Emergence of task-dependent representations in working memory circuits

Cristina Savin[1][*][†] and Jochen Triesch[1,2]

[1] Frankfurt Institute for Advanced Studies, Frankfurt am Main, Germany
[2] Physics Department, Goethe University, Frankfurt am Main, Germany

*Correspondence:
Cristina Savin, Institute of Science and Technology, Am Campus 1, Klosterneuburg 3400, Austria
e-mail: csavin@ist.ac.at

[†]Present address:
Cristina Savin, Institute of Science and Technology, Klosterneuburg, Austria

A wealth of experimental evidence suggests that working memory circuits preferentially represent information that is behaviorally relevant. Still, we are missing a mechanistic account of how these representations come about. Here we provide a simple explanation for a range of experimental findings, in light of prefrontal circuits adapting to task constraints by reward-dependent learning. In particular, we model a neural network shaped by reward-modulated spike-timing dependent plasticity (r-STDP) and homeostatic plasticity (intrinsic excitability and synaptic scaling). We show that the experimentally-observed neural representations naturally emerge in an initially unstructured circuit as it learns to solve several working memory tasks. These results point to a critical, and previously unappreciated, role for reward-dependent learning in shaping prefrontal cortex activity.

Keywords: working memory, reward-dependent learning, STDP, intrinsic plasticity, synaptic scaling, prefrontal cortex, delayed categorization

1. INTRODUCTION

Working memory is defined as the temporary storage of stimulus-specific information during a delay period. This function has been traditionally associated with circuits in prefrontal cortex (PFC). Classic work in monkeys revealed that single neurons in this region exhibit selective persistent activity during the delay period (Miyashita, 1988; Goldman-Rakic, 1990) and its disruption (by electrical stimulation, or due to distracters) leads to a decay in performance (Funahashi et al., 1989). These early observations have been interpreted as the circuit exhibiting attractor dynamics, which enable a subset of the neurons to maintain high firing throughout the delay after a brief stimulus presentation (Amit and Brunel, 1997; Brunel and Wang, 2001). This view has been revised in recent years, as it was shown that most neurons in PFC change their firing rates during the delay (Miller et al., 1996; Chafee and Goldman-Rakic, 1998; Pesaran et al., 2002; Rainer and Miller, 2002; Barak et al., 2010), suggesting that working memory circuits rely on feedforward rather than attractor dynamics (Goldman, 2009). Still, while experiments generally agree on how information is represented in working memory circuits, i.e., using spatio-temporal patterns of neural activity, exactly what information gets encoded is less clear.

An accumulation of data across different working memory experiments paints an increasingly complex picture of the features encoded in PFC. We find neurons may represent the previous stimulus, the forthcoming action, or a more complex function of the two (Durstewitz et al., 2000). When the task requires a generalization across stimuli, neurons develop category selectivity (Freedman et al., 2001). Moreover, there is a gradual shift in these representations as the number of examples per class increases, with animals switching from a stimulus-response association strategy to representing categorical distinctions directly (Antzoulatos and Miller, 2011).

Things get even more complicated when animals need to alternate between different tasks. While PFC neurons generally represent the task to be performed (Asaad et al., 1998; Cromer et al., 2010; Roy et al., 2010; Warden and Miller, 2010; Meyer et al., 2011), they can differ significantly with respect to how the information is distributed across the population in different tasks. For instance, neurons can show task-specific changes in overall firing rate, in time-dependent response profiles and in stimulus and response selectivity (Asaad et al., 2000). In some situations, the same neurons seem to participate in encoding features related to different tasks (e.g., making different category distinctions, Cromer et al., 2010), effectively multiplexing information across contexts. In other situations, however, information is encoded in different neurons for different contexts (Roy et al., 2010), and—worse still—it is unclear when one or the other coding strategy may be employed. Generally, we are missing a unifying account for PFC representations during the delay period.

Here we hypothesize that reward-dependent learning underlies the variety in PFC representations in different working memory tasks. The data itself suggest that this may be the case: the most striking feature of the above experiments is not the diversity of neural responses, but the sheer number of neurons displaying an effect. Regardless of the actual task the monkey has been trained to carry out, a significant subset of the recorded neurons are found to exhibit selectivity to the specifics of that particular task. This is a strong indication that PFC neurons adapt their responses to reflect current cognitive demands. Indeed, PFC representations change significantly over the course of training (Rainer et al., 1998b; Rainer and Miller, 2000; Baeg et al., 2003; Kennerley and Wallis, 2009). Neural responses become increasingly sparse, the tuning of the neurons narrows, and the representation becomes more robust to input noise (Rainer and Miller, 2000). These changes in neural representation parallel behavioral learning, and

allow for a better decoding of stimuli and actions (Baeg et al., 2003). Moreover, since the training-induced changes in neural responses include changes in functional connectivity (Baeg et al., 2007), it seems likely that associative learning within the circuit is responsible—at least in part—for the refinement of neural representations with learning. The specific mechanisms involved remain unclear, however.

We assume that learning in PFC is reward-dependent. This hypothesis is consistent with the observation that dopamine, a neuromodulator associated with reward prediction error (Schultz, 1998), modulates synaptic plasticity in this circuit (Otani et al., 2003). It also explains the dependence of neural representations on the magnitude of the expected reward (Kennerley and Wallis, 2009). However, the primary reason for our assumption is computational. Working memory circuits are know to operate under strict capacity constraints (Cowan, 2001), and a circuit with limited resources cannot simply encode everything. To perform well, it needs to represent the specific aspects of the stimulus that matter for the task at hand (Duncan, 2001). Hence, reward should modulate learning so as to shift representations toward task-relevant features. This points to a critical and previously unrecognized role for reward-dependent plasticity in shaping prefrontal representations.

Can reward-dependent learning alone explain the wide variety of experimental observations on PFC encoding? To address this question, we studied the effects of reward-dependent learning on the encoding properties of neurons in a working memory circuit. More specifically, we trained a generic recurrent neural network to solve tasks similar to those employed in working memory experiments. We then investigated the neural representations emerging in the circuit and compared them to neural data. We chose a simple abstract model for the network dynamics in which the output of a neuron depends only on its instantaneous inputs. Learning was implemented by reward-dependent spike timing dependent plasticity (rSTDP) (Izhikevich, 2007), supplemented by homeostatic mechanisms that stabilized the network dynamics during learning (Lazar et al., 2009). Importantly, as individual neurons have no memory themselves, the storage of information in this circuit relies exclusively on the recurrent connectivity. While this simple model cannot capture the full complexity of the temporal dynamics in PFC, it allows us to focus specifically on the reward-dependent reorganization of recurrent connections and its effects on circuit function.

We found that our model is able to capture key aspects of neuronal dynamics during working memory tasks. Neurons in the model develop specificity in space and time and, depending on the task, they preferentially encode individual stimuli, actions, or context information. In a simple delayed-response task, neurons encode stimulus identity (Miller et al., 1996; Constantinidis and Franowicz, 2001). In a delayed-categorization task, neurons learn to preferentially encode category boundaries (Freedman et al., 2001). Lastly, when learning several tasks at the same time, the degree of neural specialization depends on the specifics of the task, mirroring experimental data. When the task involves several independent category schemes, neurons act as "multiplexers," coding for different things in different contexts (Cromer et al., 2010); when the same stimuli need to be categorized differently

depending on behavioral context, the neurons segregate into distinct task-specific subpopulations (Roy et al., 2010). Furthermore, reward-dependent learning is critical for these results. A similar circuit trained by unsupervised learning shows a significant loss in working memory performance, paired with poorer neural representations. Taken together, our findings show reward-dependent learning could be a central force in the organization of working memory circuits.

