The Role of The Basal Ganglia in the Human Cognitive Architecture: A Dynamic Causal Modeling Comparison Across Tasks and Individuals

Catherine Sibert (sibert@uw.edu)
Department of Psychology, University of Washington, Seattle, WA 98195 USA

Holly Hake (hakehs@uw.edu)
Department of Psychology, University of Washington, Seattle, WA 98195 USA

John Laird (laird@umich.edu)
Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA

Christian Lebiere (cl@cmu.edu)
Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213 USA

Paul Rosenbloom (rosenblo@usc.edu)
Department of Computer Science and Institute for Creative Technologies, University of Southern California, Los Angeles, CA 90089 USA

Andrea Stocco (stocco@uw.edu)
Department of Psychology, University of Washington, Seattle, WA 98195 USA

Abstract

The basal ganglia (BG) performs an important functional role in cognition, but models disagree about the nature of the relationship between BG activity and activity in other cortical areas. Previous computational models can be categorized as implementing the effects of the BG on prefrontal cortex as either local and direct, or involving other regions and, therefore, modulatory. To test which of these two effects best represents the role of the BG, a large fMRI dataset of 200 participants performing six, representative cognitive tasks was analyzed through Dynamic Causal Modeling (DCM). To ensure that DCM models were realistic and representative of a general brain architecture, the models were implemented within the putative neural underpinnings of the Common Model of Cognition, an abstract blueprint for cognition. The comparison showed that Mixed model, including both Direct and Modulatory connectivity, consistently outperformed models that included only direct or modulatory connections. It was also found that the relative rankings of the Direct and Modulatory models depended on the specific task, suggesting that the BG is a flexible system that adapts to task demands.

Keywords: Basal ganglia, Brain architecture, Cognitive Architecture, Computational models, Dynamic Causal Modeling, fMRI.

Introduction

The BG are a set of interconnected subcortical nuclei surrounding the thalamus that receives inputs from virtually the entire cortex (Bostan, Dum, & Strick, 2018), and projects (through the dorsal and medial nuclei of the thalamus) to the prefrontal cortex. For many decades, the BG have been closely associated with motor functions, especially in neurodegenerative pathologies such as Parkinson's Disease and Huntington's Disease. Even in these cases, however, non-motor symptoms can be widespread and prevalent, including decline in working memory, executive function, and attention (Owen, 2004). Furthermore, abnormal BG function has been implied in a variety of pathologies, including obsessive compulsive disorder (Modell, Mountz, Curtis, & Greden, 1989), Tourette's syndrome (Mink, 2001) , and attention deficit disorder (Shaw et al., 2014). Finally, the scope of BG involvement in cognition has further increased with fMRI studies. At the time of writing this paper, more than 4,000 BG papers can be found in Neurosynth, spanning a variety of functions including attention (Schouwenburg, Ouden, & Cools, 2010), language (Booth, Wood, Lu, Houk, & Bitan, 2007), learning, working memory (McNab & Klingberg, 2008), and memory retrieval (Tricomi & Fiez, 2008).

In summary, neurological and neurocognitive research have shown a widespread involvement of the BG across a variety of core mental functions and disorders. This variety suggests two things. First, that understanding the function of the BG is essential for correctly understanding a variety of diseases. Second, the BG likely provides a set of core computations that are used by multiple specialized circuits in the human brain. This also suggests that we need to look at the relationship between the BG and cortical activity. To understand these core computations, a number of computational models have been put forward.

Computational Models of the BG

Many models of the BG have been proposed over the past decades, growing in complexity and biological fidelity from the early, simple actor-critic model of Barto (1995) to the complex spiking neuron system of Eliasmith et al. (2012). Despite the variety of domains

that have been tackled by these models and the differences in their approaches, models have converged on a series of assumptions that are commonly shared. For example, all models agree that multiple inputs from the cortex undergo some form of integration and selection in the striatum. They also agree that dopamine carries a reward prediction error signal, usually framed in the context of reinforcement learning theory. Additionally, all models agree that dopamine drives plasticity and results in learning. And finally, all models agree that the outputs of the BG circuit (through the thalamus) have profound and measurable effects on the neural activity of prefrontal cortex.

