized as initiating a physical response, so once a response has been made it cannot be changed and no other response can be made for that trial. The feedback cells become active once the model makes a response. If the response would elicit positive feedback, the positive feedback cell receives an input of 500 mA for the remainder of the trial. Otherwise, the negative feedback cell receives that input instead. This simulates the behavior of the cell populations observed by O'Doherty et al. (2001) discussed in Section 2.2. The positive feedback cell non-somatically excites all stimulus cell-to-PMC connections while the negative feedback cell inhibits those connections (see Fig. 4). At the end of a trial the strength of the plastic connections from the rule-set cells is based on the following equation:

$$w_{(a)\underset{ns}{\Rightarrow}\left[(b)\underset{s}{\to}(c)\right]}[n+1] = w_{(a)\underset{ns}{\Rightarrow}\left[(b)\underset{s}{\to}(c)\right]}n \tag{12.1}$$

$$- \eta_1 \cdot \sum_t O^{GABA}{}_a(t) \cdot \left[\sum_t \left[O^{Glut}{}_b(t) + P_{(b)\underset{s}{\to}(c)}(t)\right]^+ - \theta_1\right]^+$$

$$\cdot \left[w_{max} - w_{(a)\underset{ns}{\Rightarrow}\left[(b)\underset{s}{\to}(c)\right]}n\right] \tag{12.2}$$

$$+ \eta_2 \cdot \sum_t O^{GABA}{}_a(t) \cdot \left[\theta_1 - \sum_t \left[O^{Glut}{}_b(t) + P_{(b)\underset{s}{\to}(c)}(t)\right]^+\right]^+$$

$$\cdot \left[\sum_t \left[O^{Glut}{}_b(t) + P_{(b)\underset{s}{\to}(c)}(t)\right]^+ - \theta_2\right]^+$$

$$\cdot \left[w_{(a)\underset{ns}{\Rightarrow}\left[(b)\underset{s}{\to}(c)\right]}n - w_{min}\right] \tag{12.3}$$

Eq. (12) uses variables as they were previously defined where applicable, and all the new parameter values for Eq. (12) are specific to the task. However, it is important to note that θ_1 is always greater than θ_2. Segment (12.2) (starting with η_1) models the changes in the non-somatic synapse that reduce the inhibition from the rule-set cell. Reduced inhibition means that, under this rule, the range of stimulus values represented by the stimulus cell postsynaptic to the rule-set cell are associated with the response of PMC cell postsynaptic to the stimulus cell. This change should only occur when all three of the following conditions are met: the stimulus cell is active, the PMC cell is active, and the model positive feedback cell is active. Only when all three of these conditions are met is $\sum_t \left[O^{Glut}{}_b(t) + P_{(b)\underset{s}{\to}(c)}(t)\right]^+$, the input to the PMC cell, able to reach a value greater than θ_1, otherwise the entirety of (12.2) will be zero and can be ignored. This causes the association between the current stimulus and the chosen response to strengthen when that combination creates positive feedback. Segment (12.3) (starting with η_2) models what happens when only two of the three conditions are met. In this case, the connection is strengthened and the inhibitory signal becomes stronger, thus disassociating the stimulus and response. If more than one of the conditions are absent the PMC input does not reach θ_2, thus both (12.2) and (12.3) solve to zero and the connection strength does not change.

3.4. Cued rule-set selection

The architecture of the rule-set selection module mirrors that of the response selection module with cue features in place of stimulus features and concrete rule-set cells instead of response cells (Fig. 4). An abstract rule-set cell gates the connections between the cue cells and the concrete rule-set cells in the same way the concrete rule-set cells gate the connections between stimulus features and responses. Theoretically, there could be multiple abstract rule-sets. However, none of the tasks used to validate the model incorporate multiple abstract rule-sets. Thus, a 500 mA current is injected into the abstract rule cell through the entirety of each trial. The outgoing connections from the abstract rule-set cell are used to gate the connections between the cue cells and the concrete rule-set cells. Synaptic plasticity and feedback for these connections operate the same way as in the response selection component, albeit with potentially different parameters in Eq. (12). The concrete rule-set cells are the same as the connection point between the two modules and appear in Figs. 3 and 4. Concrete rule-sets cells selected by the rule-set selection module gate stimulus–response associations in the response selection module. Once a rule-set is selected a positive feedback loop with the thalamus keeps it active. In the model thalamic cells are paired with rule-set cells, resulting in one thalamic cell for each rule-set cell. Paired cells excite each other creating positive feedback loops. The activity of this loop persists across trials and only ceases when inhibition arrives from a cell representing the GPi. As previously described, cells in the globus pallidus spontaneously fire at high rates. The GPi cell is therefore inhibited by a cell representing the GPe. Without the GPe, the GPi would be perpetually active and the thalamus would be in a constant state of inhibition. The GPi receives a 200 mA
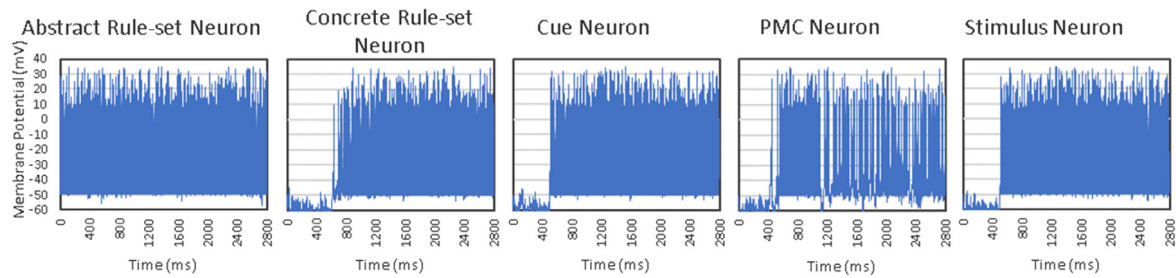
Fig. 5. Spike trains of the most active neurons from each group during a trial from the Badre et al. (2010) simulation (Section 4.1). The abstract rule cell was active throughout the trial, the cue and stimulus cells activated when the stimulus was presented, and the PMC and concrete rule-set cells activated shortly after. In this trial the model made an error and the effect of the inhibition from the feedback cell can be seen in the spike train of the PMC neuron after a response is made at 1200 ms.

excitatory current for 500 ms at the start of any trial where a task is likely (based on findings from Monsell & Mizon, 2006). This excitation comes from the pre-SMA via the subthalamic nucleus (Aron et al., 2007). This excitatory input is the only way for the GPi to overcome the inhibitory influence of the GPe.