2. MATERIALS AND METHODS

2.1. THE GENERAL TASK

The working memory tasks we investigated share a simple general structure (**Figure 1A**): at the beginning of a trial one stimulus (out of K) is briefly presented to the network. After a delay period (either fixed for a block of trials or selected at random from a given distribution) a "Go" cue is presented, after which the reward is given according to the action selected by the model (one out of M)—either $+1$ for a correct answer or -1, otherwise. Different tasks correspond to different mappings between stimuli and actions and each are described in detail in the corresponding Results section. To speed up learning, we adopt the same strategy employed in training animals for experiments, i.e., we start with the minimum delay version of the task and progressively increase the duration of the delay period during learning (Klingberg, 2010).

2.2. NETWORK MODEL

An overview of the network is shown in **Figure 1B**. The recurrent network consists of N units (unless otherwise specified, $N = 250$), 80% excitatory and 20% inhibitory, with sparse random connectivity. Input units encoding different stimuli (and possibly the context cue) activate small, non-overlapping subsets within the recurrent layer, each consisting of N_{in} excitatory neurons $N_{in} = 5$; the activation of the input unit provides a suprathreshold current which forces the corresponding subpopulation to be active for one time step. The output layer receives inputs from all excitatory units within the network and generates a decision response through a winner-take-all (WTA) mechanism. This decision outcome determines the received reward, which in turn modulates synaptic changes through r-STDP. Reward-dependent learning affects both excitatory synapses within the recurrent network and those connecting to the decision layer.

We chose a abstract model for the neural dynamics, whose simplicity allows us to focus on the essential mechanisms required for explaining the data. More specifically, we used linear threshold units to model neurons within the network, i.e., each unit has a binary output:

$$x_i(t) = I_i(t) \geq \Theta_i, \tag{1}$$

with activation depending on the total current to the neuron $I_i(t)$ and the neuron's spike threshold Θ_i (this threshold also changes over a slower time scale because of homeostatic mechanisms, see below). The activity proceeds in discrete time steps, with synchronous updates for all neurons. The input to a neuron is given by:

$$I_i = w_i^T \cdot \mathrm{x} + \epsilon, \tag{2}$$

FIGURE 1 | Schematic description of the model. (A) Delayed response task: at the beginning of each trial, one of K stimuli is presented to the network, requiring a stimulus-dependent action to be performed at the end of the delay period. When the cue appears, an action is selected yielding a corresponding reward. Initial trials have short delays, and we progressively increase the delay period during learning. **(B)** The network of threshold linear neurons receives localized, stimulus-specific inputs; the decision units determine the action to be performed by winner-take-all (WTA). The corresponding reward modulates plasticity events at synapses within the recurrent network and to the decision units. **(C)** Reward-dependent STDP is implemented using eligibility traces, with changes occurring only at the time of reward (see main text for details); additionally, **(D)** the neuron threshold is homeostatically regulated and the incoming synapses to each neuron are normalized.

where column vectors w and x describe the synaptic weights and the activity of all presynaptic neurons, respectively. The stochastic term ϵ corresponds to an unspecific background input to each unit, modeled as independent uniform random noise, $\epsilon \in [0, 0.1]$. Importantly, since the model neuron has no memory itself, working memory can develop in the model only through the network dynamics. Hence, we can use the model to study specifically reward-dependent plasticity and its effects on information storage.

The connectivity matrix was initialized randomly at the beginning of each experiment, with weights drawn from the uniform distribution $w_{ij} \in [0, 1]$, followed by a sum-to-one weight normalization of incoming synapses. The connection probabilities were $p_{ee} = 0.1$, $p_{ei} = 0.25$, $p_{ie} = 0.4$, $p_{ii} = 0$, with indices "e" and "i" marking the excitatory and inhibitory populations, respectively.

For the decision layer, the current to each neuron is computed as before, with the WTA mechanism selecting the neuron with the strongest input as the only active unit: $I_m = w_i^T \cdot x + \epsilon$, $x_m = 1$ if $m = \text{argmax}_j I_j$, and $x_m = 0$, otherwise. Decision neurons were allowed to fire during the delay period without any effect on reward.

2.3. PLASTICITY MECHANISMS

2.3.1. Reward-dependent learning

We adapted a model for r-STDP from Izhikevich (2007) (**Figure 1C**). As in the original, each synapse has an associated eligibility trace e_{ij}:

$$\frac{de_{ij}}{dt} = -\frac{e_{ij}}{\tau_e} + x_i(t) \cdot x_j(t-1) + f \cdot x_i(t-1) \cdot x_j(t) \quad (3)$$

where x_i and x_j are the output of the pre- and post-synaptic neuron, respectively, and f is a model parameter ($f = 1$ for synapses

in the recurrent layer, and $f = 0.01$ for synapses in the motor layer).

The eligibility trace stores a history of potential weight changes at the synapse, with an exponential decay, specified by the time constant τ_e ($\tau_e = 2.5$). The individual synaptic plasticity events follow a simplified STDP window: potentiation occurs when presynaptic activity is followed by a postsynaptic spike, while the reverse pattern causes depression, with a width of 1 time step (since that is the timescale of causal interactions in our network). Additionally, weights are rectified such that $w_{ij} \geq 0$ in order to respect Dale's law.

At the time of the reward synaptic weights change proportionally to the eligibility trace e_{ij} and the reward signal r:

$$w_{ij}(t+1) = w_{ij}(t) + \eta \cdot r(t) \cdot e_{ij}(t), \quad (4)$$

with learning rate η.

For simplicity, we used the absolute reward as the signal modulating synaptic modifications instead of the reward prediction error (Schultz, 1998), as done in previous models (Izhikevich, 2007). Additionally, we assumed the reward to be either positive or negative, as biological evidence from cortico-striatal synapses suggests that dopamine can induce both potentiation or depression in response to tetanic stimulation, depending on its concentration relative to baseline (Reynolds and Wickens, 2002). Specifically, at the time of the reward delivery $r(t) = 1$, if the motor output was correct and $r(t) = -1$, otherwise; $r(t) = 0$ at all other times.

To ensure that the system is given time to exploit the emerging neural representation, we assumed that changes at synapses to the decision layer occur faster than those in the recurrent network ($\eta = 10^{-5}$ for synapses in the recurrent layer and $\eta = 10^{-4}$ for those connecting to decision neurons). These changes in learning

rate were paralleled for intrinsic plasticity to ensure that the dynamics remain stable during learning (see below).

2.3.2. Homeostatic plasticity

A critical problem when optimizing recurrent networks is how to stabilize the dynamics during learning (Turrigiano and Nelson, 2004; Lazar et al., 2009). Traditionally, working memory models with attractor dynamics circumvent this problem by keeping weights fixed and fine-tuning a limit set of gain parameters by hand (Brunel and Wang, 2001). Here, we use two distinct homeostatic mechanisms to ensure stability (**Figure 1D**): synaptic scaling (Turrigiano et al., 1998) and homeostatic threshold regulation (Zhang and Linden, 2003).

