



Contents lists available at ScienceDirect

## Neuroscience and Biobehavioral Reviews

journal homepage: [www.elsevier.com/locate/neubiorev](http://www.elsevier.com/locate/neubiorev)



### Review article

# The task novelty paradox: Flexible control of inflexible neural pathways during rapid instructed task learning

Michael W. Cole<sup>a,\*</sup>, Todd S. Braver<sup>b</sup>, Nachshon Meiran<sup>c</sup>

<sup>a</sup> Center for Molecular and Behavioral Neuroscience, Rutgers University, 197 University Ave., Newark, NJ, 07102, USA

<sup>b</sup> Department of Psychology, Washington University, St. Louis, MO, 63130, USA

<sup>c</sup> Department of Psychology and Zlotowski Center for Neuroscience, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel

### ARTICLE INFO

#### Article history:

Received 24 October 2016

Received in revised form 25 January 2017

Accepted 8 February 2017

Available online xxx

#### Keywords:

Instructed learning

Cognitive control

Executive functions

Automaticity

Network science

Functional connectivity

Neuroimaging

### ABSTRACT

Rapid instructed task learning (RITL) is one of the most remarkable human abilities, when considered from both computational and evolutionary perspectives. A key feature of RITL is that it enables new goals to be immediately pursued (and shared) following formation of task representations. Although RITL is a form of cognitive control that engenders immense flexibility, it also seems to produce inflexible activation of action plans in inappropriate contexts. We argue that this “prepared reflex” effect arises because RITL is implemented in the brain via a “flexible hub” mechanism, in which top-down influences from the frontoparietal control network reroute pathways among procedure-implementing brain areas (e.g., perceptual and motor areas). Specifically, we suggest that RITL-based proactive control – the preparatory biasing of task-relevant functional network routes – results in inflexible associative processing, demanding compensation in the form of increased reactive (in-the-moment) control. Thus, RITL produces a computational trade-off, in which the top-down influences of flexible hubs increase overall cognitive flexibility, but at the cost of temporally localized inflexibility (the prepared reflex effect).

© 2017 Elsevier Ltd. All rights reserved.

### Contents

1. Introduction.....	00
2. Defining RITL.....	00
3. Flexible updating of neural pathways via PFC.....	00
4. RITL as proactive control.....	00
5. Evidence that RITL is implemented via prepared reflexes.....	00
6. RITL as proactively prepared reflexes.....	00
6.1. Flexible implementation of inflexible neural pathways.....	00
6.2. Anterior PFC as a task representation buffer.....	00
7. Conclusions.....	00
Acknowledgements.....	00
References.....	00

### 1. Introduction

Rapid instructed task learning (RITL) is the ability to quickly perform novel instructed procedures, demonstrating successful performance even on the first trial after instruction (Cole, 2009; Cole et al., 2010a, 2013a; Liefooghe et al., 2013b; Meiran et al.,

2015a). Our familiarity with this ability might make it seem mundane, yet it is remarkable from several perspectives. First, from a computational perspective RITL ability is highly non-trivial. Modern computers still require extended and tedious programming rather than rapid verbal instructions (or imitation) to produce novel instructed procedures. Second, RITL is evolutionarily remarkable, given that humans stand out in RITL abilities relative to all other species. For instance, macaque monkeys would likely require 6 months to 2 years to learn a simple task (e.g., delayed matching to sample) that even children learn quickly from instructions. Such

\* Corresponding author.

E-mail address: [mwcole@mwcole.net](mailto:mwcole@mwcole.net) (M.W. Cole).

inter-species differences cannot be solely driven by differential language skills. This can be easily seen by considering that instructions are frequently conveyed non-verbally, such as in visual-symbolic form (e.g., furniture assembly instructions; (Cole, 2009; Cole et al., 2013a)). Conversely, certain frontal brain lesions lead to “goal neglect”, in which verbal instructions are accurately understood (and can be repeated back perfectly), but are incapable of generating novel behaviors (Duncan et al., 1996; Luria, 1973b). Thus, identifying the mechanisms underlying RITL is a non-trivial undertaking that has the capacity to improve understanding of one of the most important human mental faculties.

