



Published in final edited form as:

*Cogn Affect Behav Neurosci.* 2013 March ; 13(1): 1–22. doi:10.3758/s13415-012-0125-7.

## Rapid instructed task learning: A new window into the human brain's unique capacity for flexible cognitive control

Michael W. Cole, Patryk Laurent, and Andrea Stocco

### Abstract

The human ability to flexibly adapt to novel circumstances is extraordinary. Perhaps the most illustrative yet underappreciated form of this cognitive flexibility is rapid instructed task learning (RITL) – the ability to rapidly reconfigure our minds to perform new tasks from instruction. This ability is important for everyday life (e.g., learning to use new technologies), and is used to instruct participants in nearly every study of human cognition. We review the development of RITL as a circumscribed domain of cognitive neuroscience investigation, culminating in recent demonstrations that RITL is implemented via brain circuits centered on lateral prefrontal cortex. We then build on this and other insights to develop an integrative theory of cognitive flexibility and cognitive control, identifying theoretical principles and mechanisms that may make RITL possible in the human brain. Insights gained from this new theoretical account have important implications for further developments and applications of RITL research.

### Keywords

cognitive flexibility; compositionality; cognitive control; prefrontal cortex; flexible cognitive control

### Introduction

One of the defining characteristics of human-level intelligence is the ability to rapidly restructure one's behavior into novel configurations from instruction. This ability is important in everyday life. For instance, it is essential for learning new technologies and new skills at all levels of education. Furthermore, nearly every experimental psychologist uses verbal instructions to inform participants how to perform experimental tasks, yet the mechanisms underlying this process are largely unknown (Monsell, 1996).

The neural and cognitive processes underlying this ability are the focus of an emerging area of cognitive neuroscience research. This new area investigates the neural basis of rapid instructed task learning (RITL; pronounced “rattle”) – a term we propose to describe the ability to rapidly learn task procedures from instruction. Here we interpret, distill, and build a novel theory based on recent findings regarding this key component of human cognition, helping to establish RITL as a multidisciplinary domain of scientific inquiry and thereby help accelerate further research in this area.

We also present RITL as an especially important form of cognitive flexibility, given the extraordinary speed (one trial) and adaptability (involving novel mental configurations) required by RITL. These attributes make RITL 1) an especially specific and sensitive methodology for the study of flexible cognition, and 2) an important source of constraints on the kinds of neural architectures capable of implementing flexible human cognition. We support these conclusions by reviewing cognitive, computational, and neuroscientific studies of RITL and postulating a novel neural architecture capable of implementing RITL.

This article is structured in three main sections, as follows: We first provide an overview of RITL including its definition along with a review of previous and recent research on the topic. Next, we introduce and discuss our novel neuroscientific theory of RITL (and flexible cognition generally). Finally, we take a broad view of RITL research, informed from the perspective gained from this new neuroscientific theory and looking forward to future research and potential applications.

## REVIEW OF RITL RESEARCH

### Defining RITL

RITL is the process of rapidly (typically on the first trial) learning a novel rule or task from instructions. Humans often learn new tasks using RITL, such as how to use new technologies (e.g., a new “smartphone”), how to cook new recipes (e.g., a new kind of lasagna), or how to play an unfamiliar game (e.g., the first time checkers is played). These tasks can be learned via reinforcement learning (Sutton & Barto, 1998), yet they are learned much more efficiently with RITL. Consider for a moment how difficult learning checkers would be without RITL abilities. Instead of rapidly learning the rules for how each piece moves and that the goal is to capture all of the opposing player’s pieces, you would need to randomly select from dozens of possible actions (e.g., moving your pieces to the other end of the board, having the goal of moving every piece in sequence, etc.) until an instructor rewards valid moves and eventually rewards a win. Reinforcement learning such as this has been investigated much more extensively than RITL despite the clear utility and high frequency of RITL use in everyday life.

There are two basic forms of reinforcement learning. RITL is most sharply contrasted with ‘unsupervised’ reinforcement learning, in which a task is learned based on environmental reinforcement for correct task behavior rather than from an instructor. Like problem solving, task learning can be conceptualized as search through a state space of possible tasks (see Newell & Simon, 1972). From this perspective it is apparent that in many environments an unsupervised reinforcement learning approach is equivalent to pseudo-random search, making it highly inefficient (see Figure 1).

Although reinforcement learning is the primary means of learning early in life, humans are quickly able to start receiving instruction from others (i.e., ‘supervised’ learning). The initial form of supervised learning – termed ‘shaping’ – speeds task learning substantially by reinforcing behaviors as the learner gets closer to the proper task (Figure 1A). However, even this efficient form of learning is still much slower than RITL. RITL is able to achieve first-trial learning due to the instructor using examples and/or language to directly specify the correct task in the state space of possible tasks (Figure 1B) (M. Cole, 2009). Though some tasks are so complex or nuanced that directly specifying the exact task state is difficult or impossible (e.g., performing a professional-level tennis backhand), there are many tasks in everyday life that can be immediately performed on the first try via RITL (M. Cole, 2009). RITL research investigates how the brain is able to rapidly convert the water of instructions into the wine of novel task performance.

### Defining RITL More Precisely: The Role of First-Trial RITL and Neuroscience

The distinction between RITL and reinforcement learning is typically clear, but there are some learning situations that involve aspects of both. For instance, learning can be relatively rapid but with instructions delivered in the form of post-trial feedback. Presently, it is uncertain if such multi-trial feedback-based rapid learning is truly the same as RITL. Since it is evident that first-trial (i.e., without feedback or multi-trial integration) performance on a novel task involves RITL, we suggest that RITL research should focus primarily on such “first-trial RITL” situations. Supporting this suggestion, we recently found that there is a

sharp shift in behavior between first and second encounters with novel tasks, even when other tasks are performed between those encounters (Michael W Cole & Braver, Submitted). Cohen-Kdoshay & Meiran (2009) also emphasized the importance of focusing on first trials when investigating the effects of instructions on behavior in order to rule out effects of long-term memory traces formed across multiple trials.

It will also be important, however, for future research to discover the exact boundaries of what defines RITL. We suggest that cognitive neuroscience can play a critical role here. Specifically, we suggest that identifying the neural mechanisms underlying first-trial RITL will allow precise categorization of learning episodes as either involving RITL or not. In other words, we propose that a learning episode involves RITL in so far as it involves the neural mechanisms underlying the most theoretically established form of RITL: first-trial RITL. It may be the case that there are precise boundaries surrounding first-trial RITL (e.g., completely different brain processes when feedback is involved), or it may be the case that the same neural mechanisms are involved during first-trial RITL and a variety of similar learning situations. Adjudicating among these possibilities will be an important new direction for cognitive neuroscience RITL research.

### **Previous RITL Research: Neuropsychology and The Role of Language**

Initial studies of RITL – prior to the recent advent of RITL cognitive neuroscience research – were in the areas of neuropsychology and computational modeling. Milner (1964, 1965) and Luria (1973; 1964) found that lesions in the lateral prefrontal cortex (LPFC) led to patients with normal linguistic abilities who were seemingly able to understand and remember instructions, yet had a profound inability to execute those instructions. These and other instances of ‘goal neglect’ (Duncan, Burgess, & Emslie, 1995) provided preliminary evidence that RITL does not simply rely on language or remembering instructions during execution. Instead, these results suggest the existence of processes in LPFC that convert instructions into task sets that can then be executed as necessary for accurate task performance.

These studies indicate that rather than depending on linguistic abilities per se, RITL depends on the ability to rapidly reconfigure task set. It can be further demonstrated that in many cases language is not necessary for RITL at all. For instance, instructions for building IKEA furniture can be rapidly learned despite purely iconic diagrams with no language (see Holmes, 2005 for more examples). Other forms of non-linguistic RITL involve using imitation or context (e.g., using spatial or temporal proximity to associate an arbitrary stimulus with a response) to rapidly learn new tasks.

Linguistic RITL appears to be the most powerful form of RITL, however. This is due to its ability to directly specify task states (see Figure 1B) even when they are abstract. For instance, the task ‘lift the red items’ is readily specified linguistically, but may require many trials via the other forms of task learning (e.g., lifting of a red dotted hat, lifting of a red dotted shirt [‘lift red dotted clothes?’], lifting of a red striped shoe [‘lift red clothes?’], lifting of a red striped ball [‘ah, lift the red items’]). In other words, many more tasks can be specified with linguistic RITL than is either practical or possible with other forms of task learning (M. Cole, 2009).

### **Previous RITL Research: Computational Models**

The use of punishments and rewards can at best be a part of the teaching process... It is necessary... to have some other “unemotional” channels of communication. If these are available it is possible to teach a machine by punishments and rewards to

obey orders given in some language, e.g., a symbolic language. ... The use of this language will diminish greatly the number of punishments and rewards required.

– A. Turing, in *Computing Machinery and Intelligence* (1950), in which the Turing Test was first proposed

Computational models have been used to explore the mechanisms necessary for RITL since long before neuropsychological RITL research began. Indeed, Alan Turing was inspired by the human capacity for RITL – the ability to convert instructions into novel task performance – when he made his field-defining advances in computer science. For example, his concept of the universal Turing machine demonstrated a way in which a machine could be rapidly instructed to perform any computable task (Turing, 1936). Turing also used the analogy of RITL to propose the use of high-level programming languages to make instructing a machine much easier (Turing, 1950) (see quote above). From this perspective, one of Turing’s legacies is that every modern computer is, in some sense, a computational model of the human capacity for RITL.

Programming languages were not intended to model human RITL per se, however. Therefore, computers did not yield insight into human RITL until computational models specifically attempting to model the human mind were developed. Early on, production models (consisting entirely of if-then ‘production’ rules) were built to rapidly learn from instruction (Anderson, 1976; Kieras & Bovair, 1986). However, these models required instructions to be represented in the arbitrary codes used by the models rather than emerging from known cognitive or neural mechanisms, reducing the relevance of these RITL findings to humans.

This limited relationship of computational modeling to the biological substrates of human learning was overcome, in the 1980s, with the widespread introduction of more biologically plausible ‘connectionist’ models consisting of networks of neuron-like units. However, in sharp contrast to RITL, learning in connectionist models was entirely based on either associative or error-driven algorithms, relying on trial-and-error (see Figure 1A) and typically needed thousands of trials before learning a new task (Rumelhart & McClelland, 1987). Importantly, several examples of ‘structured’ connectionist models (in which some connections are specified *a priori*) showed that simulated networks of neuron-like units could rapidly learn novel tasks from instructions (Noelle & Cottrell, 1996; Schneider & Oliver, 1991). These models used specialized units for active maintenance of if-then contingencies to allow for rapid reorganization of the model’s internal state to implement novel task parameters. Though these models were still somewhat slower (dozens of trials) than the first-trial learning humans are capable of, these computational studies demonstrated the utility of if-then rules and active maintenance of information for RITL abilities. More recently, several neural architectures have proposed more specific biological substrates for RITL computations (Doll, Jacobs, Sanfey, & Frank, 2009; Lebiere & Anderson, 1993; Ramamoorthy & Verguts, 2012; Andrea Stocco, Lebiere, & Anderson, 2010; Zylberberg, Dehaene, Roelfsema, & Sigman, 2011). Importantly, these proposed neural substrates converge with the previous neuropsychological results, since in all of these models the active maintenance of task-relevant representations is thought to occur within LPFC (Miller & Cohen, 2001).

### RITL and Human Intelligence

One of the most striking distinctions between human and non-human primate intelligence is the tremendous amount of time it takes for non-human primates to learn tasks. For instance, it takes several weeks or months for a macaque monkey to learn a simple delayed match-to-sample task (Verrico et al., 2011), while a human can learn it immediately (“press the button when two stimuli in a row match”). This striking between-species difference illustrates that

RITL may be one of the cognitive skills that have expanded most rapidly with the evolutionary changes producing *homo sapiens sapiens*, the modern human species.

These considerations suggest that we may learn something about the evolutionary process underlying RITL abilities by comparing humans to other primate species. First, however, it would be informative if we could establish the plausibility of RITL itself as the driving force of evolutionary change – as opposed to RITL emerging solely from the evolution of related cognitive abilities. Importantly, recent simulation studies have shown that RITL and related forms of social learning enhance group survival rates (L Rendell et al., 2010; Luke Rendell et al., 2011), providing a plausible selective pressure behind the evolution of RITL specifically (though independent selective pressures on related abilities certainly also played an important role).

Recent evidence suggests that macaque monkeys (despite requiring months to learn simple delayed comparison tasks) have some limited RITL abilities: they can learn very simple concrete tasks devoid of inter-stimulus conflict with a small amount (~5 trials) of practice (Cromer, Machon, & Miller, 2011) (Figure 2A & 2B). Importantly, a member of a species evolutionarily closer to humans, Kanzi the bonobo chimpanzee, was able to perform first-trial RITL (Savage-Rumbaugh et al., 1993). This remarkable non-human primate was able to understand dozens of English words such that researchers could verbally instruct Kanzi to perform arbitrary simple concrete tasks (e.g., “Put your ball on the pine needles” or “Knife the sweet potatoes”) (Figure 2C). Kanzi’s RITL ability was not perfect (74% accuracy; about the level of a 2 ½ year-old human), but this ability is striking enough to suggest that there may be some common brain difference in bonobo chimpanzees and humans relative to macaque monkeys that may help account for enhanced RITL abilities in these species.