First, as synaptic scaling constrains the total drive received by neurons by rescaling all weights in a multiplicative fashion, we implemented this mechanism by an explicit weight normalization, $\Sigma_j w_{ij} = 1$. We chose this for simplicity, although a similar outcome could in principle be achieved through a local weight-dependent rule (Gerstner and Kistler, 2002). Second, intrinsic plasticity was implemented by assuming that the threshold of excitatory neurons adapts to maintain a certain mean average firing rate, $x_0 \in (0, 1)$:

$$\Delta\Theta = \lambda_{exc} (x(t) - x_0),\qquad(5)$$

where λ_{exc} is the time constant for the threshold adaptation ($x_0 = 0.03$ within the recurrent network and $x_0 = 0.25$ in the decision layer). As mentioned above, the timescale of plasticity for the decision units is 10 times faster to match the more rapid synaptic plasticity ($\lambda_{exc} = 10^{-4}$ within the recurrent layer, and $\lambda_{exc} = 10^{-3}$ for the decision units).

We assumed a similar threshold regulation for controlling the excitability of the inhibitory neurons. The specific form was suggested by experimental evidence showing that the excitability of inhibitory neurons is determined by the overall activity of neighboring excitatory neurons, estimated via the release of diffusible messengers, such as BDNF (Rutherford et al., 1998; Turrigiano and Nelson, 2000). Specifically, we assume that the threshold of inhibitory neurons changes as:

$$\Delta\Theta_{inh} = -\lambda_{inh} (\langle x_{exc}(t)\rangle - x_0),\qquad(6)$$

with $\langle x_{exc}(t)\rangle$ denoting the population average of the activation of all excitatory neurons at time t. This is a simplification of a more realistic input-specific regulation of excitability chosen for convenience, consistent with inhibitory neurons pooling activity across a large part of the circuit. As before, x_0 is the desired average firing rate of the excitatory neurons, and λ_{inh} is the learning rate ($\lambda_{inh} = 10^{-5}$). Although this mechanism is not strictly necessary for network stability, we find it improves memory performance and ensures a fairer distribution of neuronal resources across stimuli.

2.3.3. Other simulation parameters

All trials are assumed to have fixed duration $T_{trial} = 10$ time steps, with $2 \cdot 10^4$ trials per block. We repeat each experiment five times to quantify effects different sources of variability, such as the network initialization, internal noise, etc.

3. RESULTS

3.1. A DELAYED RESPONSE TASK

The most common experimental paradigm for exploring the circuits involved in working memory is the delayed response task, where a simple stimulus-specific response needs to be delivered after a delay (Rainer et al., 1998a; Durstewitz et al., 2000). Computational models of this function assume a circuit with distinct submodules for storing the initial stimulus (the working memory component), comparing it to the sample and deciding on the action (Engel and Wang, 2011). Here we focus on the first component, and thus assume a one-to-one mapping between stimuli and actions ($M = K$). Although we neglect the intermediate step, i.e., making same-or-different judgements, nonetheless the model preserves the nature of the underlying computation. Hence this simplification should not affect our results concerning the representation within the working memory circuit.

We used two variants for the basic setup: a fixed-and a variable-delay version (**Figure 2**, right). As its name suggests, the first uses a fixed delay for all trials in a block. This version is useful for estimating the memory capacity of the network, defined as the longest delay for which performance is better than chance. However, it could potentially lead to unrealistic delay-specific representations. In the second setup, the delay for each trial is selected uniformly at random between one and a maximum delay T_{max} time steps. This version seems closer to the true constraints of the biological system, where information needs to be accessible on demand whenever the the environmental conditions call for it. Hence, we used the second version of the task to investigate the emerging neural representations.

We found that the network performance is influenced by task difficulty (**Figure 2**). As expected, it decreases with increasing delay, due to the accumulation of noise. For intermediate delays, the fixed-delay task yields slightly better results compared to the variable-delay task, consistent with it being computationally simpler. At longer delays however, the network exhibits a sharp performance decay, which signals the network reaching its memory capacity. In the variable delay task, performance degrades more gracefully, as shorter memory spans are still rewarded. In both cases, we found that recall performance increases with network size N, and decays with the number of distinct stimuli K and that the incremental learning paradigm dramatically improves network performance (not shown).

Importantly, performance is remarkably stable within a block of trials despite the constant changes induced by the different plasticity mechanisms, with the network reaching the final performance after a small number of trials (on the order of 100 trials). The critical condition to achieve such good and stable performance is a sparse representation within the recurrent layer (enforced by intrinsic plasticity), combined with balanced rSTDP. While not strictly necessary, synaptic scaling and inhibitory plasticity improve performance; additionally we found it was beneficial to reduce the LTD component for learning in the motor layer (presumably because it limits the interference due to motor activity in the delay period). Overall, the interaction between different plasticity mechanisms is needed for the circuit to maintain stable function despite variable underlying neural "hardware."

To examine the representation that emerges after learning, we measured both the spatial and the temporal selectivity of neural
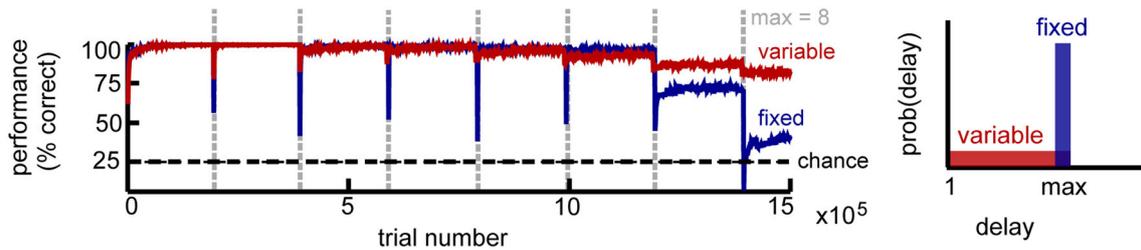
FIGURE 2 | Circuit performance for a simple delayed response task. Recall performance as a function of the number of trials for $K = 4$ stimuli. Vertical lines mark the time when the maximum delay is incremented. Within a block the maximum delay is kept fixed, with two responses. variations: in the fixed delay task (blue) all trials have the same delay, while in the variable delay task (red) the delay of each trial is drawn independently from the uniform distribution $[1, max]$. Performance estimated across 100 trial blocks.

For the spatial component, we computed the average neural activation during the delay period for each stimulus (**Figure 3A**, top). This simple measure reveals that most of the neurons respond to one of the stimuli, while remaining relatively silent for the others, as demonstrated in classic working memory experiments (Miyashita, 1988). To better quantify the effect, we used a measure called the depth of selectivity (Rainer et al., 1998b), defined as:

$$S = \frac{N_{\text{cond}} - \frac{\Sigma R_i}{R_{\text{max}}}}{N_{\text{cond}} - 1}, \tag{7}$$

where R_i is the firing rate corresponding to stimulus i, $R_{\text{max}} = \max\{R_i\}$ and N_{cond} is the number of different behavioral states considered, here the K stimuli. This measure takes the value zero when the neural response is identical for all objects and can reach the maximum of one when the neuron responds exclusively to one of the stimuli. Note that we will use this measure more generally in the following sections, to also measure the specificity to distinct actions or contexts. The depth of selectivity confirmed that most neurons exhibit stimulus-dependent activation (**Figure 3A**, bottom).