It is unusual for a model that includes the GPe and GPi to not include the striatum as well. The striatum receives input from most cortical areas and selectively inhibits both the GPi (direct pathway) and GPe (indirect pathway) (Alexander, DeLong, & Strick, 1986). These connections can give some insight into the role of the striatum regarding working memory maintenance. In the model the direct pathway enables maintenance while the indirect pathway prevents it. In tandem these pathways could dynamically constrain the items, such as task-sets, that can enter working memory, preventing distractions. This kind of functionality does not factor into the current battery of experiments and thus the striatum is omitted from this implementation of the model. However, this omission was done for simplicity of implementation only; other applications of the proposed framework could readily include the striatum.

Representational structures are determined by the task conditions. An experimenter must adjust the structure of the model to fit the desired representation. Although there is an outside chance that the model could function when every dimension has representation in both the cues cells and the stimulus cells, it was not designed with this implementation in mind, nor has it been tested using such an implementation. Any experimental task that involves the subject determining which stimulus dimensions to use in abstract vs. concrete rules require multiple implementations of this model: one for each representational structure. In each of these implementations the cue dimensions are assumed to be identified beforehand, whether from some outside instruction or an internal decision that is made by cognitive systems beyond the model's scope.

4. Simulations

In this section we used java implementations of the model to replicate three experimental results from different articles in an effort to validate the new theory. These findings cover the learning advantages of different representations, transfer effects, and priming effects. General spike train results from an archetypical trial are presented in Fig. 5. All parameter values for each simulation are listed in Appendix.

4.1. Simulation 1 — Badre et al. (2010)

The first test for the model is if it will perform better when able to use multiple rule-sets than when only able to use one. Badre et al. (2010) tested subjects in two conditions: the hierarchical condition, in which using multiple rule-sets would have

been beneficial, and the flat condition, in which it would not. Subjects in the flat condition learned slower and had a lower final accuracy than subjects in the hierarchical condition. The model was tested on these two tasks to see if it would also perform better in the hierarchical condition.

Badre et al. (2010) asked subjects to learn associations between 3 responses and 18 stimuli which varied across 3 dimensions (with $3 \times 3 \times 2$ discrete values). In one condition the associations were "flat" with no pattern to them and no abstract rule-set that would aid learning. In the other condition, the associations were "hierarchical" — the dimension with two values could be treated like a cue, signaling which of the other two dimensions were irrelevant for that trial. Subjects were not informed that any hierarchy existed, they were simply told to learn the correct response for each stimulus. Despite this obfuscation, subjects in the hierarchical condition learned faster and achieved better performance by the end of the experiment compared to subjects in the flat condition showing that subjects try to apply rule-sets to situations even when there is no prior indication that rule-sets would be useful.

4.1.1. Task

In Badre et al. (2010), subjects categorized 18 stimuli into 3 response categories. The stimuli were pictures of computer-generated 3D shapes and varied in three dimensions: the shape itself, the angle at which the shape was viewed, and the color of the border. Subjects saw each stimulus 20 times for a total of 360 trials. In the hierarchical condition, the stimuli with red borders can be accurately categorized solely by using shape while the stimuli with blue borders can be correctly categorized by using orientation. This means that subjects could represent the stimuli hierarchically with the color of the border acting as a cue to use either a rule-set on shape or a rule-set on orientation. In the flat condition, there is no exploitable structure to the stimulus–response pairings and all three dimensions must be used to correctly categorize each stimulus.

4.1.2. Human data

Subjects were given no information about any organization in the stimulus–response associations yet subjects in the hierarchical condition learned more quickly and achieved a higher final accuracy than subjects in the flat condition (Fig. 6). There are two reasons for this improvement in performance. The first is that there are fewer associations to learn when using a hierarchical representation: 2 border colors to 2 rule-sets, 3 shapes to 3 responses, and 3 orientations to 3 responses for a total of 8 associations. In the flat representation each of the 18 combinations of border color, shape, and orientation needs its own association to a response. Secondly, in the flat condition there is overlap in the relevant dimensions. Stimuli with different responses share features with each other, which causes interference. While the
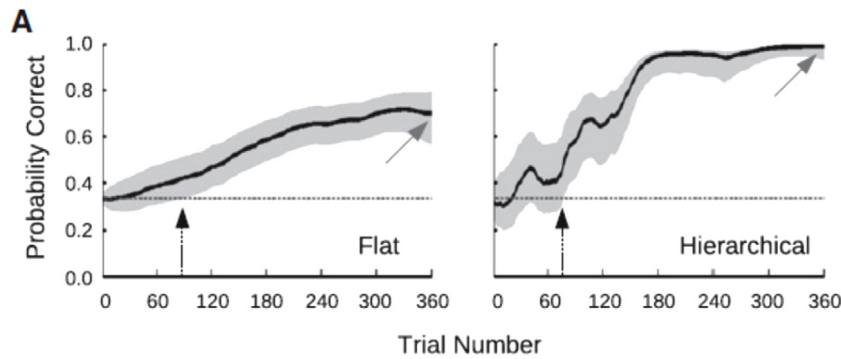
Fig. 6. Human results from Badre et al. (2010). Graphs show the learning curve estimates and 90% confidence interval of the most typical subject for each condition. *Source:* Graphs taken from Badre et al. (2010).
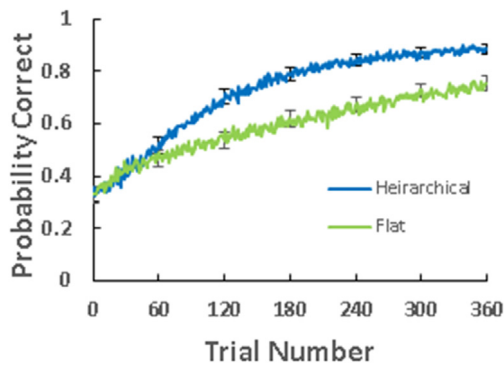

