From low to high cognition: A multi-level model of behavioral control in the primate brain

Jerald D. Kralik[1,2] (jerald.kralik@gmail.com), Dongqing Shi[1] (dongqshi@gmail.com), Omar A. El-Shroa[1] (o.a.elshroa@gmail.com), Laura E. Ray[3] (laura.e.ray@dartmouth.edu)

[1]Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH 03755 USA

[2]Department of Bio and Brain Engineering, Program of Brain and Cognitive Engineering,
College of Engineering, Korea Advanced Institute of Science and Technology (KAIST),
Daejeon, 34141, South Korea

[3]Thayer School of Engineering, Dartmouth College, Hanover, NH 03755 USA

Abstract

The basic cognitive architecture of the human brain remains unknown. However, there is evidence for the existence of distinct behavioral control systems shared by humans and nonhumans; and there is further evidence pointing to distinct higher-level problem solving systems shared by humans and other primates. To clarify the nature of these proposed systems and examine how they may interact in the brain, we present a four-level model of the primate brain and compare its performance to three other brain models in the face of a challenging foraging problem (i.e., with transparent, and thus, invisible barriers). In all manipulations (e.g., size of problem space, number of obstacles), our model never performed the best outright; however, it was always among the best, appearing to be a jack-of-all-trades. Thus, the virtues of our primate brain lie not only in the heights of thinking it can reach, but also in its range and versatility.

Keywords: cognitive control; cognitive architecture; reinforcement learning; creativity; agency; concept formation

Introduction

There is considerable evidence for the existence of two neural systems controlling mammalian behavior: (1) a goal-directed decision-making system based in the ventromedial prefrontal cortex (vmPFC), and (2) a habit system based in the striatum (Daw, Niv, & Dayan, 2005; Rangel, Camerer, & Montague, 2008). It has been argued that the goal-directed system is model-based, meaning essentially that an individual knows what will happen upon taking an action; while the habit system is not, meaning that an individual simply knows what action to take in a given state of the world (Daw et al., 2005). A model-based system potentially provides an infrastructure for mental simulation, planning, and reasoning, which in turn lead to faster learning and greater generalization across environmental conditions. Such potent abilities reduce errors when facing novel situations. However, research in engineering and computer science shows that models are notoriously brittle and therefore often break under real-world conditions. One approach to dealing with this brittleness is to have a fallback process that does not rely on the model, and this appears to be one of the main advantages of the model-free habit system. Another approach would be to have an additional system that can fix the models when they break, enabling individuals to solve these harder problems. We classify these two types of problems as *apparent* — i.e., the problem features are properly modeled by the standard model-based problem-solving system — and *nonapparent* — i.e., those that break the system.

In primates, there is neuroanatomical evidence for a distinct region of prefrontal cortex (PFC), called granular PFC (which includes lateral PFC and frontal pole), and some theorists have suggested that this region may enable primates to perform unconventional behaviors, such as looking away from a salient visual stimulus when required to do so (Passingham & Wise, 2012; Striedter, 2005). Related to this view, we hypothesize that granular PFC mediates the cognitive ability to solve *nonapparent* problems. Moreover, we believe that a detailed analysis of this ability in primates will shed light on the mechanisms that underpin creative problem solving in people.

Here, we present a computational framework and model to begin this analysis of clarifying apparent versus nonapparent problem solving, as well as to examine how these processes interrelate with the other main behavioral control systems of the primate brain. First, we focus on the classic *detour problem* in the comparative literature, in which subjects must circumvent a barrier (either via reaching or navigating) to obtain a reward item (Wynne & Udell, 2013). Most comparative research on the detour problem has focused on cases with opaque barriers as obstacles, and have generally shown that many species can solve the problem — thus, they have the basic capacity to take paths away from a goal item to reach it, at least when the obstacle is clearly defined.