Neural responses are structured also in the temporal domain, reproducing at least qualitatively the temporal specificity in experiments (Meyers et al., 2008). A post-stimulus time histogram (PSTH) of the network responses for a given stimulus reveals that, although before learning the response is highly variable (**Figure 3B**, top and **Figure 3C**, left), after learning neuronal responses become highly reproducible (**Figure 3B**; note that neuron indices were reordered as a result of sorting the neuron by the time of the peak response). Moreover, neurons respond at specific times relative to stimulus onset, pointing to a synfire chain-like representation (Aertsen et al., 1996; Prut et al., 1998). Such temporal dynamics allow neurons to remain stimulus specific, while maintaining a sparse activation enforced thorough the homeostatic regulation of neural excitability. Additionally, the network dynamics reflect the details of the task (**Figure 3C**): the delay itself is encoded much better during the fixed-delay version of the task. A low-dimensional projection of the neural activity by principal component analysis (PCA) reveals distinct vs. overlapping stimulus-specific clusters in the fixed- and the variable delay task, respectively. This reflects the intuition that the time since the stimulus presentation is important for the fixed delay task, whereas in the variable delay version the motor layer just needs to linearly separate the activity corresponding to different stimuli, irrespective of the delay. The time-dependent encoding is also reflected in the connectivity matrix, which becomes sparse and more feedforward (**Figure 3D**). More generally, learning organizes the network in largely non-overlapping feedforward chains, each starting from one of the input sub-populations and with a total size determined by the number of inputs, the size of the network, and the sparseness enforced through the homeostatic mechanisms (not shown). In summary, in a simple delayed-response task, the network uses distributed representations for encoding information about the stimuli across time and space, in a way that makes it easily accessible for decision circuits and is consistent with experiments.

3.2. A DELAYED CATEGORIZATION TASK

Neurons in PFC can encode either the initial stimulus, or the action to be taken in response to it (Brody et al., 2003). For the simple delayed-response task above there is no difference between the two, as actions simply signal stimulus identity. To investigate under which conditions the circuit learns to represent preferentially stimuli or actions, we used a delayed categorization task, inspired by experiments in monkeys, in which arbitrary categories are defined using morphed images (generated from e.g., cat and dog prototypes), see Freedman et al. (2001).

To mimic this paradigm, we constructed an arbitrary map between $K = 8$ stimuli and $M = 2$ decision outputs signaling stimulus class. Here, category boundaries are defined exclusively by the reward function (Freedman et al., 2001; Antzoulatos and Miller, 2011), unlike some experiments in which category specificity may be—to some extent—stimulus driven (Meyers et al., 2008). For illustration purposes, we define the mapping by stimulus color (**Figure 4A**, right), though in the model the random initialization of the connectivity makes any subdivision of non-overlapping stimuli to be equivalent.

Our network is able to successfully learn the task (75% correct for a delay of five time steps). The neural representations for this task show some novel characteristics compared to the simple delay task, which reflect the experimental data (Freedman et al., 2001). While some of the neurons still respond selectively

FIGURE 3 | (A) Neural selectivity in a simple delayed response task ($T_{max} = 5$, variable delay). Top: neural responses averaged across trials where one of four stimuli (different colors) was presented, for a subset of 15 randomly selected neurons. Bottom: selectivity of neural responses across the population for one example experiment; estimated using activity in 1000 trials at the end of learning. Note that the first 20 neurons receive direct inputs from the input layer. **(B)** Comparison of the post-stimulus time histogram of neural responses before and after learning for one example stimulus. Neuron indices have been reordered based on the time of maximum firing relative to stimulus onset. **(C)** A low-dimensional view of the population dynamics in response to the same stimulus before learning (left) or after training using either the fixed (right) or the variable delay (middle) paradigm. Individual points correspond to the state of the network projected along the first three principal components; color intensity marks the time since the stimulus presentation and different points of the same color correspond to different trials. **(D)** The corresponding weight matrix at the end of leaning.

FIGURE 4 | Neural selectivity in a delayed categorization task. (A) Average responses to each stimulus in 10 randomly selected example neurons. **(B)** Specificity of neural responses to one of the two categories (red or blue); colors show preferred category for each neuron. Neural selectivity was estimated using activity in 1000 trials at the end of learning. $T_{max} = 5$, variable delay.

to individual stimuli, a significant subpopulation responds now to several stimuli, and often to those belonging to the same category (**Figure 4A**). Using the depth of selectivity (with categories rather than stimuli as behaviorally relevant variable) enables us to quantify the category selectivity of neurons across the population (**Figure 4B**). Using this metric, we found that a significant fraction of the neurons (32% of excitatory neurons have

$S \geq 0.75$) exhibit category selectivity, close to the 33% reported in monkeys (Freedman et al., 2001). As in the previous experiment, their representations are time-varying; at any time, only a small fraction of neurons encode category information, with information being passed between different small subsets of neurons over the course of the trial, as shown in experiments (Meyers et al., 2008). Overall, these results confirm our hypothesis that the differences in neural selectivity in category- vs. stimulus-specific delayed response tasks could emerge due to the task-dependent reorganization of the circuit by reward-dependent learning.

3.3. MULTIPLE CATEGORY BOUNDARIES

Up to now, we have looked at representations in a circuit that specializes on one specific memory task. While this scenario is useful for describing a typical behavioral experiment in monkeys, in real-life conditions the PFC needs to flexibly (and quickly) switch across a variety of different tasks.

How exactly are multiple tasks represented in PFC circuits? The answer should not come as a surprise: "it depends on the tasks." For tasks involving non-overlapping stimuli, in particular, two independent categorization tasks (cats vs. dogs and sedans vs. sports cars), the activity of many neurons reflects both category distinctions. Thus, the neurons multitask different types of information depending on the context (Cromer et al., 2010). In

contrast, when the same stimuli need to be categorized differently depending on behavioral context, the two category boundaries are represented by largely non-overlapping neuronal populations (Roy et al., 2010).

Can a difference in task constraints explain these conflicting results? To answer this question, we constructed two versions of the multi-class delayed categorization task, similar to those used experimentally. First, to implement the multiple independent categories task we used $K = 8$ stimuli and defined two non-overlapping subsets, representing the animals and cars in the original experiment. These subsets were each split in two categories, corresponding to, e.g., the cats vs. dogs distinction, see **Figure 5A**, with $M = 4$ actions, corresponding to the different category distinctions. As in the basic task, the cue signal (now two inputs) was provided directly the decision layer; the cue was active for one time step at the end of the delay period. We found that the network was able to learn this task (average performance 85% for a variable delay task, with maximum delay $T_{max} = 5$). To assess the emerging neural representations learned for this task, we measured the average firing rate of the neurons in response to different stimuli. We found that many of the neurons responded strongly to several stimuli (**Figure 5B**). These stimuli often belonged to the same class (**Figure 5B**, e.g., for neurons 1, 3, 4, etc.), reproducing the category selectivity we have seen previously, but often neural responses are strong also for stimuli corresponding to different tasks (**Figure 5B**, e.g., the first neuron responds to category 1 and 3). Measuring the category specificity of neurons for each of the two contexts revealed that most neurons are strongly category selective (**Figure 5C**). Moreover, 33.5% of the neurons were sensitive to both category distinctions (selectivity threshold 0.75, see **Figure 5D**). This suggests that, indeed, when the tasks do not interfere with one another the circuit should multiplex information across tasks for good performance.

Second, to model the scenario involving overlapping category boundaries, we assumed $K = 8$ input stimuli that are classified, depending on the context, using two orthogonal category boundaries (**Figure 6A**, right). In this case, the context needs to be provided at the beginning of the trial, together with the stimulus (the context, i.e., which task needs to be performed in the current trial, is encoded as two non-overlapping sub-populations of the same size N_{in}, just as the stimuli). The decision layer consisted, as before, of $M = 4$ neurons, one for each category, and trials from both tasks were interleaved at random.

