Neuropsychologia 62 (2014) 375–389

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

# A neural network model of individual differences in task switching abilities

Seth A. Herd [a], Randall C. O'Reilly [a], Tom E. Hazy [a], Christopher H. Chatham [a,1], Angela M. Brant [a,2], Naomi P. Friedman [a,b,*]

[a] Department of Psychology and Neuroscience, University of Colorado Boulder, 345 UCB, Boulder, CO 80309, USA
[b] Institute for Behavioral Genetics, University of Colorado Boulder, 447 UCB, Boulder, CO 80309, USA

ARTICLE INFO

Available online 30 April 2014

*Keywords:*
Executive control
Set shifting
Computational model
Genetics

ABSTRACT

We use a biologically grounded neural network model to investigate the brain mechanisms underlying individual differences specific to the selection and instantiation of representations that exert cognitive control in task switching. Existing computational models of task switching do not focus on individual differences and so cannot explain why task switching abilities are separable from other executive function (EF) abilities (such as response inhibition). We explore hypotheses regarding neural mechanisms underlying the "Shifting-Specific" and "Common EF" components of EF proposed in the Unity/Diversity model (Miyake & Friedman, 2012) and similar components in related theoretical frameworks. We do so by adapting a well-developed neural network model of working memory (Prefrontal cortex, Basal ganglia Working Memory or PBWM; Hazy, Frank, & O'Reilly, 2007) to task switching and the Stroop task, and comparing its behavior on those tasks under a variety of individual difference manipulations. Results are consistent with the hypotheses that variation specific to task switching (i.e., Shifting-Specific) may be related to uncontrolled, automatic persistence of goal representations, whereas variation general to multiple EFs (i.e., Common EF) may be related to the strength of PFC representations and their effect on processing in the remainder of the cognitive system. Moreover, increasing signal to noise ratio in PFC, theoretically tied to levels of tonic dopamine and a genetic polymorphism in the COMT gene, reduced Stroop interference but increased switch costs. This stability–flexibility tradeoff provides an explanation for why these two EF components sometimes show opposing correlations with other variables such as attention problems and self-restraint.



## 1. Introduction

Understanding how people switch tasks (e.g., how the brain switches attention between a conversation and oncoming traffic) has obvious relevance in itself. In addition, detailed exploration of switch costs (the extra time it takes to perform a task when switching from a different task) has provided numerous insights into the mechanisms of human executive function (EF) (see reviews by Kiesel et al., 2010; Monsell, 2003; Vandierendonck, Liefooghe, & Verbruggen, 2010). A number of computational models have been proposed as explanations of the computational and neural mechanisms of task switching. Although these models have elucidated the sources of many task switching findings, they have tended to focus on mean effects that occur across subjects, without considering patterns of individual differences that might shed light on the mechanisms involved. In this paper, we present a biologically based neural network model of task switching that can explain patterns of individual differences in terms of specific neural mechanisms.

Importantly, this model is integrated with other models of executive tasks and so can elucidate what factors distinguish vs. unify switching abilities from or with other executive abilities. We focus on the largely unanswered question of why task switching abilities are separable from other EF abilities (such as response inhibition and working memory updating/capacity) in terms of individual differences (e.g., Miyake et al., 2000). Though an individual's performance on switch tasks correlates with other measures of EF such as performance on the antisaccade or Stroop tasks, task-switching scores also capture unique variance – different switch tasks correlate more closely with each other than they

* Corresponding author at: Institute for Behavioral Genetics, University of Colorado Boulder, 447 UCB, Boulder, CO 80309, USA. Tel.: +1 303 735 4457.
*E-mail address:* Naomi.friedman@colorado.edu (N.P. Friedman).
[1] Present address: Department of Cognitive, Linguistic, & Psychological Sciences, Brown University, Providence, RI 02906, USA.
[2] Present address: Department of Psychology, Pennsylvania State University, State College, PA 16802, USA.

http://dx.doi.org/10.1016/j.neuropsychologia.2014.04.014
0028-3932/

do with other EF tasks. This unique variance not only appears to be influenced by separate genes (Friedman et al., 2008), but it also appears to show some tradeoffs with general executive control; those who are better at switching, controlling for other EF abilities, seem to show more behavioral problems such as more attention and externalizing problems and lower self-restraint (Miyake & Friedman, 2012). Our model suggests variations in several neural mechanisms that may explain this separation and tradeoff between different aspects of EF.

A novel aspect of this study is that we vary many biologically based model parameters over a broad range to simulate individual differences within the normal range, and in some cases extending into ranges that would be considered pathological. This approach goes beyond typical manipulations in the computational literature that involve, for example, lesioning parts of the model to ascertain their necessity for performing a task or to simulate specific neurological insult. By focusing on more graded manipulations of key parameters, we simulate individual differences in the model performance that can explain observed patterns of correlations and anti-correlations in the literature. We focus on simulating qualitative rather than quantitative patterns. Moreover, we focus on explaining results found with normal population samples, in which individuals may fall on a spectrum of ability (and disorder), with the extremes capturing individuals who might meet the criteria for a diagnosis. Our model suggests variations in several neural mechanisms that may explain this separation and tradeoff between different aspects of EF. First, however, we situate our model in relation to previous models and theories of task switching.