However, as illustrated in Figure 1a, many nonhuman animals and human infants find the problem challenging when the barrier is transparent, as they repeatedly attempt to reach directly for the reward item, even in the face of strong, negative feedback (Diamond, 1990; Santos, Ericson, & Hauser, 1999; Wallis et al., 2001). This insensitivity to feedback is typically explained as an inability to inhibit a lower-level behavioral control system (such as a Pavlovian system). However, other experimental conditions suggest not, at least for nonhuman primates (Santos et al., 1999). For example, when first given experience with an opaque barrier, subjects tend to solve the transparent barrier problem. The ease with which they refrained from reaching directly for the food item once they had an alternative response available suggests that the major difficulty did not stem from a lack of self-control.

We propose that the difficulty results from confusion with the transparent barrier: the subjects do not readily see that there is a barrier. Although their response is blocked, there is no apparent reason for it, and so they continue to attempt the most efficient solution of reaching directly for the goal item. We suggest that this is an example in which a problem-solving system sees a clear solution and is therefore overriding feedback to the contrary. However, it is using a broken model. Put differently, it is an example of a cognitive illusion that provides insight into how cognitive systems are constructed (Kahneman, 2011).

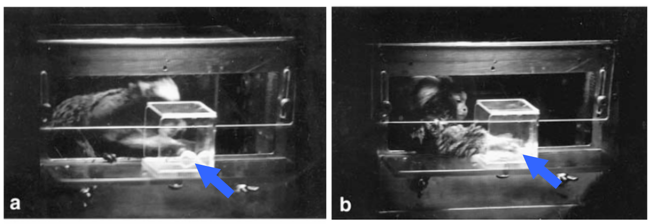

Figure 1: Illustration of a nonapparent problem. (**a**) The marmoset monkey attempts to reach directly for the marshmallow but is blocked by the transparent barrier. (**b**) The monkey learns to reach around the barrier.[1]

We further suggest that the transparent barrier problem requires a nonapparent solution. To solve it, the problem-solving system must include an obstacle in the problem formulation, thus fixing the model that did not include it. However, if an individual cannot see that an obstacle exists, it must be inferred from the effect of being blocked. Thus, taken together, we use the detour problem with a transparent barrier as an example of a nonapparent problem, and more specifically, as a case in which an individual confronts an unknown event or consequence, must infer a hidden cause, and create a (nonperceptual) concept to model it (Goswami, 2008; Holyoak & Morrison, 2012).

In addition, there is evidence that the ability to solve this nonapparent problem is subserved by a separate problem-solving system. Chiefly, it has been shown that rhesus monkeys who normally solve the transparent barrier problem lose this ability with lateral PFC lesions (Diamond, 1990). Thus, such findings implicate lateral PFC in mediating nonapparent problem solving, and suggest that it is separable from what we are calling apparent problem solving.

From these results and others, we model the primate brain with four basic levels of behavioral control. The first is based on the first main system in the vertebrate brain that controls complete goal-oriented behaviors: the hypothalamus (Swanson, 2000). That is, it is the first behavioral control system involving complete behavioral sequences that attain goals, such as goal-directed approach and ingestion behaviors when food is perceived. However, here we assume it is normally inhibited until the goal state is reached, and then is used to complete the process of actually obtaining and ingesting the food item. The second level represents the striatal-based habit system, which uses model-free reinforcement learning. The third level represents the first model-based problem solving system, i.e., that which solves *apparent* problems. This system would solve the detour problem with well-defined, opaque barriers, but would produce direct reaching with the ill-defined transparent barrier. The fourth level, then, performs *nonapparent* problem solving, and evidence suggests it is mediated by granular PFC. Evidence that granular PFC mediates nonapparent problem solving further suggests that mammalian vmPFC likely underlies the solving of apparent problems.