This version of the multiple categories experiment is significantly harder, as it requires storing information about both stimuli and the current context (because of the two extra inputs, we assumed the recurrent layer has a slightly increased firing rate $x_0 = 0.05$). Still, the network is able to perform significantly above chance (approximatively 60%, for a variable delay with $T_{max} = 3$). In contrast to the task before, however, fewer neurons develop category specificity (19.5% as opposed to 74.5%), most represent single stimuli and several neurons encode the context itself (**Figure 6B**), suggesting that the network converges to a largely input-driven solution, in which information about stimuli and task is stored separately and combined only at the level of the decision layer. Among the neurons that exhibit category specificity, almost all are selective to only one of the category boundaries (points with high selectivity cluster close to the two axes, and more so if the neural responses are context modulated, see **Figure 6C**), unlike the previous scenario. This observation is reiterated when restricting the analysis to neurons that show task specific encoding (**Figure 6C**, dark red). Thus, the network organizes into separate task-specific subpopulations, as seen experimentally (Roy et al., 2010).

Overall, we found that reward-dependent learning can account for the differences in category representation across experiments. Moreover, the emerging representations showed a strong

FIGURE 5 | Multitask categorization with non-overlapping domains. **(A)** Given a context cue, the network needs to perform one of two categorization tasks ("task 1" or "task 2"); there are eight stimuli in total (colored squares), half of each are used in each task. **(B)** Average stimulus-specific responses for 10 randomly selected neurons. **(C)** Overlap of the category selectivity in "task 1" vs. "task 2." **(D)** Correlation of the category specificity across tasks; shaded regions mark regions of high category selectivity.

FIGURE 6 | Multitask categorization with overlapping domains. The same eight stimuli need to be classified as belonging to class "blue" vs. "green" in task 1 or as "red" vs. "yellow" in task 2. The task to be performed in any given trial is determined by a context cue, provided as an extra input (in this example to neurons 41–50) during the initial stimulus presentation. **(A)** Overlap of the category specificity in the two tasks; color signals the preferred category for individual neurons. **(B)** The selectivity of the neuronal responses to the context; color marks the preferred context. **(C)** Correlation in category selectivity across the population. Colored dots correspond to neurons with high context selectivity (≥ 0.5).

task-dependent component, consistent with experimental observations (Asaad et al., 2000; Roy et al., 2010; Warden and Miller, 2010). The general match between our model and experiments is also quantitative, as shown in **Figure 7** (note that the asymmetry between tasks is a consequence of a small number of experiments). This is particularly remarkable given that these results were obtained without any tuning of the model parameters, beyond that required to obtain a good performance in the simple delayed-response task. Taken together, our results suggest task demands dramatically shape neuronal representations in working-memory circuits via reward-dependent learning.

3.4. THE IMPORTANCE OF REWARD-DEPENDENT LEARNING

To tease apart the contribution of different plasticity mechanisms to the observed effects, we compared our model to a similarly constructed network, in which weights within the recurrent layer remain fixed, or alternatively are modified by STDP independently of the obtained reward. In both cases, the readout to the decision layer was learned by r-STDP, with all homeostatic mechanisms in place.

We found that learning within the recurrent layer is critical for good memory performance, and in particular that networks with r-STDP are consistently better than those in which recurrent connectivity is fixed (**Figure 8A**). For a simple delayed-response task ($K = 4$ stimuli), reward modulation is not strictly necessary for good performance and unsupervised learning alone can improve neural representations (the performance in unsupervised learning is indistinguishable from that using reward-dependent learning; not shown), as reported elsewhere (Lazar et al., 2009). This result is expected, since when each stimulus defines an action, it is best to represent each input as distinctly as possible, something which can be done by unsupervised learning. Indeed, the emerging representations are similar for the different learning scenarios (stimulus-specific synfire chains; not shown), such that they can be exploited for reward-dependent learning at the decision units.

Importantly, we found that simple unsupervised learning by STDP is no longer sufficient once the task difficulty is increased, by introducing more stimuli and more complex decision boundaries. Indeed, a very different picture emerges when comparing reward dependent vs. unsupervised learning in a categorization task ($K = 8$ stimuli randomly mapped into $M = 2$ categories). In this case, we find that the performance of the two differs significantly (**Figure 8B**). A possible reason for this difference is that attempting to represent each different stimulus separately, via unsupervised learning, exceeds the capacity of this particular network. Because if this, unsupervised learning results in poorer performance in this task. Furthermore we found that the outcome of unsupervised learning is less robust than that of reward-dependent learning: error levels depend on the particular instantiation of the network, leading to increased across trial variability (**Figure 8B**, shaded region in red vs. blue). This dissociation is also apparent at the level of the neural representations. After reward-independent learning, the percentage of category-specific neurons is significantly lower to both our model and the experimental data (20.5% instead of 32% for reward-dependent learning and 33% in the data; see **Figure 8C**). Furthermore, the network responses appear more noisy, suggesting that the number of stimuli exceed the capacity of the network and the reward-independent learning cannot learn a robust representations for all stimuli. All in all, this suggests that for complex tasks, when the pool of available resources is indeed a limiting factor, neuronal representations need to shift toward task-relevant features for good memory performance.

4. DISCUSSION

Prefrontal circuits are shaped by a variety of task-related variables. These representations are likely to form during extensive training prior to experimental recordings, but the mechanisms underlying this development are poorly understood. Here, we have shown that representations similar to those reported experimentally naturally emerge in an initially unstructured circuit

FIGURE 7 | Summary of the results for different versions of multitask categorization; comparison between model and experiments. The category specificity at the end of learning was measured for the two variants of the multiple-category task using either overlapping or non-overlapping category boundaries. We restricted the analysis to the subset of neurons that showed any task specificity (defined as $S \geq S_{specific}$), as done for the experimental data analysis; the proportion of these neurons that are selective to one or both of the category boundaries was reported, averaging across five runs; $T_{max} = 5$ and $S_{specific} = 0.75$ for the non-overlapping version and $T_{max} = 3$ and $S_{specific} = 0.5$ for the overlapping categories discrimination, reflecting the increase in task difficulty. Experimental data reproduced from Cromer et al. (2010) for non-overlapping and from Roy et al. (2010) for overlapping categories, respectively.

through reward-dependent learning. Moreover, we found that a few generic mechanisms (rSTDP and homeostasis) are sufficient to explain a range of puzzling (and seemingly complex) experimental observations. Neurons in our model developed stimulus and action specificity, both across neurons and in time, as seen experimentally (Miller et al., 1996; Chafee and Goldman-Rakic, 1998; Rainer and Miller, 2002). The same model (with no further parameter tuning) could also account for neural representations during context-dependent tasks. For tasks involving multiple independent category sets, individual neurons multiplexed information across different contexts, matching experimental observations (Cromer et al., 2010); when the same stimuli mapped into different actions depending on the context, neurons specialized to represent single category distinctions, as in Roy et al. (2010). To the best of our knowledge, our model is the first to provide an unified account of these observations.