Fig. 7. Model data from the task used in Badre et al. (2010). Error bars show 95% confidence interval.

same is technically true in the hierarchical condition, this overlap only occurs in the dimension designated irrelevant by the cue. Once the subject has learned to use the cue to ignore one of the dimensions, they are able to disregard misleading interference from the irrelevant dimension. In the flat condition, there are no irrelevant dimensions and therefore subjects suffer disruptive interference.

4.1.3. Task-specific model parameters

As stated in the model description, problem representation is defined by the architecture of the model. The hierarchical condition was simulated using separate sets of cells for each stimulus dimension (each set with three possible values) whereas the flat condition merged the representation of the cue dimension with those of the two stimulus dimensions and represents them as a single composite dimension with 18 values (see Table A.1). Although there was only one rule-set cell in the flat condition, and therefore no need to clear the old rule-set, the GPi received input as it did in the hierarchical condition to maximize equivalence between the two simulations. In each condition, a run consisted of 20 repetitions of each of the 18 possible stimuli presented in a random order.

4.1.4. Model results

Each condition was run 1000 times and the results of each trial were averaged (Fig. 7). As a reminder Bardre, Kayser, and D'Espositio used learning curve estimates for their figures (presented in Fig. 6), not the means as used in Fig. 7. The model was able to learn the associations in both conditions but with markedly different learning curves. In the flat condition, accuracy rose linearly and reached an accuracy of approximately 0.7 by trial 300. In the Hierarchical condition accuracy rose much more

steeply until trial 180 and then slowly approached an accuracy of 0.9. To calculate the r^2 of the model, subject data was estimated from the graphs in Fig. 6 at trial increments of 30 (0, 30, 60 … 360). The results yield an r^2 of 0.966. The failure of the model to reach the perfect accuracy achieved by the representative subject in the hierarchical condition is not a universal failing. Over 70% of experimental runs reached an average accuracy of 0.9 or greater by trial 330.

4.1.5. Discussion

The model demonstrated the ability to develop associations between stimuli/response pairs and cue/rule-set pairs concurrently. This allowed the model to learn correct responses faster when using a hierarchical representation compared to a flat representation. In the human data for this condition, the accuracy started to asymptote at this point while the model's performance continued to rise linearly. An explanatory hypothesis for this difference is that the 18 unique combinations of features caused human subjects to suffer from interference effects. Although the simulation treats the images as 18 unique stimuli, each image shared many features with other images and not all of these other images had the same response association. This would cause interference in selecting the correct response. The model does not account for this kind of interference and so the asymptote does not appear in the model's results. This hypothesis could be easily tested by comparing learning rates of subjects learning the flat stimuli with the learning rate of subjects learning 18 stimuli that have no features in common.

4.2. Simulation 2 — Collins and Frank (2013)

The model is not restricted to using only one rule-set for a dimension. Multiple rule-sets can exist governing the same dimension. Neither are rule-sets limited to only using a single cue. Different cue features can be associated with the same rule-set. This allows learning to transfer into new contexts. In an experiment presented in Collins and Frank (2013), subjects categorized colored shapes. Correct categorizations were assigned such that if subjects used a particular task representation learning would transfer when subjects learned a second set of stimuli. Two implementations of the model were tested, one that used the representation conducive to learning transfer and an alternate representation that had been commonly used by subjects(see Section 1.4 for more details).

4.2.1. Task

In the Collins and Frank task, subjects were trained to associate a stimulus that varied in two dimensions: shape and color, with 4 responses. There were 4 different colors and 2 different shapes. Subjects were trained over two phases. In the first phase
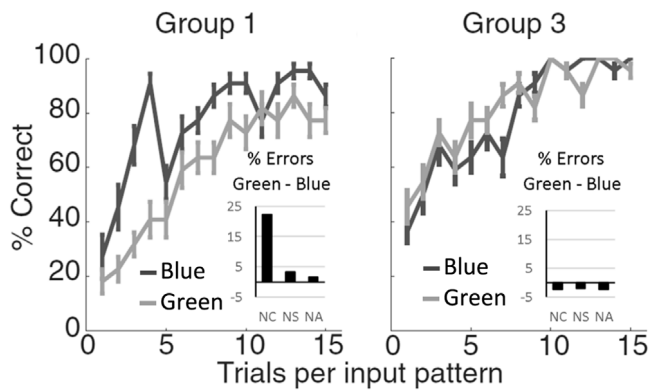
Fig. 8. Human results for Groups 1 and 3. Error bars show standard error. Inserts show green errors minus blue errors by error type: errors made due to neglecting color (NC), errors made due to neglecting shape (NS), and errors made due to neglecting both color and shape (NA).
Source: Figures adapted from Collins and Frank (2013).

two out of the four colors were used resulting in 4 stimuli, red triangles, yellow triangles, red circles, and yellow circles. Each of these stimuli was associated with a different response. If a hierarchical representation of these associations is used, there is no environmental bias toward using shapes as cues or colors as cues. In the second phase there is a shift in color — the presented shapes are now blue or green instead of red or yellow. The category associations for blue shapes are the same as for yellow shapes (so yellow triangles and blue triangles share a response) but the correct responses for green shapes follow a new pattern.