Despite their convergence, models *disagree*, however, about the *nature* of the effects of BG outputs to PFC. For the purpose of this paper, we will focus on how activity in the BG affects prefrontal activity in the context of cortical activity at large. Evaluating the activity of a target prefrontal region in this larger context is key to understanding the remainder of this paper. Here, we argue that, in general, these effects can be divided into two categories, *direct* and *modulatory*.

In the majority of models, the effect of BG activity on a prefrontal region has no implications with respect to the larger cortical activity. In other words, the effects of BG on a cortical region are local, and can be observed and modeled without any reference to the state of other cortical regions. Being local, these effects can be categorized as *direct*. Examples of these models include Gurney's model of action selection (Gurney, Prescott, & Redgrave, 2001; Humphries, Stewart, & Gurney, 2006; Redgrave, Prescott, & Gurney, 1999), Houk's model of action selection between competing motor programs (Houk, Adams, & Barto, 1995) and Ashby's models of working memory (Ashby, Ell, Valentin, & Casale, 2005).

In other models, however, the effect of BG outputs is to control how cortical activity in a different region of the brain affects cortical activity in PFC. In statistical terms, the activity of the BG is best seen as an interaction effect between the target PFC and other cortical regions. Therefore, we refer to this effect as *modulatory*. Models of this type include the PBWM model of working memory (Frank, Loughry, & O'Reilly, 2001; O'Reilly & Frank, 2006) and Chris Eliasmith's SPAUN (Eliasmith et al., 2012). In the PBWM, for example, cortical areas by default maintain information, resisting inputs from other dedicated regions. The BG switches the states of PFC cells, allowing them to cease the maintenance of the current pattern of activation and, instead, receive inputs from a posterior cortical region. In other words, the BG performs a "gating" operation, controlling the flux of inputs from posterior to prefrontal cortical regions.

The distinction between Direct and Modulatory models can be blurred, and some models include both types of effect. For example, in the Conditional Routing model (Stocco, Lebiere, & Anderson, 2010), the BG is initially used to resolve conflict and route information from posterior regions to the PFC; however, with learning, the BG can learn to directly affect PFC activity without the need from a posterior input.

Dynamic Causal Modeling

To test whether the effects of the BG are direct, modulatory, or both, one needs to examine the *functional connectivity* between the BG and other brain regions, that is, the extent to which changes in one region affect other regions. One elegant way to do this is through Dynamic Causal Modeling (DCM; Friston, Harrison, & Penny, 2003). This framework captures functional connectivity as a set of parameters that estimate the extent to which the timeseries of one brain region affects the rate of change of another. The change in the activity $\mathbf{y}$ of N regions is captured here as a differential equation:

$$d\mathbf{y}/dt = \mathbf{A}\mathbf{y} + \sum_i x_i \mathbf{B}_i \mathbf{y} + \mathbf{C}\mathbf{x} + \sum_j \mathbf{y}_j \mathbf{D}_j \mathbf{y} \qquad (1)$$

In this equation, $\mathbf{x}$ is a binary vector that defines the presence of each of M task conditions at any moment in time and $\mathbf{y}$ represents the current activity of all regions. Effective connectivity is represented by four parameter matrices: $\mathbf{A}$ is a N-by-N matrix of parameters that capture the intrinsic connectivity between regions, $\mathbf{B}$ is M-by-N-by-N matrix that defines the modulatory effects that M task conditions have on connectivity between regions; $\mathbf{C}$ is an M-by-N matrix that defines the effects the M task conditions on each region, and $\mathbf{D}$ is N-by-N-by-N matrix that defines the modulatory effects that regions have on the connections between other regions, can also reflect these computational interactions *over time*, accounting for the variability in temporal dynamics (Friston, 2011).

Essentially, this method not only fits parameters to individual regions, but also takes into account the connections between them which is important in a dynamic system like the brain. The most critical aspect of this framework is that DCM can model second-order interactions between nodes in a network, i.e., cases in which a region modulates the connectivity between two other regions (represented by $\mathbf{D}$). This is of particular interest as it plays an influential role in capturing the nature of BG functionality, whether *direct* or *modulatory*.