This brain difference may be in anterior LPFC (area 10). This area has undergone a greater evolutionary expansion in humans than almost any other brain area (Avants, Schoenemann, & Gee, 2006). Critically, bonobo chimpanzees have the largest anterior LPFC of all non-human primates (Semendeferi, Armstrong, Schleicher, Zilles, & Van Hoesen, 2001). This suggests that anterior LPFC may be especially important for computations underlying RITL abilities. Consistent with this, several RITL studies have demonstrated involvement of anterior LPFC during RITL (see discussion below). It may be that selective pressures pushing improvement of RITL abilities during evolution did so in part by driving computational improvements in anterior LPFC (see below for examples of how, e.g., greater gray matter volume in LPFC may enhance computations underlying RITL).

### Recent Methodological Innovations in RITL Research

RITL is a complex cognitive behavior and, as such, it is difficult to effectively isolate its subcomponents in laboratory-based experimental studies. For instance, one may want to separate processes that are specific to the *acquisition* of task rules (e.g., the “instructions”) from those that are specific to their *application*. Recently developed RITL experimental designs typically solve this problem by dividing each trial into two separate phases, an *encoding* phase where a new task is communicated through a prearranged notation (e.g., three words describing three consecutive mental operations to perform), and an *execution* phase where the task-specific stimuli are presented and participants can perform the instructed task.

The most challenging methodological problem with investigating RITL, however, is the need to statistically analyze novel task behaviors when even a single repetition of the same task invalidates its novelty (Rabbitt, 1997). Recent innovative experimental designs have overcome this problem by observing the first trials of a variety of different tasks, and then

pooling across these first trials to infer general properties of RITL across tasks (Cohen-Kdoshay & Meiran, 2009; M. Cole, 2009; Michael W Cole, Bagic, et al., 2010; Dumontheil, Thompson, & Duncan, 2011; Hartstra, Kühn, Verguts, & Brass, 2011; Ruge & Wolfensteller, 2010; A Stocco, Lebiere, & O'Reilly, 2010; Andrea Stocco, Lebiere, O'Reilly, & Anderson, 2012). For instance, Ruge & Wolfensteller (2010) constructed a variety of novel stimulus-response tasks by pairing novel stimuli in each task with simple responses re-used across tasks (Figure 3A). For instance, subjects might be instructed to respond with their left hand when a novel star shape or a novel spiral shape is presented, and respond with their right hand when a novel triangle-like shape or a novel circle-like shape is presented (with the tasks necessarily being novel given that the stimuli are novel). In contrast, in Cole et al. (2010)'s study, each task consisted of three mental rules (out of 12 possible) to be performed in a fixed order, such as "If the answer to 'is it sweet?' [Rule 1] is the same for both [Rule 2] words, press your left index finger [Rule 3]" (Figure 3B). By varying the included rules, the authors were able to create a pool of tasks (rule sets) that consisted of distinct procedures yet were still comparable. Similarly, Stocco et al. (2012) used tasks made of sets of mathematical operations (e.g., divide by two, multiply by three, then add one) and focused their analyses on tasks that consisted of novel combinations of such operations.

Importantly, a subset of these studies (M. Cole, 2009; Michael W Cole, Bagic, et al., 2010; Hartstra et al., 2011; Andrea Stocco et al., 2012) also utilized another methodological innovation: the inclusion of an additional set of tasks that have been practiced in a prior session. The inclusion of these control tasks provide an important baseline, enabling the assessment of processes specific to RITL, while controlling for generic processes that may be engaged prior to and during performance of practiced tasks. Note that it is possible that these contrasts pick up on differences that are also present between moderately practiced and extensively practiced task switches (Yeung & Monsell, 2003); it will be important for future research to explore this possibility. Further, it is possible that the ability to use the same stimuli among a wide variety of tasks in these paradigms (in contrast to Ruge & Wolfensteller, 2010) might add additional processes that need to be explored further (Rubin & Meiran, 2005). Nonetheless, these studies have consistently shown that anterior, middle (dorsolateral or ventrolateral), and posterior regions of LPFC all contribute to RITL. It will be important for future work to identify the unique contributions each portion of LPFC makes to RITL abilities.

### Key Distinctions in RITL Research

In this section we review some of the fundamental distinctions in RITL research covered here. We suggest that further characterization of these and other distinctions will be important for future progress in RITL research.

**Form of communication: Non-linguistic vs. linguistic RITL**—The existence of non-linguistic RITL (see above), and the observation that lesions in LPFC can impair RITL without impairing linguistic abilities (Luria, 1973), demonstrate that language is not necessary for RITL. Non-linguistic RITL involves learning through examples. For instance, imitation leads to RITL via simply copying the observed behavior of someone else, and may be the most extensively studied form of RITL (Heyes, 2001). Imitation is highly related to 'emulation', in which the intentions/goals are copied rather than the specific motor movements (Byrne & Russon, 1998). In contrast, linguistic instructions utilize symbolic representations to communicate task procedures. These forms of RITL communication likely involve some shared and some distinct cognitive and brain processes. Further, there are advantages to each form of RITL communication, depending on the particular task to be learned (see below).

**Level of abstraction: Concrete vs. abstract RITL**—Ultimately, we are embodied creatures. Therefore, concrete (i.e., embodied/sensory-motor) tasks are likely processed differently from abstract tasks. Concrete tasks specify a limited but exact set of percept-action associations. In contrast, abstract tasks specify a less limited but typically inexact set of percept-action associations. For example, the relatively concrete task “point left when you see a red hat” is quite exact (i.e., nearly the entire percept-action scenario is specified), yet those instructions are limited to one small set of scenarios. In contrast, the relatively abstract task “point left when you see clothing” is less exact but it applies to a wider variety of scenarios. Thus, abstract task instructions and representations can be considered compressed [in an information theoretic sense (Gray & Tall, 2007)]. Importantly, concrete tasks are more readily communicated via imitation (non-linguistic RITL), while abstract tasks are more readily communicated via language. See Ruge & Wolfensteller (2010) for an example of concrete RITL, and Cole et al. (2010) for an example of abstract RITL. Note that the use of novel stimuli for concrete RITL may allow for less proactive interference (negative transfer) than the reuse of rules across task contexts in abstract RITL. However, such reuse of rules may alternatively result in positive transfer (from greater practice with them), effectively facilitating performance. Further research is necessary to explore this issue.

**Level of complexity: Simple vs. complex RITL**—Task complexity is another potentially important distinction between the Ruge & Wolfensteller (2010) and Cole et al. (2010) studies. Ruge & Wolfensteller (2010) only included non-integrated rules (i.e., rules that could be executed independent of one another), while Cole et al. (2010) included three integrated rules for each task. This additional complexity likely led to an additional multi-rule integration process for Cole et al. (2010). Future work will be necessary to dissociate complexity from abstraction effects in RITL, which were confounded across the two studies. It is possible to de-confound these two dimensions, for instance, using novel multi-step spatial routes (an example of concrete complex RITL). Note, however, that the Ruge & Wolfensteller (2010) paradigm likely involved an integration process between the constituent stimulus and response representations of each task, such that the differences with Cole et al. (2010) may have been more of degree (greater complexity) than kind (simple vs. complex).

**Task preparation stage: Instruction vs. initial implementation**—Recent research has suggested there are distinct phases of task preparation during RITL (Michael W Cole, Bagic, et al., 2010; Andrea Stocco et al., 2012). These stages are similar to the ‘cue’ and ‘target’ stages in the task switching literature (Monsell, 2003; Ruge, Jamadar, Zimmermann, & Karayanidis, 2011), though there is strong evidence that RITL is distinct from typical forms of task switching (see below). During the instruction stage, a novel task set must be communicated by an instructor and interpreted by the learner. This interpretation process likely involves activation of the proper task semantics (e.g., ‘point left’ motor representations and ‘red hat’ visual representations). During the initial implementation stage the task is executed in response to a stimulus for the first time. In some cases (especially during abstract RITL) preparation likely continues during initial task implementation, as the stimulus allows for more concrete specification of the task procedure that is to be executed.

**RITL vs. typical task switching: Task set formation vs. task set retrieval**—The vast majority of task switching experiments have involved switching among a very small number of highly practiced tasks. These typical task switching paradigms are thought to involve a task set retrieval process from long-term memory (Mayr & Kliegl, 2000), yet such a process is impossible with RITL since the tasks are novel by definition. There may be other “task reconfiguration” processes shared between RITL and typical task switching, however. Such processes are important constituent parts of RITL, yet identifying the

component processes that are unique to RITL is essential for understanding RITL as an independent cognitive construct. Recently developed RITL paradigms have sought to identify such components, utilizing practiced in contrast to strictly novel tasks to effectively isolate RITL from processes that are not unique to RITL (Michael W Cole, Bagic, et al., 2010; Andrea Stocco et al., 2012). In addition to generic task switching processes, these paradigms also controlled for stimulus novelty and peculiarities about the specific task rules used. These designs are able to control for these processes by using the same rules across novel and practiced tasks, yet using novel (i.e., never seen or executed before) combinations of those rules for RITL trials and practiced (i.e., repeatedly performed) combinations for the control condition. Recent work (Michael W Cole, Bagic, et al., 2010; Michael W Cole & Braver, Submitted) suggests that RITL involves a unique ‘task set formation’ process that leads to an integrated task set that can then be later retrieved during practiced task preparation (i.e., during typical task switching). Further research is necessary to fully characterize the similarities and differences between these two modes of task preparation. For instance, it will be important to explore differences in the amount of between-task interference (caused from stimulus re-use) across RITL and typical task switching (Ruge et al., 2011), as well as between different RITL experimental paradigms.

### LPFC Processes Underlying RITL

The original neuropsychological studies of RITL raised the important question of what specialized processes are implemented in LPFC during RITL. Modern neuroimaging is allowing us to begin addressing this question. For instance, Ruge & Wolfensteller (2010) found that a large portion of LPFC is involved in RITL, but that regions within LPFC differentiate as initially novel stimulus-response rules become practiced. Specifically, they found several anterior LPFC regions (among others) highly active in the initial trials but decreasing with rule repetition, while more posterior LPFC regions (among others) increased their activity with rule repetitions. This suggests an anterior-to-posterior shift in task control with practice.

This observation may indicate an anterior-to-posterior gradient of abstraction (and/or complexity) within LPFC (David Badre & D’Esposito, 2009; Fuster, 2001; Koechlin, Ody, & Kouneiher, 2003). Under this interpretation, task representations begin as highly abstract rules in anterior LPFC and middle LPFC (i.e., dorsolateral PFC), which are converted to concrete representations for implementation by posterior LPFC (i.e., pre-motor cortex) and sensory-motor regions. In the case of Ruge & Wolfensteller (2010) this shift could go back quite far towards the concrete end of the abstraction gradient (into posterior LPFC, primary motor, and sensory cortices), likely because the tasks consisted of highly concrete stimulus-response associations (Figure 3A). In contrast, tasks used by Cole et al. (2009; 2010) and Stocco et al. (2012) were highly abstract. For instance, Cole et al.’s tasks consisted of combinations of three rule types that had to generalize across dozens of stimuli (Figure 3B). Consistent with an anterior-to-posterior gradient of abstraction within LPFC, Cole et al. (2010) found that LPFC activity during task implementation shifted from anterior LPFC to middle LPFC with practice. In other words, there was still an anterior-to-posterior shift in task control with practice, but this shift occurred between more anterior LPFC regions than in Ruge & Wolfensteller (2010), perhaps because the tasks were more abstract in Cole et al. (2010) (Figure 4).

Also consistent with the anterior-to-posterior gradient of abstraction within LPFC, Cole et al. (2010) looked at a finer (within-trial) timescale and found using both magnetoencephalography and functional MRI that activity flowed from middle to anterior LPFC during RITL. This could reflect relatively concrete task information within middle LPFC (e.g., ‘is it sweet?’, ‘are they the same?’, and ‘press left index finger’) being compositionally combined into more abstract/integrative task information within anterior



PFC (e.g., the ‘press your left index finger if both objects are sweet’ task) during novel task preparation in order to coordinate task rules. Importantly, Cole et al. (2010) found that this pattern reversed once a task became practiced: activity flowed from anterior to middle LPFC during practiced task preparation. This suggests that familiarity with a task changes preparation, such that an abstract/integrative task representation (likely recalled from long-term memory) is used to activate and coordinate more concrete representations for task implementation.

### **Subcortical Contributions to RITL**

RITL processes are also present outside LPFC. One potentially important set of locations for RITL-related processes is the basal ganglia. The caudate nucleus, midbrain dopamine nuclei, and other parts of the basal ganglia have been associated with procedural learning and the acquisition of new skills. Single cell recording in primates learning new rule-based tasks, for instance, suggests that caudate responses anticipate responses in LPFC and mediate the acquisition of new behaviors (Pasupathy & Miller, 2005). In humans, Ruge and Wolfensteller (2010) found that caudate activity during novel trials predicts the amount of learning (measured by decreases in reaction time) for subsequent repetitions of the same task.