The shift during phase 2 caused a dissociation between subjects that used color as a cue and subjects that used shape as a cue. If subjects used shape as a cue, there should have been no difference in the rate at which stimulus–response associations are learned for blue stimuli vs. green stimuli since associations for both needed to be added to the rule-sets cued by shape. If instead subjects used color as a cue, they would have been able to associate blue with the same rule-set used for yellow stimuli. By doing this they took advantage of concrete rules that had already been learned. Green stimuli, on the other hand, required a new set of rules and thus a new rule-set, all of which needed to be learned. In this case, the acquisition of the response pairings for the blue stimuli was expected to occur faster than for the green stimuli.

4.2.2. Human data

For each subject, Collins and Frank calculated two reaction-time switch-costs, one for switches in stimulus color and one for switches in stimulus shape. They divided subjects into three groups based on the difference between these switch costs. Group 1 was made up of the top third of subjects with the highest color-switch cost minus shape-shift cost. Group 3 was made up of the bottom third of subjects with the lowest color-switch cost minus shape-shift cost. Group 2 was made of the remaining third. Collins and Frank then examined the learning curves and error patterns across these groups (Fig. 8). They looked at the types of errors made, whether subjects neglected the stimulus' shape when choosing a response, the stimulus' color, or both (Fig. 8 insets). Group 1's learning curve and error patterns matched the predicted pattern for using color as a cue while Group 3's learning curve and error patterns matched the predicted pattern for using shape as a cue. For subjects in Group 3, green and blue stimuli were learned at statistically similar rates and the types of errors made did not significantly differ between stimuli of those two

colors. In Group 1, blue stimuli were learned more quickly than green stimuli and there were significantly more 'neglect color' errors made with green stimuli than with blue. Since Group 2 represented subjects whose task representations were harder to infer and who may have used alternative representations, their results are not considered.

4.2.3. Task-specific model parameters

Group 1 was simulated by using color as a cue while Group 3 was simulated by using shape as a cue. For both group simulations, a run consisted of 15 instances of each of the four stimuli that were colored red or yellow, presented in random order. This was followed by 15 instances of each of the four stimuli that were colored blue or green, again presented in random order.

4.2.4. Model results

The mean results for the simulations across 1000 repetitions of each group are shown in Fig. 9. Simulation of Group 3 (Shape Cue condition) achieved a maximum performance accuracy of 0.845 for stimuli of both colors. The errors made in this condition are also the same across all colors and error types (see insets of Fig. 9). In the simulation of Group 1 (Color Cue condition), the blue stimuli reached an accuracy of 0.807 and the green stimuli reached an accuracy of 0.749. The r^2 of the model results was 0.937. While errors caused by neglecting shape or neglecting both color and shape were the same between the two colors, there were far more neglect color errors made with the green stimuli than with the blue stimuli.

4.2.5. Discussion

The model learned the task well and showed itself capable of concurrently learning multiple rule-sets on the same dimension. The faster acquisition of blue stimuli vs. green stimuli when color was used as the cue shows that the model can transfer previously learned rule-sets into new environments instead of having to relearn them. Importantly, the model made more errors by neglecting color when categorizing green stimuli compared to blue stimuli in the color cue condition. These are the same error patterns that were found in human data so it can be concluded that the model was, in some ways at least, using a method similar to that of human subjects.

The model has shown that it can successfully learn associations between stimuli and responses while simultaneously learning associations between cues and rule-sets. This has been demonstrated in a few different situations while reproducing several key advantages of using rule-sets. Next, the model was tested on its ability to account for reaction-time data in priming experiments.

4.3. Simulation 3 — van t Wout et al. (2015)

Response time can be reduced through priming as shown in Sudevan and Taylor (1987). van t Wout et al. (2015) expanded upon this finding by manipulating the number of potential rule-sets that subjects might have to use and examining the interaction with priming. The effect of the number of rule-sets was minimal in all cases but their data presents a clear example of the effect of priming and thus is a good source of human data with which to test the model. In this experiment, the cue was presented in advance of the stimulus with a long or short interval between the two. With more time to process the cue, subjects responded more quickly.
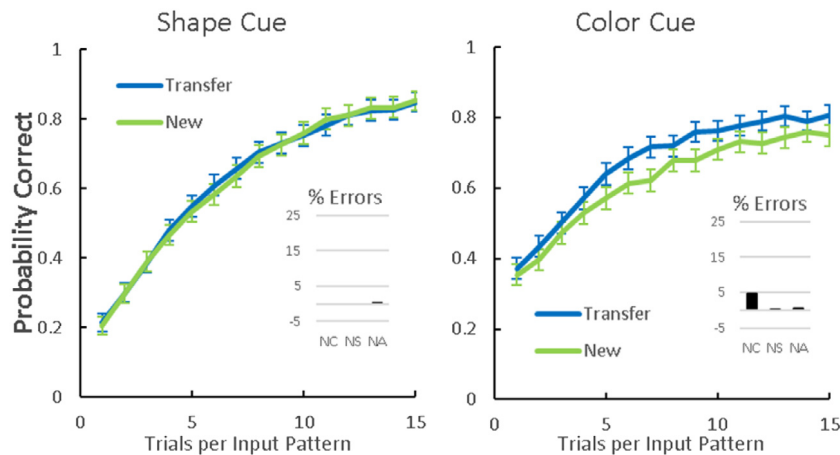
Fig. 9. Model performance in the Collins and Frank task. The graphs show average accuracy and the inserts show green errors minus blue errors by error type: errors made due to neglecting color (NC), errors made due to neglecting shape (NS), and errors made due to neglecting both color and shape (NA). Error bars show 95% confidence interval.