The Common Model of Cognition

One of the downsides of DCM is that, instead of providing data-driven estimates of network structure, it can only measure connectivity within a given network model. In turn, this means that the precise measure of connectivity between two regions might change depending on the larger network context in which they are embedded. For this reason, it is essential to apply DCM within a biologically-plausible network architecture. Because, in this paper, the function of the BG will be tested across a variety of tasks, such a network architecture would need to be large enough to

span the most common regions yet simple enough to make DCM computationally tractable.

This role is precisely fit by the so-called Common Model of Cognition (CMC; Laird, Lebiere, & Rosenbloom, 2017). The CMC is a computational framework for cognitive components, designed to distill the lessons learned from 40 years of cognitive science research on the essential computations of the mind and their relationships. Abstract computations are categorized into five functional components (long-term memory, working memory, procedural memory, perception systems, and action systems) and their relationships are specified in a graph structure (Fig 1A).

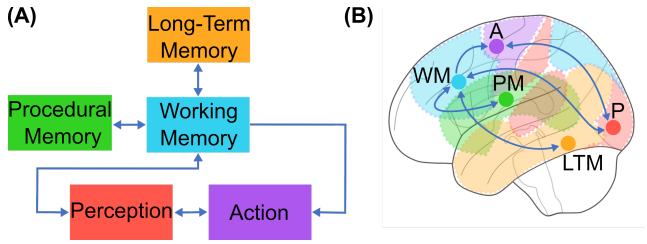

Figure 1: (A) The Common Model of Cognition (CMC); (B) Proposed associations between components and anatomical brain regions.

Although it was not proposed specifically as a *brain* architecture, a number of studies have found that the CMC is surprisingly effective at modeling brain activity across tasks and individuals (Steine-Hanson, Koh, & Stocco, 2018; Stocco, Laird, Lebiere, & Rosenbloom, 2018; Stocco et al., 2021). In this interpretation, the CMC's functional components are mapped onto large-scale brain regions (Fig 1B) and their relations are translated into predicted patterns of functional connectivity. In other words, the neural counterparts of the functional components and their connections serve as a simplified architecture for the human brain, not only the human mind. Importantly, the neural interpretation of the CMC explicitly includes the BG as one of its components: it serves as the neural implementation of Procedural Memory.

Recent work has shown that the CMC can provide a remarkably good fit to fMRI data from over 200 participants, across a variety of representative tasks. Inspired by its success in accounting for functional connectivity, we decided to extend this approach to investigate the role of the BG by creating three different variants of the CMC architecture, corresponding to three different characterizations of the basal ganglia. In the *Direct* model (Fig 2A), the basal ganglia directly affects the prefrontal cortex (corresponding to the "Working Memory" component of the CMC). In the *Modulatory* model (Fig 2B), the BG does not affect PFC directly, but instead mediate the connectivity between perceptual inputs and long-term memory to PFC. Finally, a third, *Mixed* model was examined (Fig 2C); in this model, the BG can exert both direct and modulatory effects on PFC. Because the Mixed model contains all of the connections of the other two, the Direct and Modulatory models can be said to be nested under the Mixed model.

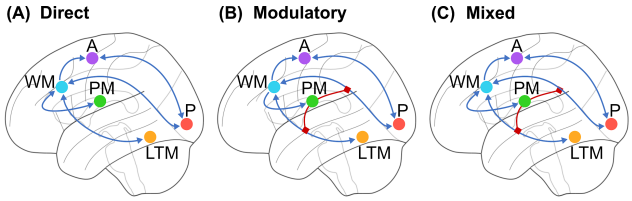

Figure 2: Structure of all three DCM network models.

The Human Connectome Project Dataset

A few previous studies have reported analysis of direct vs., modulatory connections in the BG (Stocco, 2018; Steine-Hanson et al., 2018), but only within a few similar tasks. Because these models embody different hypotheses about the *general* functional role of the BG in cognition, it is important to test them across multiple tasks. This is particularly important when considering the Mixed model, as the importance of direct and modulatory connections might adaptively change depending on the situation, and testing a single task might yield biased results that are not representative of the whole functionality. Although previous attempts have found that the Modulatory model fits the data better than the Direct model within a specific task (Steine-Hanson et al., 2018), and in healthy participants over BG-impaired patients (Ketola, Thompson, Madhyastha, Grabowski, & Stocco, 2020; Prat, Stocco, Neuhaus, & Kleinhans, 2016), a true test of the models requires an examination of a larger variety of tasks that span different cognitive domains.