When comparing our model to a network using reward-independent learning we found reward-dependent plasticity to be critical for solving hard tasks, such as the categorization of many stimuli. This finding is consistent with the notion that reward-dependent learning should be particularly important when resources are limited, either in terms of the amount of information that can be stored (unsupervised learning can be used to store four stimuli for the required time, but not eight), or in terms of the computations allowed for retrieving it (the readout

is linear). In such scenarios, separately representing each stimulus and then mapping the neural activity into the correct output becomes unfeasible (because the resources may not suffice for representing all stimuli individually or because reading out the answer becomes too complicated). Instead, the circuit needs to compute some parts of the map between stimuli and actions during the delay, by clustering together stimuli which should yield the same behavioral response. Given generally recognized resource limitations in working memory circuits (Cowan, 2001), this finding suggests that PFC needs to be malleable, with experience shaping the sensitivity of neurons to reflect current behavior.

Here we chose a very simple model for the network dynamics, known to have small memory capacity (Büsing et al., 2010), because we wanted to focus on the recurrent circuitry and its changes during learning. It should in principle be possible to extend the memory capacity of the network closer to the biologically-relevant range (order of seconds) by using larger networks, a more realistic model of the neural dynamics and including slow time-constants, e.g., NMDA receptors (Durstewitz et al., 2000; Brunel and Wang, 2001) or short-term facilitation (Mongillo et al., 2008). Nonetheless, as the restrictions enforced by resource limitations are likely general, we expect the main features of the representations emerging in the model to be preserved, at least qualitatively, in a detailed circuit. Thus, we predict reward-dependent learning should play a general role in the formation and task-specific tuning of working memory circuits.

From a developmental perspective, it is tempting to hypothesize that reward-dependent learning may play a role in the age-dependent improvement of working memory (estimated to be approximately four-fold between the ages of 4 and 14) (Luciana and Nelson, 1998), complementing other known factors such as the maturation of the underlying cortical architecture, a better representation of the inputs, the development of attention, or the usage of memorization strategies such as rehearsal and chunking (Gathercole, 1999). This suggestion is consistent with the known dependence of PFC function on dopamine in early life (Diamond and Baddeley, 1996). Furthermore, the same mechanisms may account for training-induced improvements in working memory in adults (Klingberg, 2010).

From a broader computational perspective, our work is also relevant in the context of reservoir computing (Lukoševičius and Jaeger, 2009). While this framework traditionally assumes fixed recurrent connectivity, recent work has increasingly argued for the importance of learning in shaping reservoir properties (Schmidhuber et al., 2007; Haeusler et al., 2009; Lazar et al., 2009). Previous work used general-purpose optimization through unsupervised learning. Here, however, the network is shaped directly by the task, which improves performance significantly compared to static networks or networks shaped by reward-independent learning. Thus, our model provides a stepping stone toward general task-specific optimization of recurrent networks.

Time-dependent representations are preferred to traditional attractor-based solutions (Amit and Brunel, 1997; Brunel and Wang, 2001; Mongillo et al., 2008) in our model, consistent with recent experimental observations (Miller et al., 1996; Chafee and Goldman-Rakic, 1998; Pesaran et al., 2002; Rainer and Miller, 2002; Barak et al., 2010) and previous theoretical predictions

FIGURE 8 | The importance of reward-dependent learning for shaping circuit dynamics. (A) Performance comparison of networks shaped by r-STDP (blue) vs. circuits where the recurrent connectivity is static (green) in a simple delayed response task. **(B)** Performance comparison of networks shaped by r-STDP (blue) vs. circuits where the recurrent connectivity is shaped by reward-independent STDP (red) in a 2-class categorization task; $K = 8$, fixed delay. For all conditions, the learning of the decision output was reward-modulated. Dark colors mark averages across five repetitions; light colors show the standard deviation around this mean. **(C)** Neuronal selectivity to stimulus category when learning in the recurrent circuit is reward independent; $K = 8$, fixed delay (compare to **Figure 3D**).

(Goldman, 2009). This effect is a consequence of intrinsic plasticity, which discourages neurons from remaining active for a long time (Horn and Usher, 1989). Given that homeostasis plays a critical role in stabilizing the circuit dynamics during learning (Turrigiano and Nelson, 2004), the fact that the emerging representation is time-varying is not really surprising. While our model emphasizes the temporal component of this representation, it is likely that the patterns of activity seen experimentally emerge through the interaction between feedforward and feedback dynamics, which would require a more detailed model of the neural dynamics. Although the homeostatic mechanisms acting in PFC circuits have yet to be characterized experimentally, it is tempting to assume that the sparsification of activity and increase in robustness observed experimentally after training (Rainer and Miller, 2000) may be signatures of the interaction between hebbian and homeostatic plasticity as shown in our model. More generally, similar mechanisms could play a role in developing feedforward dynamics in other recurrent circuits (see also Levy et al., 2001; Buonomano, 2005; Gilson et al., 2009; Fiete et al., 2010), for instance in other areas known to exhibit delay period responses, such as the perirhinal cortex, inferotemporal cortex,

or the hippocampus (Miller et al., 1993; Quintana and Fuster, 1999).

Our model combines both hebbian (r-STDP) and homeostatic (intrinsic plasticity, synaptic scaling) forms of plasticity, lending further support to the notion that the interaction between different forms of plasticity is critical for circuit computation (Triesch, 2007; Lazar et al., 2009; Savin et al., 2010). In particular, our results confirm the computational importance of intrinsic plasticity and synaptic scaling in excitatory neurons (Savin et al., 2010; Keck et al., 2012). To this, we add the role of inhibitory plasticity, which we found improved both neural representations and memory performance.

We view this model as a starting point for investigating reward-dependent learning in working memory circuits, to which many additions can be made. While the abstract network model used here allowed us to focus on the essential mechanisms underlying PFC coding, it would be important to investigate reward-dependent learning in more realistic spiking neural networks. Furthermore, the model for different plasticity mechanisms operating in the network could be refined as well. First, reward-dependent learning could be improved by using recent

extensions of r-STDP to spiking neuron populations (Urbanczik and Senn, 2009). Second, the simplistic regulation of inhibition should be replaced by realistic inhibitory plasticity (Castillo et al., 2011), which is expected to also aid network selectivity (Vogels et al., 2011). Third, activity-dependent structural plasticity could optimize the cortical connectivity to best encode the task-specific information (Savin and Triesch, 2010; Bourjaily and Miller, 2011), consistent with experimental observations that working memory training alters circuit connectivity (Takeuchi et al., 2010). Lastly, preliminary work, supported by recent observations about the effects of neuromodulation on inhibitory and homeostatic plasticity (Seamans et al., 2001; Di Pietro and Seamans, 2011), suggests that the homeostatic plasticity mechanisms themselves may be reward-dependent.

AUTHOR CONTRIBUTIONS

ACKNOWLEDGMENTS

REFERENCES

Aertsen, A., Diesmann, M., and Gewaltig, M. O. (1996). Propagation of synchronous spiking activity in feedforward neural networks. *J. Physiol. Paris* 90, 243–247. doi: 10.1016/S0928-4257(97)81432-5

Amit, D., and Brunel, N. (1997). Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cereb. Cortex* 7, 237–252. doi: 10.1093/cercor/7.3.237

Antzoulatos, E. G., and Miller, E. K. (2011). Differences between neural activity in prefrontal cortex and striatum during learning of novel abstract categories. *Neuron* 71, 243–249. doi: 10.1016/j.neuron.2011.05.040

Asaad, W. F., Rainer, G., and Miller, E. K. (1998). Neural activity in the primate prefrontal cortex during associative learning. *Neuron* 21, 1399–1407. doi: 10.1016/S0896-6273(00)80658-3

Asaad, W. F., Rainer, G., and Miller, E. K. (2000). Task-specific neural activity in the primate prefrontal cortex. *J. Neurophysiol.* 84, 451–459.