Because these systems of the primate mind/brain evolved under specific selection pressures, we also attempted to model these pressures to best understand the utility of each system and their interactions. More specifically, we used a foraging problem, and tested parameters to mimic basic selection pressures, such as size of the foraging environment, number of obstacles, and changing conditions, e.g., changes in the goal state position. In addition, because the primate brain evolved along a specific phylogenetic trajectory, rather than studying each system in isolation, we examine them from this phylogenetic perspective, in which each system appears to be added to a previous combination. Thus, very roughly, we compared the four-level primate brain model to an ancestral vertebrate brain, consisting of the first two levels, and an ancestral mammalian brain, consisting of the first three levels. We also compared the model to an alternative primate brain consisting of Levels 1, 2, and 4 (i.e., without the simpler *apparent* problem-solving system).

In this paper, we focus on the main difference between the levels and therefore only use transparent barriers (i.e., no opaque ones). Thus, when no obstacles are in the direct path, the apparent system can solve the problem; otherwise, the nonapparent system must be utilized.

In sum, the aims of the current project were (1) to begin specifying more clearly how nonapparent problems can be solved when the simpler model-based problem solving system fails; (2) to examine the potential advantages and disadvantages of a four-level system; and (3) to determine the foraging conditions that would provide a selective advantage for this brain architecture.

Methods

In what follows, we describe the modeling environment, and then layout the details of our model; we next present the four competing models, and then describe how the models were assessed.

Modeling environment

To focus on the main features of our model, we sought to keep the testing environment as simple and straightforward as possible. Therefore, we tested the model with a *foraging problem* using a *2D grid world*, in which the agent must learn a path from its current location to the goal location that avoids all obstacles (Figure 2).

[1]With permission from Wallis, Dias, Robbins, & Roberts (2001)

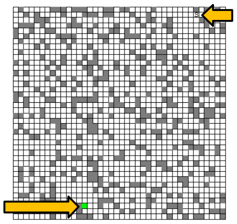

Figure 2: Example grid world, with the 'Initial state' denoted as "S" (Start) and the 'Goal state' in green. Obstacles are denoted in black, however, they are actually transparent (and thus hidden to the problem solver visually).

Model description

Our model consists of the four levels and a cognitive control mechanism that passes control between them.

Level 1 enables actual *goal attainment* once in view. For foraging, it completes the act of food consumption. This level is assumed in the four-level model, and is not explicitly modeled or tested here. With respect to the brain, Level 1 represents hypothalamus, which is considered the first behavioral control system involving complete behavioral sequences that attain goals (Swanson, 2000).

Level 2 represents the striatal-based habit system. It uses a *Markov Decision Process (MDP)* and *Reinforcement-Learning (RL) framework* (Sutton & Barto, 1998). Thus, there is a set of states, i.e., the (x, y) positions in the grid world; a set of possible actions, i.e., all eight reachable positions from a given grid world position; and a reward function that assigns values (called *action* or Q values) to the actions based on environmental feedback, with the Q values representing expected future reward. The agent then learns to choose the best action (i.e., policy) in a given state that leads to the highest future reward. Thus, rather than having a model of the world, i.e., an understanding of how the states relate to each other in the larger grid world, Level 2 sees the states independently, simply using a Q-Table to determine action values in each state (Daw et al., 2005). Level 2 is composed of *learning* and *acting* components. The *learner* updates the action values using the following Q-learning algorithm (a form of RL learning):

$$Q_{t+1}(s_t, a_t) = Q_t(s_t, a_t) + \alpha\,[r_{t+1} + \gamma \max_a(s_{t+1}, a) - Q_t(s_t, a_t)]$$

where α is the learning rate $\alpha \in [0, 1]$; r_{t+1} is the actual reward received at episode $t+1$; and γ is the discount factor $\gamma \in (0, 1)$ (Sutton & Barto, 1998). The *actor* selects an action according to the Boltzmann distribution of Q values:

$$\pi(a|s_t) = \frac{e^{Q_t(s_t,a)/\tau}}{\sum_b e^{Q_t(s_t,b)/\tau}}$$, where τ is the temperature that controls the degree of action exploration.