To this extent, we used a subset of the fMRI data available on the Human Connectome Project (Van Essen et al., 2013), one of the largest, high-quality human neuroimaging datasets. The HCP includes data from 1,200 young adults performing different tasks, six of which were selected to cover different domains of cognition. Table 1 lists these six paradigms and their relevant experimental conditions. The three architectural models described above were translated into equivalent DCM network models, and data from the six tasks were used to fit network connections that best predicted the observed human brain activity. These predictions were compared across the alternate model structures.

Materials and Methods

The study presented herein consists of an analysis of a large sample (N=200) of neuroimaging data from the Human Connectome Project. The analysis was restricted to the task-based fMRI data, which consisted of two sessions of each of seven paradigms.

Table 1: Paradigms used in the HCP dataset.

Paradigm (Reference)	Condition (Baseline)
Emotion Processing	Fearful and angry faces (Neutral faces)
Incentive Processing	Winning blocks of choices (Losing blocks)
Language and Math	Answering questions about stories or math (Passive listening)
Relational Reasoning	Inferring the rule in relational arrays (No rule in control arrays)
Social Cognition	Socially interacting shapes (randomly moving shapes)
Working Memory	2-Back blocks of stimuli (0-Back)

Tasks fMRI Data

The HCP task-fMRI data encompasses seven different paradigms designed to capture a wide range of cognitive capabilities. Of these paradigms, six were included in our analysis (the seventh was a motor localization task). A full description of these tasks and the rationale for their selection can be found in the original HCP papers (Barch et al., 2013; Van Essen et al., 2013).

Data Processing and Analysis

Imaging Acquisition Parameters. As reported in Barch et al., (2013), functional neuroimages were acquired with a 32-channel head coil on a 3T Siemens Skyra with TR = 720 ms, TE = 33.1 ms, FA = 52°, FOV = 208 × 180 mm. Each image consisted of 72 2.0mm oblique slices with 0-mm gap in-between. Each slice had an in-plane resolution of 2.0 × 2.0 mm. Images were acquired with a multi-band acceleration factor of 8X.

Image Preprocessing. Images were acquired in the "minimally preprocessed" format (Van Essen et al., 2013), which includes unwarping to correct for magnetic field distortion, motion realignment, and normalization to the MNI template. The images were then smoothed with an isotropic 8.0 mm FWHM Gaussian kernel.

Canonical GLM. A canonical GLM analysis was conducted on the smoothed minimally preprocessed data using a mass-univariate approach, as implemented in the SPM12 software package. First-level (i.e., individual-level) models were created for each participant. The model regressors were obtained by convolving a design matrix with a hemodynamic response function; the design matrix replicated the analysis of Barch et al., (2013), and included regressors for the specific conditions of interest described in Table 1. Second-level (i.e., group-level) models were created using the brain-wise parametric images generated for each participant as input.

DCM-specific GLM. A second GLM analysis was carried out to define the event matrix **x** that is used in the DCM equation (Eq. 1). Because these models are not used to perform a standalone data analysis, the experimental events and conditions are allowed to be collinear. All of the HCP paradigms include a critical condition and a baseline (Table 1). As is common in DCM analysis, these two experiment conditions were modeled in a layered, rather than orthogonal fashion, with all trials included in a generic regressor and trials from the critical condition representing a special regressor that further drives neural activity. Following the procedures of (Stocco et al., 2021), the association between regressors and ROIs was kept constant across all tasks, with the generic regressor affecting the perceptual component and the critical regressor affecting WM (capturing the greater mental effort that is common to all critical conditions).

Regions of Interest Definition. Regions of Interest (ROIs) for each task and participant were defined using the method described in Stocco et al. (2021). For each CMC component, a group-level centroid was first identified by running a canonical GLM analysis that compared the stimuli against their baseline (see Table 1) and then locating the peak of a statistical parametric map within the general areas associated with that CMC component (Fig 1). Because all tasks show stronger activation in the left hemisphere than in the right, all the group-level centroids were located in the left hemisphere.