The basal ganglia (including the striatum and the midbrain dopamine nuclei), however, are also thought to play an important role in controlling/modulating LPFC activity. In particular, it has been suggested that these regions rapidly select and gate the flow of signals from posterior sensory and motor cortical areas to LPFC (Braver & Cohen, 2000; Mcnab & Klingberg, 2008; see R O’Reilly & Frank, 2006 for a detailed computational model of this process; Andrea Stocco et al., 2010). Thus, basal ganglia involvement in RITL may extend beyond a simple associative role in learning. The capacity to rapidly gate information to LPFC becomes particularly useful when novel tasks must be learned and executed in a single trial, as with first-trial RITL. Supporting this view, Stocco et al. (2012) analyzed a first-trial RITL experiment explicitly searching for regions whose activity selectively increased during the execution of novel tasks, and carefully excluding those regions whose involvement could be ascribed to either stimulus novelty or task difficulty alone. In addition to LPFC and posterior parietal cortex, the results identified the basal ganglia as a key contributor to the execution of newly instructed trials.

The key role of another subcortical region – hippocampus – in long-term memory encoding makes it likely to be important for the transition from novel to practiced task performance. This region might also have a more direct role in RITL, however: Some theories have suggested that hippocampus may also be important for working memory of task sets, especially when novel (Hasselmo & Stern, 2006) and when consisting of conjunctions of rules or sensory/motor representations (O’Reilly, Braver, & Cohen, 1999). This seems to suggest a critical role for hippocampus in RITL, yet it is clear that hippocampus is unnecessary for RITL given that hippocampal lesion patients can perform RITL. For instance, patient H.M. was able to use RITL to learn and coordinate the rules of a mirror-tracing task despite an inability to encode those rules in long-term memory (Squire, 2009). Further, it was shown in a large group of lesion patients that those with hippocampal lesions (along with patients with a variety of other lesions) could use RITL to learn and coordinate a complex set of rules for a visual maze task, while only those with LPFC lesions could not (Milner, 1965).

### **The Cognitive Control Network’s Role in RITL**

LPFC is strongly connected with a set of cortical regions sometimes referred to as the cognitive control network (Michael W Cole & Schneider, 2007; N. U. F. Dosenbach et al.,

2006; Duncan, 2010; Wager, Jonides, & Reading, 2004). This network is thought to be composed of the majority of LPFC, anterior cingulate/pre-supplementary motor area, posterior parietal cortex, anterior insula cortex, and sometimes posterior middle temporal cortex. These regions are co-active across a wide variety of studies (Wager et al., 2004) and are more correlated at rest than the whole brain on average (Michael W Cole & Schneider, 2007). Although there are ways to dissociate the regions using task functional MRI (Michael W Cole & Schneider, 2007) and high connectivity thresholds (N. Dosenbach et al., 2007), they are more often co-active and connected with each other than with sensory-motor or “default-mode” regions (Fox et al., 2005). This suggests that if LPFC is central to RITL, then all or most of the cognitive control network may be as well.

In surprising contrast to this argument, RITL functional MRI studies to date indicate that little of the control network is involved in RITL specifically. For instance, when looking relative to resting baseline (see M. Cole, 2009) the entire control network was active during novel and practiced conditions for Cole, Bagic, et al. (2010). However, of the control network regions, only LPFC and posterior parietal cortex were selectively active for RITL relative to practiced tasks. Hartstra et al. (2011) found an even more restricted portion of the control network – left posterior LPFC – when looking for RITL vs. practiced activations. In contrast, Ruge & Wolfensteller (2010) found a change in the entire control network between RITL and practiced task performance. Importantly, however, they found that activity in only LPFC and posterior parietal cortex (along with the caudate) correlated with behavioral improvement between RITL and practiced task performance, suggesting these were especially important regions during the learning process. Note that Dumontheil et al. (2011) found increases in activity for the entire control network for large vs. small instruction sets during RITL, yet it is difficult to interpret this result given that this contrast did not control for short-term memory load (the number of task rules) and several other factors controlled for in the other RITL studies. Together these studies suggest that only a portion of the control network – including LPFC and posterior parietal cortex– involves processes specific to RITL. It will be important for future research to more directly test this conclusion, however, using region-of-interest analysis and statistical dissociations (Henson, 2005). It will also be important to assess the shared and distinct contributions that these two regions make to RITL and other forms of flexible cognitive control.

## INTEGRATIVE THEORETICAL ACCOUNT OF RITL

### A Compositional Theory of Flexible Cognitive Control

It is not clear exactly how the existing cognitive neuroscientific studies of RITL relate to one another. In an attempt to help unify these cognitive neuroscientific observations, we describe a new theoretical model of flexible cognitive control that provides a mechanistic account of RITL. We focus primarily on **core principles** (in bold; Table 1) underlying this theory, with the expectation that future work will make more concrete (i.e., mathematical or computational) implementations of the theory based on these principles.

The key principle of the theory, based primarily on observations by Cole et al. (2011) and Reverberi et al. (2011), is **compositionality** (also see RC O'Reilly, Braver, & Cohen, 1999). This is the ability to reuse representations in concert with a variety of other representations (Figure 5). This ability leads to immense flexibility as it allows for massive combinatorics of possible representational sets. For instance, just 100 concepts can be combined in 4950

possible pairs or 161700 triplets (formula for combinations:  $\frac{n!}{k!(n-k)!}$ ). Healthy adult humans have tens of thousands of concepts ready to be combined (Biederman, 1987), suggesting billions of combinations are possible. An ability to access such a large set of possible

conceptual and procedural configurations is essential for the proposed architecture given the immense variety of possible tasks that healthy humans are capable of learning via RITL.

What possible neural architecture could provide the rapid compositional updating necessary to account for RITL? Some evidence comes from Cole et al. (2011), who found compositional coding within human LPFC during RITL (Figure 5C). They found this by training multivariate classifiers (c.f., Norman, Polyn, Detre, & Haxby, 2006) on LPFC functional MRI activity patterns to identify task rules during practiced task performance, then showing that these classifiers could identify the constituent rules involved during RITL (novel rule combinations). This finding in LPFC is consistent with the unanimous involvement of LPFC in recent RITL functional MRI studies (Michael W Cole, Bagic, et al., 2010; Dumontheil et al., 2011; Hartstra, Kühn, Verguts, & Brass, 2011; Ruge & Wolfensteller, 2010). This finding points to another important principle of the theory (directly related to the principle of compositionality): **immediate transfer** (Figure 6). This concept emphasizes the benefits of prior experience with the constituent task-relevant rules in rapidly learning a new task (Michael W Cole et al., 2011; Kieras & Bovair, 1986; Singley & Anderson, 1989). Here it is the compositional reuse of relevant sets of practiced task features that facilitates RITL.

We suggest two properties of LPFC make it ideal for implementing the rapid compositionality necessary for RITL: 1) **rapid updating** of activity and connectivity patterns due to gating by dopamine and/or other basal ganglia signals (Braver & Cohen, 2000; Mcnab & Klingberg, 2008; R O'Reilly & Frank, 2006; Andrea Stocco et al., 2010; Andrea Stocco et al., 2012), and 2) high **global connectivity** (M W Cole, Anticevic, Repovs, & Barch, 2011; Michael W Cole, Pathak, & Schneider, 2010) resulting in **latent connectivity** (previously unused connections and connection patterns that can quickly come into use when necessary; Figure 6) allowing for a combinatorial explosion of possible active connectivity patterns across a wide range of possible task semantics.

Another important aspect of LPFC is its ability to represent rules and other abstract concepts (Haynes et al., 2007; Wallis, Anderson, & Miller, 2001). **Abstraction** is defined here as the grouping of representations or representational features into categories, allowing for the activation of a subset of representational features of one concept in representing a different concept. For instance, the abstract concept 'circle' is a category defined by a common set of sub-features that have been extracted across many instances of perceiving specific imperfect circles (e.g., uniform roundness). Alternatively, a different sort of abstraction (a "policy abstraction"; David Badre & D'Esposito, 2009) such as 'make coffee' is a category defined by several related sets of action representations, with each set able to lead you to make coffee in a different situation (e.g., one action set might involve grinding coffee beans while another might not). These examples can also be conceptualized in terms of LPFC having broad (categorical) receptive fields (Seger & Miller, 2010). Abstraction is clearly a critical feature of a compositional architecture, given that its definition is nearly identical to that of compositionality itself (i.e., the ability for a representation to be meaningfully applied across multiple situations). Abstractions are further related to compositionality in that abstractions are likely built from the compositional combination of features constituting the range of a given abstraction-representing neuron's receptive field (e.g., all relevant cat-like features for representing "cat").

Note that posterior cortical regions outside LPFC (i.e., in the temporal, parietal, and occipital lobes) build and represent abstractions as well (Kiehl et al., 1999), and that both concrete (Fuster, Bauer, & Jervey, 1985; Pouget, Emeric, Stuphorn, Reis, & Schall, 2005) and abstract (Muhammad, Wallis, & Miller, 2006) representations are projected from posterior cortex to LPFC, such that there is extensive representational redundancy across

these posterior and LPFC systems. Critically, the theory differentiates these systems by characterizing posterior cortex as representing the semantics of the external world, in contrast to LPFC organizing representations in terms of task/goal relevance. This, along with rapid updating and latent connectivity, makes novel task-relevant cognitive configurations (largely unconstrained by established semantics/experience) readily available in LPFC during RITL and other situations requiring flexible cognition, while allowing established semantics of the external world to remain intact within posterior cortex. Due to bidirectional connectivity between LPFC and posterior representations (Fuster et al., 1985), activated representations are distributed across both systems, allowing for simultaneous activation of rich (yet over-constrained) semantics in posterior cortex and flexible (yet under-constrained) representations within LPFC.

Abstraction, compositionality, and transfer are highly related to the concept of **analogy** – “the perception of like relational patterns across different contexts” (Gentner & Colhoun, 2010). As the circle example above illustrates, analogy among multiple instances of imperfect circles can lead to mappings among the common elements between the circles, creating an abstract representation of circle that can generalize to new instances, allowing knowledge learned about circles to transfer to new contexts (i.e., new circles or things similar to circles, e.g., spheres). We suggest that such analogical mapping and resulting abstract representation (Gentner & Medina, 1998) can occur within LPFC (in concert with other brain regions), leading to the ability to immediately transfer knowledge and skills during RITL. Identification of analogical similarity between existing abstract representations and a novel task (e.g., via key words during instruction) is also an important part of this process.

There is evidence that LPFC neuron receptive fields include a variety of nonintuitive variants of task-relevant rules (Jun et al., 2010; Rigotti, Rubin, Wang, & Fusi, 2010), such as “is green” being partially represented by neurons that fire most to “not red”. This suggests that the flexibility necessary for RITL may arise in part from **coarse-coded conjunctive representations** (R. C. O’Reilly, Busby, & Soto, 2003; Rigotti et al., 2010) formed from the combination of various semi-task-relevant abstract representations into task sets. It has been suggested that this kind of representational binding occurs via **synchrony/co-activation** of neurons with relevant receptive fields (Fries, 2005). Another account suggests representational binding occurs via activation of (and feedback from) higher-level conjunction neurons (R. C. O’Reilly & Rudy, 2001). We posit that these two mechanisms of binding – synchrony and conjunction – are in fact complementary mechanisms in which feedforward synchrony can activate higher-level conjunctions and feedback activation of conjunctions can lead to lower-level synchrony in a representational hierarchy (see Table 1), resulting in the binding of representations via both mechanisms. These principles of the theory are similar to those used in a recent computational model of “compositional connectionism” (Hummel et al., 2004). It will be important for future research to verify the exact mechanisms underlying rapid feature binding (sometimes called ‘variable binding’) during RITL, however.

In sum, this theoretical model can be conceptualized as a specific form of domain-general working memory that emphasizes rapid updating, compositionality, and combinatorics of the representations within task sets. Another way to conceptualize the proposed model is as a projection of many posterior cortical representations (i.e., perceptual, motor, semantic, and long-term memories) to LPFC (see Dehaene, Kerszberg, & Changeux, 1998), where billions of combinations of those representations are functionally available for selection and goal-directed sustained processing at a moment’s notice. The theory postulates that having this extra space for representations to combinatorially interact and coalesce provides the human

brain with an immensely flexible architecture capable of such computational feats as first-trial RITL.

### Specific Mechanisms of the Compositional Theory

Of the principles outlined above, rapid updating and global connectivity of LPFC are perhaps the most mechanistic. Building on these mechanisms (Table 2), the theory proposes that RITL starts with a working memory encoding event in which 1) dopamine and/or basal ganglia signals (e.g., from reward prediction) interrupt the current task state and allow **rapid updating** of LPFC representations, and 2) instructions are converted into task semantics via distributed domain-specific semantic representations in posterior cortex that activate sets of equivalent (and/or more abstract/complex) semantics within LPFC via its extensive **multi-system global connectivity** (Michael W Cole, Pathak, & Schneider, 2010; M. W. Cole, Yarkoni, Repovs, Anticevic, & Braver, 2012; Power et al., 2011).

The many sets of abstract/complex representations are also made possible due to extensive **within-LPFC global connectivity** – which has only recently been investigated for the first time (M W Cole, Anticevic, Repovs, & Barch, 2011) – that likely allows for the building of sets of abstract/complex features. This principle of the model is consistent with recent non-human primate work demonstrating immense variability in LPFC single-neuron receptive fields (Jun et al., 2010), given that such observations could reflect the building of abstract/complex representations via extensive within-LPFC connectivity.