Cognitive control, i.e., control between levels, in the current model is generally hierarchical and modulatory (Kahneman, 2011; Miller & Cohen, 2001). Control begins at Level 2; however, when Level 2 fails (here, when the maximum Q value for an action is not unique), control passes to Level 3. Levels 3 & 4, representing *apparent* and *nonapparent* problem solving respectively, derive their behavioral control by modulating the actor in Level 2 (by modifying the action values and then passing control back to the Level 2 actor when a solution is found). If Level 3 fails to find a solution, it passes control to Level 4; and if Level 4 fails, Level 2 selects an action randomly. Future work will examine a more sophisticated controller based on, e.g., cost/benefit analysis (Daw et al., 2005; Kowaguchi et al., *accepted*).

Level 3 represents explicit problem solving when problem components are apparent, passing control to Level 4 when they are ill-defined and a solution cannot be found. Both Levels 3 and 4 look for solutions based on a cognitive model—i.e., a cognitive map—of the problem environment that is built in the background as the problem solver moves through it and experiences the state transitions: i.e., the subsequent state reached when an action is taken in a given state (with all transition probabilities = 1), e.g., $T(s_i, a_j) \rightarrow s'$, where $s_i = (1, 1)$, a_j = *move right*; and $s' = (2, 1)$, and thus, placing s' to the right of s_i in the map/model (Sutton & Barto, 1998; Daw et al., 2005). This model thus has an understanding of the relationships among the states in the grid world that Level 2 cannot see. Currently, we have one cognitive model that is built; however, Levels 3 and 4 contribute different elements to it and utilize what they have access to.

More specifically, for both Levels 3 and 4, all problems are considered *multi-agent problems*, with three classes of agents: self, others, and goals (Holyoak & Morrison, 2012; Shi et al., 2010; Wooldridge, 2009). All are considered agents because they could theoretically invoke a state change by virtue of their actions and functional relationships with other agents. (This would occur for the goal item if, for example, it were moving prey; however, in the current case, the goal item is stationary.) Thus, the cognitive model used by Levels 3 and 4 consists of four main components: (1) an x, y coordinate frame that defines each location in the grid world; (2) the identification and location of every agent in the problem; (3) the set of available actions each agent could take; and (4) an understanding of functional interactions among the agents (Goswami, 2008): i.e., f_i(*agent*$_j$, *target agent*$_k$), where f_i could be *acquiring* the goal item by the problem solver or *blocking* of the problem solver by an obstacle (also considered an agent by virtue of this blocking effect). For the current study, there is only one type of obstacle with only one available action: blocking.

For any novel problem in the grid world, the problem solver cannot see the entire problem immediately — the world is too large — and so a cognitive model must be developed via initial experience with each state. Model building entails developing the cognitive map of (x, y) coordinates for each state as well as whether an agent resides in each state. Again, because Level 3 can only see apparent obstacles, not invisible ones, it cannot see any of the transparent obstacles, and thus assumes there aren't any.

For problem solving, Level 3 uses the cognitive model to find a path to the goal. Since Level 3's view of the grid-world problem appears clear of obstacles, it always finds a direct path to the goal. The potential advantage of this is that when there are clear paths to the goal, a brain that contains this system would provide fast and efficient solutions. Since Level 3 always sees a clear path, it will continue producing this solution (via hill climbing), leaving the problem solver in a potential phase of perseveration (i.e., continuing to attempt the same direct path solution). This impasse, then, is what causes cognitive control to pass to Level 4.

Level 4 assumes control when Level 3's model breaks. Level 4 therefore represents the system that attempts to find these nonapparent solutions. For the current project, this occurs every time a hidden cause, i.e., an 'invisible' obstacle, blocks the direct path to the goal. In this case, the obstacle is literally nonapparent to Level 3.