To account for individual-level variability in functional neuroanatomy, the group-level coordinates were then used as the starting point to search in 3D space for the closest activation peak within each individual statistical parameter map. Figure 3 illustrates the distribution of the individual coordinates of each region for each task, overlaid over a corresponding group-level statistical map of task-related activity (as in Stocco et al., 2021). Each individual coordinate is represented by a point; the ~200 points form a cloud that captures the spatial variability in the distribution of coordinates. Next, the individualized ROI coordinates were used as the center of a spherical ROI. All voxels within the sphere whose response was significant at a minimal threshold of $p < .50$ (50%) were included as part of the ROI.

Finally, for each ROI of every participant in every task, a representative time course of neural activity was extracted as the first principal component of the time series of all of the voxels within the sphere.

Model Fitting. Once the time-series for each ROI was extracted, different networks were created by connecting all of the individually-defined ROIs according to the specifications of each model (Fig 2). The predicted neural activity for each model was then calculated using Equation 1, and the predicted time course of BOLD signal was then generated by applying a biologically-plausible model of neurovascular coupling to the simulated neural activity of each region. All of the

model parameters were estimated through an expectation-maximization procedure (Friston et al., 2003) to reduce the difference between the predicted and observed time course of the BOLD signal in each ROI.

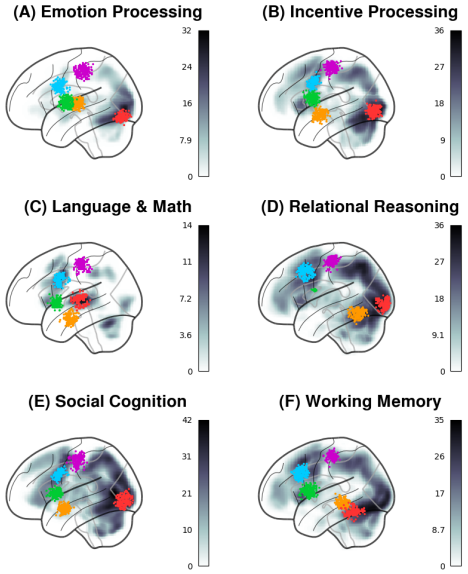

Figure 3: Location of ROI centroids; variations account for individual differences in functional anatomy.

Model Comparison. Given two or more generative models that produce the same data **x**, it is possible to compare them by estimating a *likelihood function* $L(m \mid \mathbf{x})$ and select the model with the highest likelihood. A model's likelihood is the posterior probability of a model m given the observed data **x**, and is defined as the probability that m would generate **x**,. that is, $L(m \mid \mathbf{x}) = P(\mathbf{x} \mid m)$. Group-level likelihood values for a model m can then be expressed as the product of the likelihood of that model fitting each participant p, i.e., $\prod_p L(m \mid \mathbf{x}_p)$ (Stephan, Penny, Daunizeau, Moran, & Friston, 2009). Because probability values might become vanishingly small (leading to rounding errors), it is customary to compare models in terms of their log-likelihood, with the group-level log-likelihood then becoming the sum of all of the individual log-likelihoods: $\sum_p \log L(m \mid \mathbf{x}_p)$. Although more sophisticated model comparison procedures have been proposed (Stephan et al., 2009), the log-likelihood based metric used here is not only the most easily interpretable, but also the most relevant, as it specifically applies to cases in which it is assumed that the model is constant or architectural across individuals (Kasess et al., 2010).

Results

Figure 4 illustrates the group-level log-likelihoods of the three models in each of the six tasks. Because the group-level log-likelihoods vary dramatically across tasks and models, the figure presents them as *relative* log-likelihoods: within each task, the lowest log-likelihood is subtracted from all the others. As a result, the worst-fitting model always has a relative log-likelihood value of zero, while the other two are positive. In all tasks, the Mixed model significantly outperforms the other two.

