

A Tweaking Principle for Executive Control: Neuronal Circuit Mechanism for Rule-Based Task Switching and Conflict Resolution

Salva Ardid^{1,2} and Xiao-Jing Wang^{1,3}

¹Department of Neurobiology and Kavli Institute for Neuroscience, Yale University School of Medicine, New Haven, Connecticut 06510, ²Department of Biology, Centre for Vision Research, York University, Toronto, Ontario M3J 1P3, Canada, and ³Center for Neural Science, New York University, New York, New York 10003

A hallmark of executive control is the brain's agility to shift between different tasks depending on the behavioral rule currently in play. In this work, we propose a “tweaking hypothesis” for task switching: a weak rule signal provides a small bias that is dramatically amplified by reverberating attractor dynamics in neural circuits for stimulus categorization and action selection, leading to an all-or-none reconfiguration of sensory-motor mapping. Based on this principle, we developed a biologically realistic model with multiple modules for task switching. We found that the model quantitatively accounts for complex task switching behavior: switch cost, congruency effect, and task-response interaction; as well as monkey's single-neuron activity associated with task switching. The model yields several testable predictions, in particular, that category-selective neurons play a key role in resolving sensory-motor conflict. This work represents a neural circuit model for task switching and sheds insights in the brain mechanism of a fundamental cognitive capability.

Key words: flexible behavior; sensory-motor conflict; top-down control; attractor dynamics; switch cost; congruency effect

Introduction

Humans and other animals exhibit a remarkable ability to flexibly select an appropriate response to a sensory input, and rapidly switch to another sensory-response mapping when task rule or goal changes (Allport et al., 1994; Rogers and Monsell, 1995; Meiran, 2000; Logan and Bundesen, 2003; Mayr and Kliegl, 2003; Monsell, 2003; Koch and Allport, 2006; Altmann and Gray, 2008; Gallagher, 2009; Klingberg, 2009; Collins and Frank, 2013). In experiments designed to study task switching, typically a rule indicates to a subject which feature of a given stimulus (e.g., color or shape of a visual object) determines a correct response. Task rule changes from trial to trial, or in blocks of trials, and sensory-response mapping must switch accordingly. An increasing number of monkey experiments have been performed using task-switching paradigms, combined with single-neuron recording from sensory (Mirabella et al., 2007; Sasaki and Uka, 2009; Cohen and Maunsell, 2011; Mante et al., 2013), parietal (Stoet and Snyder, 2004, 2007a; Kamigaki et al., 2009), and prefrontal (Mansouri et al., 2006, 2007, 2009; Hussar and Pasternak, 2009; Mante

et al., 2013; Buschman et al., 2012; Wendelken et al., 2012) cortical areas. Physiological evidence from these studies suggests that modulation of neural activity by task rule is typically weak. By contrast, most previous models commonly assume that a rule signal is similarly strong as sensory stimulation in affecting activity of cortical neurons (Cohen et al., 1990; Rougier et al., 2005; Brown et al., 2007). How can small rule modulation explain large (binary) behavioral changes in task switching?

In this work, we propose a solution to this puzzle, which we will refer to as “the tweaking hypothesis”. The basic idea is that network reconfiguration underlying task switching can be realized by very weak top-down signals from rule neurons. This is because a weak input can be greatly amplified through reverberating “attractor” dynamics in categorization and decision circuits, ultimately leading to circuit selection in favor of one sensory-motor mapping over another. We tested the tweaking hypothesis by developing a neural circuit model for task switching that consists of several basic and interacting circuit modules for sensory coding, rule representation, categorization of stimulus features, and action selection, respectively. The model was validated by reproducing salient single-neuron physiological observations and behavioral effects associated with task switching. Notably, the model identifies specific circuit mechanisms with testable predictions, in terms of neural dynamics and reward-dependent synaptic plasticity, that explain salient and widely observed behavioral effects associated with task switching: (1) Switch cost: response time (RT) and error rate (ER) increase in trials following a task switch. Switch cost splits into a component that decreases with a longer time for preparation and a residual component that remains (Rogers and Monsell, 1995; Brown et al., 2007; Vandierendonck et al., 2010). (2) Task-response interaction: on task repeat trials RT is shorter if the same response is

Received March 31, 2013; revised Oct. 17, 2013; accepted Nov. 2, 2013.

Author contributions: S.A. and X.-J.W. designed research; S.A. performed research; S.A. analyzed data; S.A. and X.-J.W. wrote the paper.

This work was supported by the Office of Naval Research Grant N00014-13-1-0297, The Swartz Foundation Fellowship (S.A.), and John Simon Guggenheim Foundation Fellowship (X.-J.W.). We thank T.A. Engel for fruitful discussions, and A. Compte, J.B. Morton, W. Wei, and T. Womelsdorf for comments on a previous version of the paper. We also thank the reviewers for their thoughtful comments and suggestions.

The authors declare no competing financial interests.

Correspondence should be addressed to Xiao-Jing Wang, Center for Neural Science, New York University, 4 Washington Place, New York, NY 10003. E-mail: xjwang@nyu.edu.

DOI:10.1523/JNEUROSCI.1356-13.2013

Copyright © 2013 the authors 0270-6474/13/3319504-14\$15.00/0

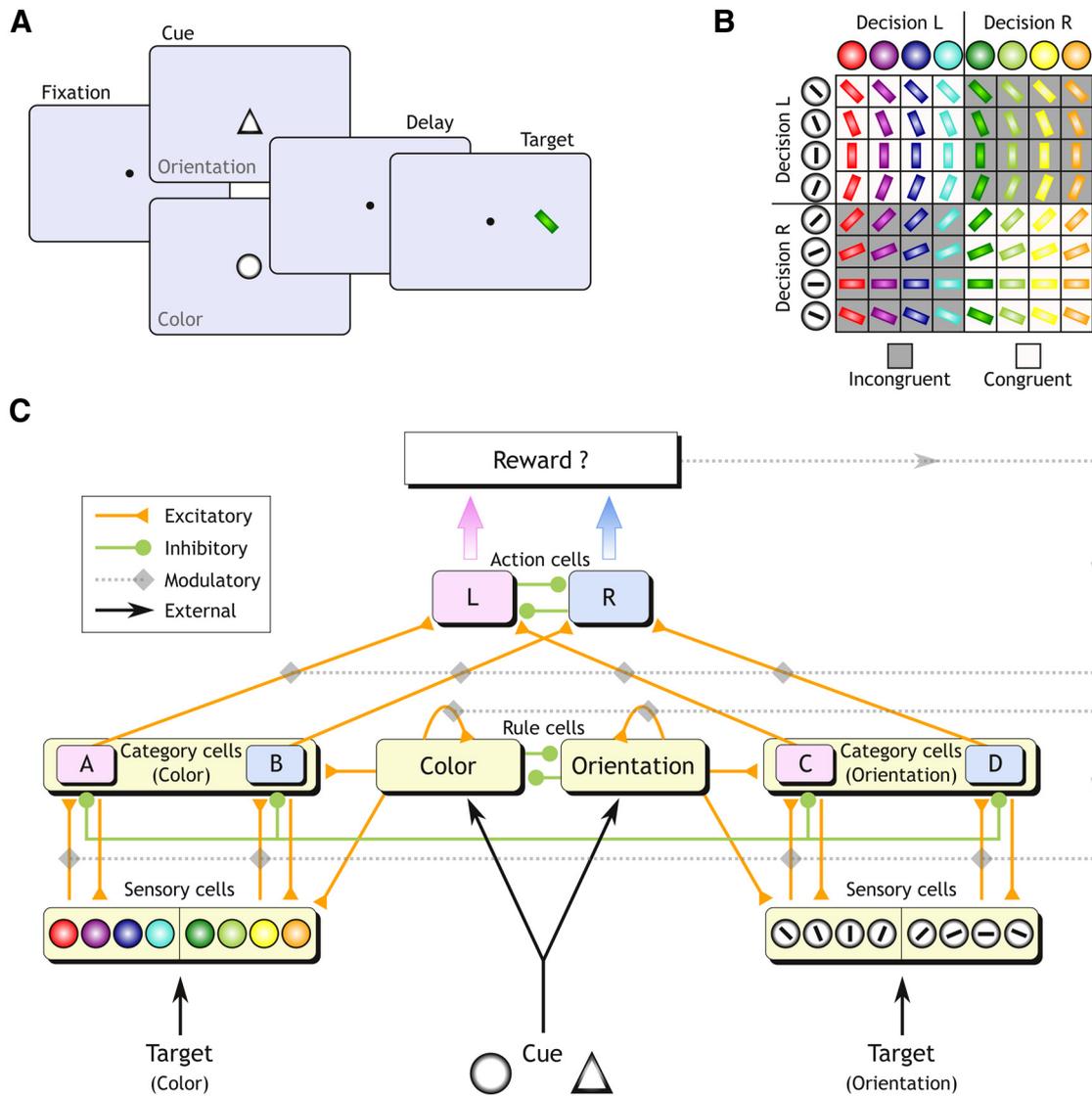


Figure 1. Neural circuit model for rule-based task switching. **A**, Cued task switch paradigm. At the beginning of a trial, a cue indicates the task to be performed (circle, color; triangle, orientation). After a delay, a target stimulus (oriented color bar) is presented. A response (left or right) has to be selected as fast as possible. **B**, Sensory-motor association. Target features in each dimension are classified into two categories that are mapped to opposite behavioral responses. Congruent (marked in white background) and incongruent (in gray) stimuli are defined based on whether the mapped behavioral response is the same or different according to the two rules (e.g., the green stimulus oriented at 135° maps to either right or left if the task rule is color or orientation, respectively). **C**, Model architecture with four interacting neural circuits that instantiate distinct cognitive processes: rule encoding, sensory processing, feature categorization, and action selection. The three nonsensory (rule, category, action) modules are modeled similarly, endowed with attractor dynamics underlying winner-take-all competition. A rule neural pool maintains the task context in the form of persistent activity, which can be switched to the other rule neural pool if instructed by a cue. Sensory cells independently process color and orientation features and receive top-down attentional modulations from rule and category cells. Category cells group feature values into classes; they project to action cells that compete with each other and select a motor response. Action cells display firing-rate adaptation (see Materials and Methods). The model is endowed with reward-dependent plasticity at recurrent synapses in the rule circuit and at feedforward synapses from sensory to category cells and from category to action cells.

selected than if it is different; by contrast, the opposite trend holds on switch trials (Rogers and Monsell, 1995; Meiran, 2000; Altmann, 2011). (3) Congruency effect: RTs and ER are larger when the stimulus is incongruent than when it is congruent (depending on whether the mapped behavioral response is different or the same according to the two rules; Rogers and Monsell, 1995; Meiran, 2000; Brown et al., 2007).

Materials and Methods

Simulation protocol. For concreteness, the model was designed for the cued task-switching experiment of Mirabella et al. (2007) and Caselli and Chelazzi (2011). In this experiment, a cue (circle or triangle) indicates which of two features (color or shape, respectively) must be mapped to a motor response. The two tasks are nearly symmetrical, in the sense that

neither feature is prepotent and the performance is similar for orientation- and color-based behavior (Mirabella et al., 2007; Caselli and Chelazzi, 2011). Each simulated trial (Fig. 1A) starts with a fixation period (250 ms). One of the two alternative cue stimuli is then randomly selected and presented for 250 ms, so the task can repeat or switch with equal probability. A target is presented after a delay (cue-target interval; CTI) of 1.75 s, although we also varied the delay to assess the dependence of the switch cost on CTI (see Fig. 6A, C). The target stimulus is a bar that includes the two features (color and orientation). Feature values map to one of the two alternative responses (left or right) according to a previously established sensory-motor association (Fig. 1B). A target stimulus can be “congruent” or “incongruent” depending on whether its two features point to the same response or to different responses, respectively. The feature values are randomly selected between 0°, 45°, 90°, and

135° for orientation; and red, blue, green, and yellow for color. Thus, congruent and incongruent stimuli appear with same probability. The response to a target have to be executed as fast as possible (with the only exception of simulations for Fig. 3D, where we fixed the duration of the target presentation to 1.5 s). In a trial, a response is produced when a neural pool in the action selection module reaches the threshold level at 20 Hz. Once a decision is made, the target stimulus is removed, which is followed by an intertrial interval of 500 ms before the following trial begins. In a small number of trials (2%), the response was premature (RT < 250 ms) or no response was initiated during the target presentation. These trials were considered invalid and discarded.