Level 4's main contributions occur with both the building of the overall cognitive model of the grid-world problem and utilizing it to solve the nonapparent problems. First, as the problem solver is moving in the problem space (the grid world) via the actor module, when it is blocked, Level 4 uses this information conceptually, inferring that there is an agent doing the blocking (Goswami, 2008; Holyoak & Morrison, 2012; Tenenbaum et al., 2011; Wynne & Udell, 2013). That is, it infers $f_i(obstacle, self)$, where f_i is *blocking*. From this inference, Level 4 places a *blocking* agent at the grid location in the cognitive model. Unfortunately, Level 3 cannot see this nonvisual conceptual obstacle; it is beyond Level 3's comprehension.

Second, when cognitive control is passed to Level 4, it uses the complete cognitive model (including the transparent barriers) to find a path to the goal. To achieve this, Level 4 currently uses the planning algorithm A* (called 'A star') to find an efficient path around the obstacles to the goal (Russell & Norvig, 2010). Once a path is found, Level 4 then modifies the Level 2 action values so that the actor module will use the path. The main advantage of Level 4 over Level 2 (simple model-free RL) is that it can perform inductive inference, and use the internal cognitive map for mental simulation and planning, leading to rapid, one-trial learning via problem solving (as well as greater generalization to novel problems in future model development) (Passingham & Wise, 2012).

Figure 3 summarizes the key characteristics of each level.

Levels

1. See goal → obtain it
(e.g., Approach food → ingest)

2. (a) Actor (action selection via Boltzmann distribution)
(b) Learner (Q-learning algorithm)

3. (a) Builds internal cognitive model
(b) Hill-climbing search
(c) Change Level 2 Q-values for Actor

4. (a) Adds transparent obstacles to the internal model
(b) A* search
(c) Change Level 2 Q-values for Actor

Figure 3: Description of model levels.

Brain models

We compared our model to three others, thus testing four different multi-level models:

(1) Model 1: consisting of levels 1 & 2
(2) Model 2: levels 1, 2, & 3
(3) Model 3: levels 1, 2, & 4
(4) Our model, Model 4: levels 1, 2, 3, & 4

All four models assume the existence of Level 1. Model 1 simply uses model-free reinforcement learning, and roughly speaking, perhaps represents an ancestral vertebrate. Model 2 combines model-free RL with the ability to solve more straightforward, *apparent* problems, perhaps representing the ancestral mammalian brain. Model 3 represents a brain that contains both lower level model-free RL and the higher level *nonapparent* problem-solving system that one might argue should replace the simpler *apparent* problem-solving system altogether. Model 4 is our multi-level model of the primate brain.

Model assessment

We examined the effects of (1) grid world size, (2) number of obstacles, (3) changing initial states, (4) changing goal states, and (5) changing obstacles. The models were assessed via two measures (average from ~50 iterations): (1) *Cumulative number of steps* to reach the goal after 200 learning episodes (i.e., 200 times in which the goal item was reached); and (2) *Cumulative computational cost* needed per action, measured as the amount of processing time per action. The two measures were combined to obtain an overall *performance score*.

Simulation results

Because Level 1 was not explicitly modeled or examined here, results are from Levels 2-4. We maintain the color-coding of the model names to help keep them straight.

Grid world size

To examine the effects of grid world size, no obstacles were included. As seen in Figure 4a, performance was best for Model 1 with the smaller world and worse with the largest grid size. As the world size increased, Model 2 and Model 4 performed the best. Figure 4b shows the learning rates, and Figure 4c the cumulative computational costs for all four brain models for the largest grid world size (40x40). Model 1 was slower to learn a path to the goal (and progressively more so as the world size increased); in contrast, the other models all continued to learn very quickly. For all grid world sizes, the computational cost was greatest for Model 3 and lowest for Model 1.

Number of transparent obstacles

All remaining analyses used the large grid size. As the number of obstacles increased, Model 1 and Model 2 were slower to learn a path to the goal, while Model 3 and Model 4 continued to learn quickly. As shown in Figure 5, Model 4 performed relatively well across all numbers of obstacles, Model 2 performed relatively well until the largest number, Model 3 was relatively better with the largest number of obstacles, while Model 1 performed the poorest at every number of obstacles.