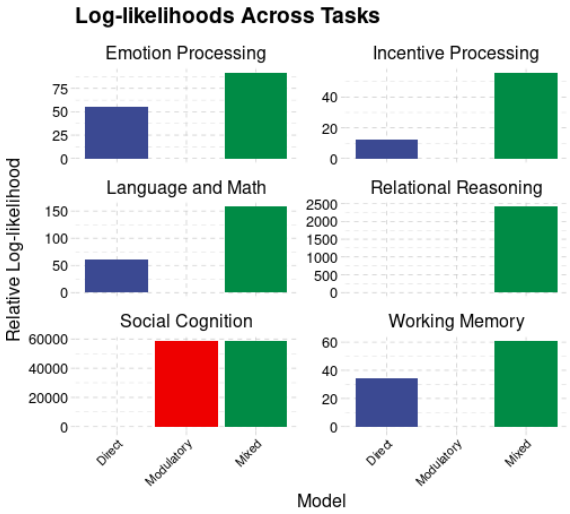

Figure 4: Relative log-likelihoods across tasks.

The results also show significant differences across tasks. In the Relational Reasoning task, for example, the Mixed model dramatically outperforms the other two. This is to be expected, given that this task requires integration of knowledge. It is also interesting to note the relative fit of the other two models is dependent on the task; while the Direct model is generally preferred over the Modulatory model, the Modulatory vastly outperforms the Direct model in the Social Cognition task, being almost as effective as the Mixed model.

Since the three models differ also in terms of the number of parameters, it is possible that the Mixed model's greater likelihood is due to it simply having more degrees of freedom to fit the data. Although common corrections can be applied (such as BIC and AIC) to provide a correction for the number of parameters, the fact that the Direct and Modulatory models are nested within the Mixed model allows us to use Wilks' theorem (1938), which accounts for the different number of parameters and translates the log-likelihood difference into interpretable p-values. The theorem states that, for two models of which one is nested, the probability that the fit of the more complex is due to chance (its p-value) approximates the probability of obtaining the value of 2λ (twice the log-likelihood difference) in a χ^2 distribution with degrees of freedom corresponding to the difference in the number of parameters. Using this theorem, we calculated the probability that the greater fit of the Mixed model is due to chance (note that this comparison accounts for the greater complexity of the Mixed model in the χ^2 distribution). Figure 5 illustrates the performance of both

models against the Mixed model; the dashed lines correspond to the differences in log-likelihood of each model (i.e, 2λ) against the Mixed model, while the shaded areas depict the corresponding χ^2 distributions (note that the distributions are the same for all tasks, but the scale of the *x*-axis changes to accommodate the likelihood differences). All of the differences in log-likelihood are all far to the right of χ^2 distributions, corresponding to $p < .0001$ for all comparisons in all tasks. This implies that the Mixed model's superiority cannot be simply due to its greater complexity.

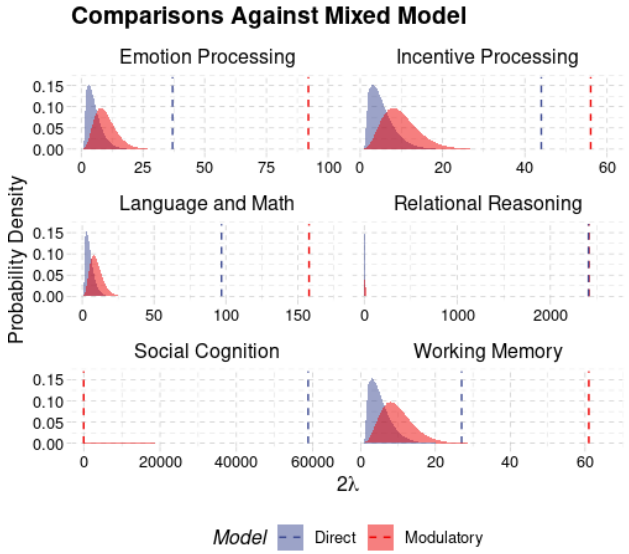

Figure 5: Comparisons of the Direct and Modulatory model against the Mixed model.