There are clear capacity limits on working memory (Conway & Engle, 1996) (and RITL and LPFC by proxy), such that the tremendous combinatorics of possible sets of co-activated features within LPFC may overwhelm the system's limited capacity as the appropriate configuration is being searched for. The theory deals with this by allowing sets of features distributed between LPFC and posterior cortex to be **incrementally selected** and connections among them strengthened over many trials during prior experiences (i.e., 'chunked' via repeated use) (Hebb, 1949; Lynch, 2004) and then **rapidly selected** (and coordinated within LPFC) with a limited number of other features via activation of instruction semantics during RITL. The reactivation of incrementally selected sets of features allows for immediate transfer of previously learned abilities as novel combinations of such feature sets are rapidly selected during RITL. One important example of this process is the incremental selection of associations between words and rule meanings (i.e., selected and strengthened connections from language regions to LPFC), which can then be rapidly selected by incoming linguistic instructions during RITL. This aspect of the theory is consistent with a recent formulation of working memory in which there is a distinction between activated long-term memory (activation of representations that were incrementally selected and/or strengthened/refined during consolidation) and a 'region of direct access' that is able to flexibly select and bind a variety of possible novel representations (Meiran, Cole, & Braver, 2012; Oberauer, 2009).

It is theoretically possible for the building of abstract/complex representations within LPFC to emerge from random connectivity built upon sensory/motor primitives from posterior cortex (Rigotti et al., 2010). There is evidence, however, that there is a posterior-to-anterior hierarchy of processing or representation in LPFC (see previous sections). It is possible that this hierarchy supports efficient building of abstract/complex representations used during RITL. However, there is currently a controversy regarding the exact nature of this LPFC hierarchy (D Badre, 2008; Reynolds, O'Reilly, Cohen, & Braver, 2012). Some studies suggest it is a processing hierarchy organized by time or action (M. M. Botvinick, 2008; Koechlin et al., 2003), while others suggest it is a representational hierarchy organized by abstraction (David Badre & D'Esposito, 2007). We suggest that LPFC builds abstract and complex representations – which become processes when activated (due to downstream

effects of connectivity) – in a **general processing hierarchy**. This avoids the current controversy by subsuming the two camps: gradients of abstraction, complexity, action, and time are all built using conjunctions of random sets (and sets of sets) of sensory/motor primitives (ultimately from primary sensory/motor regions). Consider, for instance, the general processing hierarchy illustrated in Figure 6. Based on several primitives at the bottom (already somewhat built up from primitives in, e.g., V1), both abstractions (e.g., ‘is green’) and complex task representations (e.g., “press the left button when you see red”) are built. Representational hierarchies of time and action are possible due to the existence of temporal and motor primitives (i.e., neurons that fire for particular event timings or motor movements) for building upon in the hierarchy.

The observed anterior-to-posterior LPFC activation shift with practice (see Figure 4) can be accounted for by positing a **hierarchical conservation** principle for the theoretical model. This principle suggests that incremental selection occurring during practice is biased toward selecting and strengthening representations/connections lower in the general processing hierarchy. Thus, while initial RITL rapid selection likely involves both high-level and low-level representations, the involved representations can be whittled down over time by incrementally selecting posterior representations to more efficiently represent the task set. The theory suggests that higher-level representations in anterior LPFC are involved during RITL for two reasons: 1) to allow for transfer via abstract representations in anterior LPFC (see above) that can readily transfer rules across task contexts, and 2) activation of a wide variety of ad hoc coarse-coded representations that can together represent the task rapidly but inefficiently during RITL. With practice, abstract representations (in anterior LPFC) can become less involved as more task-specific (in posterior LPFC) representations/connections become tuned and incrementally selected to perform the task. In the case of an abstract or complex task, the posterior shift cannot go very far, given that anterior representations are necessary to represent such task sets even after connections are selected and tuned (see Figure 4B). In contrast, concrete stimulus-response associations (see Ruge & Wolfensteller, 2010) can become fully automatic with enough practice (Schneider & Shiffrin, 1977), such that they can go all the way down the hierarchy to sensory-motor cortices (J. Chein & Schneider, 2005; Schneider & Chein, 2003). This suggests there may be three learning stages for concrete tasks, based on the state of incrementally selected task representations: 1) RITL, 2) controlled, and 3) automatic (J. M. Chein & Schneider, 2012). The first stage involves instruction interpretation, ad hoc coarse-coded representation, and transfer, while controlled processing involves incremental selection and tuning of representations, eventually resulting in highly efficient automatic processing. It will be important for future research to test these predictions and better characterize the transition between each stage of skill acquisition.

The combinatorial explosion of possible tasks is a major issue for neural theories of RITL, and several principles postulated above may help. To illustrate the issue, consider that a conservative estimate of 10000 concepts available for humans (Biederman, 1987) would result in over 160 billion possible triplets for RITL (see above for combination equation). The human neocortex has only 16 billion neurons (Azevedo et al., 2009), with only a fraction of these within LPFC, such that it is impossible for each conceptual combination to have a dedicated neuron. Coarse-coded conjunctions and synchrony binding (see above) would allow for reuse of neurons across contexts, such that LPFC could represent more combinations than the number of neurons within it, since these mechanisms would allow concepts to be built from sets of reusable sub-features (R. C. O’Reilly et al., 2003). Similarly, the general processing hierarchy can allow for compositional reuse of lower-level concepts via various higher-level representations within the hierarchy. Importantly, these principles help deal with the combinatorial explosion of possible tasks while allowing for efficient compositional transfer of rules during RITL.

## The Compositional Theory vs. The Adaptive Coding Theory

The present compositional theory is compatible with a variety of existing theories, as outlined above. However, the compositional theory appears to be incompatible with the adaptive coding theory (Duncan, 2001). That theory posits that neurons within LPFC are adaptive and change their receptive fields across task contexts, yet the compositional theory requires that receptive fields are relatively rigid to allow for transfer (see Figure 6). The adaptive coding theory is based on the observation of LPFC being active in humans across many task contexts (Duncan & Owen, 2000), and the observation of macaque monkey LPFC neurons representing whatever task rule was used during training (Freedman, Riesenhuber, Poggio, & Miller, 2001). Supporting the compositional theory, however, are the human functional MRI studies considered above that found that constituent rule representations remain stable within LPFC despite changes in task context (see Figure 5).

Also incompatible with the adaptive coding theory, a recent study with macaque monkeys found that different categories are represented in separate LPFC neural populations (Roy, Riesenhuber, Poggio, & Miller, 2010). Importantly, that study used the exact same stimuli for both categories (e.g., a large cat could be categorized using either cat vs. dog or large animal vs. small animal) such that the categories were in conflict. Another recent study showed that non-conflicting categories involving distinct stimuli (e.g., cars for sedan vs. sports car and animals for cat vs. dog) are represented in the same LPFC neurons (Cromer, Roy, & Miller, 2010). This appears to support the adaptive coding theory yet, when considered along with Roy et al. (2010), it is actually compatible with the compositional theory. Specifically, in contrast to the adaptive coding theory, these studies suggest a neuron uses a complex static receptive field of, e.g., “cat OR sedan” to represent both categorical distinctions cat vs. dog and sedan vs. sports car (rather than shift its receptive field depending on the context) when the categories are not in conflict. When the categories are in conflict, however, LPFC uses neurons with non-overlapping static receptive fields to reduce interference.

In other words, it appears that receptive fields are not adaptive so much as complex – such that they appear to be adaptive in certain contexts. Thus, data and theory suggest that representations within LPFC are consistent across contexts, allowing for compositional transfer of LPFC representations between related tasks. More specifically, unlike the adaptive coding theory, the compositional theory suggests that the receptive field properties of LPFC neurons change only slowly, allowing experience-dependent tuning of representations via connection strength changes to incrementally improve task-specific performance, which can then rapidly transfer to new related tasks during RITL.

It may be possible to make the compositional theory compatible with a variant of the adaptive coding theory – **population adaptive coding**. We suggest that individual neurons have relatively static receptive fields (Roy et al., 2010), allowing for transfer, but that LPFC as a whole is highly adaptive (compatible with Duncan & Owen, 2000). The compositional theory suggests this is possible due to the great variety of intermixed receptive fields within LPFC (Jun et al., 2010; Rigotti et al., 2010). Specifically, the great variety of static coarse-coded representations within LPFC can be conceptualized as a large set of “basis functions” that can be rapidly selected such that they together “fit” the task parameters specified by instructions during RITL. The very large number of possible sets of such basis functions allows for highly adaptive population coding within LPFC, while allowing for compositional transfer due to static coding of individual neurons.

It is important to consider that the same compositional coarse coding described above that results in abstract representations would also result in the kinds of complex receptive fields described by Cromer et al. (2010). Random variations in compositional combinations of

representations can result in standard abstractions like “red OR orange” (equivalent to “hot” colors), but they can also result in more counter-intuitive complex representations like “cat OR sedan”. The compositional theory suggests such combinations of unrelated concepts provide two functions in LPFC: 1) Allowing LPFC to represent a wider variety of concepts without increasing the number of neurons, and 2) Providing a mechanism for “far transfer”, in which similarities between seemingly unrelated concepts allow learning in one context to transfer to another. To illustrate, consider the possibility that you recently learned to sell your sedan on a new online marketplace (like eBay) and you now want to use it to sell your cat. With a set of “cat OR sedan” neurons tuned to the online marketplace concepts and procedure you can readily transfer them to allow for RITL when selling your cat.

Consider, however, that it is also very important for LPFC to have neurons with very distinct/orthogonal receptive fields in order to reduce interference when transfer is not possible (e.g., if the online marketplace has different procedures for selling cars and selling animals). There are also other forms of conflict during RITL (i.e., negative transfer) from previous associations. The compositional theory emphasizes selection of the correct activity/connectivity pattern in LPFC for novel task performance, with suppression of previous associations occurring through activation of orthogonal representations. Identifying the specific mechanisms for selecting and adaptively increasing activation of orthogonal representations will be an important area for future RITL (and general cognitive control) research. Two possible mechanisms include 1) conflict detection by medial prefrontal cortex increasing activation of orthogonal representations and/or suppressing non-orthogonal representations (M. Botvinick, Braver, Barch, Carter, & Cohen, 2001; Michael W Cole, Yeung, Freiwald, & Botvinick, 2009), and 2) the activation of LPFC neurons during RITL (as part of the complex set of co-active neurons specifying a given task set) that suppress irrelevant previous associations – possibly via LPFC projections to inhibitory neurons in thalamus (Barbas & Zikopoulos, 2007) and/or gating by basal ganglia (Andrea Stocco et al., 2010).

### Predictions of the Compositional Theory

The compositional theory makes a variety of predictions about neural and behavioral factors that should lead to increased RITL abilities, and increased cognitive flexibility generally (e.g., set-shifting, divergent thinking, fluid intelligence). Our expectation that increased RITL abilities will correspond to increases in general cognitive flexibility comes from the extraordinary speed (one trial) and adaptability (involving complex novel brain configurations) required for RITL – attributes that are shared yet typically taxed less by other forms of flexible cognition. Also supporting this unified view of cognitive flexibility is recent work demonstrating that fluid intelligence (related to RITL; Dumontheil et al., 2011; Duncan, Schramm, Thompson, & Dumontheil, 2012) and creativity – typically considered to be uncorrelated abilities – are actually highly correlated once less noisy ‘latent’ measures are used (Nusbaum & Silvia, 2011; Silvia & Beaty, 2012). The predictions postulated below can be tested in a variety of ways, such as using variability between individuals, groups, species, or cognitive/brain states. This is not an exhaustive list of predictions of the theory, but is rather a general outline of possible predictions. We expect that more explicit computational or mathematical implementations of the theory will make predictions that are more specific and critical for the theory in the future.

One important prediction is that greater global connectivity, both within LPFC and between LPFC and the rest of the brain, should result in greater RITL abilities. Greater multi-system global connectivity would allow LPFC to better access a variety of potentially task-relevant systems, allowing it to receive and influence more task-relevant information during RITL and other situations requiring flexible cognitive control (M. W. Cole et al., 2012). Similarly, the within-LPFC global connectivity prediction is based on the resulting increase in latent



connectivity, which would likely result in greater representational capacity for novel conceptual configurations (as discussed above).

Having more neurons within LPFC (measured, e.g., as LPFC gray matter volume) should also increase representational capacity, resulting in greater RITL abilities. This would increase latent connectivity substantially (exponentially with increases in the number of neurons), yet it would increase representational capacity in other ways as well. Rather than just reflecting the number of possible configurations like latent connectivity, having more neurons can increase the orthogonality of those configurations. The theory predicts that increased orthogonality/distinctness should reduce between-rule interference and allow for activation of multiple rules simultaneously during transfer of practiced rules to novel contexts. It will be important, therefore, to assess if having more LPFC neurons corresponds with greater representational orthogonality within LPFC, and if that in turn corresponds with better RITL abilities. Importantly, there is already evidence supporting this prediction of the theory. Specifically, the theory is compatible with the finding (described above) that macaque monkeys can perform RITL-like behavior (despite having low LPFC representational capacity) only once between-task interference is eliminated (Cromer et al., 2011) (Figure 2A). The theory suggests that humans are better at non-linguistic RITL than monkeys primarily due to greater representational capacity, which reduces between-task interference. This not only reduces catastrophic interference during RITL but also improves transfer of shared concepts between tasks to facilitate RITL further.