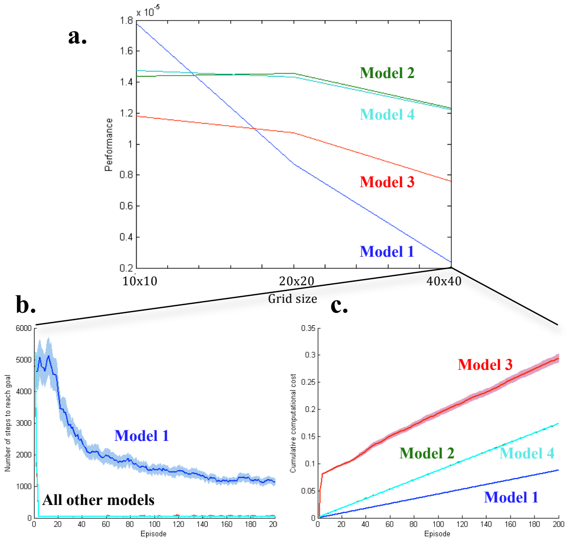

Figure 4: (**a**) Performance of the four brain models as a function of grid world size, (**b**) Learning rates, and (**c**) Cumulative computational costs for all four brain models for the largest grid world size (40x40). Bands around the curves represent standard error of the mean (SEM).

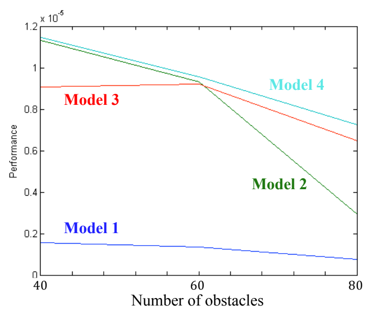

Figure 5: Model performance as a function of the number of transparent obstacles in the grid world.

Changing world

To examine the effects of a changing initial state and changing goal state, no obstacles were used.

Changing initial state The initial state of the problem solver was changed once every 50 episodes or once every 10 episodes. The models were generally robust with the changing initial states, although computational costs increased, especially for Model 3. As shown in Figure 6, overall performance was best for Model 2 and Model 4.

Changing goal state As seen in Figure 7a, rises in path length occurred when the goal state changed, especially for Model 1. With a changing goal state (and no obstacles), the computational costs for Model 3 were relatively high (Fig. 7b).

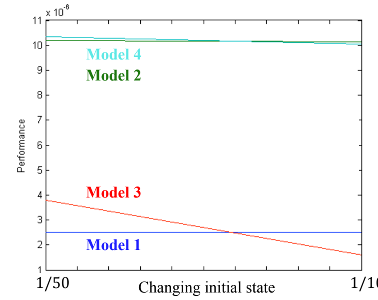

Figure 6: Model performance as a function of the rate of change of the initial state: once every 50 episodes or once every 10 episodes.

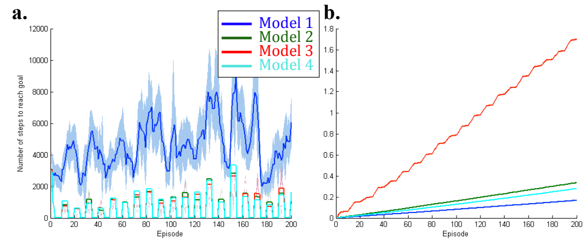

Figure 7: (**a**) Learning rates and (**b**) Cumulative computational costs for the highest rate of goal state location change (once every 10 episodes). Bands around the curves are SEMs.

As seen in Figure 8, Model 2 and Model 4 outperformed the others.