Recent evidence suggests that a particular set of large-scale neural mechanisms may underlie RITL abilities. These mechanisms are integrated within the “flexible hub” theory (Cole et al., 2014) – a cognitive neuroscientific framework that has been tested using the tools of network science (Sporns, 2012). Extending from previous accounts (e.g., Miller and Cohen, 2001), the flexible hub framework postulates that frontoparietal cognitive control networks implement top-down cognitive control via global network interactions with task-implementing networks (e.g., visual and motor networks for a visual-motor task). The flexible hub framework integrates two basic mechanisms. First, “global connectivity”: control network regions have been shown to be connector hubs (Cole et al., 2010b, 2015; Power et al., 2011) – brain regions with extensive inter-network functional connectivity. Second, “flexible connectivity”: control network regions have been shown to shift their task functional connectivity interactions globally depending on the current task being learned during RITL (Cole et al., 2013b). Together these findings suggest that control networks utilize global flexible connectivity to encode new task instructions and bias task-implementing networks to follow those instructions during task performance.

Another important component of RITL is that it produces constraints on cognitive functioning when it is implemented. In particular, although RITL increases overall cognitive flexibility, it may also have the paradoxical side-effect of limiting flexibility in certain contexts. In particular, it has been shown that holding the intention in mind to execute a newly instructed task interferes with related (and incompatible) task performance (Meiran et al., 2015a), which we have described as an example of the classic prepared reflex effect (e.g., Hommel, 2000). We postulate that this interference – which we term the intention-based reflexivity interference effect – reveals a computational trade-off between proactive control and in-the-moment flexibility (which is aided by reactive control). Proactive control is the capacity to prepare for the implementation of control processes prior to the time they are needed. In contrast, reactive control is the capacity to implement control on-the-fly, in response to a detected increase in ongoing control demands. RITL is likely to be highly dependent on proactive control, utilizing the instruction period as preparation to form and implement task configurations prior to novel task performance. However, as we discuss below, our findings with a specific prepared reflex paradigm (NEXT), suggests that this strong proactive component can interfere with related processes, requiring enhanced reactive control to resolve this interference.

In the sections that follow, we provide a more detailed review of both the computational benefits – and trade-offs – produced by RITL. We begin by elaborating on the concept of RITL: how it is defined, its operating characteristics, and why it appears to be such a remarkable human capability. Next, we discuss some of the constraints that come along with implementing RITL as a prepared reflex. Specifically, we discuss recent findings that document some of the potentially counter-intuitive interference effects – classically associated with expertise rather than task novelty – that arise when individuals perform novel tasks. We highlight our recently developed NEXT paradigm as an attractive platform for isolating

and identifying such interference effects. In subsequent sections, we describe how RITL might be implemented in the human brain, drawing on our recent work providing evidence for the flexible hub theory. We describe evidence that there is a strong relationship between flexible hub mechanisms, RITL, and proactive control.

In the final sections, we lay out our primary thesis that the proactive control capabilities implemented by flexible hubs create temporary inflexibility in the brain networks responsible for task implementation, via the transient reconfiguration of functional pathways. We suggest that reactive control can dynamically compensate for the temporary reduction in flexibility. Thus, optimal flexibility during novel task performance likely requires involvement of both proactive and reactive control mechanisms. We speculate that the anterior prefrontal cortex (PFC) might play a particularly important role in balancing these two types of control process, by enabling task-preparation to occur while protecting the system from actual task set implementation, so as to reduce interference in some circumstances. We describe supportive evidence that anterior PFC may be critical for RITL more broadly, especially in the formation of complex novel task sets.

## 2. Defining RITL

Proper investigation of RITL requires a precise definition. As discussed elsewhere (Cole et al., 2013a), RITL is the ability to rapidly perform novel instructed procedures. This definition leaves several aspects of the cognitive construct underspecified, however. For instance, what does it mean for learning to be “rapid”? What forms can “instruction” take? What exactly does it mean to “learn” something when it does not involve multiple exposures? Finally, what exactly is a “task” or “procedure”, and how does this differentiate RITL from other forms of rapid learning (e.g., one-shot learning of non-procedural content, such as semantic knowledge)?