Perhaps the most counter-intuitive prediction of the theory is that a higher “learning rate” (the rate at which connection weights change between neurons, changing those neurons’ receptive fields) in LPFC should result in slower task learning via a reduction in RITL abilities. More precisely, the theory predicts that a learning rate that is optimal for most reinforcement learning (or other forms of incremental learning) situations should be higher than the optimal learning rate for RITL. We expect that this prediction can be tested using computational models manipulating learning rates, or using single-unit recording or neuroimaging measuring differences in the rate of receptive field change across individuals. This prediction is based on the theoretical claim that a higher learning rate would result in over-fitting of LPFC connectivity to practiced task contexts, paradoxically reducing the ability to generalize practiced rules to novel contexts. Thus, the compositional theory directly argues against theories of cognitive flexibility that posit fast weight-based learning in LPFC as a key mechanism (e.g., Bugmann, 2011). Note that a higher learning rate might be effective for situations in which prior learning has no relevance, or is incompatible with the to-be-learned task (i.e., negative transfer), but we suggest that such situations are rare once a sufficiently large set of generally relevant abstract rules have been learned.

Another potentially counterintuitive prediction of the theory is that using a rule in a variety of task contexts should improve RITL performance with that rule. One might expect that using a rule in many contexts would increase the number of associations formed with that rule, increasing between-rule interference in novel contexts. The theory, however, predicts that using the rule in many contexts would reduce over-fitting of the rule to any one context, increasing the ability of the rule to generalize to new contexts. The theory’s specific mechanism for this involves strengthening of rule-consistent connectivity and weakening of rule-inconsistent connectivity, such that using the rule in more contexts reduces between-rule connectivity (and thus interference). This leads to the surprising prediction that a paradigm involving a rule in many task contexts would involve only negligible reductions in performance relative to a paradigm with only a few task contexts for the rule (which would promote over-fitting). A related prediction is that using a rule with a variety of other rules should reduce between-rule interference within LPFC during RITL rather than increase it.

This may be one reason that RITL performance was so high (> 90% accuracy) for Cole et al. (2010) despite the use of each rule with many others.

## IMPLICATIONS AND FUTURE DIRECTIONS FOR RITL RESEARCH

### Potential Applications of the Compositional Theory, and RITL Research Generally

Our society relies heavily upon the human capacity for RITL. For instance, instructed learning is the dominant form of learning in scholastic education and professional training, putting those with lower RITL abilities at a disadvantage. The prominence of RITL in everyday life, along with variation in RITL abilities across individuals and groups (e.g., younger vs. older adults), suggests future advances in RITL research will have important practical applications.

We present several predictions of the compositional theory as illustrations of potential future applications of RITL research. In the domain of education, the compositional theory predicts that students should be able to increase RITL abilities by practicing a general strategy of identifying common constituent concepts across multiple tasks, and trying to apply familiar concepts to new problems whenever possible. This prediction is not completely new, as others have emphasized the importance of metaphor, analogy, and self-explanation for transfer in education (Billing, 2007; Gentner, Loewenstein, & Thompson, 2003; VanLehn, Jones, & Chi, 1992; Wormeli, 2009). However, the mechanistic grounding that the compositional theory provides may lead to elaboration and refinement of strategies promoting between-task transfer, as well as other approaches to improve RITL.

Similar to students seeking between-task transfer, the compositional theory predicts that RITL should be enhanced by instructors (or cognitive tutors; Ritter, Anderson, Koedinger, & Corbett, 2007) emphasizing common concepts among tasks. For instance, instructors could label and repeatedly point out ‘deep structure’ common to solving several word problems in a mathematics course. Also, lesson plans should emphasize constituent concepts available to all or most students when teaching new concepts/tasks.

The compositional theory also predicts that there is an optimal set of abstract concepts that can be recombined to allow for RITL of most tasks that anyone is likely to learn in a lifetime. An important application of future RITL research may be to identify these abstract concepts (perhaps through careful analysis of tasks and rule frequency) and teach them to students to facilitate RITL abilities. The compositional theory also predicts that it will be important to practice using the abstractions repeatedly in many unique contexts to strengthen their representational connectivity within LPFC while avoiding over-fitting to a small subset of contexts. One major example of this is mathematics – a set of procedural abstractions already identified as important and taught universally – yet there are likely other sets of important common concepts (even in mathematics) that are not being taught at present.

Even with the availability of optimal strategies and lesson plans, there will always be individual differences in RITL abilities. This puts some individuals or groups at a disadvantage. For instance, older adults have difficulty with RITL relative to young adults, such as when learning to use new technology (Hickman, Rogers, & Fisk, 2007). Similarly, deficits in flexible cognitive control – as indexed by fluid reasoning abilities – are present in a wide variety of mental illnesses (Gale, Batty, Tynelius, Deary, & Rasmussen, 2010; Gottfredson & Saklofske, 2009; Koenen et al., 2009). It will be important to identify the severity of RITL (and general flexible cognition) deficits in each of these groups, and look for ways to alleviate the detrimental effects of these deficits (e.g., on education and employment).

The compositional theory may facilitate the transition from identification to treatment of flexible cognition deficits by suggesting mechanisms of action. For example, the theory predicts that drugs affecting dopamine should affect rapid updating in LPFC, potentially improving RITL abilities during mental illness. The theory also predicts that drugs targeting other neurotransmitters within LPFC (e.g., acetylcholine; Crosson, Kyriazis, & Baxter, 2011) may enhance other LPFC mechanisms supporting RITL, such as orthogonality of representations to facilitate transfer. It should be possible to also enhance transfer during mental illness using cognitive strategies similar to those suggested for education, such as emphasizing between-task similarities during RITL. It will be important for detailed computational implementations of the compositional model to make more nuanced predictions of ways that RITL deficits can be alleviated in a variety of mental illnesses.

### Future Directions for RITL Research

The cognitive neuroscience of learning is currently dominated by reinforcement learning research. However, RITL is a much more powerful form of learning in many instances (see Figure 1), and the basic cognitive and neural mechanisms underlying this ability are in need of further investigation. Uncovering the basic mechanisms of RITL will likely also provide important insights into flexible cognition generally, as rapidly learning a never-performed task is one of the best demonstrations of cognitive flexibility possible.

It will be especially important for future RITL research to investigate the role of RITL in mental illnesses. This need arises not just from scientific curiosity, but also from the increasing debilitation of various mental diseases likely affecting RITL (and thus the ability to rapidly adapt) as technological innovation increases the rate of change in the world. It is currently unclear exactly which mental illnesses affect RITL abilities. However, the observation that LPFC lesions decimate RITL abilities (Luria, 1973) and that diseases such as schizophrenia (which involve LPFC disruption; Barch et al., 2001) limit the ease of cognitive task learning (Barch, Braver, Carter, Poldrack, & Robbins, 2009; Young & Freyslinger, 1995) suggest RITL is affected in a variety of mental illnesses. The link between general fluid intelligence and RITL (Dumontheil et al., 2011; Duncan et al., 2012) – and the widespread association of mental illness with impaired fluid intelligence (Koenen et al., 2009) – further suggests there are widespread RITL deficits. Research into whether and which mental illnesses involve RITL deficits could yield important new insights into the nature of those deficits, especially with regard to cognitive flexibility. Furthermore, the mechanisms by which RITL is impaired may differ across mental illnesses, mandating different therapeutic strategies to improve RITL for different mental diseases.

Recent innovations in RITL research promise a bevy of new insights regarding human learning and intelligence. Cognitive paradigm designs that permute rule combinations to investigate novel relative to practiced tasks appear especially promising for isolating RITL processes from stimulus novelty and general task switching processes (Michael W Cole, Bagic, et al., 2010; Andrea Stocco et al., 2012). It will be important for future research to investigate how such rule-combination approaches – which lend themselves to investigating ‘abstract’ task learning – differ from ‘concrete’ stimulus-response rule learning (see Figure 3). It will be especially important for such research to differentiate RITL-specific processes from stimulus novelty and general task switching processes in these ‘concrete’ rule-learning paradigms, in addition to exploring the possibility of greater between-task interference in ‘abstract’ relative to ‘concrete’ paradigms.

RITL research provides unprecedented access to the kind of cognitive flexibility that makes human cognition unique. Unlike studies utilizing planning, problem solving, or reinforcement learning, RITL paradigms directly reference novel mental states (rather than forcing pseudo-random exploration) and so are typically better controlled by the

experimenter. This increased control allows for more precise and efficient experimental paradigms. For instance, Cole et al. (2010) were able to investigate cognitive flexibility across 64 tasks while carefully controlling for extraneous factors. We expect that this combination of increased experimental control and rapid access to a virtually infinite variety of possible mental configurations will lead to new insights into the impressive human capacity for flexible cognitive control.

RITL is something we encounter every day, yet we understand surprisingly little about it. Beyond providing an understanding of this basic ability, we suggest the RITL framework provides unprecedented access to an even more fundamental cognitive ability: flexibly controlling cognition and behavior according to task demands. We expect the emergence of new insights into such flexible cognitive control as the RITL framework, and the compositional theory presented here, are developed further and applied in new research contexts.

## Acknowledgments

We would especially like to thank Todd Braver, as well as Jeff Zacks, Alan Anticevic, Grega Repovs, and Jeremy Reynolds, for invaluable feedback and suggestions during preparation of this manuscript. This work was supported by National Institutes of Health grant MH096801.

## References

- Anderson, JR. Language, memory, and thought. Lawrence Erlbaum; 1976.
- Avants B, Schoenemann P, Gee J. Lagrangian frame diffeomorphic image registration: Morphometric comparison of human and chimpanzee cortex. *Medical Image Analysis*. 2006; 10(3):397–412.10.1016/j.media.2005.03.005 [PubMed: 15948659]
- Azevedo FAC, Carvalho LRB, Grinberg LT, Farfel JM, Ferretti REL, Leite REP, Herculano-Houzel S. Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *The Journal of Comparative Neurology*. 2009; 513(5):532–541.10.1002/cne.21974 [PubMed: 19226510]
- Badre D. Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends in Cognitive Sciences*. 2008; 12(5):193–200.10.1016/j.tics.2008.02.004 [PubMed: 18403252]
- Badre D, D'Esposito M. Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex. *Journal of Cognitive Neuroscience*. 2007; 19(12):2082–2099.10.1162/jocn.2007.19.12.2082 [PubMed: 17892391]
- Badre D, D'Esposito M. Is the rostro-caudal axis of the frontal lobe hierarchical? [Review]. *Nature Reviews Neuroscience*. 2009; 10(9):659–669.10.1038/nrn2667
- Barbas H, Zikopoulos B. The Prefrontal Cortex and Flexible Behavior. *The Neuroscientist: a review journal bringing neurobiology, neurology and psychiatry*. 2007; 13(5):532–545.10.1177/1073858407301369
- Barch DM, Braver TS, Carter CS, Poldrack RA, Robbins TW. CNTRICS Final Task Selection: Executive Control. *Schizophrenia Bulletin*. 2009; 35(1):115–135.10.1093/schbul/sbn154 [PubMed: 19011235]
- Barch DM, Carter CS, Braver TS, Sabb FW, Macdonald A, Noll DC, Cohen JD. Selective deficits in prefrontal cortex function in medication-naive patients with schizophrenia. *Archives of general psychiatry*. 2001; 58(3):280–288. [PubMed: 11231835]
- Biederman I. Recognition-by-components: a theory of human image understanding. *Psychological Review*. 1987; 94(2):115. [PubMed: 3575582]
- Billing D. Teaching for transfer of core/key skills in higher education: Cognitive skills. *Higher Education*. 2007; 53(4):483–516.10.1007/s10734-005-5628-5
- Botvinick M, Braver T, Barch D, Carter C, Cohen J. Conflict monitoring and cognitive control. *Psychological Review*. 2001; 108(3):624–652. [PubMed: 11488380]