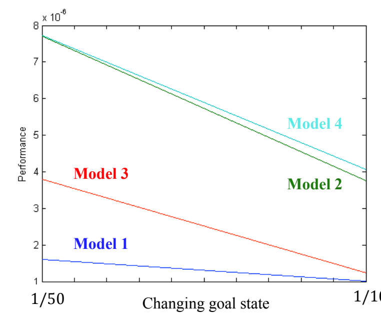

Figure 8: Performance as a function of changing goal state.

Changing (transparent) obstacles As shown in Figure 9, when the number of obstacles (600) and frequency of change (1/10) were high, Model 3 and Model 4 found a path to the goal most quickly (Fig. 9a), however, the computational demands were relatively steep (Fig. 9b).

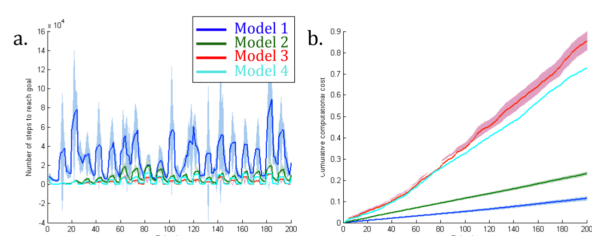

Figure 9: (**a**) Learning rates and (**b**) Cumulative computational costs for changing obstacles (600 total, once every 10 episodes). Bands around the curves are SEMs.

Discussion

Model-based problem solving potentially provides great advantages, such as lowering the number of errors during learning and generalizing to novel problems (Daw et al., 2005; Holyoak & Morrison, 2012). However, models of the real world are notoriously brittle, and thus require other approaches to problem solving when they break. It has been hypothesized that primates have evolved granular prefrontal cortex to cope with these more challenging, nonapparent problems. Here, the aims of our study were threefold: (1) to begin clarifying the mechanisms of nonapparent problem solving, (2) to examine the potential advantages and disadvantages of having four types of behavioral control systems, and (3) to determine the foraging conditions that would provide a selective advantage for this brain architecture. We used the classic detour problem in the comparative literature, and in particular, focused on the use of transparent barriers, which prove challenging for nonhuman animals. We suggest that an *apparent* problem-solving system, likely mediated by vmPFC, can solve detour problems with well-defined obstacles; while a *nonapparent* system, mediated by granular PFC, can solve the problem with 'invisible', transparent barriers. It does so by inferring the existence of a barrier from its effects on the problem solver. Thus, this system may underlie the powerful ability of humans to infer hidden causes from given events and consequences (Holyoak & Morrison, 2012; Tenenbaum et al., 2011). Other behavioral research we have conducted with monkeys suggests further possible mechanisms for the nonapparent system that help solve insight problems by both nonhuman primates and people (Murray, Kralik, Wise, 2005; Kowaguchi et al., *accepted*; Kralik, 2005, 2011). We plan to incorporate these findings in the future.

With respect to our second and third aims, the advantages of greater cognitive abilities must outweigh the disadvantages, and the advantages of a multi-level brain architecture appears to lie in its versatility. In all manipulations conducted here, our model (Levels 1-4) never performed the best outright; however, it was always among the best, appearing to be a jack-of-all-trades. Thus, rather than fitting a high-level cognitive niche best, our brain model appears to best fit a niche with problems of varying levels of complexity: a low-to-high cognitive niche. Thus, it may be useful to have multiple behavioral control systems at different levels of sophistication, which allow computational savings when facing simpler problems, and more elaborate capabilities when faced with more challenging ones (Kahneman, 2011; Rangel et al., 2008).

More theoretical development, however, is required to better understand the characteristics of such multi-level systems. For example, we plan further developments that include using a dynamic environment, different classes of agents that can both hinder or aid the problem solver in goal attainment, more sophisticated inductive reasoning and cognitive-control mechanisms, and further levels of abstraction (Botvinick, Niv, & Barto, 2009; Daw et al., 2005; Shi et al., 2010; Tenenbaum et al., 2011).