Identifying ideal examples of RITL may help locate its conceptual boundaries. One such example is the immediate successful implementation of the task instruction, “Verbally name the letters for the word ‘instruction’, backwards.” This written instruction would immediately (within several seconds) produce the procedure: 1) saccade to the last letter (“n”), 2) identify the letter from memory and associated English label, 3) produce the verbal utterance for that letter (i.e., “en”), and 4) saccade to the letter to the left of the previously uttered letter (i.e., “o”) and 5) repeat steps 2 through 4 until a blank space is encountered. Note the simplicity of the instruction, the novelty of the procedure (for most individuals), the immediacy of highly accurate performance, along with the complex visual, motor, and abstract procedural (step 5 involves recursion) processing involved. This scenario is near the center of the conceptual space constituting RITL, and is thus a good starting point for accurately defining RITL.

The above letter-naming task takes several seconds to prepare, and several more seconds to execute. Of these components, the preparation phase is what is meant by “rapid”. It is difficult to identify a definite limit on how long preparation can be for task learning to qualify as rapid. However, to qualify as a bona fide instance of RITL, performance should be correct on the first attempt to execute the instructed task. One might also use working memory capacity limits – in number of items, complexity of inter-item relations, and duration of the memory traces – to predict RITL-related limits. Thus, RITL likely involves active maintenance in working memory to enable successful task implementation, rather than requiring the construction of procedural representations piecemeal over multiple trials and over extended periods of time. For example, RITL is used repeatedly to learn a complex game like tennis (e.g., learning the rules), but the entire skill set to be obtained in learning tennis

is clearly beyond working memory capacity, and would therefore not be considered an instance of RITL.

Most forms of learning, such as reinforcement learning (Botvinick et al., 2008), involve multiple trials. This might lead some to question whether RITL is learning at all. RITL is consistent with previously established concepts of “one-shot” learning (Lee et al., 2015) and “first trial”/“zero trial” learning (Gick and Holyoak, 1980; Palmer, 2012). These rapid forms of learning all involve some information that is not present in the mind, but which rapidly enters the mind and is used at a later time. The delay between encoding and retrieval can be long or short (e.g., less than a second later), but there will typically be a delay (reflecting critical RITL cognitive processes) to indicate that learning has occurred.

RITL is more specific than other forms of rapid learning, due to the requirement that the information be transformed (rather than simply recalled in its original form) into a novel procedure. This transformation process is non-trivial, as shown by the phenomenon of goal neglect (Bhandari and Duncan, 2014; Duncan et al., 1996). Goal neglect involves the successful perception, encoding, and retrieval of instructions without the ability to actually carry out those instructions. There can even be instances in which an individual can recognize that the new task is not being performed correctly, and might even verbalize what he or she should be doing, but still be unable to do it (Duncan et al., 2008). Goal neglect can occur in healthy individuals, but is much more common in frontal lobe lesion patients (Luria, 1973a) and older adults (De Jong, 2001). The existence of goal neglect in frontal lobe patients suggests that frontoparietal brain networks may play a prominent role in RITL. Consistent with this, a recent functional MRI (fMRI) study demonstrated that when the same information is framed as instructions (rather than merely to-be-remembered stimuli) the frontoparietal control network represents that information in anticipation of task performance (Muhle-Karbe et al., 2016). The important role of hippocampus in one-shot learning generally (Eichenbaum et al., 2007) suggests this brain structure may be important for RITL as well.