4.3.1. Task

In the task used in van t Wout et al. (2015), subjects were presented with stimuli featuring stylized animals or trees. Both animals and trees varied in five dimensions (e.g. head size, fruit shape, etc.). Each dimension had only two possible values, which corresponded to one of the two possible responses. Subjects were given a verbal cue prior to the presentation of the stimulus that indicated which feature should guide the response. Subjects were trained for an extended period before data collection to minimize the effect of learning on the results. Once the extensive training was complete, subjects were tested by varying two factors. The first factor was that the relevant dimension was either limited to one of three dimensions or could be any one of the five possible dimensions. Subjects in the three-rule-set condition were aware that they were in this condition and understood the effect of the condition on the task. The second factor varied the interval between the onset of the verbal cue and the presentation of the stimulus (Cue–Stimulus Interval – CSI) to be either 100 ms or 1300 ms. These two factors were fully crossed.

4.3.2. Human data

The mean response time for each condition is plotted above (Fig. 10). The most relevant effect for the current model is the effect of manipulating CSI. Subjects that were given longer CSIs responded about 450 ms after the stimulus was presented. Subjects that saw the shorter, control CSI responded after around 700 ms. There is also a small effect of rule-set number: a larger pool of tasks slows reaction time slightly. However, this effect was not found to be significant in either condition. Only reaction times from correct responses were used.

4.3.3. Task-specific model parameters

The subjects were fully aware of and well versed in the associations between cues/stimuli and rule-sets/responses. As such, the weights of the non-somatic inhibition from the abstract and concrete rule-set cells were set to the extreme values for the task and learning was disabled. Although van t Wout et al. described the cue as presented 100 ms or 1300 ms before the stimulus, remember that it was a verbal cue. Speech has a time component and it arguably makes more sense to present the cue to the model at the off-set of the cue rather than the on-set. van t Wout et al. stated that all of the verbal cues last 350 ms and so in the model the cue was presented 950 ms before the stimulus presentation or 250 ms afterward. In the case where the cue was presented after the stimulus, the presentation of the stimulus to the model
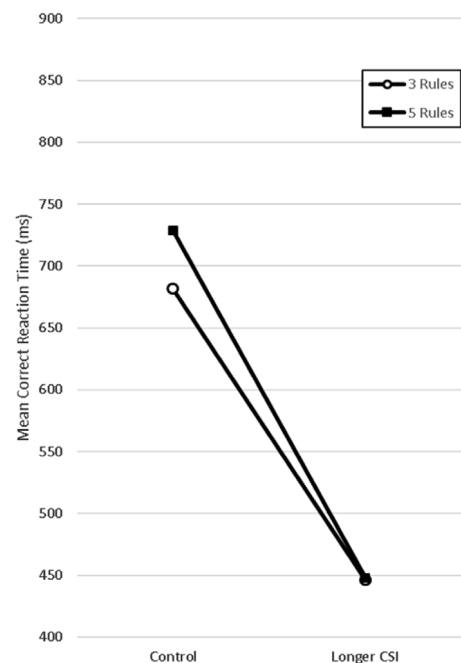

Fig. 10. Human data from van t Wout et al. (2015).
Source: Reproduced using figures from van t Wout et al. (2015).

was delayed to coincide with the presentation of the cue due to limited attentional resources of the subjects and their knowledge that processing the stimulus would be of no use without knowing which dimension was relevant.

A run consisted of 576 trials that used only 3 possible relevant dimensions (6 of each possible stimulus and relevant dimension combination) with short CSI, 576 trials that used only 3 possible relevant dimensions with long CSI, 640 trials that used all 5 possible relevant dimensions (4 of each possible stimulus and relevant dimension combination) with short CSI, and 640 trials that all used 5 possible relevant with long CSI. The mean response time for correct trials was calculated for each condition using 1000 runs.
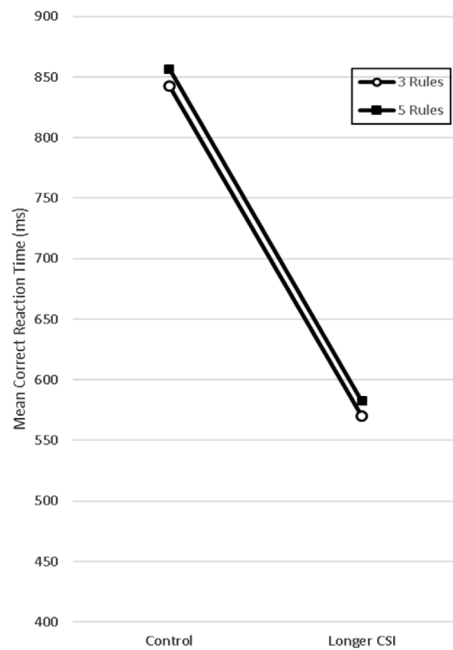
Fig. 11. Model data for the van t Wout et al. (2015) task. The 95% confidence interval is too small to be seen with error bars.

4.3.4. Model results

The model reproduced the main effect of longer mean response time in short CSI conditions. The magnitude of the difference, a 250 ms increase is reproduced as well (Fig. 11). However, the model results show a longer response time than the human subjects show for both conditions. Although the model is slower than the human subjects, the pattern of results is well reproduced with an r^2 of 0.99. The insignificant effect of more possible rule-sets slightly increasing reaction time was reproduced as well, which is only notable because the model was not designed to reproduce it.

4.3.5. Discussion

The relevant effects are well reproduced by the model. However, for both conditions the results show a longer response time than the human subjects. A possible explanation is that well practiced tasks have lower response thresholds. The observant reader will note that the response threshold was lowered for this simulation. However, any further reduction of the threshold increased the rate of errors due to activity in the PMC neurons occurring before a concrete rule-set is chosen. Future versions of this model could include an 'undecided' state that prevents activation of the PMC cells until a concrete rule-set has been chosen (see Section 5.2).

5. General discussion

The goal of this research was to create a model that only used realistic neurons to simulate rule learning and rule-set switching via pre-synaptic inhibition. Care was taken to avoid contradicting current neurological findings in the design of the model. The results show that the model can account for a variety of data in both learning and task-switching paradigms. However, the most important contribution of a model is making new testable predictions.