Model architecture. The model includes four reciprocally interacting neuronal circuits dedicated to distinct basic cognitive functions (Fig. 1C): sensory processing, rule encoding, feature categorization, and action selection. The sensory module constitutes two separate networks based on the evidence that color- and orientation-selective neurons are largely segregated in the ventral pathway of visual cortex: color selective domains in V4 have been identified in several electrophysiological (Zeki, 1973; Conway et al., 2007; Conway and Tsao, 2009; Harada et al., 2009) and fMRI (Conway and Tsao, 2006; Conway et al., 2007; Harada et al., 2009; Tanigawa et al., 2010) studies. Orientation-selective domains have also been shown (Ghose and Ts'o, 1997). Interestingly, a recent study has found that feature preference for color and orientation is segregated in V4 bands (Tanigawa et al., 2010). This does not imply a complete segregation of feature selectivity at a single-neuron level, but that the activity of the local population is more pronounced toward color or orientation representation (Tanigawa et al., 2010). Neural pools in our model represent averaged local activity. Thus, the selectivity of a pool of neurons encoding (e.g., green) is assumed to represent all local neurons in a color band that are selective to greenish colors, and in average not strongly tuned to orientation. This argument similarly applies to orientation-selective pools. Color is simplified to hue in the model, which can be mapped to a one-dimensional periodic variable (Conway et al., 2007). This allows us to use a ring implementation for each sensory network (Ben-Yishai et al., 1995; Compte et al., 2000; Ardid et al., 2007). Sensory circuits receive bottom-up inputs and are subject to top-down modulations from rule and category cells.

Rule, categorization, and action selection modules are all assumed to be attractor circuits (Wang, 2001, 2002; Wong and Wang, 2006). Note that, for the sake of simplicity, we used different neural pools for color categorization and orientation categorization. Category neurons for color and orientation may well be part of a single and general categorization system; indeed this is likely if category selectivity is formed through learning, by virtue of plasticity of input synapses onto neurons in such a categorization system (Engel and Wang, 2010). Attractor dynamics enables each module to integrate inputs over time, and generates winner-take-all competition between the neural pools. Rule cells compete to establish a proper context, which is maintained by self-sustained persistent activity, although it can be reconfigured by a cue input. Based on rule-biased sensory inputs, category cells determine through competition to which category the relevant target feature belongs. Action cells receive inputs from category cells and produce the final behavioral response.

Neuronal dynamics. We have used a reduced firing-rate implementation (Wong and Wang, 2006; Engel and Wang, 2011) of a spiking neural network model (Compte et al., 2000; Brunel and Wang, 2001; Wang, 2001, 2002; Ardid et al., 2007). This simplification facilitates an efficient exploration of parameters, and makes it computationally feasible to execute a circuit model running thousands of consecutive trials, which is required to analyze trial history effects in task switching. In this simplified model, inhibitory cells are only described implicitly, leading to effective inhibition between two selective neural pools. The neural dynamics is governed by the activation variable s for the synaptic current of the NMDA receptor:

$$\frac{ds}{dt} = -s/\tau_s + (1 - s)\gamma r,$$

where $\tau_s = 60$ ms and $\gamma = 0.641$. The firing rate r is given by a function of total input I (Abbott and Chance, 2005):

$$r = f(I) = \frac{aI - b}{1 - \exp[-d(aI - b)]}$$

For the sake of simplicity, we assume the same input–output relationship for rule, category and action selection neural pools, with $a = 518$ Hz/nA, $b = 208$ Hz, and $d = 123$ ms. For sensory circuits, the parameters are $a = 272$ Hz/nA, $b = 230$ Hz, and $d = 21.2$ ms. This modification increases the exponent of the power-law $f-I$ curve that has been proposed to approximate multiplicative gain by attention in the sensory cortex (Ardid et al., 2007). As a result, sensory neurons that receive stronger bottom-up input are also more sensitive to top-down modulations (Ardid et al., 2007). This is efficient for processing nonspecific top-down inputs, such the rule signal.

The total input I consists of two terms: synaptic inputs between neural pools in the model and external inputs.

Synaptic inputs. The synaptic input from neural population A to neural population B is described by:

$$I_{syn,j}^{A \rightarrow B} = \sum_{i \in A} g_{ij}^{A \rightarrow B} s_i^A,$$

where $g_{ij}^{A \rightarrow B}$ is the synaptic strength between neuron i in population A and neuron j in population B .

The rule circuit encodes two contexts that are represented by two neural pools, color and orientation (Fig. 1C). Local connections are excitatory $g = 0.158$ nA for recurrent inputs, and inhibitory $g = -0.05$ nA from one pool to the other. The strong intrapool excitation and interpool inhibition give rise to reverberatory attractor dynamics (Brunel and Wang, 2001; Wang, 2001, 2002). Rule cells selectively project to the sensory ($g = 0.075$ nA) and category ($g = 0.001$ nA) circuits.

Neurons in sensory modules are organized (Fig. 1C) by their preferred stimulus feature θ_i (from 0° to 180°). Populations for color and orientation features are both simulated by 64 equally distributed neural pools ($\theta_{i+1} - \theta_i = 180^\circ/64$). Within each network, the synaptic strength g_{ij} between neurons with preferred directions θ_i and θ_j presents a Gaussian profile:

$$g_{ij}(\theta_i - \theta_j) = g_- + g_+ \exp(-(\theta_i - \theta_j)^2/2\sigma^2),$$

with $\sigma = 18^\circ$. Parameters $g_- = -1$ nA, and $g_+ = 0.2$ nA determine the amount of recurrent global inhibition and local excitation in the circuit. This strong global inhibition mediates divisive normalization (Heeger et al., 1996; Carandini and Heeger, 2012) in sensory circuits (Ardid et al., 2007).

The connectivity between sensory and category cells is shaped by learning (Engel and Wang, 2010), which we do not model explicitly in this work. We assume that the synaptic weight is a periodic sigmoid function of the presynaptic feature θ , so for instance neurons selective to feature values belonging to category C significantly contributed to it, but barely to category D (Fig. 1C):

$$g_{ij}(\theta_i) = g/(1 + \exp(-\cos(2(\theta_i - \theta_{0,j}))/\sigma)),$$

where i denotes sensory cells encoding a particular stimulus feature (e.g., orientation θ_i) and j refers to each of the associated two categories (categories C and D in the example). The cosine function is introduced in the sigmoid function to generate periodicity. Category boundaries are 67.5° and 157.5°. $\theta_{0,j}$ represents the center within each category domain (22.5° and 112.5°, respectively). σ controls the slope of the sigmoid, so the lower its value, the closer the curve is to the step function. We use $\sigma = 0.18$, as it generates a smooth but sharp categorical connectivity. g represents the maximum strength and is equal to 0.07 nA from sensory to category cells, and 0.2 nA from category to sensory cells.

Similarly to the rule circuit, circuits underlying categorization and action planning also obey winner-take-all attractor dynamics, with strong intrapool excitation ($g = 0.17$ nA, for the two circuits), and interpool inhibition ($g = -0.065$ nA for the categorization circuit and $g = -0.05$ nA for the action selection circuit). Finally, category pools selectively project to action pools ($g = 0.075$ nA; Fig. 1C).

In addition to synaptic inputs, action pools are subject to spike-rate adaptation I_A . The aim for introducing adaptation (and only in action

pools) was to computationally test our proposed mechanism for the task-response interaction: a particular combined effect of plasticity and adaptation taking place in action cells (see Fig. 7C,D). Adaptation is highly likely to happen in other modules, including for instance the sensory circuits in this model; however, we explicitly avoided including adaptation in the rest of our model for the sake of simplicity, as it is not particularly implicated in any other task switching phenomena. We have used a phenomenological implementation because our aim has been to center on behavioral interactions rather than to capture the precise biophysical mechanism of response suppression in these cells. The dynamics of the adaptation current was described by Wang et al. (2003) and Engel and Wang (2011):

$$dI_A/dt = -I_A/\tau_A + g_A r,$$

where r represents the neuron's firing rate, $\tau_A = 5$ s, and $g_A = 0.0015$ nA.

Plastic synapses. The model is endowed with reward-dependent plasticity under the assumption that reward information is used by the brain to up- or down-regulate the relevance of competing task rules and sensory-motor pathways, based on experience. However because of the low amount of errors (5.6%), indistinguishable results were obtained in the model with a non-reward-dependent, otherwise equivalent plasticity rule. Plasticity in the model underlies the residual switch cost (plasticity in recurrent synapses of the rule circuit) and stimulus priming effects (plasticity in feedforward synapses from sensory to category cells and from category to action cells). Plastic inputs add to the nonplastic synaptic inputs described above. Each pair of pre- and postsynaptic cells is connected by a set of bistable synapses that are in either a potentiated or a depressed state (Fusi, 2002; Soltani and Wang, 2006; Engel and Wang, 2011). The fraction c of potentiated synapses quantifies the strength of the plastic connection. Plastic synaptic strengths are expressed as $g_{ij} = g \cdot c_{ij}$. Parameter g represents the maximum strength and is equal to 0.02 nA for recurrent synapses of the rule circuit, 0.01 nA for inputs from sensory to category cells, and 0.011 nA for inputs from category to action cells.

All plastic synapses are updated according to a Hebbian plasticity rule that is gated by reward (Soltani and Wang, 2006; Fusi et al., 2007; Engel and Wang, 2011). Thus, synaptic changes occur only on rewarded trials in our model. However, due to the low error rate (5.6%), results remain the same when Hebbian plasticity is reversed in the absence of reward. Plastic synapses are either potentiated or depressed at the end of the trial depending on pre- and postsynaptic activities, computed by mean firing rates within the interval $\Delta t = 50$ ms, right before a response is selected. Potentiation occurs when presynaptic and postsynaptic activities are above a firing-rate threshold, otherwise the synaptic strength is depressed. The firing-rate threshold for plasticity is chosen based on neuronal activation relative to spontaneous activity. The threshold is 25 Hz for sensory cells and 10 Hz for the rest. Potentiation is described as a transition of depressed synapses to the potentiated state: $c \rightarrow c + q(1 - c)$. Analogously, depression is described as the transition of potentiated synapses to the depressed state: $c \rightarrow c - q \cdot c$. q refers to the plasticity rate and is equal to 0.25 in both transitions.

External inputs. There are three different sources for external inputs: the noisy current I_N , the sensory input I_S , and the reset input I_R . I_N represents the background synaptic input to each circuit and obeys:

$$\tau_N dI_N/dt = -(I_N - I_0) + \sqrt{\tau_N} \sigma_N \eta(t),$$

where $\tau_N = 2$ ms and $\eta(t)$ represents a Gaussian white noise process with $\langle \eta(t) \rangle = 0$ and $\langle \eta(t)\eta(t') \rangle = \delta(t - t')$. Variance of background noise is manipulated with σ_N that is equal to 0.03 nA for sensory cells, and 0.013 nA for the rest. Mean input I_0 is 0.369, 0.3, 0.367, and 0.355 nA respectively for rule, sensory, category, and action cells.

I_S refers to two sensory inputs that represent the cue and the target stimuli. There are two alternative cue inputs. In each trial, one of the two is randomly selected and specifically excites one pool of rule cells ($I_S = 0.0025$ nA) and inhibits the other ($I_S = -0.05$ nA). This combination of excitation and inhibition presumably involves balanced excitation and inhibition through feedforward inhibition, which is not explicitly included in the model. The target stimulus impacts on sensory neurons that encode a particular target feature with the following Gaussian profile:

$$I_s(\theta_i) = I_M \exp(-(\theta_s - \theta_i)^2/2\sigma_s^2),$$

where the target feature is represented by θ_s , θ_i ranges the selectivity of sensory cells encoding the feature (e.g., orientation), $\sigma_s = 18^\circ$ specifies the width, and $I_M = 0.07$ nA is the maximum input current, input that neurons encoding the stimulus feature θ_s receive.

A reset input $I_R = -0.25$ nA is introduced right after a response is selected. This inhibitory input lasts 500 ms and represents a corollary discharge (Crapse and Sommer, 2008) that suppresses activity of cells involved in decision making processing (action and category cells).