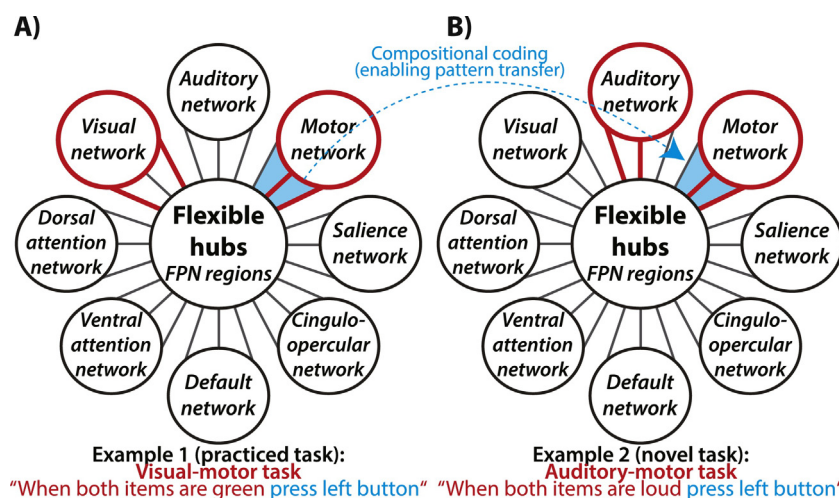
### 3. Flexible updating of neural pathways via PFC

A prominent theory of the neural basis of cognitive control – the guided activation theory – postulates that lateral PFC biases

activity in other brain regions to implement currently relevant task goals (Miller and Cohen, 2001). These PFC biasing effects are thought to increase cognitive and behavioral flexibility by overriding automatic/habitual associations. Demonstrating this increase in flexibility enabled by PFC, patients with PFC lesions often carry out automatic behaviors that are driven by external stimuli but are inappropriate in the current circumstance (Lhermitte, 1983; Luria, 1973a). For instance, a patient might put on glasses placed in front of him/her, even though he/she is already wearing glasses. Thus, PFC is essential for freeing the human mind from inflexible responding to stimulus-triggered associations.

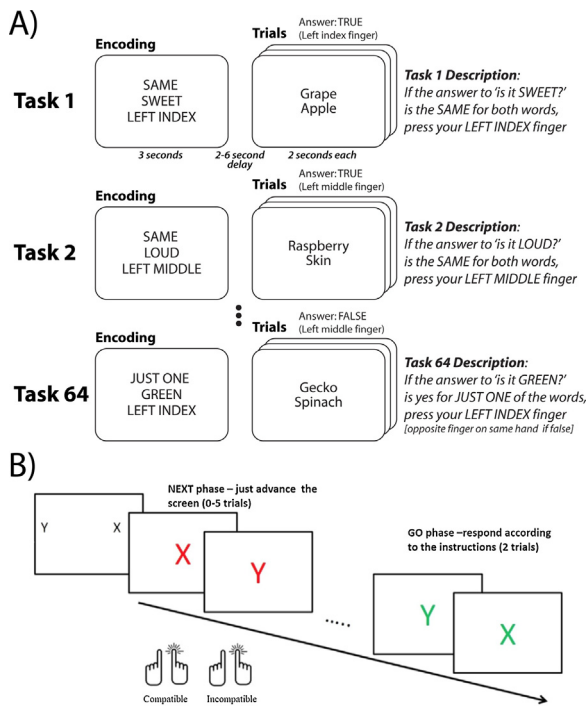
Building on the guided activation theory – which in turn built on other theories of cognitive control and attention (Cohen et al., 1990; Desimone et al., 1990; Desimone and Duncan, 1995) – we developed the flexible hub theory. This framework builds on the previous theory by: 1) extending the properties assigned to lateral PFC to a distributed frontoparietal cognitive control/multiple demand network (Cole and Schneider, 2007; Duncan, 2010); 2) being explicit about hub-related network mechanisms underlying guided activation by the frontoparietal network; and 3) linking these mechanisms to RITL and other forms of instructed learning.

The flexible hub theory predicts that frontoparietal regions are hubs and, moreover, that their functional connections are flexible across task contexts (Fig. 1). The first evidence for this claim came from resting-state functional connectivity fMRI studies, which found that frontoparietal regions have among the highest global connectivity in the brain (Cole et al., 2010b). However, in this initial study a relatively simple measure of hub connectivity was used – the overall number (weighted by strength) of functional connections with each brain region. This global brain connectivity measure can be biased toward identifying regions with connectivity primarily within a single network, rather than having truly global connectivity across networks (Wig et al., 2011). In contrast, graph theoretical measures such as participation coefficient are able to identify “connector” hubs (with extensive connectivity with multiple networks) separately from “provincial” hubs (with primarily within-network connectivity) (Guimera et al., 2005). Participation coefficient can be used to identify connector hubs, given that it reflects the degree to which a node's connections are uniformly distributed across