5.1. Implications and predictions

The most important predictions made by this model follow from the role of the basal ganglia and the effect of storing task-sets in working memory. For example, if salient, attention-grabbing distractions were occasionally paired with stimuli presented in a task-switching experiment the model predicts that the distractor would enter working memory, ousting the task-set. When the subject refocuses on the task, the task-set would have to be re-acquired. This would increase reaction time but moreover there would be no reaction time difference between switch trials and stay trials.

Another, more speculative prediction involves the basal ganglia. Both the direct pathway and the indirect pathway originate in the striatum. In the indirect pathway the GPe is inhibited, freeing the GPi to inhibit the thalamus. In the direct pathway the striatum inhibits the GPi, freeing the thalamus. Combining this with the model's hypothesis that the thalamus maintains working memory creates the prediction that both pathways work together to constrain which items can enter WM. Signals along the indirect pathway send general inhibition to the thalamus, making most items impossible to maintain in WM. At the same time, the direct pathway is activated, sending more localized signals, protecting a few select WM items allowing them to be maintained. Testing this hypothesis would be challenging but one could try conducting a single cell recording study comparing indirect pathway activation in a familiar task or environment with activation in a wholly unfamiliar task or environment.

5.2. Extensions, improvements, and future work

The model was shown to be robust and versatile, but it is by no means perfect. Tasks like the WCST and intradimensional/extradimensional task involve rule-set switching based on task performance instead of cues. Even the HICL model (Hélie et al., 2015b) was designed to perform un-cued rule-set switching. Yet this model does not include a mechanism for cue-less rule-set switching. Some of the pieces are already in place. Un-cued task switching requires that the switches be infrequent, making the thalamic maintenance mechanism of the model very useful. When a rule-set switch occurs, a reset signal can be sent by, for instance, a performance monitoring system. The biggest challenge is to prevent the abandoned rule-set from being immediately reselected. Future work should focus on calibrating this system.

A more ambitious way to add performance monitoring would be to incorporate a comprehensive model of memory. Not only would this implement performance monitoring, it would also allow a more meaningful exploration of cross-task interference in task-switching and priming studies. A confidence mechanic could also be useful should these phenomena be revisited.

Another potential improvement to the model is the addition of a concrete rule-set cell (or similar) that represents the lack of a selected rule-set. Currently when a winning rule-set has yet to be selected none of the rule-set cells are active enough to substantially gate the stimulus/response connections. The addition of a "no rule-set" cell allows that cell to send blanket inhibition to every stimulus/response connection in the absence of a winning rule-set cell.

A variation of the Badre et al. (2010) task where the cue is presented after the stimulus would be interesting in several ways. Will subjects develop the same hierarchical task representation with border as the cue or could the shape and orientation be the cues and the border color be the stimulus feature? In such a case, would the cue dimensions be separated resulting in two simultaneously active rule-sets or would the dimensions be combined resulting in 9 different rules? Would priming effects occur?

Table A.1

Parameter values associated with different task representations across the various simulations. The parameters that were not part of the experimental procedure were found using grid-search.

	Badre et al.		Collins and Frank		van t Wout et al.	
	Hierarchical	Flat	Color Cue	Shape Cue	Control	Primed
$w_{(stim)\xrightarrow{s}(PMC)}$	55	55	55	55	50	50
$w_{(externalcue)\xrightarrow{s}(concreterule)}$	25	25	25	25	25	25
$w_{(concreterule)\xrightarrow{s}(thalamus)}$	40	40	40	40	40	40
$w_{(thalamus)\xrightarrow{s}(concreterule)}$	40	40	40	40	40	40
$w_{(GPi)\underset{s}{\Rightarrow}(thalamus)}$	100	100	100	100	100	100
$w_{(GPe)\underset{s}{\Rightarrow}(GPi)}$	40	40	40	40	40	40
$w_{(PMC)\underset{ns}{\Rightarrow}\left((stim)\xrightarrow{s}(PMC)\right)}$	3	3	3	3	3	3
$w_{(pos)\underset{ns}{\Rightarrow}\left((stim)\xrightarrow{s}(PMC)\right)}$	0.8	0.8	1.0	1.0	0	0
$w_{(neg)\underset{ns}{\Rightarrow}\left((stim)\xrightarrow{s}(PMC)\right)}$	0.4	0.4	0.4	0.4	0	0
$w_{(concreterule)\underset{ns}{\Rightarrow}\left((externalcue)\xrightarrow{s}(concreterule)\right)}$	6	6	6	6	6	6
$w_{(pos)\underset{ns}{\Rightarrow}\left((externalcue)\xrightarrow{s}(concreterule)\right)}$	0.8	0.8	1.0	1.0	0	0
$w_{(neg)\underset{ns}{\Rightarrow}\left((externalcue)\xrightarrow{s}(concreterule)\right)}$	0.4	0.4	0.4	0.4	0	0
$w_{(concreterule)\underset{ns}{\Rightarrow}\left((outstim)\xrightarrow{s}(PMC)\right)}$	6	6	6	6	6	6
Stimulus dimensions	2	1	1	1	5	5
Features per dimension	3	18	2	4	2	2
Cues/Rule-set cells	2	1	4	2	3 & 5	3 & 5
Responses	3	3	4	4	2	2
Stimulus onset	500 ms	500 ms	500 ms	500 ms	1750 ms	1400 ms
Cue onset	500 ms	500 ms	500 ms	500 ms	1750 ms	450 ms
Response Threshold	7000	7000	7000	7000	5000	5000
Trial length	2800 ms	2800 ms	2800 ms	2800 ms	3500 ms	3500 ms
Response time = 0	500 ms	500 ms	500 ms	500 ms	1400 ms	1400 ms

Table A.2

Parameters for the model's plastic connections $w_{(concreterule)\underset{ns}{\Rightarrow}\left((instim)\xrightarrow{s}(PMC)\right)}$ and $w_{(abstractrule)\underset{ns}{\Rightarrow}\left((internalcue)\xrightarrow{s}(concreterule)\right)}$ across the various simulations. These parameters were found using grid-search.