Computer simulations. For numerical simulations of the firing-rate model, we used the Euler integration method with a time step of 1 ms. Simulations were initialized with different random seeds and, unless otherwise indicated, the model was run on $n = 2 \times 10^4$ consecutive trials. The first 1000 trials were always discarded so that the results did not depend on the initial conditions. In Figure 6B, the RT residual switch cost for each q value was averaged from 100 different realizations to diminish its variability. For analysis of trial history dependence (Fig. 8), the number of trials was increased to $n = 10^5$, because three repetitions in a row were unlikely events.

To evidence variability in the model, we ran $n = 100$ realizations of a same sequence of $n = 16$ trials. Rule and target features in each trial of the sequence were randomly selected in advance. To avoid any dependence of the results on the initial conditions, each realization was initialized with 100 randomly selected trials and random seed. The first 100 trials were then discarded and only the sequence of 16 trials was kept for the analysis shown in Figure 5A,B.

Results

A neural circuit model for rule-based task switching

We have adapted a cued task-switching experiment designed by Mirabella et al. (2007) and Caselli and Chelazzi (2011), which provides behavioral and single-neuron data that has been used to constrain our model. In this task (Fig. 1A), a cue is shown, indicating whether the attended feature is color or orientation. The target, a color oriented bar, appears after a delay (CTI), until a behavioral response is selected (left vs right). Incongruent and congruent targets (Fig. 1B) are defined depending on whether the mapped behavioral response is different or the same based on the two target features. Our model circuit consists of four neural modules (Fig. 1C). In the sensory module, distinct neural populations encode color and orientation. Each population is described by a continuous (or "ring") circuit in which neurons are selective for a stimulus feature as an analog quantity (Ben-Yishai et al., 1995; Compte et al., 2000; Ardid et al., 2007). All the other three modules are modeled similarly according to our previous strongly recurrent network model (Brunel and Wang, 2001; Wang, 2001, 2002; Wong and Wang, 2006), which is characterized by attractor dynamics underlying winner-take-all competition among selective neural pools. In the category module, neurons can discriminate color and orientation features (which are analog quantities) into binary categories. In the rule module, distinct neural pools compete to establish a proper context that is maintained by persistent activity. Sensory cells are subject to top-down modulation from both rule and category cells. Category neurons combine bottom-up inputs from sensory cells and a weak top-down biasing signal from rule cells to decide on the target feature category. This information is sent to the action selection module, where one of the competing neural pools (selective for left vs right) ultimately produces a motor response. See Material and Methods for a detailed description of the model and simulation protocol.

Figure 2 shows how different modules in our model perform different computations yet work cooperatively in task switching. Two representative trials are shown: a switch trial with congruent stimulus (Fig. 2, left), and a repeat trial with incongruent stimu-

lus (Fig. 2, right). The rule in play is maintained by elevated persistent activity in one of the two rule-selective neural pools, say color rule cells (Fig. 2, top). In a switch trial, a cue input is presented, leading to a switch of persistent activity pattern (Fig. 2A), whereas in a repeat trial the cue input barely changes the activity pattern (Fig. 2B).

The spatiotemporal activity pattern of sensory neurons is shown for the two sample trials in Figures 2C,D. A bar stimulus activates a neural group in the color coding network and another in the orientation color network (with category boundaries of each feature marked by horizontal dashed lines). The model reproduces two salient experimental observations. First, top-down modulation is present but weak in the activity of neurons that are selective to the relevant target feature (Fig. 2C, orientation neurons $\sim 45^\circ$; D, color neurons encoding green show slightly darker orange color indicative of higher activity caused by top-down modulation; see also Fig. 3A). Second, top-down modulation increases over time notably on incongruent stimulation (Fig. 2D, darker orange color toward the end of the trial; see also Fig. 3C), but not much on the congruent condition (Figs. 2C, 3C; Mirabella et al., 2007). This effect is explained in the model by the time course of activity of category-selective neurons that is much more pronounced in incongruent trials (Figs. 2F, 3B) than in congruent trials (Figs. 2E, 3B), leading to a buildup of modulation of sensory neurons by category cells on the incongruent condition.

The model assumes that there is competition between all four category-selective neural pools, which ultimately leads to a single winner (“selected category”). In a congruent trial (Fig. 2, left), selecting the relevant category is not essential, because the two activated category pools (B: color-right and D: orientation-right) in Figure 2E project to the same action pool (Fig. 2G). In contrast, in an incongruent trial (Fig. 2, right), the two activated category pools (B: color-right, C: orientation-left) in Figure 2F send inputs to conflicting action pools. The two category neural pools compete with each other, but only the relevant category is slightly favored by a small modulation from rule cells. Once a category neural pool wins the competition, its associated action pool ramps up, and a response is triggered at the time its activity crosses 20 Hz (Fig. 2H).

It is known that V4 neuronal responses are modestly modulated by selective attention (Moran and Desimone, 1985; McAdams and Maunsell, 2000; Reynolds and Chelazzi, 2004; Maunsell and Treue, 2006). Experimental observations (Mirabella et al., 2007; Cohen and Maunsell, 2011) suggest that selective attention

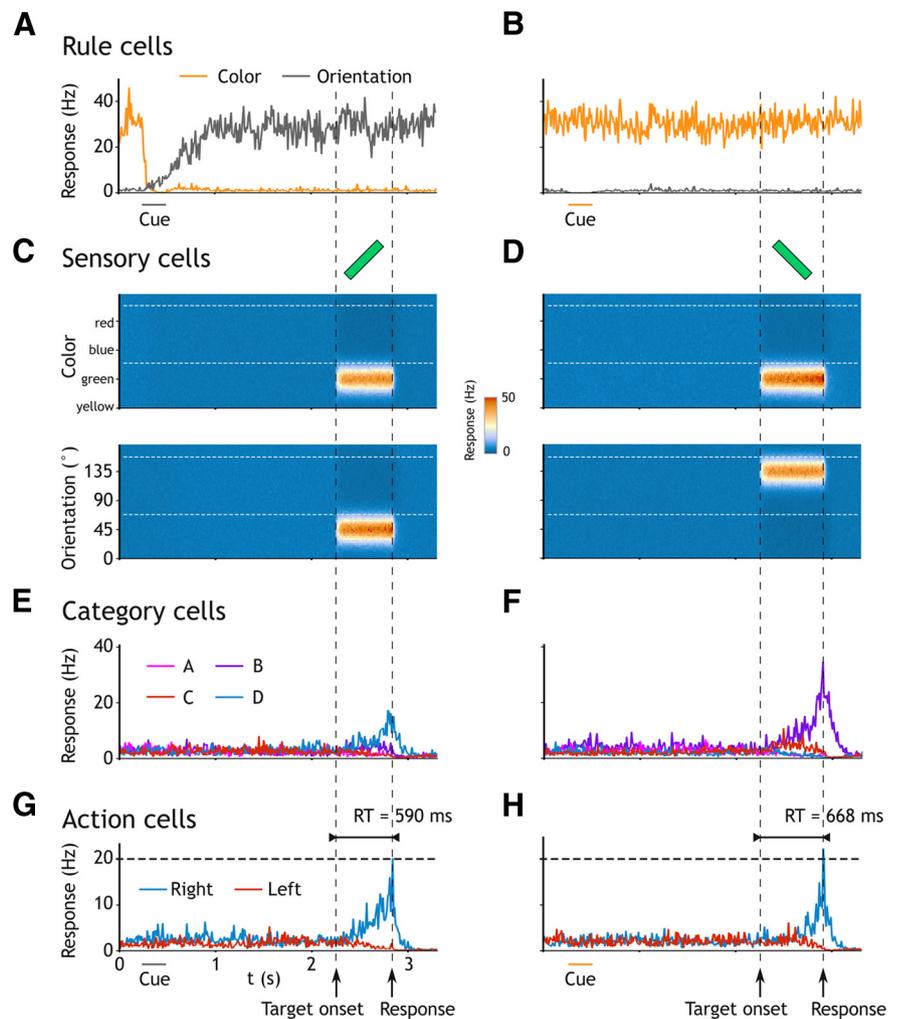


Figure 2. Neural activity in the model. **A, B,** The rule in play is represented by elevated persistent activity in one of the two rule-selective neural pools. **A,** A switch trial. Persistent activity is switched from color-selective neural pool (orange) to orientation-selective neural pool (gray) by a cue signal. Note the slow time course of the switching dynamics that gives rise to the preparatory switch cost. **B,** A repeat trial. Persistent activity remains in the color-selective neural pool. Its activity barely changes with the cue input. **C, D,** Spatiotemporal activity pattern of sensory neurons labeled by their preferred feature (y -axis). Firing rate is color-coded. Dashed vertical black lines: RT from target onset to response. Dashed white horizontal lines: category boundaries. **C,** With a congruent target stimulus (green bar oriented 45°), orientation- and color-selective sensory cells point to the same action. Relevant pools (neurons selective for 45° orientation) are slightly enhanced by attentional inputs. **D,** With an incongruent target stimulus (green bar oriented 135°), orientation- and color-selective sensory cells point to opposite actions. Relevant pools (neurons selective for green color) are enhanced by attentional inputs. **E, F,** Category cells increase their activity after target onset. Rule cells bias the competition in the category circuit so eventually one category neural pool wins the competition (D category cell in blue that represents orientation-right, in **E**; B category cell in purple, that represents color-right, in **F**). **G, H,** Activity of action cells. With a congruent stimulus **G,** the activity of right-selective neurons (in blue) grows immediately after the target onset, while the activity for left-selective neurons (in red) rapidly decreases. In contrast, with an incongruent stimulus **H,** activities of the two competing action-selective neural pools remain approximately the same until one of the category neural pools wins the competition. The activity of the corresponding action pool (right response in blue) ramps up very quickly at that moment.

can be extended to rule-based attention (e.g., orientation vs color) and feature categorization (e.g., Category 1 vs Category 2), in addition to feature value (e.g., red vs green). In our model, rule neurons provide a weak top-down biasing signal, consistent with the rule-based attentional modulation observed in V4 neurons already during CTI, before the target presentation (Mirabella et al., 2007). To quantify modulation in sensory neurons in our model, we define a “modulation ratio” as the ratio between the activity of the sensory pool that processes the cued relevant feature and that of the sensory pool that processes the other (irrelevant) target feature. The attentional modulation for sensory