- Botvinick MM. Hierarchical models of behavior and prefrontal function. *Trends in Cognitive Sciences*. 2008; 12(5):201–208.10.1016/j.tics.2008.02.009 [PubMed: 18420448]
- Braver TS, Cohen JD. On the control of control: The role of dopamine in regulating prefrontal function and working memory. *Control of cognitive processes: Attention and performance*. 2000; XVIII: 713–737.
- Bugmann G. Modeling fast stimulus–response association learning along the occipito-parieto-frontal pathway following rule instructions. *Brain Research*. 2011:1–17.10.1016/j.brainres.2011.09.028
- Byrne RW, Russon AE. Learning by imitation: a hierarchical approach. [Review]. *The Behavioral and brain sciences*. 1998; 21(5):667–684. discussion 684–721. [PubMed: 10097023]
- Chein J, Schneider W. Neuroimaging studies of practice-related change: fMRI and meta-analytic evidence of a domain-general control network for learning. *Brain Res Cogn Brain Res*. 2005; 25(3):607–623. [PubMed: 16242923]
- Chein JM, Schneider W. The Brain's Learning and Control Architecture. *Current Directions in Psychological Science*. 2012; 21(2):78–84.10.1177/0963721411434977
- Cohen-Kdoshay O, Meiran N. The Representation of Instructions Operates Like a Prepared Reflex. *Experimental Psychology (formerly "Zeitschrift für Experimentelle Psychologie")*. 2009; 56(2): 128–133.10.1027/1618-3169.56.2.128
- Cole, M. Dissertation. 2009. The Biological Basis of Rapid Instructed Task Learning.
- Cole MW, Anticevic A, Repovs G, Barch D. Variable Global Dysconnectivity and Individual Differences in Schizophrenia. *BPS*. 2011; 70(1):43–50.10.1016/j.biopsycho.2011.02.010
- Cole MW, Bagic A, Kass R, Schneider W. Prefrontal dynamics underlying rapid instructed task learning reverse with practice. *J Neurosci*. 2010; 30(42):14245–14254.10.1523/JNEUROSCI.1662-10.2010 [PubMed: 20962245]
- Cole MW, Braver TS. Switching Between Novel Tasks: Evidence for a Distinct Task Set Formation Process. Submitted.
- Cole MW, Etzel JA, Zacks JM, Schneider W, Braver TS. Rapid transfer of abstract rules to novel contexts in human lateral prefrontal cortex. *Frontiers in Human Neuroscience*. 2011; 5(142)10.3389/fnhum.2011.00142
- Cole MW, Pathak S, Schneider W. Identifying the brain's most globally connected regions. *NeuroImage*. 2010; 49(4):3132–3148.10.1016/j.neuroimage.2009.11.001 [PubMed: 19909818]
- Cole MW, Schneider W. The cognitive control network: Integrated cortical regions with dissociable functions. *Neuroimage*. 2007; 37(1):343–360.10.1016/j.neuroimage.2007.03.071 [PubMed: 17553704]
- Cole MW, Yarkoni T, Repovs G, Anticevic A, Braver TS. Global Connectivity of Prefrontal Cortex Predicts Cognitive Control and Intelligence. *Journal of Neuroscience*. 2012; 32(26):8988–8999.10.1523/Jneurosci.0536-12.2012 [PubMed: 22745498]
- Cole MW, Yeung N, Freiwald WA, Botvinick M. Cingulate cortex: diverging data from humans and monkeys. *Trends in Neurosciences*. 2009; 32(11):566–574.10.1016/j.tins.2009.07.001 [PubMed: 19781794]
- Conway AR, Engle RW. Individual differences in working memory capacity: more evidence for a general capacity theory. [Clinical Trial]. *Memory (Hove, England)*. 1996; 4(6):577–590.10.1080/741940997
- Cromer JA, Machon M, Miller EK. Rapid association learning in the primate prefrontal cortex in the absence of behavioral reversals. *Journal of Cognitive Neuroscience*. 2011; 23(7):1823–1828.10.1162/jocn.2010.21555 [PubMed: 20666598]
- Cromer JA, Roy JE, Miller EK. Representation of Multiple, Independent Categories in the Primate Prefrontal Cortex. *Neuron*. 2010; 66(5):796–807.10.1016/j.neuron.2010.05.005 [PubMed: 20547135]
- Crosson PL, Kyriazis DA, Baxter MG. Cholinergic modulation of a specific memory function of prefrontal cortex. *Nature neuroscience*. 2011; 14(12):1510–1512.10.1038/nn.2971
- Dehaene S, Kerszberg M, Changeux JP. A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences of the United States of America*. 1998; 95(24):14529–14534. [PubMed: 9826734]

- Doll BB, Jacobs WJ, Sanfey AG, Frank MJ. Instructional control of reinforcement learning: A behavioral and neurocomputational investigation. *Brain Research*. 2009; 1299(C):74–94.10.1016/j.brainres.2009.07.007 [PubMed: 19595993]
- Dosenbach N, Fair D, Miezin F, Cohen A, Wenger K, Dosenbach R, Raichle M. Distinct brain networks for adaptive and stable task control in humans. *Proceedings of the National Academy of Sciences*. 2007; 104(26):11073–11078.10.1073/pnas.0704320104
- Dosenbach NUF, Visscher KM, Palmer ED, Miezin FM, Wenger KK, Kang HC, Petersen SE. A core system for the implementation of task sets. *Neuron*. 2006; 50(5):799–812.10.1016/j.neuron.2006.04.031 [PubMed: 16731517]
- Dumontheil I, Thompson R, Duncan J. Assembly and use of new task rules in fronto-parietal cortex. *Journal of Cognitive Neuroscience*. 2011; 23(1):168–182.10.1162/jocn.2010.21439 [PubMed: 20146600]
- Duncan J. An adaptive coding model of neural function in prefrontal cortex. [Review]. *Nature Reviews Neuroscience*. 2001; 2(11):820–829.10.1038/35097575
- Duncan J. The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. *Trends in Cognitive Sciences*. 2010; 14(4):172–179.10.1016/j.tics.2010.01.004 [PubMed: 20171926]
- Duncan J, Burgess P, Emslie H. Fluid intelligence after frontal lobe lesions. *Neuropsychologia*. 1995; 33(3):261–268. [PubMed: 7791994]
- Duncan J, Owen A. Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in neurosciences*. 2000; 23(10):475–483. [PubMed: 11006464]
- Duncan J, Schramm M, Thompson R, Dumontheil I. Task rules, working memory, and fluid intelligence. *Psychonomic bulletin & review*. 2012.10.3758/s13423-012-0225-y
- Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. [Comparative Study]. *Proceedings of the National Academy of Sciences of the United States of America*. 2005; 102(27):9673–9678.10.1073/pnas.0504136102 [PubMed: 15976020]
- Freedman DJ, Riesenhuber M, Poggio T, Miller EK. Categorical representation of visual stimuli in the primate prefrontal cortex. *Science (New York, NY)*. 2001; 291(5502):312–316.10.1126/science.291.5502.312
- Fries P. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*. 2005; 9(10):474–480.10.1016/j.tics.2005.08.011 [PubMed: 16150631]
- Fuster J. The prefrontal cortex--an update: time is of the essence. *Neuron*. 2001; 30(2):319–333. [PubMed: 11394996]
- Fuster J, Bauer R, Jervey J. Functional interactions between inferotemporal and prefrontal cortex in a cognitive task. *Brain Research*. 1985; 330(2):299–307.10.1016/0006-8993(85)90689-4 [PubMed: 3986545]
- Gale CR, Batty GD, Tynelius P, Deary IJ, Rasmussen F. Intelligence in early adulthood and subsequent hospitalization for mental disorders. *Epidemiology*. 2010; 21(1):70–77.10.1097/EDE.0b013e3181c17da8 [PubMed: 19907333]
- Gentner D, Colhoun J. Analogical processes in human thinking and learning. Towards a theory of thinking. 2010:35–48.10.1007/978-3-642-03129-8\_3
- Gentner D, Loewenstein J, Thompson L. Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology*. 2003; 95(2):393.10.1037/0022-0663.95.2.393
- Gentner D, Medina J. Similarity and the development of rules. [Review]. *Cognition*. 1998; 65(2–3):263–297. [PubMed: 9557385]
- Gottfredson L, Saklofske DH. Intelligence: Foundations and issues in assessment. *Canadian Psychology/Psychologie canadienne*. 2009; 50(3):183–195.10.1037/a0016641
- Gray E, Tall D. Abstraction as a natural process of mental compression. *Mathematics Education Research Journal*. 2007
- Hartstra E, Kühn S, Verguts T, Brass M. The implementation of verbal instructions: An fMRI study. *Human Brain Mapping*. 2011; 32(11):1811–1824.10.1002/hbm.21152 [PubMed: 21140434]

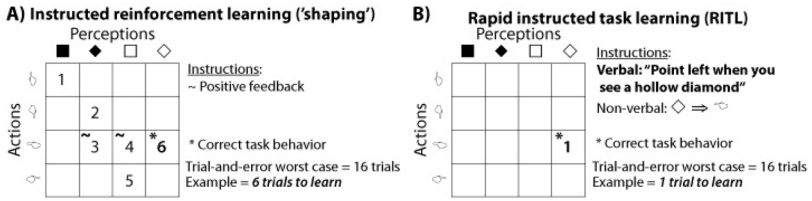
- Hasselmo ME, Stern CE. Mechanisms underlying working memory for novel information. *Trends in Cognitive Sciences*. 2006; 10(11):487–493.10.1016/j.tics.2006.09.005 [PubMed: 17015030]
- Haynes J-D, Sakai K, Rees G, Gilbert S, Frith C, Passingham RE. Reading hidden intentions in the human brain. [Comparative Study]. *Current biology: CB*. 2007; 17(4):323–328.10.1016/j.cub.2006.11.072 [PubMed: 17291759]
- Hebb, DO. *The organization of behavior: A neuropsychological theory*. Lawrence Erlbaum; 1949.
- Henson R. What can functional neuroimaging tell the experimental psychologist? *The Quarterly Journal of Experimental Psychology Section A*. 2005; 58(2):193–233.
- Heyes C. Causes and consequences of imitation. *Trends in Cognitive Sciences*. 2001; 5(6):253–261. [PubMed: 11390296]
- Hickman JM, Rogers WA, Fisk AD. Training older adults to use new technology. [Comparative Study]. *The journals of gerontology Series B, Psychological sciences and social sciences*. 2007; 62(Spec No 1):77–84.
- Holmes, N. *Wordless Diagrams*. Bloomsbury USA: 2005.
- Hummel, JE.; Holyoak, KJ.; Green, C.; Doumas, LAA.; Devnich, D.; Kittur, A.; Kalar, DJ. A solution to the binding problem for compositional connectionism. In: Levy, SD.; Gayler, R., editors. *Compositional connectionism in cognitive science: Papers from the AAAI Fall Symposium*. 2004. p. 31-34.
- Jun JK, Miller P, Hernandez A, Zainos A, Lemus L, Brody CD, Romo R. Heterogenous Population Coding of a Short-Term Memory and Decision Task. *Journal of Neuroscience*. 2010; 30(3):916–929.10.1523/JNEUROSCI.2062-09.2010 [PubMed: 20089900]
- Kiehl K, Liddle P, Smith A, Mendrek A, Forster B, Hare R. Neural pathways involved in the processing of concrete and abstract words. *Human Brain Mapping*. 1999; 7(4):225–233. [PubMed: 10408766]
- Kieras D, Bovair S. The acquisition of procedures from text: A production-system analysis of transfer of training. *Journal of Memory and Language*. 1986; 25(5):507–524.
- Koechlin E, Ody C, Kouneiher F. *The Architecture of Cognitive Control in the Human Prefrontal Cortex*. Science (New York, NY). 2003
- Koenen KC, Moffitt TE, Roberts AL, Martin LT, Kubzansky L, Harrington H, Caspi A. Childhood IQ and Adult Mental Disorders: A Test of the Cognitive Reserve Hypothesis. *The American journal of psychiatry*. 2009; 166(1):50–57.10.1176/appi.ajp.2008.08030343 [PubMed: 19047325]
- Lebiere, C.; Anderson, JR. A connectionist implementation of the ACT-R production system. *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society*; 1993. p. 635-640.
- Luria A. The frontal lobes and the regulation of behavior. *Psychophysiology of the frontal lobes*. 1973:3–28.
- Luria A, Pribram K, Homskaya E. An experimental analysis of the behavioral disturbance produced by a left frontal arachnoidal endothelioma (meningioma). *Neuropsychologia*. 1964; 2(4):257–280.
- Lynch M. Long-term potentiation and memory. *Physiological reviews*. 2004; 84(1):87–136. [PubMed: 14715912]
- Mayr U, Kliegl R. Task-set switching and long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2000; 26(5):1124–1140.
- Mcnab F, Klingberg T. Prefrontal cortex and basal ganglia control access to working memory. *Nature neuroscience*. 2008; 11(1):103–107.10.1038/nn2024
- Meiran N, Cole MW, Braver TS. When Planning Results in Loss of Control: Intention-Based Reflexivity and Working-Memory. *Front Hum Neurosci*. 2012; 6:104.10.3389/fnhum.2012.00104 [PubMed: 22586382]
- Miller E, Cohen J. An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*. 2001; 24:167–202.
- Milner, B. Some effects of frontal lobectomy in man. In: Warren, JM.; Akert, K., editors. *The frontal granular cortex and behavior*. New York: McGraw Hill; 1964.
- Milner B. Visually-guided maze learning in man: effects of bilateral hippocampal, bilateral frontal, and unilateral cerebral lesions. *Neuropsychologia*. 1965; 3(4):317–338.