	Badre et al.		Collins and Frank	van t Wout et al.
$w_{(concreterule)\underset{ns}{\Rightarrow}\left((instim)\xrightarrow{s}(PMC)\right)}$	Hierarchical	Flat	–	–
η_1	$1.4\cdot10^{-9}$	$3.0\cdot10^{-10}$	$1.4\cdot10^{-9}$	0
η_2	$1.8\cdot10^{-13}$	$1.3\cdot10^{-15}$	$1.2\cdot10^{-13}$	0
θ_1	15000	34000	15000	10000
θ_2	6000	2000	6000	2000
w_{max}	4	4	4	4
w_{min}	0	0	0	0
w_{init}	0.7	0.7	0.7	Correct: 0; All others: 4
$w_{(abstractrule)\underset{ns}{\Rightarrow}\left((externalcue)\xrightarrow{s}(concreterule)\right)}$				
η_1	$4.0\cdot10^{-10}$	0	$3.0\cdot10^{-9}$	0
η_2	$2.5\cdot10^{-14}$	0	$1.6\cdot10^{-13}$	0
θ_1	20000	15000	25000	35000
θ_2	8000	5000	13000	2000
w_{max}	4	4	4	4
w_{min}	0	0	0	0
w_{init}	0.7	0.0	0.7	Correct: 0; All others: 4

Exploring these possibilities would also guide the development of the model.

Finally, another direction would be to replicate single-cell recording data. Ideally little of the design of the model would change but some capabilities may need to be scaled back to mirror the cognitive limitations of animal subjects, but the use of spiking neurons in the model means that the model is already able to produce spike trains. Hence, this extension should be pursued in the near future.

Acknowledgment

Appendix

There are several parameters that are dependent on the task that the model is to perform. Supplemental Table A.1 lists the

parameters used in each experimental task condition. "Out stim" refers to all stimulus cells that code features that are outside the dimension associated with the concrete rule-set cell. There are two plastic non-somatic connections: $(concreterule) \underset{ns}{\Rightarrow} ((instim)$ $\underset{s}{\rightarrow} (PMC))$ and $(abstractrule) \underset{ns}{\Rightarrow} ((externalcue) \underset{s}{\rightarrow} (concreterule))$. Parameters for the learning rule for these connections can be found in Supplemental Table A.2. In the plastic connection denoted by $(concreterule) \underset{ns}{\Rightarrow} ((instim) \underset{s}{\rightarrow} (PMC))$ "in stim" refers to those stimulus cells that represent features in the dimension of the rule-set associated with the concrete rule-set cell. In this way, it is the complement of "out stim" described above. Response time = 0 indicates the point in the model's runtime where, if a response is made at that time step, the response time would be recorded as 0. This corresponds to the moment at which the physical stimulus is presented.

References

Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience*, 9(1), 357–381.

Alport, A., Styles, E. A., & Hsieh, S. (1994). *17 Shifting intentional set: Exploring the dynamic control of tasks*.

Ardid, S., & Wang, X. J. (2013). A tweaking principle for executive control: neuronal circuit mechanism for rule-based task switching and conflict resolution. *Journal of Neuroscience*, 33(50), 19504–19517.

Aron, A. R., Behrens, T. E., Smith, S., Frank, M. J., & Poldrack, R. A. (2007). Triangulating a cognitive control network using diffusion-weighted magnetic resonance imaging (MRI) and functional MRI. *Journal of Neuroscience*, 27(14), 3743–3752.

Ashby, F. G., Ell, S. W., Valentin, V. V., & Casale, M. B. (2005). FROST: a distributed neurocomputational model of working memory maintenance. *Journal of Cognitive Neuroscience*, 17(11), 1728–1743.

Ashby, F. G., & Helie, S. (2011). A tutorial on computational cognitive neuroscience: Modeling the neurodynamics of cognition. *Journal of Mathematical Psychology*, 55(4), 273–289.

Badre, D. (2008). Cognitive control, hierarchy, and the rostro–caudal organization of the frontal lobes. *Trends in cognitive sciences*, 12(5), 193–200.

Badre, D., Hoffman, J., Cooney, J. W., & D'Esposito, M. (2009). Hierarchical cognitive control deficits following damage to the human frontal lobe. *Nature Neuroscience*, 12(4), 515.

Badre, D., Kayser, A. S., & D'Esposito, M. (2010). Frontal cortex and the discovery of abstract action rules. *Neuron*, 66(2), 315–326.

Badre, D., & Nee, D. E. (2018). Frontal cortex and the hierarchical control of behavior. *Trends in Cognitive Sciences*, 22(2), 170–188.

Badre, D., & Wagner, A. D. (2004). Selection, integration, and conflict monitoring: assessing the nature and generality of prefrontal cognitive control mechanisms. *Neuron*, 41(3), 473–487.

Bamber, D. (1969). Reaction times and error rates for same-different judgments of multidimensional stimuli. *Perception & Psychophysics*, 6(3), 169–174.

Buckner, R. L. (2003). Functional–anatomic correlates of control processes in memory. *Journal of Neuroscience*, 23(10), 3999–4004.

Bunge, S. A. (2004). How we use rules to select actions: a review of evidence from cognitive neuroscience. *Cognitive, Affective, & Behavioral Neuroscience*, 4(4), 564–579.

Bunge, S. A., & Zelazo, P. D. (2006). A brain-based account of the development of rule use in childhood. *Current Directions in Psychological Science*, 15(3), 118–121.