### 1.1. Previous models of task switching

Theoretical and computational models of task switching have focused on two possible sources of switch costs (Kiesel et al., 2010; Vandierendonck et al., 2010). While these possibilities were considered competing explanations in early work, more recent theoretical work holds that both are likely true (Vandierendonck et al., 2010). The first possibility is that switch costs reflect time needed to resolve interference from prior conflicting task sets (rules for mapping stimuli to responses on a given task, often thought to be instantiated neurally as representations of continuously firing neurons in prefrontal cortex or PFC). This task-set interference explanation includes, in some formulations, activation triggered by previous stimulus-task associations (i.e., "task-set inertia"; Allport, Styles, & Hsieh, 1994). The second possibility is that switch costs reflect a time cost for active reconfiguration of task sets (e.g., Rogers & Monsell, 1995) arising, in some accounts, from retrieving the task set from long-term memory (Altmann & Gray, 2008, Logan & Schneider, 2010). Task sets are representations guiding the performance of a particular stimulus–response mapping when stimuli have been linked to many possible responses. They likely include rules about categorization of particular stimuli, response mappings, attention orientation, and response threshold (Vandierendonck et al., 2010). Mechanistically detailed models of task switching and executive control usually hold that these task set representations consist of persistent neural firing in PFC (e.g., Cohen, Dunbar, & McClelland, 1990; Herd, Banich, & O'Reilly, 2006; Reynolds, Braver, Brown, & Van der Stigchel, 2006). This mechanism is based on the abundant empirical evidence from monkey electrophysiology in working memory tasks, and human neuroimaging during task switching and other EF tasks. Most abstract mathematical models are also consistent with this hypothesis (Logan & Gordon, 2001; Meiran, Kessler, & Adi-Japha, 2008; Sohn & Anderson, 2001).

Although both explanations of switch costs are likely correct in part, Vandierendonck et al. (2010) point out that an integrated view has not been implemented in a computational model (although Brown, Reynolds, and Braver (2007) have come close with a model that incorporates a task-set representation system and interference control). Here we build on prior work by modeling task switching with a biologically based model of working memory. This model produces switch costs through both task set reconfiguration (i.e., updating the contents of working memory with new task set information based on a cue) and interference resolution (through top-down biasing and competitive inhibition). In this account, the two sources of costs are inextricably intertwined (at the level of task set representations; interference in stimulus–response mappings is separable, as we discuss later) – reconfiguring a task set takes time in part because of interference with the previous task set.

The current model is consistent with previously proposed proximal neural mechanisms of cognitive control, as explored in some detail in neural network models (Cohen et al., 1990; Herd et al., 2006). However, we further explore their origins and interactions with other brain mechanisms crucial to making cognitive control flexible and appropriate to each situation. These earlier models show how a task set representation, in the form of persistent neural firing in prefrontal cortex (PFC), influences the neural processing that takes place in other brain regions. A variety of common neural learning mechanisms (both error-driven and associative in nature) can create strong connections between neurons in the prefrontal task set and those in motor and parietal areas, thereby causing them to drive those neurons that carry out the correct mappings necessary for that task whenever the task set is active.

Crucially, those models consider task sets to be a kind of working memory. By using a well-developed model of working memory and training it to perform task switching, we follow and test this hypothesis. However, most of those previous neural network task switching models simply assumed that a task set representation is appropriately maintained during the performance of a task – that is, a specific set of prefrontal neurons remains active while a given task is performed, and that set switches when the task to be performed switches. Only a few models have included realistic mechanisms that learn task sets, (e.g., Collins & Frank, 2013; Rougier, Noelle, Braver, Cohen, & O'Reilly, 2005), and these studies do not investigate resulting switch costs in detail. Here, we explore a set of brain mechanisms that could appropriately perform that type of switching, including learning task set representations, and how to switch them appropriately. The inclusion of these mechanisms gives our model enough depth to explain not only mean effects across the population, but individual differences in task performance.

### 1.2. Individual differences in task switching are separable from other cognitive control abilities

Previous work has repeatedly demonstrated that EFs can be characterized as a family of related but separable abilities, both at behavioral and neural levels (e.g., Collette et al., 2005; Friedman et al., 2006; Hedden & Yoon, 2006; Lehto, Juujärvi, Kooistra, & Pulkkinen, 2003; Miyake et al., 2000). That is, these abilities show unity and diversity (Teuber, 1972; Miyake et al., 2000). Here we explore some possible mechanistic underpinnings of two of these factors, one common to many or most EF tasks, and one specific to switching tasks. These two factors appear to be somewhat anti-correlated, so that those better at EF in general show a greater RT difference between switch and repeat trials in two-task switching paradigms than those who are weaker on general EF (although they still tend to be faster in both conditions in absolute terms). As described below, the evidence for this hypothesis is indirect but strong.