Baeg, E., Kim, Y., Huh, K., Mook-Jung, I., Kim, H., and Jung, M. (2003). Dynamics of population code for working memory in the prefrontal cortex. *Neuron* 40, 177–188. doi: 10.1016/S0896-6273(03)00597-X

Baeg, E., Kim, Y., Kim, J., Ghim, J., Kim, J., and Jung, M. (2007). Learning-induced enduring changes in functional connectivity among prefrontal cortical neurons. *J. Neurosci.* 27, 909–918. doi: 10.1523/JNEUROSCI.4759-06.2007

Barak, O., Tsodyks, M., and Romo, R. (2010). Neuronal population coding of parametric working memory. *J. Neurosci.* 30, 9424–9430. doi: 10.1523/JNEUROSCI.1875-10.2010

Bourjaily, M. A., and Miller, P. (2011). Excitatory, inhibitory, and structural plasticity produce correlated connectivity in random networks trained to solve paired-stimulus tasks. *Front. Comput. Neurosci.* 5:37. doi: 10.3389/fncom.2011.00037

Brody, C., Romo, R., and Kepecs, K. (2003). Basic mechanisms for graded persistent activity: discrete attractors, continuous attractors, and dynamic representations. *Curr. Opin. Neurobiol.* 13, 204–211. doi: 10.1016/S0959-4388(03)00050-3

Brunel, N., and Wang, X. (2001). Effects of neuromodulation in a cortical model of object working memory dominated by recurrent inhibition. *J. Comput. Neurosci.* 11, 63–85. doi: 10.1023/A:1011204814320

Buonomano, D. V. (2005). A learning rule for the emergence of stable dynamics and timing in recurrent networks. *J. Neurophysiol.* 94, 2275–2283. doi: 10.1152/jn.01250.2004

Büsing, L., Schrauwen, B., and Legenstein, R. (2010). Connectivity, dynamics, and memory in reservoir computing with binary and analog neurons. *Neural Comput.* 22, 1272–1311. doi: 10.1162/neco.2009.01-09-947

Castillo, P. E., Chiu, C. Q., and Carroll, R. C. (2011). Long-term plasticity at inhibitory synapses. *Curr. Opin. Neurobiol.* 21, 328–338. doi: 10.1016/j.conb.2011.01.006

Chafee, M., and Goldman-Rakic, P. (1998). Matching patterns of activity in primate prefrontal area 8a and parietal area 7ip neurons during a spatial working memory task. *J. Neurophysiol.* 79, 2919–2940.

Constantinidis, C., and Franowicz, M. (2001). The sensory nature of mnemonic representation in the primate prefrontal cortex. *Nature* 4, 311–316. doi: 10.1038/85119

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of menta storage capacity. *Behav. Brain Sci.* 24, 87–114. doi: 10.1017/S0140525X01003922

Cromer, J. A., Roy, J. E., and Miller, E. K. (2010). Representation of multiple, independent categories in the primate prefrontal cortex. *Neuron* 66, 796–807. doi: 10.1016/j.neuron.2010.05.005

Di Pietro, N. C., and Seamans, J. K. (2011). Dopamine and serotonin interactively modulate prefrontal cortex neurons *in vitro*. *Biol. Psychiatry* 69, 1204–1211. doi: 10.1016/j.biopsych.2010.08.007

Diamond, A., and Baddeley, A. (1996). Evidence for the importance of dopamine for prefrontal cortex functions early in life. *Philos. Trans. R. Soc. B Biol. Sci.* 351, 1483–1494. doi: 10.1098/rstb.1996.0134

Duncan, J. (2001). An adaptive coding model of neural function in prefrontal cortex. *Nat. Rev. Neurosci.* 2, 820–829. doi: 10.1038/35097575

Durstewitz, D., Seamans, J., and Sejnowski, T. (2000). Neurocomputational models of working memory. *Nat. Neurosci.* 3, 1184–1191. doi: 10.1038/81460

Engel, T. A., and Wang, X.-J. (2011). Same or different? A neural circuit mechanism of similarity-based pattern match decision making. *J. Neurosci.* 31, 6982–6996. doi: 10.1523/JNEUROSCI.6150-10.2011

Fiete, I. R., Senn, W., Wang, C. Z. H., and Hahnloser, R. H. R. (2010). Spike-Time-Dependent plasticity and heterosynaptic competition organize networks to produce long scale-free sequences of neural activity. *Neuron* 65, 563–576. doi: 10.1016/j.neuron.2010.02.003

Freedman, D. J., Riesenhuber, M., Poggio, T., and Miller, E. K. (2001). Categorical representation of visual stimuli in the primate prefrontal cortex. *Science* 291, 312–316. doi: 10.1126/science.291.5502.312

Funahashi, S., Bruce, C. J., and Goldman-Rakic, P. S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J. Neurophysiol.* 61, 331–349.

Gathercole, S. (1999). Cognitive approaches to the development of short-term memory. *Trends Cogn. Sci.* 3, 410–419. doi: 10.1016/S1364-6613(99)01388-1

Gerstner, W., and Kistler, W. (2002). Mathematical formulations of Hebbian learning. *Biol. Cybern.* 87, 404–415. doi: 10.1007/s00422-002-0353-y

Gilson, M., Burkitt, A. N., Grayden, D. B., Thomas, D. A., and Hemmen, J. L. (2009). Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks IV. *Biol. Cybern.* 101, 427–444. doi: 10.1007/s00422-009-0343-4

Goldman, M. S. (2009). Memory without feedback in a neural network. *Neuron* 61, 621–634. doi: 10.1016/j.neuron.2008.12.012

Goldman-Rakic, P. S. (1990). Cellular and circuit basis of working memory in prefrontal cortex of nonhuman primates. *Prog. Brain Res.* 85, 325–335. doi: 10.1016/S0079-6123(08)62688-6

Haeusler, S., Schuch, K., and Maass, W. (2009). Motif distribution and computational performance of two data-based cortical microcircuit templates. *J. Physiol. Paris* 103, 73–87. doi: 10.1016/j.jphysparis.2009.05.006

Horn, D., and Usher, M. (1989). Neural networks with dynamical thresholds. *Phys. Rev. A* 40, 1036–1044. doi: 10.1103/PhysRevA.40.1036

Izhikevich, E. (2007). Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cereb. Cortex* 17, 2443–2452. doi: 10.1093/cercor/bhl152

Keck, C., Savin, C., and Lücke, J. (2012). Feedforward inhibition and synaptic scaling – two sides of the same coin? *PLoS Comput. Biol.* 8:e1002432. doi: 10.1371/journal.pcbi.1002432

Kennerley, S., and Wallis, J. (2009). Reward-dependent modulation of working memory in lateral prefrontal cortex. *J. Neurosci.* 29, 3259–3270. doi: 10.1523/JNEUROSCI.5353-08.2009

Klingberg, T. (2010). Training and plasticity of working memory. *Trends Cogn. Sci.* 14, 317–324. doi: 10.1016/j.tics.2010.05.002

Lazar, A., Pipa, G., and Triesch, J. (2009). Sorn: a self-organizing recurrent neural network. *Front. Comput. Neurosci.* 3:23. doi: 10.3389/neuro.10.023.2009