Cass, W. A., & Gerhardt, G. A. (1995). In vivo assessment of dopamine uptake in rat medial prefrontal cortex: comparison with dorsal striatum and nucleus accumbens. *Journal of Neurochemistry*, 65(1), 201–207.

Collins, A. G., & Frank, M. J. (2013). Cognitive control over learning: Creating, clustering, and generalizing task-set structure. *Psychological Review*, 120(1), 190.

Crawford, E., Gingerich, M., & Eliasmith, C. (2016). Biologically plausible, human-scale knowledge representation. *Cognitive Science*, 40(4), 782–821.

Crone, E. A., Wendelken, C., Donohue, S. E., & Bunge, S. A. (2005). Neural evidence for dissociable components of task-switching. *Cerebral Cortex*, 16(4), 475–486.

Dove, A., Pollmann, S., Schubert, T., Wiggins, C. J., & von Cramon, D. Y. (2000). Prefrontal cortex activation in task switching: an event-related fMRI study. *Cognitive Brain Research*, 9(1), 103–109.

Freedman, D., Riesenhuber, M., Poggio, T., & Miller, E. K. (2003). A comparison of primate prefrontal and inferior temporal cortices during visual categorization. *Journal of Neuroscience*, 23, 5235–5246.

Fuster, J. M. (2008). *The prefrontal cortex*. Academic Press.

Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*, 38(4), 404.

Hélie, S., Ell, S. W., & Ashby, F. G. (2015a). Learning robust cortico-frontal associations with the basal ganglia: An integrative review. *Cortex*, 64, 123–135.

Hélie, S., Ell, S. W., Filoteo, J. V., & Maddox, W. T. (2015b). Criterion learning in rule-based categorization: Simulation of neural mechanism and new data. *Brain and Cognition*, 95, 19–34.

Hélie, S., Roeder, J. L., & Ashby, F. G. (2010). Evidence for cortical automaticity in rule-based categorization. *Journal of Neuroscience*, 30(42), 14225–14234.

Hübner, R., Futterer, T., & Steinhauser, M. (2001). On attentional control as a source of residual shift costs: Evidence from two-component task shifts. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 27(3), 640.

Izhikevich, E. M. (2007a). Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cerebral Cortex*, 17(10), 2443–2452.

Izhikevich, E. M. (2007b). *Dynamical systems in neuroscience*. MIT press.

Jersild, A. T. (1927). Mental set and shift. *Archives of Psychology*.

Koch, I., & Allport, A. (2006). Cue-based preparation and stimulus-based priming of tasks in task switching. *Memory & Cognition*, 34(2), 433–444.

Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science*, 3025648, 1181–1185.

Logan, G. D., & Bundesen, C. (2003). Clever homunculus: Is there an endogenous act of control in the explicit task-cuing procedure? *Journal of Experimental Psychology: Human Perception and Performance*, 29(3), 575.

Miller, E. K., Freedman, D. J., & Wallis, J. D. (2002). The prefrontal cortex: categories, concepts and cognition. *Philosophical Transactions of the Royal Society, Series B (Biological Sciences)*, 3571424, 1123–1136.

Monsell, S., & Mizon, G. A. (2006). Can the task-cuing paradigm measure an endogenous task-set reconfiguration process? *Journal of Experimental Psychology: Human Perception and Performance*, 32(3), 493.

Nee, D. E., & D'Esposito, M. (2016). The hierarchical organization of the lateral prefrontal cortex. *Elife*, 5, e12112.

O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience*, 4(1), 95.

O'Reilly, R. C., Noelle, D. C., Braver, T. S., & Cohen, J. D. (2002). Prefrontal cortex and dynamic categorization tasks: representational organization and neuromodulatory control. *Cerebral Cortex*, 12(3), 246–257.

Reynolds, J. R., Braver, T. S., Brown, J. W., & Van der Stigchel, S. (2006). Computational and neural mechanisms of task switching. *Neurocomputing*, 69(10–12), 1332–1336.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, 124(2), 207.

Rumelhart, D. E., & Zipser, D. (1985). Feature discovery by competitive learning. *Cognitive Science*, 9(1), 75–112.

Rushworth, M. F. S., Hadland, K. A., Paus, T., & Sipila, P. K. (2002). Role of the human medial frontal cortex in task switching: a combined fMRI and TMS study. *Journal of Neurophysiology*, 87(5), 2577–2592.

Rutishauser, U., Douglas, R. J., & Slotine, J. J. (2011). Collective stability of networks of winner-take-all circuits. *Neural Computation*, 23(3), 735–773.

Seamans, J. K., & Robbins, T. W. (2010). Dopamine modulation of the prefrontal cortex and cognitive function. In *The dopamine receptors* (pp. 373–398). Totowa, NJ: Humana Press.

Shepherd, G. M. (2004). Introduction to synaptic circuits. In G. M. Shepherd (Ed.), *The synaptic organization of the brain* (5th ed.). (pp. 1–38). Oxford University Press.

Spector, A., & Biederman, I. (1976). Mental set and mental shift revisited. *The American Journal of Psychology*, 66, 9–679.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18(6), 643.

Sudevan, P., & Taylor, D. A. (1987). The cuing and priming of cognitive operations. *Journal of Experimental Psychology: Human Perception and Performance*, 13(1), 89.

Wallis, J. D., & Miller, E. K. (2003). From rule to response: Neuronal processes in the premotor and prefrontal cortex. *Journal of Neurophysiology*, 90, 1790–1806.

van t Wout, F., Lavric, A., & Monsell, S. (2015). Is it harder to switch among a larger set of tasks? *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 41(2), 363.

Zanolie, K., Van Leijenhorst, L., Rombouts, S. A. R. B., & Crone, E. A. (2008). Separable neural mechanisms contribute to feedback processing in a rule-learning task. *Neuropsychologia*, 46(1), 117–126.

Zarr, N., & Brown, J. W. (2016). Hierarchical error representation in medial prefrontal cortex. *NeuroImage*, 124, 238–247.