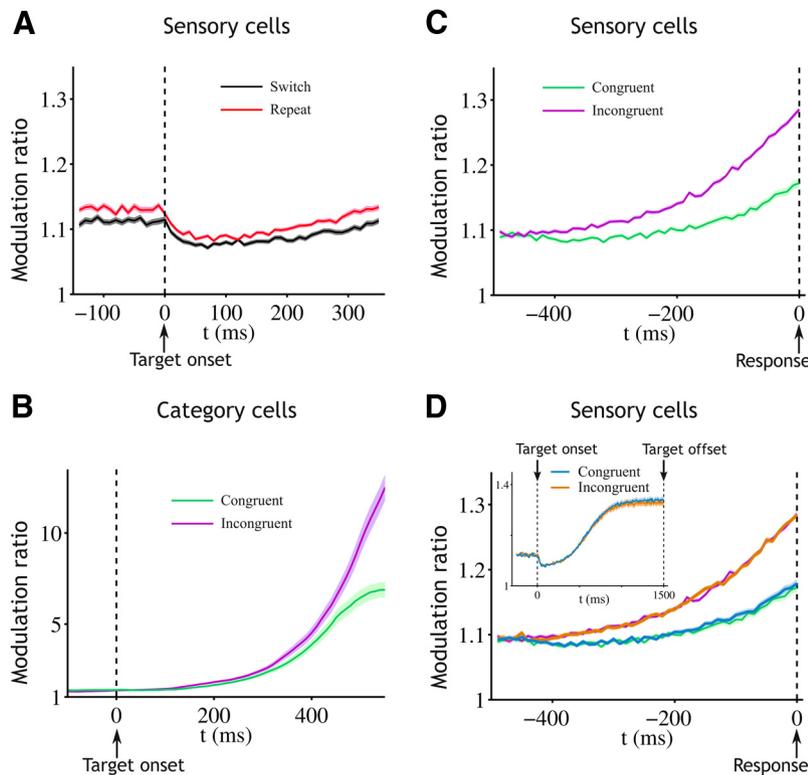


Figure 3. Task switching is routed by weak top-down modulation. **A**, Modulation ratio between sensory circuits relative to the target onset. The modulation ratio is defined as the ratio of the activity of neurons selective for the relevant target feature and that of neurons selective for the irrelevant feature. The modulation ratio deviates from 1 (by $\sim 10\%$) before the target onset (CTI), which is explained by weak top-down inputs originated at the rule circuit. The rule bias is significantly higher after a single task repeat (red) compared with switch (black). **B**, Modulation ratio between category pools. The relevant category pool receives a small top-down input from rule neurons that barely enhances its activity before target onset. This weak modulation is not enough to trigger winner-take-all competition in the category circuit. However, it acts in conjunction with the bottom-up sensory input to bias the winner-take-all competition during target presentation. Once the attractor dynamics is triggered, the modulation ratio ramps up very fast. The different modulation ratio that is reached between congruent (green) and incongruent (magenta) conditions is due to shorter RTs on the congruent condition. **C**, Modulation ratio between sensory circuits relative to response. The modulation ratio increases during the target period toward the behavioral response due to moderate top-down inputs from category cells (from ~ 10 to $\sim 20\%$ modulation, averaging both conditions together). This modulation is higher under the incongruent condition ($\sim 30\%$, magenta) relative to the congruent condition ($\sim 15\%$, green). **D**, Fixed duration task. The task design was modified, so that the behavioral response occurs after a fixed-target duration of 1.5 s. This allowed us to prove that the categorical modulation is the same for the congruent and incongruent conditions (inset). The same traces from inset panel are replotted in main panel but aligned by the average RT of the respective stimulus condition. The overlap between these traces and those from the middle panel proves that the different modulation ratio between conditions is only due to a difference in RT.

neurons before target onset is indeed modest ($\sim 10\%$ increase of spontaneous activity that corresponds to a modulation ratio of ~ 1.1 , Fig. 3A). Interestingly, our model predicts that reward-dependent plasticity in the rule circuit affects this modulation, due to the activation of rule pools that have been previously reinforced (task repetition) or weakened (task switch) by reward-dependent plasticity in recurrent synapses. Thus, this modulation is higher for a repeat trial compared with a switch trial (Fig. 3A).

Activity of V4 neurons is also influenced by behaviorally relevant category signals during target presentation (Mirabella et al., 2007). The signal may reflect top-down input from category neurons in downstream regions (Mirabella et al., 2007), such as prefrontal cortex (PFC; Freedman et al., 2003; Cromer et al., 2010; Roy et al., 2010) or lateral intraparietal area (LIP; Freedman and Assad, 2006; Fitzgerald et al., 2011). In agreement with neural data (Mirabella et al., 2007), the attentional modulation in sensory neurons modestly increases through the target presentation (from ~ 10 to $\sim 20\%$), and more significantly under the incon-

gruent (magenta) compared with the congruent (green) condition ($\sim 10\%$ difference; Fig. 3C). This difference may be explained either because dynamics in the category circuit depends on stimulus congruency, or due to different RTs for the two stimulus conditions. To test this, we ran a set of simulations using a fixed duration (1.5 s) for the target presentation, which revealed that the effect is independent of stimulus condition (Fig. 3D, inset) and solely because category cells ramp up for a longer time and to a higher level in incongruent trials (Fig. 3D).

The source of the category signal in the model is originated from the attractor dynamics in the category circuit. To illustrate its dynamics, we repeated the analysis of the modulation ratio for category neurons (Fig. 3B). As of sensory neurons, category neurons also receive a moderate top-down input from rule neurons that slightly enhances the low spontaneous activity before target onset by $\sim 35\%$ (from ~ 2.4 to ~ 3.2 Hz), which is barely distinguishable from activity fluctuations (Fig. 2E,F), unless averaged through many trials (Fig. 3B). This modulation is not enough to trigger the winner-take-all competition in the attractor network. It only becomes fully visible in conjunction with the bottom-up sensory input when the target stimulus is presented (Fig. 3B). This mechanism is what we call the tweaking hypothesis for executive control: a weak rule bias that is able to efficiently reconfigure the sensory-motor mapping based on context. Attentional modulation to both sensory and category neurons is biologically plausible ($\sim 30\%$; Fig. 3), yet sufficient to properly bias the winner-take-all competition in the category circuit during target presentation.

Our model identifies a new mechanism for resolving conflicts caused by incongruent stimulation between competing sensory features that are equally salient: a category circuit, at an intermediate stage between sensory processing and action selection, that is endowed with winner-take-all competition. First, such implementation accounts for modest attentional modulations of sensory neurons (an increase in activity from ~ 40 to ~ 50 Hz; Fig. 2C,D). Second, both types of stimulus (congruent and incongruent) equally project to competing category pools. The symmetry between the two activated category pools is only broken by the rule modulation, or from stimulus priming (an implicit memory effect that influences behavioral responses under stimulus repetition, see Material and Methods). Thus, the congruency effect is limited only to the readout of category inputs in the action circuit: whether the two competing categories project to a same action pool or not.

To shed insights into the tweaking principle, we compare three simplified versions of the model that only differ in the dynamics of the category circuit (Fig. 4). The first version of the

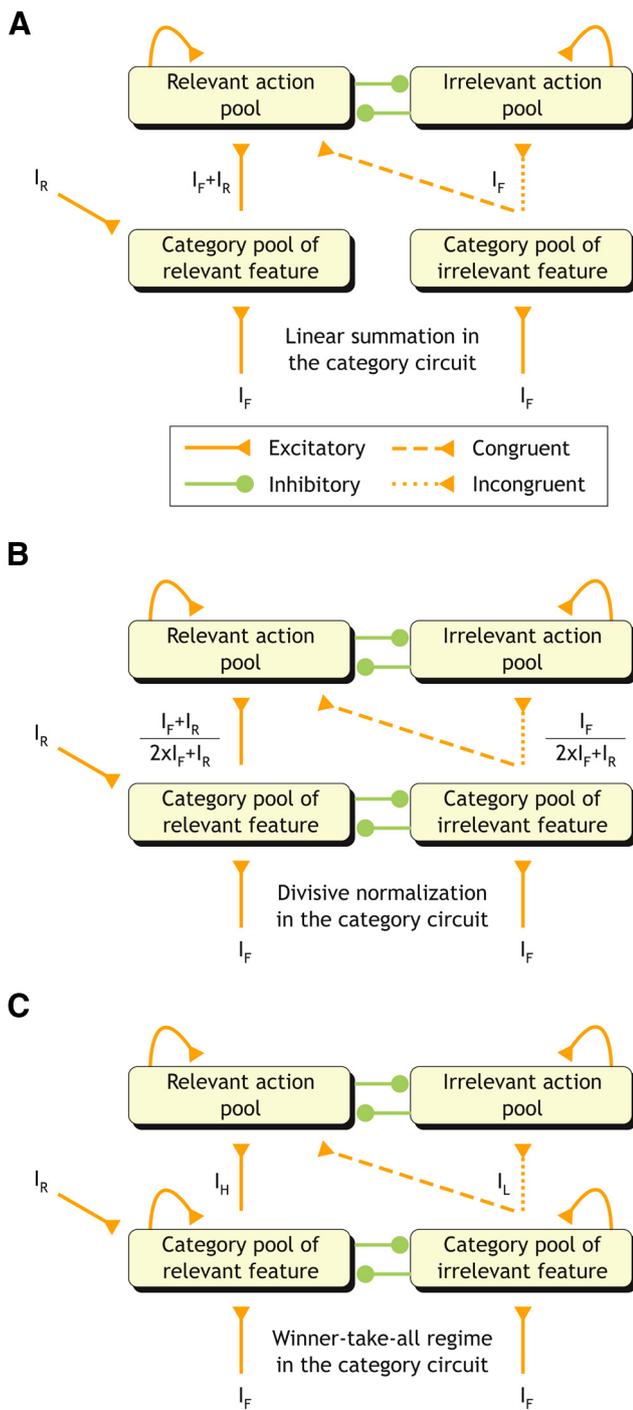


Figure 4. Schematic models for explaining a modest congruency effect in the task switching model. The three scenarios differ only in the assumed dynamics of the category circuit. **A**, A category circuit that operates by linear summation. The stimulus input (I_F) and rule bias (I_R) add together linearly in the category pools. **B**, A category circuit endowed with divisive normalization. Category neural pools are normalized in proportion to the total input received by the circuit. **C**, A category circuit in winner-take-all regime. The stimulus input (I_F) is sufficiently strong to trigger attractor dynamics in the category circuit. In that regime, recurrent excitation and lateral inhibition work together, so that only the input difference (I_R) is amplified. The third, but not the first or second, scenario gives rise to a modest congruency effect as observed experimentally, under the assumption that top-down rule signaling is much weaker than sensory stimulation. For detailed explanation, please see the Results section.

model consists of a category circuit that linearly summates their inputs, I_F (for each of two stimulus features) and I_R (for rule bias; Fig. 4A). In the action selection circuit, the neural pool selective for the relevant action receives an input $I = 2 \times I_F + I_R$ on the congruent condition, and the other neural pool receives no input; the difference of the two inputs (which determines the ultimate motor response; Wang, 2002) is $\Delta I = 2 \times I_F + I_R$. By contrast, on the incongruent condition, the inputs for the two neural pools are $I = I_F + I_R$ and $I = I_F$; the difference is $\Delta I = I_R$. In order for the congruency (or conflict) effect to be small, the input difference must be similar under the two conditions. This is possible only if I_R is much larger than I_F , which is not biologically realistic.

The second version of the model relies on divisive normalization (Heeger et al., 1996; Carandini and Heeger, 2012). In this case, activity of category neurons is normalized in proportion to the total input that the category circuit receives, $2 \times I_F + I_R$ (Fig. 4B). This means that, in the action selection circuit, the neural pool selective for the relevant action on the congruent condition receives an input $I = \frac{I_F + I_R}{2 \times I_F + I_R} + \frac{I_F}{2 \times I_F + I_R}$, and the other pool receives zero input. On the incongruent condition each action pool receives $\frac{I_F + I_R}{2 \times I_F + I_R}$ or $\frac{I_F}{2 \times I_F + I_R}$, and the difference is $\frac{I_R}{2 \times I_F + I_R}$. Even for $I_R \sim I_F$ the input difference on the incongruent condition is still much smaller ($\Delta I = 1/3$) than in the congruent case ($\Delta I = 1$), therefore the congruency effect predicted by the model remains much larger than that observed experimentally. Although divisive normalization is proposed to increase sensitivity (Carandini and Heeger, 2012) by scaling inputs to a particular interval, which may as well contribute to enhance stability, it is not capable of resolving conflicting inputs in task switching.

The third version of the model assumes a winner-take-all mechanism as in our model. The mechanism is based on attractor dynamics in the category circuit endowed with two important properties: a strong input (I_F) is needed, but only the input difference (I_R) is amplified. The output I_L from the category pool that does not win the competition becomes almost negligible ($I_F \rightarrow I_L$) compared with the output I_H from the category pool that wins the competition ($I_F + I_R \rightarrow I_H$), even for a weak rule bias and in both the congruent and incongruent conditions (Fig. 4C). The congruency effect is explained at the level of the action circuit. On the congruent condition, the neural pool selective for the relevant action receives a total input of $I = I_H + I_L$, and the other receives no input; whereas on the incongruent condition, the relevant action pool receives $I = I_H$ and the other pool receives $I = I_L$. The input difference between the two pools is $\Delta I = I_H + I_L$ for the congruent condition and $\Delta I = I_H - I_L$ for the incongruent condition. Because I_L is much smaller than I_H , the input difference is very similar in the two cases, and the congruency effect stays modest.

This simplified analysis provides an intuitive explanation of how our proposed mechanism, based on the tweaking principle, resolves the conflict and reproduces the congruency effect observed in behavior (Rogers and Monsell, 1995; Meiran, 2000; Brown et al., 2007). Also in this regime, a conflict monitor, suggested to modestly enhance the task rule bias (I_R) under demanding or conflicting situations (Botvinick et al., 2004; Botvinick, 2007; Brown et al., 2007), is especially suited for affecting behavior (Fig. 7A).