Discussion

This paper has examined the functional role of the basal ganglia in brain and cognition. Based on a review of existing computational models, the effects of the BG on prefrontal cortex were categorized as either local and *direct* or involving other regions and, therefore, *modulatory*. To test which best represents how neural activity is propagated from the BG to the cortex, a large fMRI dataset of 200 participants performing six, representative cognitive tasks was analyzed through Dynamic Causal Modeling. The DCM models were implemented within the Common Model of Cognition, an abstract blueprint for cognition that has proven surprisingly effective at capturing neural activity. The comparison showed that a Mixed model that includes both direct and modulatory connectivity consistently outperformed models that include only direct or only modulatory connections. It was also found that the relative rankings of the direct and modulatory models depended on the specific task, suggesting that BG is a flexible system that adapts to task demands.

The results are silent about the exact nature of the BG computations and the relationship between direct and modulatory connectivity. The Conditional model (Stocco et al., 2010) provides a useful framework, in which tasks initially require modulation but, with practice, switch to a more direct effect. These predictions could potentially be addressed in future analysis of the data. Finally, the results show the value of the CMC as a potential, high-level brain architecture. The CMC has provided a way to capture different tasks within a single, consistent network architecture; this was key to our success.

Acknowledgments

This research was supported by grant FA9550-19-1-0299 from the Air Force Office of Scientific Research to AS.

References

Ashby, F. G., Ell, S. W., Valentin, V. V., & Casale, M. B. (2005). FROST: A Distributed Neurocomputational Model of Working Memory Maintenance. *Journal of Cognitive Neuroscience*, *17*, 1728–1743.

Barch, D. M., Burgess, G. C., Harms, M. P., Petersen, S. E., Schlaggar, B. L., Corbetta, M., & WU-Minn HCP Consortium. (2013). Function in the human connectome: Task-fMRI and individual differences in behavior. *NeuroImage*, *80*, 169–189.

Barto, A. G. (1995). Adaptive critics and the basal ganglia. In *Computational Neuroscience. Models of information processing in the basal ganglia* (pp. 215–232). Cambridge, MA, US: The MIT Press.

Booth, J. R., Wood, L., Lu, D., Houk, J. C., & Bitan, T. (2007). The role of the basal ganglia and cerebellum in language processing. *Brain Research*, *1133*, 136–144.

Bostan, A. C., Dum, R. P., & Strick, P. L. (2018). Functional Anatomy of Basal Ganglia Circuits with the Cerebral Cortex and the Cerebellum. *Progress in Neurological Surgery*, *33*, 50–61.

Eliasmith, C., Stewart, T. C., Choo, X., Bekolay, T., DeWolf, T., Tang, Y., & Rasmussen, D. (2012). A Large-Scale Model of the Functioning Brain. *Science*, *338*, 1202–1205.

Frank, M. J., Loughry, B., & O'Reilly, R. C. (2001). Interactions between frontal cortex and basal ganglia in working memory: A computational model. *Cognitive, Affective & Behavioral Neuroscience*, *1*, 137–160.

Friston, K. J. (2011). Functional and effective connectivity: A review. *Brain Connectivity*, *1*, 13–36.

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *NeuroImage*, *19*, 1273–1302.

Gurney, K., Prescott, T. J., & Redgrave, P. (2001). A computational model of action selection in the basal ganglia. I. A new functional anatomy. *Biological Cybernetics*, *84*, 401–410.

Houk, J. C., Adams, J. L., & Barto, A. G. (1995). A model of how the basal ganglia generate and use neural signals that predict reinforcement. In *Computational Neuroscience. Models of information processing in the basal ganglia* (pp. 249–270). Cambridge, MA, US: The MIT Press.

Humphries, M. D., Stewart, R. D., & Gurney, K. N. (2006). A physiologically plausible model of action selection and oscillatory activity in the basal ganglia. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *26*, 12921–12942.

Kasess, C. H., Stephan, K. E., Weissenbacher, A., Pezawas, L., Moser, E., & Windischberger, C. (2010). Multi-Subject Analyses with Dynamic Causal Modeling. *NeuroImage*, *49*, 3065–3074.

Ketola, M., Thompson, S., Madhyastha, T., Grabowski, T., & Stocco, A. (2020). *Applying the Common Model of Cognition to Resting-State fMRI Leads to the Identification of Abnormal Functional Connectivity in Parkinson's Disease*.