- Monsell S. Control of mental processes. *Unsolved mysteries of the mind: Tutorial essays in cognition*. 1996:93–148.
- Monsell S. Task switching. *Trends in Cognitive Sciences*. 2003
- Muhammad R, Wallis JD, Miller EK. A comparison of abstract rules in the prefrontal cortex, premotor cortex, inferior temporal cortex, and striatum. *Journal of Cognitive Neuroscience*. 2006; 18(6): 974–989. [PubMed: 16839304]
- Newell, A.; Simon, HA. *Human problem solving*. Prentice-Hall Englewood Cliffs; NJ: 1972.
- Noelle, D.; Cottrell, G. Modeling interference effects in instructed category learning. *Proceedings of the 18th Annual Conference of the Cognitive Science Society*; 1996. p. 475-480.
- Norman K, Polyn S, Detre G, Haxby J. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*. 2006; 10(9):424–430.10.1016/j.tics.2006.07.005 [PubMed: 16899397]
- Nusbaum EC, Silvia PJ. Are intelligence and creativity really so different? *Intelligence*. 2011; 39(1): 36–45.10.1016/j.intell.2010.11.002
- O'Reilly R, Braver T, Cohen J. A biologically based computational model of working memory. *Models of working memory: mechanisms of active maintenance and executive control*. 1999:375–411.
- O'Reilly R, Frank M. Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. *Neural computation*. 2006; 18(2):283–328. [PubMed: 16378516]
- O'Reilly RC, Busby RS, Soto R. Three forms of binding and their neural substrates: Alternatives to temporal synchrony. *The unity of consciousness: Binding, integration, and dissociation*. 2003:168–192.
- O'Reilly RC, Rudy JW. Conjunctive representations in learning and memory: principles of cortical and hippocampal function. [Review]. *Psychological Review*. 2001; 108(2):311–345. [PubMed: 11381832]
- O'Reilly R, Braver T, Cohen J. A biologically based computational model of working memory. *Models of working memory: mechanisms of active maintenance and executive control*. 1999:375–411.
- Oberauer K. Design for a working memory. *Psychology of learning and motivation*. 2009
- Pasupathy A, Miller EK. Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature*. 2005; 433(7028):873–876.10.1038/nature03287 [PubMed: 15729344]
- Pouget P, Emeric EE, Stuphorn V, Reis K, Schall JD. Chronometry of visual responses in frontal eye field, supplementary eye field, and anterior cingulate cortex. [Comparative Study]. *Journal of Neurophysiology*. 2005; 94(3):2086–2092.10.1152/jn.01097.2004 [PubMed: 15944228]
- Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA, Petersen SE. Functional Network Organization of the Human Brain. *Neuron*. 2011; 72(4):665–678.10.1016/j.neuron.2011.09.006 [PubMed: 22099467]
- Rabbitt P. *Methodology of Frontal and Executive Function*. Psychology Press. 1997
- Ramamoorthy A, Verguts T. Word and deed: A computational model of instruction following. *Brain Research*. 2012:1–12.10.1016/j.brainres.2011.12.025
- Rendell L, Boyd R, Cownden D, Enquist M, Eriksson K, Feldman MW, Laland KN. Why Copy Others? Insights from the Social Learning Strategies Tournament. *Science (New York, NY)*. 2010; 328(5975):208–213.10.1126/science.1184719
- Rendell L, Fogarty L, Hoppitt WJE, Morgan TJH, Webster MM, Laland KN. Cognitive culture: theoretical and empirical insights into social learning strategies. *Trends in Cognitive Sciences*. 2011; 15(2):68–76.10.1016/j.tics.2010.12.002 [PubMed: 21215677]
- Reverberi C, Görgen K, Haynes J-D. Compositionality of Rule Representations in Human Prefrontal Cortex. *Cerebral Cortex*. 201110.1093/cercor/bhr200
- Reynolds JR, O'Reilly RC, Cohen JD, Braver TS. The Function and Organization of Lateral Prefrontal Cortex: A Test of Competing Hypotheses. *PLoS ONE*. 2012; 7(2):e30284.10.1371/journal.pone.0030284.t002 [PubMed: 22355309]



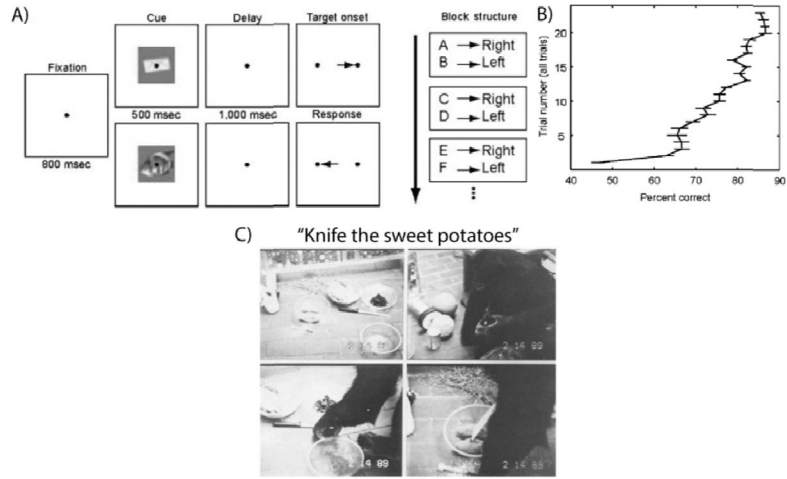
- Rigotti M, Rubin DBD, Wang X-J, Fusi S. Internal representation of task rules by recurrent dynamics: the importance of the diversity of neural responses. *Frontiers in Computational Neuroscience*. 2010; 4:24.10.3389/fncom.2010.00024 [PubMed: 21048899]
- Ritter S, Anderson JR, Koedinger KR, Corbett A. Cognitive Tutor: Applied research in mathematics education. *Psychonomic bulletin & review*. 2007; 14(2):249–255. [PubMed: 17694909]
- Roy JE, Riesenhuber M, Poggio T, Miller EK. Prefrontal Cortex Activity during Flexible Categorization. *Journal of Neuroscience*. 2010; 30(25):8519–8528.10.1523/JNEUROSCI.4837-09.2010 [PubMed: 20573899]
- Rubin O, Meiran N. On the Origins of the Task Mixing Cost in the Cuing Task-Switching Paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2005; 31(6):1477–1491.10.1037/0278-7393.31.6.1477
- Ruge H, Jamadar S, Zimmermann U, Karayanidis F. The many faces of preparatory control in task switching: Reviewing a decade of fMRI research. *Human Brain Mapping*. 2011:n/a–n/a.10.1002/hbm.21420
- Ruge H, Wolfensteller U. Rapid formation of pragmatic rule representations in the human brain during instruction-based learning. *Cereb Cortex*. 2010; 20(7):1656–1667.10.1093/cercor/bhp228 [PubMed: 19889712]
- Rumelhart, DE.; McClelland, J. *Parallel Distributed Processing*. Vol. 1. Bradford: 1987.
- Savage-Rumbaugh E, Murphy J, Sevcik R, Brakke K, Williams S, Rumbaugh D, Bates E. *Language Comprehension in Ape and Child*. Monographs of the Society for Research in Child Development. 1993; 58(3/4):1–222. [PubMed: 8366872]
- Schneider W, Chein J. Controlled & automatic processing: behavior, theory, and biological mechanisms. *Cognitive Science*. 2003; 27(3):525–559.
- Schneider, W.; Oliver, W. *An Instructable Connectionist/Control Architecture: Using Rule-Based Instructions to Accomplish Connectionist Learning in a Human Time Scale*. Architectures for intelligence: the twenty-second Carnegie Mellon Symposium on Cognition; 1991. p. 113
- Schneider W, Shiffrin R. Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*. 1977; (84):1–66.
- Seger CA, Miller EK. Category Learning in the Brain. *Annual Review of Neuroscience*. 2010; 33(1): 203–219.10.1146/annurev.neuro.051508.135546
- Semendeferi K, Armstrong E, Schleicher A, Zilles K, Van Hoesen GW. Prefrontal cortex in humans and apes: a comparative study of area 10. *American Journal of Physical Anthropology*. 2001; 114(3):224–241.10.1002/1096-8644(200103)114:3<224::AID-AJPA1022>3.0.CO;2-I [PubMed: 11241188]
- Silvia PJ, Beaty RE. Making creative metaphors: The importance of fluid intelligence for creative thought. *Intelligence*. 2012; 40(4):343–351.10.1016/j.intell.2012.02.005
- Singley, MK.; Anderson, JR. *The transfer of cognitive skill*. Vol. 9. Harvard University Press; 1989.
- Squire LR. The legacy of patient HM for neuroscience. *Neuron*. 2009; 61(1):6–9. [PubMed: 19146808]
- Stocco A, Lebiere C, Anderson JR. Conditional routing of information to the cortex: A model of the basal ganglia's role in cognitive coordination. *Psychological Review*. 2010; 117(2):541–574.10.1037/a0019077 [PubMed: 20438237]
- Stocco A, Lebiere C, O'Reilly R. The Role of the Basal Ganglia–Anterior Prefrontal Circuit as a Biological Instruction Interpreter. *Biologically Inspired Cognitive Architectures 2010. Frontiers in Artificial Intelligence and Applications*. 2010
- Stocco A, Lebiere C, O'Reilly RC, Anderson JR. Distinct contributions of the caudate nucleus, rostral prefrontal cortex, and parietal cortex to the execution of instructed tasks. *Cognitive, affective, & behavioral neuroscience*. 2012.10.3758/s13415-012-0117-7
- Sutton, R.; Barto, A. *Reinforcement Learning: an introduction*. MIT Press; 1998.
- Turing A. On computable numbers, with an application to the Entscheidungsproblem. *Proc Lond Math Soc*. 1936; 42(2):230–265.
- Turing A. Computing machinery and intelligence. *Mind*. 1950

- VanLehn K, Jones RM, Chi MTH. A model of the self-explanation effect. *The Journal of the Learning Sciences*. 1992; 2(1):1–59.
- Verrico CD, Liu S, Asafu-Adjei JK, Sampson AR, Bradberry CW, Lewis DA. Acquisition and baseline performance of working memory tasks by adolescent rhesus monkeys. *Brain Research*. 2011; 1378(C):91–104.10.1016/j.brainres.2010.12.081 [PubMed: 21215729]
- Wager T, Jonides J, Reading S. Neuroimaging studies of shifting attention: a meta-analysis. *NeuroImage*. 2004; 22(4):1679–1693. [PubMed: 15275924]
- Wallis JD, Anderson KC, Miller EK. Single neurons in prefrontal cortex encode abstract rules. *Nature*. 2001; 411(6840):953–956.10.1038/35082081 [PubMed: 11418860]
- Wormeli, R. *Metaphors & Analogies: Power Tools for Teaching Any Subject*. Stenhouse Pub; 2009.
- Yeung N, Monsell S. The effects of recent practice on task switching. [Clinical Trial]. *Journal of experimental psychology Human perception and performance*. 2003; 29(5):919–936.10.1037/0096-1523.29.5.919 [PubMed: 14585014]
- Young DA, Freyslinger MG. Scaffolded instruction and the remediation of Wisconsin Card Sorting Test deficits in chronic schizophrenia. [Clinical Trial]. *Schizophrenia Research*. 1995; 16(3): 199–207. [PubMed: 7488565]
- Zylberberg A, Dehaene S, Roelfsema PR, Sigman M. The human Turing machine: a neural framework for mental programs. *Trends in Cognitive Sciences*. 2011; 15(7):293–300.10.1016/j.tics.2011.05.007 [PubMed: 21696998]



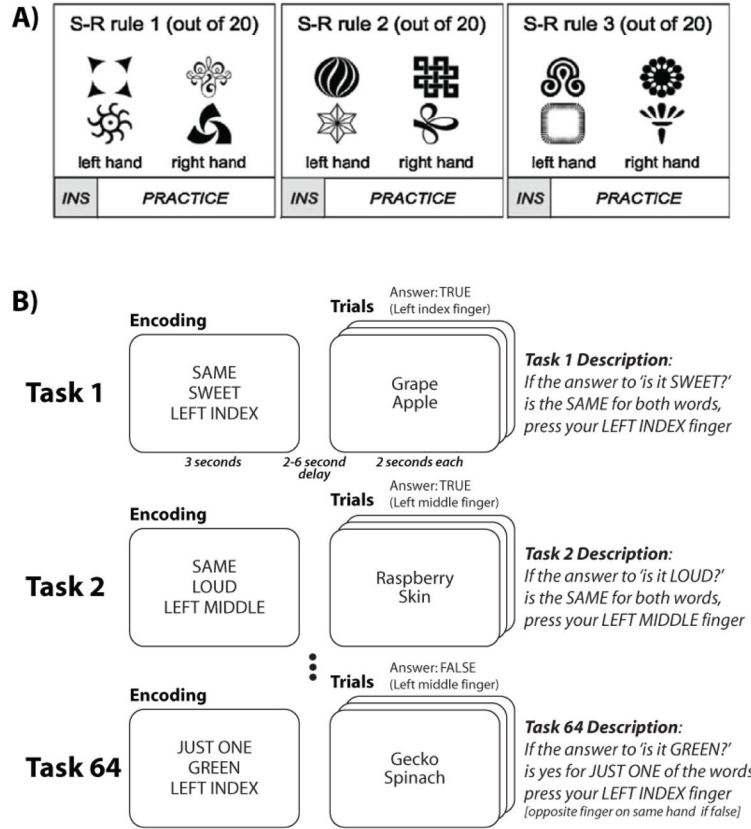
**Figure 1. Task learning as search through space of possible tasks**

Task learning can be conceptualized – in a manner similar to problem solving (Newell & Simon, 1972) – as search through a state space of possible tasks. The speed of learning depends on the ability to efficiently traverse this space. The power of RITL lies in the ability for the instructor to directly specify the appropriate goal state (or one nearby) using examples and/or linguistic instruction. **A)** Typical studies of learning in psychology and neuroscience have focused on unsupervised reinforcement learning, which is akin to trial-and-error learning with reinforcement at the goal state. ‘Shaping’ is a form of supervised reinforcement learning that can speed learning substantially (though not as much as RITL) using rewards to coax the learner closer to the goal. **B)** RITL involves directly referring to the goal state (either verbally or non-verbally), such that the learner can immediately execute the instructed task.