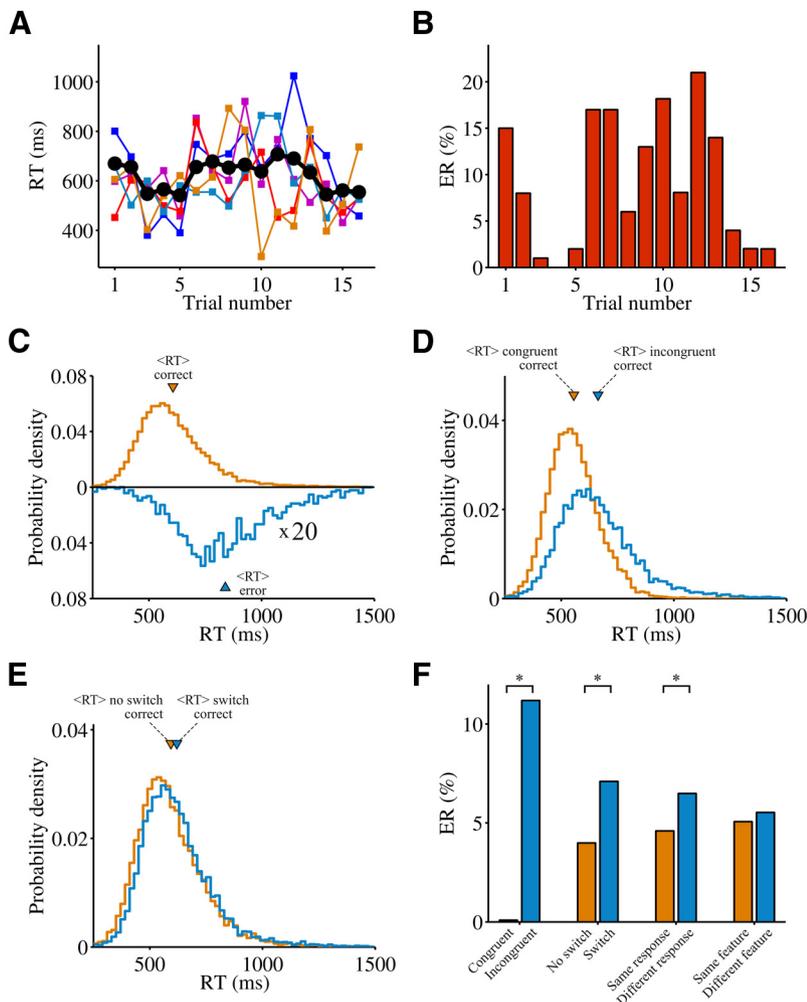


Figure 5. Task switching accuracy and RT distributions in the model. **A, B**, Stochastic behavior. **A**, RT varies across a sequence of 16 trials due to two factors: intrinsic stochasticity in the model (five realizations are shown in colored thin lines), and trial conditions, e.g., congruent versus incongruent stimulus or rule repeat versus switch (black thick line: averaged RT over 100 realizations). **B**, ER for each trial in the sequence computed across the 100 realizations. Fluctuations in ER reflect changes in trial condition. **C**, RT distributions in correct and error trials. Mean RT values are indicated by colored triangles. RT distribution for error trials is magnified $20\times$ for better comparison. RTs for error trials are significantly longer compared with correct trials. The error rate is 5.6%. **D**, RT distributions for congruent and incongruent trials. Only correct trials are included. The mean RT is larger with incongruent stimuli (the congruency effect). **E**, RT distributions for rule repeat and switch. Only correct trials are included. Given the long CTI = 1.75 s in these trials, the difference in mean RTs represents the residual switch cost. **F**, ER under different conditions of stimulus congruency, task rule, behavioral response, and feature priming. The proportion of error trials is sensitive to the congruency effect, rule switch and response alternation, but not to feature priming.

From neural data to behavior

Neural fluctuations in our model give rise to high variability of RTs (Fig. 5A) and ER (Fig. 5B; overall ER = 5.6%). However, correct trials present systematically shorter RTs compared with error trials (mean RT: 606.45 ms and 838.88 ms, respectively; $p < 0.001$ Mann–Whitney–Wilcoxon test; Fig. 5C). In perceptual decision-making (Gold and Shadlen, 2007; Wang, 2008; Luce, 2009), when the signal-to-noise is low, RTs are longer and ER is higher. We found that the same principle applies to task switching on the incongruent condition, except that the signal does not merely reflect the quality of external sensory information, but the relative input difference between the alternative action pools, which is determined by the stochastic dynamics of rule cells, and the strength of their top-down inputs.

For correct trials, we analyzed the RT distributions under different conditions of stimulus congruency, task rule, behavioral

response, and feature priming. RT is longer for incongruent than for congruent stimuli (mean RT: 663.73 ms and 555.94 ms, respectively; $p < 0.001$ Mann–Whitney–Wilcoxon test; Fig. 5D). The difference represents the congruency effect (Rogers and Monsell, 1995; Meiran, 2000; Brown et al., 2007). Difference in RTs between switch versus repeat trials represents the switch cost (mean RT: 620.03 ms and 593.41 ms, respectively; $p < 0.001$ Mann–Whitney–Wilcoxon test; Fig. 5E). Given the large delay used in these simulations (CTI = 1.75 s), this switch cost mostly reflects its residual component (Monsell, 2003). The effect emerges from the activation of rule pools that have been previously reinforced (repetition) or weakened (switch) by reward-dependent plasticity in recurrent synapses. RT distributions for same versus different behavioral response, and same (primed) versus different (not primed) target feature, are not distinguishable (data not shown). Proportion of errors under different conditions of congruency, task, response, and priming is shown in Figure 5F. Performance is mostly impaired on incongruent stimuli, on switch trials and on response alternation, which significance is confirmed by logistic regression analysis ($p < 0.001$).

To assess the switch cost as a function of preparation time between the cue and target stimulus, we varied CTI from 0 s to 3.75 s, and found two components in the RT switch cost (Fig. 6A). In the model, the preparatory component of the switch cost (Rogers and Monsell, 1995; Brown et al., 2007; Vandierendonck et al., 2010) results from attractor dynamics in the rule circuit (Fig. 2, top). The rule is actively maintained from one trial to the next by the activity of a rule-selective neural pool. We designed the rule circuit in this way based on the evidence of sustained rule-selective activity in PFC across trials (Mansouri et

al., 2006; Durstewitz et al., 2010; Yamada et al., 2010). On a switch trial, rule updating requires a transition between two different attractor states, its time course then gives rise to the preparatory cost, but this cost is not reflected in the RT for sufficiently long CTIs (Fig. 6A). The residual component is the asymptote of switch cost for long CTIs and indicates that there is part of the cost that cannot be avoided even after long time for preparation. In the model, neural interactions are modulated by reward-dependent synaptic plasticity: on correct trials, activated pathways are facilitated through long-term potentiation (LTP) and inactivated pathways get depressed through long-term depression. On repeat trials this reinforcement leads to behavioral melioration, but on switch trials depressed pathways become relevant and compete with facilitated pathways, leading to the residual switch cost that remains for arbitrarily long CTIs (Fig. 6A). Figure 6B shows that the magnitude of the residual switch

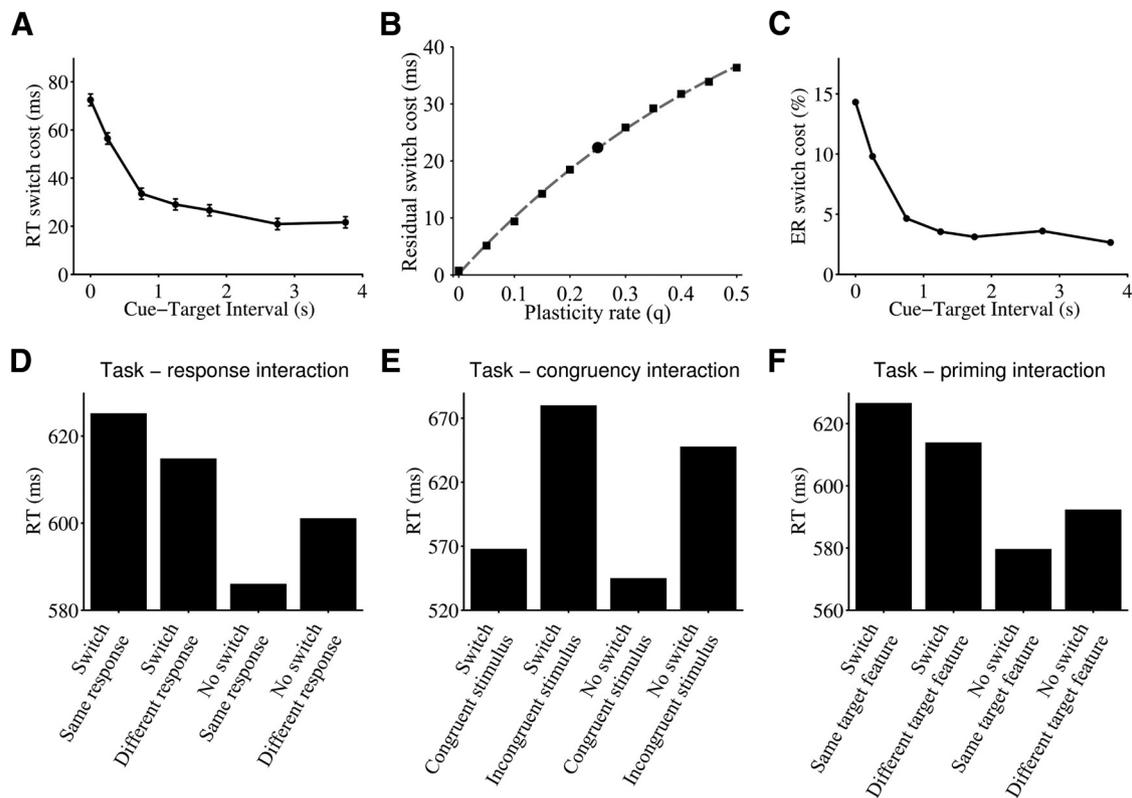


Figure 6. Switch cost and task switching interactions in the model. **A**, RT switch cost. The cost is computed as the difference in RT between switch and repeat trials. The difference is displayed as a function of time interval between cue and target stimuli (CTI). The total switch cost consists of two components: a preparatory component that decays with CTI, and a residual component that represents the asymptotic level for long CTIs. Error bars represent SE. **B**, The magnitude of the residual switch cost depends on the plasticity rate q (see Materials and Methods). Dots represent averaged residual switch costs over 100 different realizations. The reference value ($q = 0.25$) is indicated by a circle. Data were fitted with an exponential function to guide eye inspection. **C**, ER switch cost. A similar temporal profile to the RT switch cost is observed in the ER difference between switch and repeat trials. **D**, Task-response interaction. Response repetition emphasizes the impact of the switch cost: RT is shorter for repeat trials, but it gets longer for switch trials. **E**, Task-congruency interaction. Difference in RTs between repeat vs switch trials is increased on incongruent compared with congruent condition. **F**, Task-priming interaction. RT is shorter if the presently relevant feature was also relevant in the previous trial, compared with the case that it did not appear (positive priming). The trend is the opposite if the feature was relevant in the previous trial, but becomes irrelevant in the current trial (negative priming).

cost in the model clearly depends on a sufficiently high learning rate q in the plasticity rule (see Materials and Methods). Importantly, a difference in the strength of the learning rate could underlie across species (human–monkey) differences in the magnitude of the residual switch cost. Monkeys learn more slowly than humans and present lower residual switch costs (Stoet and Snyder, 2007b). A lower plasticity rate implies slower learning and, according to our model, reduces the magnitude of the residual switch costs. Similarly to the RT, ER also decreases with longer time for preparation, but even after long time, part of ER switch cost is asymptotically not zero (Fig. 6C).