Levy, N., Horn, D., Meilijson, I., and Ruppin, E. (2001). Distributed synchrony in a cell assembly of spiking neurons. *Neural Netw.* 14, 815–824. doi: 10.1016/S0893-6080(01)00044-2

Luciana, M., and Nelson, C. (1998). The functional emergence of prefrontally-guided working memory systems in four- to eight-year-old children. *Neuropsychologia* 36, 273–293. doi: 10.1016/S0028-3932(97)00109-7

Lukoševičius, M., and Jaeger, H. (2009). Reservoir computing approaches to recurrent neural network training. *Comput. Sci. Rev.* 3, 127–149. doi: 10.1016/j.cosrev.2009.03.005

Meyer, T., Qi, X.-L., Stanford, T. R., and Constantinidis, C. (2011). Stimulus selectivity in dorsal and ventral prefrontal cortex after training in working memory tasks. *J. Neurosci.* 31, 6266–6276. doi: 10.1523/JNEUROSCI.6798-10.2011

Meyers, E. M., Freedman, D. J., Kreiman, G., Miller, E. K., and Poggio, T. (2008). Dynamic population coding of category information in inferior temporal and prefrontal cortex. *J. Neurophysiol.* 100, 1407–1419. doi: 10.1152/jn.90248.2008

Miller, E., Erickson, C., and Desimone, R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *J. Neurosci.* 16, 5154–5167.

Miller, E., Li, L., and Desimone, R. (1993). Activity of neurons in anterior inferior temporal cortex during a short-term memory task. *J. Neurosci.* 13, 1460–1478.

Miyashita, Y. (1988). Neuronal correlate of visual associative long-term memory in the primate temporal cortex. *Nature* 335, 817–820. doi: 10.1038/335817a0

Mongillo, G., Barak, O., and Tsodyks, M. (2008). Synaptic theory of working memory. *Science* 319, 1543–1546. doi: 10.1126/science.1150769

Otani, S., Daniel, H., Roisin, M.-P., and Crepel, F. (2003). Dopaminergic modulation of long-term synaptic plasticity in rat prefrontal neurons. *Cereb. Cortex* 13, 1251–1256. doi: 10.1093/cercor/bhg092

Pesaran, B., Pezaris, J., Sahani, M., Mitra, P., and Andersen, R. (2002). Temporal structure in neuronal activity during working memory in macaque parietal cortex. *Nat. Neurosci.* 5, 805–811. doi: 10.1038/nn890

Prut, Y., Vaadia, E., Bergman, H., Haalman, I., Slovin, H., and Abeles, M. (1998). Spatiotemporal structure of cortical activity: properties and behavioral relevance. *J. Neurophysiol.* 79, 2857–2874.

Quintana, J., and Fuster, J. (1999). From perception to action: temporal integrative functions of prefrontal and parietal neurons. *Cereb. Cortex* 9, 213–221. doi: 10.1093/cercor/9.3.213

Rainer, G., Asaad, W. F., and Miller, E. K. (1998a). Memory fields of neurons in the primate prefrontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 95, 15008–15013. doi: 10.1073/pnas.95.25.15008

Rainer, G., Asaad, W. F., and Miller, E. K. (1998b). Selective representation of relevant information by neurons in the primate prefrontal cortex. *Nature* 393, 577–579. doi: 10.1038/31235

Rainer, G., and Miller, E. (2000). Effects of visual experience on the representation of objects in the prefrontal cortex. *Neuron* 27, 179–189. doi: 10.1016/S0896-6273(00)00019-2

Rainer, G., and Miller, E. (2002). Timecourse of object-related neural activity in the primate prefrontal cortex during a short-term memory task. *Eur. J. Neurosci.* 15, 1244–1254. doi: 10.1046/j.1460-9568.2002.01958.x

Reynolds, J., and Wickens, J. (2002). Dopamine-dependent plasticity of corticostriatal synapses. *Neural Netw.* 15, 507–521. doi: 10.1016/S0893-6080(02)00045-X

Roy, J. E., Riesenhuber, M., Poggio, T., and Miller, E. K. (2010). Prefrontal cortex activity during flexible categorization. *J. Neurosci.* 30, 8519–8528. doi: 10.1523/JNEUROSCI.4837-09.2010

Rutherford, L. C., Nelson, S. B., and Turrigiano, G. G. (1998). BDNF has opposite effects on the quantal amplitude of pyramidal neuron and interneuron excitatory synapses. *Neuron* 21, 521–530. doi: 10.1016/S0896-6273(00)80563-2

Savin, C., Joshi, P., and Triesch, J. (2010). Independent component analysis in spiking neurons. *PLoS Comput. Biol.* 6:e1000757. doi: 10.1371/journal.pcbi.1000757

Savin, C., and Triesch, J. (2010). "Structural plasticity improves stimulus encoding in a working memory model," in *Frontiers in Computational Neuroscience. Conference Abstract: Computational and systems neuroscience* (Salt Lake City, UT).

Schmidhuber, J., Wierstra, D., Gagliolo, M., and Gomez, F. (2007). Training recurrent networks by Evolino. *Neural Comput.* 19, 757–779. doi: 10.1162/neco.2007.19.3.757

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *J. Neurophysiol.* 80, 1–27.

Seamans, J., Gorelova, N., Durstewitz, D., and Yang, C. (2001). Bidirectional dopamine modulation of GABAergic inhibition in prefrontal cortical pyramidal neurons. *J. Neurosci.* 21, 3628–3638.

Takeuchi, H., Sekiguchi, A., and Taki, Y. (2010). Training of working memory impacts structural connectivity. *J. Neurosci.* 30, 3297–3303. doi: 10.1523/JNEUROSCI.4611-09.2010

Triesch, J. (2007). Synergies between intrinsic and synaptic plasticity mechanisms. *Neural Comput.* 19, 885–909. doi: 10.1162/neco.2007.19.4.885

Turrigiano, G., Leslie, K., Desai, N., Rutherford, L., and Nelson, S. (1998). Activity-dependent scaling of quantal amplitude in neocortical neurons. *Nature* 391, 892–896. doi: 10.1038/36103

Turrigiano, G., and Nelson, S. (2000). Hebb and homeostasis in neuronal plasticity. *Curr. Opin. Neurobiol.* 10, 358–364. doi: 10.1016/S0959-4388(00)00091-X

Turrigiano, G., and Nelson, S. (2004). Homeostatic plasticity in the developing nervous system. *Nat. Rev. Neurosci.* 5, 97–107. doi: 10.1038/nrn1327

Urbanczik, R., and Senn, W. (2009). Reinforcement learning in populations of spiking neurons. *Nat. Neurosci.* 12, 250–252. doi: 10.1038/nn.2264

Vogels, T. P., Sprekeler, H., Zenke, F., Clopath, C., and Gerstner, W. (2011). Inhibitory plasticity balances excitation and inhibition in sensory pathways and memory networks. *Science* 334, 1569–1573. doi: 10.1126/science.1211095

Warden, M. R., and Miller, E. K. (2010). Task-dependent changes in short-term memory in the prefrontal cortex. *J. Neurosci.* 30, 15801–15810. doi: 10.1523/JNEUROSCI.1569-10.2010

Zhang, W., and Linden, D. (2003). The other side of the engram: experience-dependent changes in neuronal intrinsic excitability. *Nat. Rev. Neurosci.* 4, 885–900. doi: 10.1038/nrn1248