The ability to solve problems creatively across a wide range of domains embedded in complex, physical environments remains out of reach for current artificial systems; but we are extending their reach (Hélie & Sun, 2010). A detailed analysis of how it evolved in the human lineage should help to further demystify the creative process. Such an analysis can also help to clarify how primates in general, and humans in particular, have come to fill the low-to-high cognitive niche. Creative thinking represents the pinnacle of high-level cognition and underlies many of our greatest achievements. This success not only derives from the heights of thinking we can attain, but also the diversity of challenges we can master.

Acknowledgments

ONR Grant N00014-08-1-0693 supported the project.

References

Botvinick, M. M., Niv, Y., & Barto, A. C. (2009). Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective. *Cognition*, 113, 262-280.

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, *8*(12), 1704–1711.

Diamond, A. (1990). Developmental time course in human infants and infant monkeys, and the neural bases of, inhibitory control in reaching. *Ann New York Ac Sci*, 608, 637-676.

Goswami, U. (2008). *Cognitive Development*. London, UK: Routledge.

Hélie, S. & Sun, R. (2010). Incubation, Insight, and Creative Problem Solving: A Unified Theory and a Connectionist Model. *Psychological Review*, 117, 994-1024.

Holyoak, K. J. and Morrison, R. G. (2012). *The Oxford Handbook of Thinking and Reasoning*. Oxford: Oxford University Press.

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus, and Giroux.

Kowaguchi, M., Patel, N. P., Bunnell, M. E., & Kralik, J. D. (*Accepted*). Competitive control of cognition in rhesus monkeys. *Cognition*.

Kralik, J.D. (2005). Inhibitory control and response selection in problem solving: How cotton-top tamarins (*Saguinus oedipus*) overcome a bias for selecting the larger quantity of food. *Journal of Comparative Psychology,* 119(1), 78-89.

Kralik, J.D. (2011). Rhesus Macaques (*Macaca mulatta*) spontaneously generalize to novel quantities in a reverse-reward contingency task. *Journal of Comparative Psychology*, 255-262.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*, 167-202.

Murray, E. A., Kralik, J. D., & Wise, S. P. (2005). Learning to inhibit prepotent responses: successful performance by rhesus macaques, Macaca mulatta, on the reversed-contingency task. *Animal Behavior,* 69, 991-998.

Passingham, R. E., & Wise, S. P. (2012). *The neurobiology of the prefrontal cortex: Anatomy, evolution, and the origin of insight*. Oxford, UK: Oxford University Press.

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, *9*(7), 545–556.

Russell, S., & Norvig, P. (2010). *Artificial Intelligence: A Modern Approach*. Upper Saddle River, NJ: Prentice Hall

Santos, L. R., Ericson, B. N., & Hauser, M. D. (1999). Constraints on problem solving and inhibition: Object retrieval in cotton-top tamarins. *J. Comparative Psychology*, 113, 186-193.

Shi, D., Sauter, M. Z., Sun, X., Ray, L. E., & Kralik, J. D. (2010). An extension of Bayesian game approximation to partially observable stochastic games with competition and cooperation. *Proceedings of the International Conference on Artificial Intelligence (ICAI)*.

Sutton, R. S. & Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press.

Striedter, G. F. (2005). *Principles of Brain Evolution*. Sunderland: Sinaur Assoc.

Swanson, L. W. (2000). Cerebral hemisphere regulation of motivated behavior. *Brain Research*, *886*(1-2), 113–164.

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, Structure, and Abstraction. *Science*, 331, 1279-1285.

Wallis, J. D., Dias, R., Robbins, T. W., & Roberts, A. C. (2001). Dissociable contributions of the orbitofrontal and lateral prefrontal cortex of the marmoset to performance on a detour reaching task. *European Journal of Neuroscience*, *13*(9), 1797–1808.

Wooldridge, M. (2009). *An Introduction to MultiAgent Systems*. Chichester, UK: John Wiley & Sons.

Wynne, C., & Udell, M. A. R. (2013). *Animal Cognition*. London, UK: Palgrave Macmillan.