We then examined whether task rule interacts with behavioral response, stimulus congruency or feature priming. Figure 6D shows the task-response interaction. The model captures an intriguing effect commonly observed in humans (Rogers and Monsell, 1995; Meiran, 2000; Altmann, 2011): RT switch cost increases on response repetition, so selecting the same behavioral response shortens RT under rule repeat, but lengthens RT under rule switch. The interaction is explained in the model by a combination of two counteracting factors. On the one hand, a same response under rule repeat requires activation of a category-to-action pathway that has been facilitated through reward-dependent plasticity, which shortens RT. On the other hand, response repetition leads to intrinsic adaptation in action cells, which makes RT longer. The combination of the two effects explains the task-response interaction, under the condition that the

benefit from plasticity overcomes the effect of adaptation (Fig. 7C,D).

The RT switch cost is longer for incongruent stimuli than for congruent stimuli (Fig. 6E), as observed experimentally (Goschke, 2000; Meiran, 2000; Liston et al., 2006; Wendelken et al., 2012). In the model, this is because activated category pools cooperate on congruent stimuli leading to faster responses, so that the role of the rule bias is less important. The model also accounts for an interaction between task rule and feature priming (Koch and Allport, 2006; Fig. 6F). A bias from feature repetition on consecutive trials is caused by plasticity at synapses from sensory cells to category cells. The feature bias explains why RT is shorter in a given trial if the relevant feature was also relevant in the previous trial, compared with the case in which that feature did not appear (positive priming). To control that this is indeed a genuine effect of feature priming, our analysis considered only trials where dissimilar features belonged to a same category. Thus, rule, category, and response were fixed in consecutive trials, whereas the target feature might repeat or change. In addition, the feature bias is responsible for an increased RT when a feature was relevant in the previous trial but becomes thereafter irrelevant (negative priming).

We tested statistical significance of RT differences across conditions with a N -way ANOVA analysis. In advance, we log-transformed RT distributions to accommodate normality limitations of the ANOVA analysis. Among main factors (task rule,

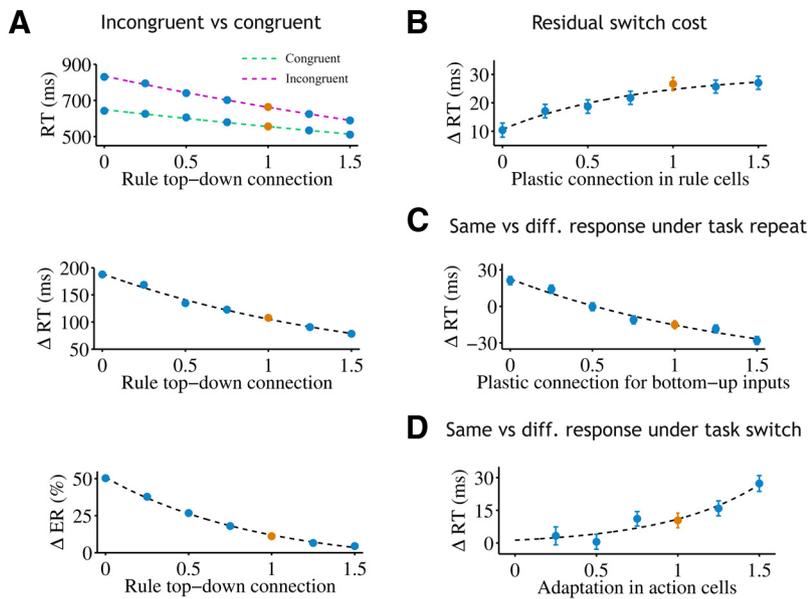


Figure 7. Gradual dependence of task switching behavior on model parameters. The strength of each neural mechanism is parametrically varied with a multiplicative factor g_{norm} in (0, 1.5), the reference value (with $g_{\text{norm}} = 1$) is indicated by an orange circle. Data were fitted with exponential functions to guide eye inspection. Error bars represent SE. **A**, Rule modulation. Congruent and incongruent conditions are affected by rule bias, but the incongruent condition (magenta) is more sensitive. The congruency effect is parametrically reduced with a stronger top-down input from the rule circuit to the sensory and the category circuits (middle, RT cost; bottom, ER cost). **B**, Residual switch cost is enhanced by an increase in the strength of plastic synapses in the rule circuit. **C**, Task-response interaction depends on plastic synapses from sensory to category cells and from category to action cells. When they are absent, adaptation in action cells leads to longer RTs on response repetition under task repeat, but this effect is reduced and eventually reversed with sufficiently strong plastic connections. **D**, Adaptation in action cells. When adaptation is absent, no increase in the RT switch cost is found on response repetition. Only when adaptation scales up, a significant increase appears.

stimulus congruency, behavioral response, and stimulus priming), only stimulus congruency and task rule reached significance ($p < 0.001$). In addition, interactions between task rule and behavioral response, and task rule and feature priming also reached significance ($p < 0.001$ and $p = 0.0014$, respectively). However, the ANOVA analysis on log-transformed RT distributions clarified that the apparent interaction between stimulus congruency and task rule (Fig. 6E) did not reach significance, a prediction from the model that could be tested experimentally by applying the same log-transformation to data from other studies (Goschke, 2000; Meiran, 2000; Liston et al., 2006; Wendelken et al., 2012). Regarding ER, once interactions were taken into account in the logistic regression analysis, only stimulus congruency and behavioral response reached significance as main factors ($p < 0.001$), and in addition, the interaction between task and congruency also became significant ($p < 0.001$). This analysis confirmed that errors predominantly occur with incongruent stimuli, and more so if the rule switched or the response changed (Fig. 5F).

Neuronal mechanisms underlying task switching observations

To better illustrate the neuronal processes underlying task switching behavior in the model, we varied the strength of each specific neuronal mechanism (such as the synaptic weight of top-down signaling from rule cells, and so on) by a multiplicative factor (g_{norm}) between 0 and 1.5. Figure 7A shows that the rule bias to sensory and category cells has an impact on the congruency effect, which is parametrically reduced with stronger activation of the rule bias. This effect is due to a higher sensitivity of the incongruent condition to the rule bias. The residual switch cost is

caused by plasticity in recurrent inputs of rule cells and Figure 7B proves that the higher its synaptic strength, the stronger the recurrent input on rule repeat compared with switch, and the higher the residual cost. Note that even when plastic synapses in the rule circuit are inactivated, part of the residual switch cost remains (~ 10 ms). This is due to plastic synapses in bottom-up inputs underlying priming effects in the model. The task-response interaction depends on two counteracting mechanisms: the plastic synapses at bottom-up inputs (from sensory to category cells and from category to action cells) and the intrinsic adaptation in action cells. When plastic synapses are absent, there is a response repetition cost under task repeat because of firing-rate adaptation in action cells, but this pattern is reduced and eventually reversed with stronger plastic connections (Fig. 7C). On the other hand, if adaptation is absent, no significant response repetition cost is found under task switch. Only when adaptation scales up, an increased switch cost by response repetition appears (Fig. 7D). Alternatively, Figure 7 can be explained by the fact that the different task switching effects in the model are highly robust with respect to changes (up to $\sim 50\%$) in the parameters of the model.

A recent study has found that task switching affects behavior beyond a single trial (Brown et al., 2007). We suggest that this is indicative of a graded plastic process rather than a bistable switch and our model is originally built under such assumption: we use a plasticity rate ($q = 0.25$), which maintains memory traces in plastic synapses during few consecutive trials. As a result, neural mechanisms in which plasticity is involved accumulate their effect on RTs and ER under successive presentations. In the analysis, we consider subsets of trials with 1–3 repetitions of a given task rule, relevant category or relevant stimulus feature. Results show that increasing the number of repetitions, RTs become progressively shorter (Fig. 8A). However, RTs increase if the previously relevant category, or relevant stimulus feature, is still present after a rule switch (Fig. 8A). The explanation is simple: activated pathways have been potentiated through LTP for the previous 1, 2, or 3 trials but this pathway becomes irrelevant in the current trial because the rule has switched. Similar patterns are observed for ER (Fig. 8B): lower ER occurs if a relevant condition repeats, but ER increases when such a condition is concomitant with a rule switch, which is indicative of increased behavioral perseveration.

Discussion

Humans are excellent at resolving conflicting sources of information and at adapting to new contexts (Sakai, 2008). Congruency and switch costs associated with task switching are typically rather modest (Rogers and Monsell, 1995; Meiran, 2000; Brown et al., 2007; Vandierendonck et al., 2010; Zylberberg et al., 2010). Based on the results we presented, we suggest that the tweaking principle may be an efficient mechanism used by the brain to resolve sensory-motor conflicts in task switching. Our approach

builds on four basic cognitive functions and our model computationally captures the properties and the dynamics of the implicated neural circuits: rule selectivity in PFC, sensory processing in V4, feature categorization (and we predict conflict resolution) in PFC/LIP, and action selection in premotor neurons. Compared with more abstract models (Cohen et al., 1990; Gilbert and Shallice, 2002; Morton and Munakata, 2002), our biologically realistic model is capable of reproducing single-neuron activity associated with task switching, and sheds insights into its underlying circuit mechanism. At a conceptual level, previous models assumed that top-down bias by task rule is powerful (twice as strong as bottom-up sensory inputs in Cohen et al., 1990). By contrast, in our model the rule bias modulates neuronal activity in the sensory and category circuits only modestly; but this small signal is highly amplified through strongly recurrent dynamics in local circuits, leading to a complete rerouting of information (by reconfiguring the sensory-motor mapping) that underlies task switching.

Our model identifies potential neural correlates for task switching behavior: first, the model assumes strongly recurrent neural dynamics for decision making (Wang, 2002) and to maintain the rule in play in working memory (Amit, 1995; Wang, 2001). Behaviorally, internal maintenance in working memory of behavioral rules is particularly relevant for the preparatory switch cost. Attractor dynamics has been used before in task-switching models (Rougier et al., 2005), but only for the internal maintenance of rule representation and not for amplification of a weak rule bias. Therefore, the present model is the first to propose the tweaking principle and implement it by strong neural reverberation: attractor dynamics is essential in the category circuit to optimally resolve the sensory-motor conflict in a way that is consistent with the congruency effect observed behaviorally (Rogers and Monsell, 1995; Meiran, 2000; Brown et al., 2007). Second, the model quantitatively accounts for the strength and time course of the top-down attentional modulations from rule- and category-selective neurons observed in sensory neurons (Mirabella et al., 2007). These modulations enhance the processing of the behaviorally relevant information already in the sensory cortex. Third, our model employs reward-dependent Hebbian learning for potentiating the pathways in use and for depressing those that are inactive. Hebbian synaptic plasticity in the model underpins residual switch cost and priming effects, as well as it contributes to the task-response interaction.

It has been proposed that rule-based flexible behavior generally requires neurons with mixed selectivity for a combination of task relevant information, such as stimulus, rule, and response (Rigotti et al., 2010), and this proposal has recently received experimental support (Rigotti et al., 2013). Our model also contains neurons with mixed selectivity; in particular, category selective neurons are sensitive to both sensory inputs and rule signals, although we implemented them through a structured hierarchical organization rather than using randomly connected neurons (RCNs; Rigotti et al., 2010). Another difference is that here we used reward-dependent plasticity as a plausible reinforcement learning

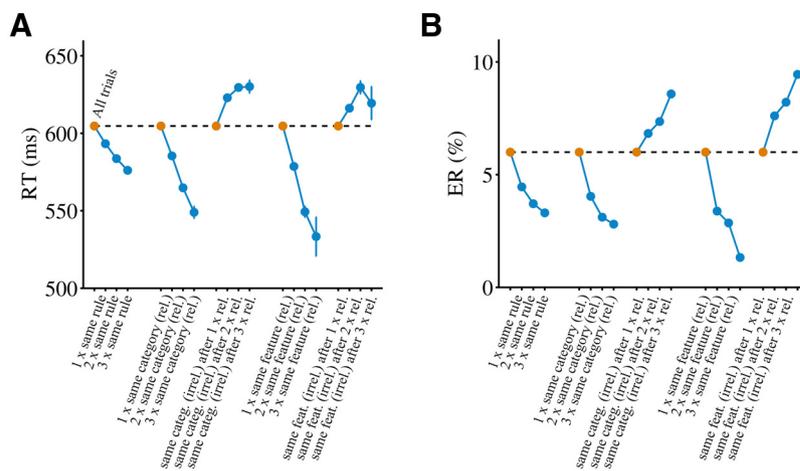


Figure 8. RT and performance are sensitive to trial history. **A**, Memory in RTs across trials. Task-switching effects accumulate through consecutive trials. The timescale is controlled by the learning rate ($q = 0.25$) of plastic synapses (corresponding to a timescale of $\sim 1/q = 4$ trials). We analyzed subsets of trials with 1, 2, and 3 repetitions of a given task rule, relevant category, or relevant stimulus feature. Progressively shorter RTs (compared with control, shown in orange) appear for successive repetitions that are relevant. However, if category or feature repetitions occur under rule switch, RTs become longer. Error bars represent SE. **B**, Memory in ER across trials. Similarly, lower ER appears for relevant repetitions, but ER is higher when a repeated category or stimulus feature becomes irrelevant, which represents increased behavioral perseveration.

mechanism, whereas Rigotti et al. (2010) set the network connectivity through supervised learning using a perceptron algorithm. There is no reason to believe that using RCNs would alter our conclusion in support of the tweaking principle, which is the focus of the present work. Our tweaking principle should be applied to a more general framework, possibly endowed with RCNs, that is not specially designed for a particular task but capable of adaptive coding (Duncan, 2001) to flexibly perform a number of decision tasks.