Laird, J. E., Lebiere, C., & Rosenbloom, P. S. (2017). A Standard Model of the Mind: Toward a Common Computational Framework across Artificial Intelligence, Cognitive Science, Neuroscience, and Robotics. *AI Magazine*, *38*, 13–26.

McNab, F., & Klingberg, T. (2008). Prefrontal cortex and basal ganglia control access to working memory. *Nature Neuroscience*, *11*, 103–107.

Mink, J. W. (2001). Basal ganglia dysfunction in Tourette's syndrome: A new hypothesis. *Pediatric Neurology*, *25*, 190–198.

Modell, J. G., Mountz, J. M., Curtis, G. C., & Greden, J. F. (1989). Neurophysiologic dysfunction in basal ganglia/limbic striatal and thalamocortical circuits as a pathogenetic mechanism of obsessive- compulsive disorder. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *1*, 27–36.

O'Reilly, R. C., & Frank, M. J. (2006). Making Working Memory Work: A Computational Model of Learning in the Prefrontal Cortex and Basal Ganglia. *Neural Computation*, *18*, 283–328.

Owen, A. M. (2004). Cognitive dysfunction in Parkinson's disease: The role of frontostriatal circuitry. *The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, *10*, 525–537.

Prat, C. S., Stocco, A., Neuhaus, E., & Kleinhans, N. M. (2016). Basal ganglia impairments in autism spectrum disorder are related to abnormal signal gating to prefrontal cortex. *Neuropsychologia*, *91*, 268–281.

Redgrave, P., Prescott, T. J., & Gurney, K. (1999). The basal ganglia: A vertebrate solution to the selection problem? *Neuroscience*, *89*, 1009–1023.

Schouwenburg, M. R. van, Ouden, H. E. M. den, & Cools, R. (2010). The Human Basal Ganglia Modulate Frontal-Posterior Connectivity during Attention Shifting. *Journal of Neuroscience*, *30*, 9910–9918.

Shaw, P., De Rossi, P., Watson, B., Wharton, A., Greenstein, D., Raznahan, A., … Chakravarty, M. M. (2014). Mapping the development of the basal ganglia in children with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *53*, 780-789.e11.

Steine-Hanson, Z., Koh, N., & Stocco, A. (2018). Refining the Common Model of Cognition Through Large Neuroscience Data. *Procedia Computer Science*, *145*, 813–820.

Stephan, K. E., Penny, W. D., Daunizeau, J., Moran, R. J., & Friston, K. J. (2009). Bayesian model selection for group studies. *NeuroImage*, *46*, 1004–1017.

Stocco, A. (2018). A Biologically Plausible Action Selection System for Cognitive Architectures: Implications of Basal Ganglia Anatomy for Learning and Decision-Making Models. *Cognitive Science*, *42*, 457–490.

Stocco, A., Laird, J., Lebiere, C., & Rosenbloom, P. (2018). *Empirical Evidence from Neuroimaging Data for a Standard Model of the Mind*. In C. Kalish, M. Rau, J. Zhou, and T. T. Rogers (Ed.), Proceedings of the 40th Annual Meeting of the Cognitive Science Society (pp. 1094–1099).

Stocco, A., Lebiere, C., & Anderson, J. R. (2010). Conditional routing of information to the cortex: A model of the basal ganglia's role in cognitive coordination. *Psychological Review*, *117*, 541–574.

Stocco, A., Sibert, C., Steine-Hanson, Z., Koh, N., Laird, J. E., Lebiere, C. J., & Rosenbloom, P. (2021). Analysis of the human connectome data supports the notion of a "Common Model of Cognition" for human and human-like intelligence across domains. *NeuroImage*, *235*, 118035.

Tricomi, E., & Fiez, J. A. (2008). Feedback signals in the caudate reflect goal achievement on a declarative memory task. *NeuroImage*, *41*, 1154–1167.

Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E. J., Yacoub, E., Ugurbil, K., & WU-Minn HCP Consortium. (2013). The WU-Minn Human Connectome Project: An overview. *NeuroImage*, *80*, 62–79.

Wilks, S. S. (1938). The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *Annals of Mathematical Statistics*, *9*, 60–62.