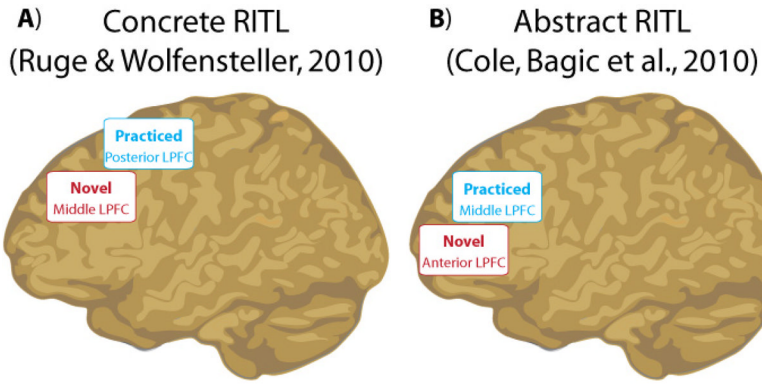
**Figure 2. RITL is present but limited in non-human primates**

**A)** Cromer et al. (2011) recently showed that macaque monkeys could perform something similar to concrete RITL. The figure illustrates two stimulus-response (object-saccade) associations the monkeys had to learn across multiple trials. They were able to rapidly learn these and other new stimulus-response associations via feedback. However, they were only able to do so when there was no interference between stimulus-response associations (i.e., stimuli were never reused). Note that it is unclear if this is truly RITL since there were no instructions given before presenting the first stimulus of each task, yet it suggests monkeys may be capable of concrete RITL. **B)** Further, unlike humans, macaque monkeys took 20 trials to reach 90% accuracy. Humans were at 90% on the first trial in Ruge & Wolfensteller (2010) and Cole et al. (2010). Also, note that the macaques required months of training on the ‘meta-task’ specifying timing, kinds of stimuli, etc., while this took seconds to minutes with humans for Cole et al. (2010; 2011). **C)** In contrast to macaque monkeys, a bonobo chimpanzee named Kanzi was able to perform first-trial RITL at 70% accuracy using verbal instructions (Savage-Rumbaugh et al., 1993), suggesting evolutionary changes toward RITL abilities developed gradually on the path to the evolution of *homo sapiens sapiens*.



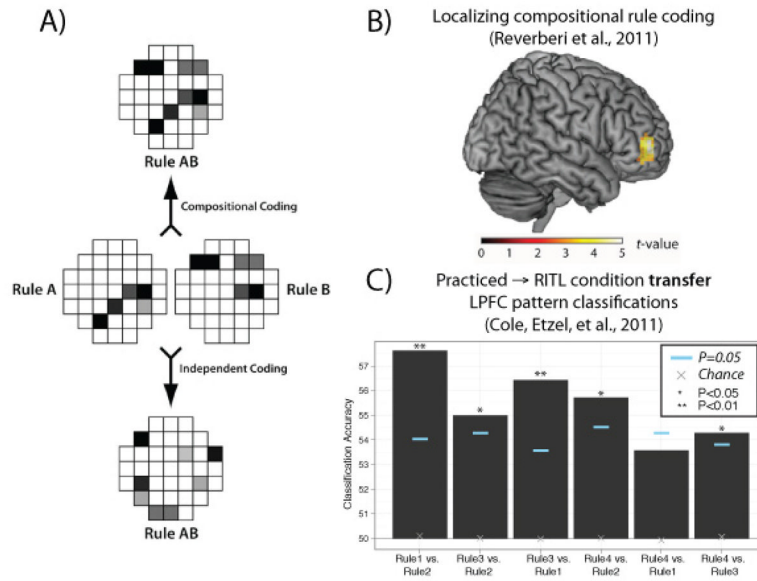
**Figure 3. ‘Concrete’ and ‘abstract’ approaches to investigating RITL**

Two major approaches have been developed for investigating RITL. Both overcome the inherent difficulty of investigating task novelty when a single task repetition invalidates a task’s novelty. **A)** Ruge & Wolfensteller (2010) achieved repeated task novelty using a large set of unique stimuli (with a small set of responses), making it possible to compositionally build a large set of novel stimulus-response associations. Hartstra et al. (2011) used a similar method using object stimuli, and also object-color pairings. These approaches rely on concrete stimulus-response pairings rather than more abstract concepts, and thus constitute what might be called ‘concrete RITL’. Note that the Ruge & Wolfensteller (2010) paradigm could not differentiate responses due to stimulus novelty and/or task switching from RITL (though their analysis correlating instruction activity with later performance was helpful in this respect). **B)** Cole et al. (2010) achieved repeated task novelty by combining 12 rules into many unique sets (i.e., 64 tasks). This approach allows for an important ‘practiced’ control condition, in which a subset of tasks is highly practiced. Contrasting novel to practiced conditions controls for rule/stimulus novelty and task switching, allowing the experimenter to isolate processes specific to task novelty (i.e., the unique combination of rules). Stocco et al. (2012) used a similar approach with algebraic rules. These approaches use unique combinations of abstract rules to investigate RITL, and so could be called ‘abstract RITL’.



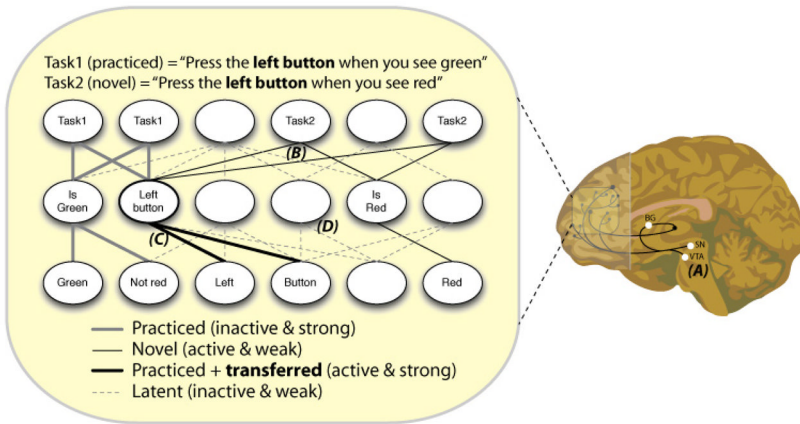
**Figure 4. Evidence for an anterior-to-posterior LPFC gradient in RITL**

**A)** Ruge & Wolfensteller (2010) demonstrated that concrete RITL implementation-related activity shifted anterior-to-posterior with practice. This may reflect the need to integrate task-relevant representations within more anterior LPFC regions during RITL, which can then shift more posteriorly as specific task-relevant connections are selected and strengthened with practice. Alternatively, this may be conceptualized as instruction-driven (novel task) processes being more abstract and being converted to more concrete/pragmatic representations with practice – consistent with an anterior-to-posterior gradient of abstraction within LPFC. **B)** Similar to concrete RITL, Cole et al. (2010) demonstrated that abstract RITL implementation-related activity also shifted anterior-to-posterior with practice. In contrast to concrete RITL, however, the activity started (before practice) and finished (after practice) in more anterior portions of LPFC. This may reflect the greater overall abstraction of the learned tasks, consistent with an anterior-to-posterior gradient of abstraction within LPFC. Note that the “abstract” RITL paradigm was also more complex than the “concrete” RITL paradigm, leaving open the possibility that the anterior-to-posterior gradient reflects complexity rather than abstraction.



**Figure 5. Compositional coding of task rules within LPFC**

**A)** Compositional coding allows activity patterns to remain intact when combined (top), in contrast to independent coding (bottom). Compositional coding may allow constituent rule practice – which likely shaped activity patterns to effectively implement the rules – to rapidly transfer to novel rule combinations during RITL. Figure adapted from Reverber et al. (2011). **B)** In order to localize compositional rule coding, Reverber et al. (2011) used constituent rule activity patterns (e.g., Rule A and Rule B) to predict compound rule activity patterns (e.g., Rule AB). Of the entire brain (using searchlight analysis), only LPFC showed statistically significant compositional coding. Figure adapted from Reverber et al. (2011). **C)** The hypothesis that compositional coding in LPFC allows for transfer from practiced to novel/RITL conditions was tested directly by Cole, Etzel, et al. (2011). See Figure 3B for the cognitive paradigm used in this study. LPFC activity pattern classifiers were trained to discriminate 4 rules (6 comparisons) using a set of highly practiced tasks (rule combinations). These classifiers were then tested using a set of completely novel tasks, and 5 of the 6 classifications were statistically significant ( $p < .05$ ). This result suggests LPFC coding is compositional and is transferred from practiced to novel task contexts. Figure adapted from Cole, Etzel, et al. (2011).



**Figure 6. RITL-capable theoretical model**

A portion of LPFC (receiving midbrain dopamine and other basal ganglia [BG] projections) is depicted with groups of neurons illustrated as ovals, some of which are labeled with their receptive fields (representations). Each level is a distinct portion of LPFC (top=anterior, bottom=posterior) in a **general processing hierarchy**. Task1 has been extensively practiced while Task2 is novel – requiring RITL capabilities. Note that connection activation levels are important for task set activation, and stronger connections are more readily activated and maintained. RITL (Task2) is implemented by (A, right) the dopamine system (substantia nigra [SN] & ventral tegmental area [VTA]) and the rest of the BG predicting reward and signaling LPFC to update its active connections (**rapid updating**), incoming cortico-cortical connections activated by instructions (not depicted) activate the appropriate task semantics and (B) latent integrating representations become active (via **latent connectivity**) to facilitate synchrony/binding of representations. **Compositional** reuse of previously practiced task rules (C) facilitates RITL via **transfer** of connection strengths (and connection accuracy) to facilitate task set activation. RITL with a variety of other tasks is possible due to extensive combinatorics of the latent connections (D). Note that many details were simplified here for illustrative purposes (e.g., receptive fields are likely quite complex; there should be relational units that specify green→left button rather than just an association; coding should be more coarse [not necessarily fully dedicated neurons for each task at the top]).



**Table 1**

Basic principles of the compositional theory of flexible cognitive control.

<b>Principle</b>	<b>Description</b>
Compositionality	The ability to reuse representations with a variety of others, resulting in massive combinatorics of possible representational combinations for task learning. All of the theory's principles are ultimately tied to compositionality of representations and how this benefits cognitive flexibility.
Immediate transfer	The combination of practiced rules into novel configurations resulting in the benefiting of novel task performance from previous practice.
Abstraction	The compositional grouping of representations (including feature subsets of full representations) into categories. Activation of such feature subsets of one concept can then take part in representing a different concept. This is highly related to compositionality, and facilitates transfer.
Analogy	The recognition of similarity between two or more representations. This allows selection of features common among those representations to form an abstraction, which can then transfer to new tasks during RITL.
Compositional hierarchy	A series of representations, ultimately based on simple sensory-motor features, building upon each other in stages with increasing abstraction and/or complexity at each stage. Lower-level representations in such a hierarchy can be compositionally combined and coordinated by higher-level (abstract and/or complex) representations to create novel task sets during RITL.

**Table 2**  
**Mechanistic principles of the compositional theory of flexible cognitive control**

These principles differ from those presented in Table 1 in that these are less abstract, such that we consider these to be readily implementable in computational models.

Principle	Description
Multi-system global connectivity	LPFC connectivity with many content-specific systems throughout the brain, giving access to many potentially task-relevant representations.
Rapid updating	A fast change of active content within LPFC, likely via a mechanism (basal ganglia) that gates instruction information (from posterior cortex).
Within-LPFC global connectivity	Extensive connectivity between neurons within LPFC, allowing for complex processing and latent connectivity (see below).
Latent connectivity	Unused connections and connectivity patterns that can become used as necessary by novel tasks during RITL.
Coarse-coded conjunctive representations	A large set of neurons with broad receptive fields that receive inputs from (potentially random) combinations of each other to produce many conjunctive receptive fields. This allows for representational binding, general processing (see below), and other principles considered here.
Synchrony/co-activation	Binding via synchronous co-activation of multiple representations, allowing for rapid selection (see below) of sets of representations to achieve massive combinatorics during RITL.
Incremental selection	Slow, multi-trial selection and tuning of task-representing neurons and connections for optimizing task performance from practice. Transfer of subsets of these neurons and connections to new tasks facilitates RITL.
Rapid selection	Fast selection of novel and (previously incrementally selected) representations from practiced tasks during RITL.
General processing hierarchy	Specific instantiations of a “compositional hierarchy” (Table 1). It consists of many connected neural populations building a wide variety of representations via conjunctions, unions, and other set theory operations, ultimately based on primitives in primary sensory-motor cortices. Due to wiring costs promoting short-distance connectivity, this results in multiple hierarchies of processing starting from primary cortices and going outward anatomically in terms of complexity and abstraction. We focus on the general processing hierarchy within LPFC.
Hierarchical conservation	A bias to incrementally select and strengthen more posterior (lower-level) representations of a task during practice.
Population adaptive coding	The ability of LPFC as a whole to represent a wide variety of possible tasks. This is accomplished by compositionally selecting sets of individual neurons, each with relatively static and coarse coding, that together specify processes necessary to implement each specific task.