Furthermore, rule representation could also involve rhythmic coordinated activity in neuronal subnetworks (Womelsdorf et al., 2013). A recent work found that subgroups of cells in the PFC show task-rule selective beta band synchronization (Buschman et al., 2012). Such oscillatory dynamics may contribute to the formation of a selective cell assembly for task rule representation in a flexible manner. Interestingly, it has recently been shown that beta synchrony and firing-rate modulation are very closely related to each other (Canolty et al., 2012). The role of oscillation in a rule representation circuit can be investigated computationally in the future by extending our model with the incorporation of beta band synchrony.

Categorization plays an important role in the model. In general, categorization represents a key computational step because it is the basis to generalize sensory-motor associations on features that have not been presented yet (Seger and Miller, 2010). Moreover, if a category is primed from experience, this priming effectively impacts on all its constitutive features. Therefore, categorization does not only support the capability of generalizing but also the ability to efficiently process novel features, and our model uses these properties. In addition to that, our model proposes a dynamical regime for the category circuit, in the context of the tweaking principle that is key to resolve conflicts. It may be argued that this mechanism fails when task rules are not evident, or during early stages of practice because behaviorally relevant categories are supposed to appear after training. Interestingly, it has been shown that even without explicit knowledge of rules, learning task-related feature categories may be inferred from practice (Dreisbach et al., 2006). On the other hand, we suggest that in the course of training, category cells may be first selective to stimulus features, and later through practice their selec-

tivity evolves to categories. Under this assumption, the mechanism for sensory-conflict resolution remains valid if attractor dynamics in category cells emerges at the time of their initial feature selectivity. The fact that category cells display highly adaptable selectivity in LIP (Freedman and Assad, 2006; Fitzgerald et al., 2011) and PFC (Cromer et al., 2010; Roy et al., 2010) supports this hypothesis. Note that categorization of a stimulus attribute is generally independent of any motor response. In any particular experimental paradigm, subjects are required to learn a particular mapping between categories and motor responses, as modeled in the present study. However, such mapping can be arbitrary and our work can readily be extended by incorporating learning of such arbitrary mapping as we have done in a previous work (Fusi et al., 2007).

Our model presents a limitation, it has been designed for cued tasks in which the processing of different stimulus features (e.g., color and orientation) is behaviorally similar, as measured by response times and error rates. Previously described mechanisms are then generic for near-symmetric cued tasks, and applicable to other designs different than color versus orientation, as for instance the letter versus digit paradigm, used in human studies (Brown et al., 2007). However, in some task designs, the processing of one feature is prepotent (e.g., when the word “red” is printed in blue ink, the name red prevails over the color blue, which is known as the Stroop effect; Stroop, 1935; Besner et al., 1997). We believe that extending our model to such asymmetrical tasks will not be difficult, with two additional ingredients: (1) to assume asymmetrical pathway strengths for the two stimulus attributes, the prepotent one is significantly stronger than the other; (2) to include a module for response inhibitory control (Logan and Cowan, 1984; Boucher et al., 2007; Lo et al., 2009), which is required to suppress a prepotent response when inappropriate.

Our model has a number of testable predictions, both in behavior and in physiology. First, under incongruent stimulation, the winner-take-all dynamics in the category circuit and response selection take place serially, so that the former needs to be completed before the latter begins. This result can be tested in cued task switching paradigms by recording category neurons in LIP/PFC and action selection neurons in premotor areas. Second, long-term plasticity underlying residual cost may be tested behaviorally by systematically increasing the CTI and the intertrial interval. If the residual cost disappears for sufficiently long intervals, this would imply that the residual switch cost is not truly independent of CTI but originates from a slowly decaying process. However, two studies have shown significant residual costs after preparatory times longer than 5 s (Kimberg et al., 2000; Sohn et al., 2000). Physiologically, a prediction from the model is that plasticity in rule cells generates stronger rule modulation on sensory cells even after a single task repetition. In support of this idea, it has been recently shown that dopaminergic receptor activation in prefrontal neurons modulates the amplitude of top-down inputs to visual cortex (Noudoost and Moore, 2011). We suggest that a reward-related release of dopamine in the rule circuit gives rise to behavioral benefits under task rule repetition. Third, plasticity at bottom-up synapses explains feature and category priming. Behaviorally, stimulus priming (Koch and Allport, 2006) and category switch cost (Dreisbach et al., 2006) have already been reported, yet the neural correlates remain to be found. The mechanism that we propose can be tested physiologically in monkeys performing a rule-based task switching paradigm (Mirabella et al., 2007; Caselli and Chelazzi, 2011). Fourth, task-switching observables that depend on plasticity show in the model short-lasting memory traces, which accumulate across few trials. This is explained by gradual changes in the strength of

plastic synapses. A complete absence of trial history would instead indicate sudden transitions: plastic synapses would be either potentiated or depressed. Mixed evidence is available (Brown et al., 2007; Caselli and Chelazzi, 2011) and more systematic studies and improved statistics are necessary to check the presence of memory traces in task switching. Fifth, the preparatory cost depends on persistent rule-selective activity across trials (Mansouri et al., 2006; Durstewitz et al., 2010; Yamada et al., 2010). This result predicts that a “preparatory cost” should appear for task repetition if rule cells are inactivated during the intertrial interval. Furthermore, single-neuron physiological experiments and human fMRI studies can directly test the prediction that rule-selective persistent activity is maintained across intertrial intervals, either only in block designs (when the rule remains the same in a block of trials but varies from block to block) or also in cued designs (when the rule is cued on single trials).

In conclusion, we have presented a biologically based neural circuit model for task switching that provides a synthesis and mechanistic explanation for many neurophysiological and behavioral observations. At the conceptual level, it demonstrates that all-or-none network reconfiguration underlying task switching can be accomplished through dynamical amplification of weak top-down control signals. The notion of attractor networks has been developed for local circuits, for instance as a plausible mechanism for the generation of persistent activity underlying working memory. The present study demonstrates that this basic mechanism is computationally quite versatile and can be used as a “building block” in models for much more complex cognitive processes. The tweaking principle reveals that attractor dynamics within local brain regions have broader implications for communication and executive control across a large-scale system of interconnected brain areas that cooperatively give rise to task switching.

References

- Abbott LF, Chance FS (2005) Drivers and modulators from push-pull and balanced synaptic input. *Prog Brain Res* 149:147–155. [CrossRef Medline](#)
- Allport DA, Styles EA, Hsieh S (1994) Shifting intentional set: exploring the dynamic control of tasks. In: *Attention and performance XV*, (Umiltà C, Moscovitch M, eds), pp 421–452. Cambridge, MA: MIT.
- Altmann EM (2011) Testing probability matching and episodic retrieval accounts of response repetition effects in task switching. *J Exp Psychol Learn Mem Cogn* 37:935–951. [CrossRef Medline](#)
- Altmann EM, Gray WD (2008) An integrated model of cognitive control in task switching. *Psychol Rev* 115:602–639. [CrossRef Medline](#)
- Amit DJ (1995) The Hebbian paradigm reintegrated: local reverberations as internal representations. *Behav Brain Sci* 18:617–626. [CrossRef](#)
- Ardid S, Wang X-J, Compte A (2007) An integrated microcircuit model of attentional processing in the neocortex. *J Neurosci* 27:8486–8495. [CrossRef Medline](#)
- Ben-Yishai R, Bar-Or RL, Sompolinsky H (1995) Theory of orientation tuning in visual cortex. *Proc Natl Acad Sci U S A* 92:3844–3848. [CrossRef Medline](#)
- Besner D, Stolz JA, Boutilier C (1997) The Stroop effect and the myth of automaticity. *Psychon Bull Rev* 4:221–225. [CrossRef Medline](#)
- Botvinick MM (2007) Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. *Cogn Affect Behav Neurosci* 7:356–366. [CrossRef Medline](#)
- Botvinick MM, Cohen JD, Carter CS (2004) Conflict monitoring and anterior cingulate cortex: an update. *Trends Cogn Sci* 8:539–546. [CrossRef Medline](#)
- Boucher L, Palmeri TJ, Logan GD, Schall JD (2007) Inhibitory control in mind and brain: an interactive race model of countermanding saccades. *Psychol Rev* 114:376–397. [CrossRef Medline](#)
- Brown JW, Reynolds JR, Braver TS (2007) A computational model of frac-

- tionated conflict-control mechanisms in task-switching. *Cogn Psychol* 55:37–85. [CrossRef Medline](#)
- Brunel N, Wang X-J (2001) Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. *J Comput Neurosci* 11:63–85. [CrossRef Medline](#)
- Buschman TJ, Denovellis EL, Diogo C, Bullock D, Miller EK (2012) Synchronous oscillatory neural ensembles for rules in the prefrontal cortex. *Neuron* 76:838–846. [CrossRef Medline](#)
- Canolty RT, Ganguly K, Carmena JM (2012) Task-dependent changes in cross-level coupling between single neurons and oscillatory activity in multiscale networks. *PLoS Comput Biol* 8:e1002809. [CrossRef Medline](#)
- Carandini M, Heeger DJ (2012) Normalization as a canonical neural computation. *Nat Rev Neurosci* 13:51–62. [CrossRef Medline](#)
- Caselli L, Chelazzi L (2011) Does the macaque monkey provide a good model for studying human executive control? A comparative behavioral study of task switching. *PLoS ONE* 6:e21489. [CrossRef Medline](#)
- Cohen JD, Dunbar K, McClelland JL (1990) On the control of automatic processes: a parallel distributed processing account of the Stroop effect. *Psychol Rev* 97:332–361. [CrossRef Medline](#)
- Cohen MR, Maunsell JH (2011) Using neuronal populations to study the mechanisms underlying spatial and feature attention. *Neuron* 70:1192–1204. [CrossRef Medline](#)
- Collins AG, Frank MJ (2013) Cognitive control over learning: creating, clustering, and generalizing task-set structure. *Psychol Rev* 120:190–229. [CrossRef Medline](#)
- Compte A, Brunel N, Goldman-Rakic PS, Wang X-J (2000) Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb Cortex* 10:910–923. [CrossRef Medline](#)
- Conway BR, Tsao DY (2006) Color architecture in alert macaque cortex revealed by fMRI. *Cereb Cortex* 16:1604–1613. [CrossRef Medline](#)
- Conway BR, Tsao DY (2009) Color-tuned neurons are spatially clustered according to color preference within alert macaque posterior inferior temporal cortex. *Proc Natl Acad Sci U S A* 106:18034–18039. [CrossRef Medline](#)
- Conway BR, Moeller S, Tsao DY (2007) Specialized color modules in macaque extrastriate cortex. *Neuron* 56:560–573. [CrossRef Medline](#)
- Crapse TB, Sommer MA (2008) Corollary discharge circuits in the primate brain. *Curr Opin Neurobiol* 18:552–557. [CrossRef Medline](#)
- Cromer JA, Roy JE, Miller EK (2010) Representation of multiple, independent categories in the primate prefrontal cortex. *Neuron* 66:796–807. [CrossRef Medline](#)
- Dreisbach G, Goschke T, Haider H (2006) Implicit task sets in task switching? *J Exp Psychol Learn Mem Cogn* 32:1221–1233. [CrossRef Medline](#)
- Duncan J (2001) An adaptive coding model of neural function in prefrontal cortex. *Nat Rev Neurosci* 2:820–829. [CrossRef Medline](#)
- Durstewitz D, Vittoz NM, Floresco SB, Seamans JK (2010) Abrupt transitions between prefrontal neural ensemble states accompany behavioral transitions during rule learning. *Neuron* 66:438–448. [CrossRef Medline](#)
- Engel TA, Wang X-J (2010) Category learning from examples: how plastic synapses encode exemplars and compute similarity. *Soc Neurosci Abstr* 40:279.4.
- Engel TA, Wang X-J (2011) Same or different? A neural circuit mechanism of similarity-based pattern match decision making. *J Neurosci* 31:6982–6996. [CrossRef Medline](#)
- Fitzgerald JK, Freedman DJ, Assad JA (2011) Generalized associative representations in parietal cortex. *Nat Neurosci* 14:1075–1079. [CrossRef Medline](#)
- Freedman DJ, Assad JA (2006) Experience-dependent representation of visual categories in parietal cortex. *Nature* 443:85–88. [CrossRef Medline](#)
- Freedman DJ, Riesenhuber M, Poggio T, Miller EK (2003) A comparison of primate prefrontal and inferior temporal cortices during visual categorization. *J Neurosci* 23:5235–5246. [CrossRef Medline](#)
- Fusi S (2002) Hebbian spike-driven synaptic plasticity for learning patterns of mean firing rates. *Biol Cybern* 87:459–470. [CrossRef Medline](#)
- Fusi S, Asaad WF, Miller EK, Wang X-J (2007) A neural circuit model of flexible sensorimotor mapping: learning and forgetting on multiple time-scales. *Neuron* 54:319–333. [CrossRef Medline](#)
- Gallagher W (2009) *Rapt: attention and the focused life*. New York: Penguin.
- Ghose GM, Ts'o DY (1997) Form processing modules in primate area V4. *J Neurophysiol* 77:2191–2196. [Medline](#)
- Gilbert SJ, Shallice T (2002) Task switching: a PDP model. *Cogn Psychol* 44:297–337. [CrossRef Medline](#)
- Gold JL, Shadlen MN (2007) The neural basis of decision making. *Annu Rev Neurosci* 30:535–574. [CrossRef Medline](#)
- Goschke T (2000) Intentional reconfiguration and involuntary persistence in task set switching. In: *Control of cognitive processes: attention and performance XVIII*, (Monsell S, Driver J, eds), pp331–355. Cambridge, MA: MIT.
- Harada T, Goda N, Ogawa T, Ito M, Toyoda H, Sadato N, Komatsu H (2009) Distribution of colour-selective activity in the monkey inferior temporal cortex revealed by functional magnetic resonance imaging. *Eur J Neurosci* 30:1960–1970. [CrossRef Medline](#)
- Heeger DJ, Simoncelli EP, Movshon JA (1996) Computational models of cortical visual processing. *Proc Natl Acad Sci U S A* 93:623–627. [CrossRef Medline](#)
- Hussar CR, Pasternak T (2009) Flexibility of sensory representations in prefrontal cortex depends on cell type. *Neuron* 64:730–743. [CrossRef Medline](#)
- Kamigaki T, Fukushima T, Miyashita Y (2009) Cognitive set reconfiguration signaled by macaque posterior parietal neurons. *Neuron* 61:941–951. [CrossRef Medline](#)
- Kimberg DY, Aguirre GK, D'Esposito M (2000) Modulation of task-related neural activity in task-switching: an fMRI study. *Brain Res Cogn Brain Res* 10:189–196. [CrossRef Medline](#)
- Klingberg T (2009) *The overflowing brain: Information overload and the limits of working memory*. New York: Oxford UP.
- Koch I, Allport A (2006) Cue-based preparation and stimulus-based priming of tasks in task switching. *Mem Cognit* 34:433–444. [CrossRef Medline](#)
- Liston C, Matalon S, Hare TA, Davidson MC, Casey BJ (2006) Anterior cingulate and posterior parietal cortices are sensitive to dissociable forms of conflict in a task-switching paradigm. *Neuron* 50:643–653. [CrossRef Medline](#)
- Lo CC, Boucher L, Paré M, Schall JD, Wang X-J (2009) Proactive inhibitory control and attractor dynamics in countermanding action: a spiking neural circuit model. *J Neurosci* 29:9059–9071. [CrossRef Medline](#)
- Logan GD, Bundesen C (2003) Clever homunculus: is there an endogenous act of control in the explicit task-cuing procedure? *J Exp Psychol Hum Percept Perform* 29:575–599. [CrossRef Medline](#)
- Logan GD, Cowan WB (1984) On the ability to inhibit thought and action: a theory of an act of control. *Psych Rev* 91:295–327. [CrossRef](#)
- Luce RD (2009) *Response times: their role in inferring elementary mental organization*. New York: Oxford UP.
- Mansouri FA, Matsumoto K, Tanaka K (2006) Prefrontal cell activities related to monkeys' success and failure in adapting to rule changes in a Wisconsin card sorting test analog. *J Neurosci* 26:2745–2756. [CrossRef Medline](#)
- Mansouri FA, Buckley MJ, Tanaka K (2007) Mnemonic function of the dorsolateral prefrontal cortex in conflict-induced behavioral adjustment. *Science* 318:987–990. [CrossRef Medline](#)
- Mansouri FA, Tanaka K, Buckley MJ (2009) Conflict-induced behavioural adjustment: a clue to the executive functions of the prefrontal cortex. *Nat Rev Neurosci* 10:141–152. [CrossRef Medline](#)
- Mante V, Sussillo D, Shenoy KV, Newsome WT (2013) Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* 503:78–84. [CrossRef Medline](#)
- Maunsell JH, Treue S (2006) Feature-based attention in visual cortex. *Trends Neurosci* 29:317–322. [CrossRef Medline](#)
- Mayr U, Kliegl R (2003) Differential effects of cue changes and task changes on task-set selection costs. *J Exp Psychol Learn Mem Cogn* 29:362–372. [CrossRef Medline](#)
- McAdams CJ, Maunsell JH (2000) Attention to both space and feature modulates neuronal responses in macaque area V4. *J Neurophysiol* 83:1751–1755. [Medline](#)
- Meiran N (2000) Modeling cognitive control in task-switching. *Psychol Res* 63:234–249. [CrossRef Medline](#)
- Mirabella G, Bertini G, Samengo I, Kilavik BE, Frilli D, Della Libera C, Chelazzi L (2007) Neurons in area V4 of the macaque translate attended visual features into behaviorally relevant categories. *Neuron* 54:303–318. [CrossRef Medline](#)
- Monsell S (2003) Task switching. *Trends Cogn Sci* 7:134–140. [CrossRef Medline](#)

- Moran J, Desimone R (1985) Selective attention gates visual processing in the extrastriate cortex. *Science* 229:782–784. [CrossRef Medline](#)
- Morton JB, Munakata Y (2002) Active versus latent representations: a neural network model of perseveration, dissociation, and decalage. *Dev Psychobiol* 40:255–265. [CrossRef Medline](#)
- Noudouost B, Moore T (2011) Control of visual cortical signals by prefrontal dopamine. *Nature* 474:372–375. [CrossRef Medline](#)
- Reynolds JH, Chelazzi L (2004) Attentional modulation of visual processing. *Annu Rev Neurosci* 27:611–647. [CrossRef Medline](#)
- Rigotti M, Rubin DB, Wang X-J, Fusi S (2010) Internal representation of task rules by recurrent dynamics: the importance of the diversity of neural responses. *Front Comput Neurosci* 4:24. [CrossRef Medline](#)
- Rigotti M, Barak O, Warden MR, Wang X-J, Daw ND, Miller EK, Fusi S (2013) The importance of mixed selectivity in complex cognitive tasks. *Nature* 497:585–590. [CrossRef Medline](#)
- Rogers RD, Monsell S (1995) Costs of a predictable switch between simple cognitive tasks. *J Exp Psychol G* 124:207–231. [CrossRef](#)
- Rougier NP, Noelle DC, Braver TS, Cohen JD, O'Reilly RC (2005) Prefrontal cortex and flexible cognitive control: rules without symbols. *Proc Natl Acad Sci U S A* 102:7338–7343. [CrossRef Medline](#)
- Roy JE, Riesenhuber M, Poggio T, Miller EK (2010) Prefrontal cortex activity during flexible categorization. *J Neurosci* 30:8519–8528. [CrossRef Medline](#)
- Sakai K (2008) Task set and prefrontal cortex. *Annu Rev Neurosci* 31:219–245. [CrossRef Medline](#)
- Sasaki R, Uka T (2009) Dynamic readout of behaviorally relevant signals from area MT during task switching. *Neuron* 62:147–157. [CrossRef Medline](#)
- Seeger CA, Miller EK (2010) Category learning in the brain. *Annu Rev Neurosci* 33:203–219. [CrossRef Medline](#)
- Sohn MH, Ursu S, Anderson JR, Stenger VA, Carter CS (2000) The role of prefrontal cortex and posterior parietal cortex in task switching. *Proc Natl Acad Sci U S A* 97:13448–13453. [CrossRef Medline](#)
- Soltani A, Wang X-J (2006) A biophysically based neural model of matching law behavior: melioration by stochastic synapses. *J Neurosci* 26:3731–3744. [CrossRef Medline](#)
- Stoet G, Snyder LH (2004) Single neurons in posterior parietal cortex of monkeys encode cognitive set. *Neuron* 42:1003–1012. [CrossRef Medline](#)
- Stoet G, Snyder LH (2007a) Correlates of stimulus-response congruence in the posterior parietal cortex. *J Cogn Neurosci* 19:194–203. [CrossRef Medline](#)
- Stoet G, Snyder LH (2007b) Extensive practice does not eliminate human switch costs. *Cogn Affect Behav Neurosci* 7:192–197. [CrossRef Medline](#)
- Stroop JR (1935) Studies of interference in serial verbal reactions. *J Exp Psychol* 18:643–662. [CrossRef](#)
- Tanigawa H, Lu HD, Roe AW (2010) Functional organization for color and orientation in macaque V4. *Nat Neurosci* 13:1542–1548. [CrossRef Medline](#)
- Vandierendonck A, Liefvooghe B, Verbruggen F (2010) Task switching: interplay of reconfiguration and interference control. *Psychol Bull* 136:601–626. [CrossRef Medline](#)
- Wang X-J (2001) Synaptic reverberation underlying mnemonic persistent activity. *Trends Neurosci* 24:455–463. [CrossRef Medline](#)
- Wang X-J (2002) Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* 36:955–968. [CrossRef Medline](#)
- Wang X-J (2008) Decision making in recurrent neuronal circuits. *Neuron* 60:215–234. [CrossRef Medline](#)
- Wang X-J, Liu Y, Sanchez-Vives MV, McCormick DA (2003) Adaptation and temporal decorrelation by single neurons in the primary visual cortex. *J Neurophysiol* 89:3279–3293. [CrossRef Medline](#)
- Wendelken C, Munakata Y, Baym C, Souza M, Bunge SA (2012) Flexible rule use: common neural substrates in children and adults. *Dev Cogn Neurosci* 2:329–339. [CrossRef Medline](#)
- Womelsdorf T, Westendorff S, Ardid S (2013) Subnetwork selection in deep cortical layers is mediated by beta-oscillation dependent firing. *Front Syst Neurosci* 7:25. [CrossRef Medline](#)
- Wong KF, Wang X-J (2006) A recurrent network mechanism of time integration in perceptual decisions. *J Neurosci* 26:1314–1328. [CrossRef Medline](#)
- Yamada M, Pita MC, Iijima T, Tsutsui K (2010) Rule-dependent anticipatory activity in prefrontal neurons. *Neurosci Res* 67:162–171. [CrossRef Medline](#)
- Zeki SM (1973) Colour coding in rhesus monkey prestriate cortex. *Brain Res* 53:422–427. [CrossRef Medline](#)
- Zylberberg A, Fernández Slezak D, Roelfsema PR, Dehaene S, Sigman M (2010) The brain's router: a cortical network model of serial processing in the primate brain. *PLoS Comput Biol* 6:e1000765. [CrossRef Medline](#)