**Fig. 1.** Conceptual illustration of task-control flexible hubs in the frontoparietal network (FPN). (A,B) Task-control flexible hubs are schematically illustrated as brain regions in the FPN that exhibit global variable connectivity (A) and compositional coding (B). These mechanisms may explain how the FPN contributes to a wide variety of tasks, including novel tasks. Global variable connectivity is depicted by the shifting connectivity pattern (red lines connecting FPN to other brain networks) across multiple networks across the two example tasks. Compositional coding (enabling task skill transfer to novel tasks) is depicted by the reuse of a subset of the red connectivity pattern corresponding to the reuse of the “press left button” task component. These mechanisms would likely allow the FPN to meaningfully contribute to a wide variety of task contexts by allowing rapid reconfiguration of information flow across multiple task-relevant networks via reuse of previously learned sets of connectivity patterns. Figure adapted from (Cole et al., 2013b). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)





**Fig. 2.** RITL cognitive paradigms. A) The permutated rule operations (PRO) paradigm involves performance of novel combinations of 12 elementary task rules. Each task involves three elementary rules including a logical decision rule, a sensory semantic rule, and a motor response rule. Four of the tasks (each, a unique combination of three elementary rules) counterbalanced across participants) are practiced, while the remaining 60 tasks are novel. Adapted from (Cole et al., 2011). B) The NEXT paradigm begins with instruction of two novel arbitrary mappings, one stimulus (in this case, a letter) indicating a left pointer finger button press and another stimulus indicating a right pointer finger button press. An intermediate "NEXT phase" involves pressing one of the buttons to simply advance the screen. However, pressing the button incompatible with the maintained task set produces slowing (interference) with the screen advancement. The GO phase involves implementation of the learned task set. Adapted from (Meiran et al., 2016a).

networks (see Guimera et al., 2005 for exact equation). The hub status of the frontoparietal network was verified by Power et al. (2011) using participation coefficient. These results support the conclusion that frontoparietal regions are indeed connector hubs, consistent with the flexible hub theory.

A corollary to the flexible hub theory's hub prediction is that such hub status should have an important role in facilitating cognitive ability. Consistent with this, a follow-up study found that individual differences in the global connectivity of a particular frontoparietal region (within left lateral PFC) correlated with general fluid intelligence and cognitive control abilities (Cole et al., 2012). Further, the level of fluid intelligence was best predicted by the degree to which this hub exhibited connector hub properties (as opposed to provincial hub properties) (Cole et al., 2015). Together these results further support the hub status of frontoparietal regions, as well as their contribution to flexible cognition consistent with the flexible hub theory.

The other major prediction of the flexible hub theory is that frontoparietal regions should exhibit extensive flexible connectivity, in terms of reconfiguration patterns observed across distinct cognitive control procedures. This prediction was tested using a RITL paradigm – the permutated rule operations (PRO) paradigm (Fig. 2A) – optimized for examining flexible cognitive control (Cole et al., 2010a). Specifically, 12 elementary task rules were combined into 64 novel task sets, each task set consisting of a unique combination of 3 elementary rules. Functional connectivity fMRI (Pearson correlations between pairs of brain activity time series) was employed to test whether task-specific reconfiguration was present (Cole et al.,

2013b). Consistent with our primary hypothesis, the frontoparietal cognitive control network was found to shift its global functional connectivity pattern across distinct task sets. Further, these global connectivity patterns were distinct for each task – the functional connectivity pattern between the frontoparietal network and the rest of the brain could be used to decode which task each participant was performing at any given time. Together these findings suggest that the frontoparietal network consists of flexible hubs that help implement flexible changes in cognitive programs during RITL.

#### 4. RITL as proactive control

An important corollary of RITL in our account is that it depends preferentially on proactive control, drawing upon a distinction made within the Dual Mechanisms of Control (DMC) framework (Braver et al., 2007; Braver, 2012). In the DMC framework, cognitive control can be flexibly utilized in two distinct operating modes that vary in terms of their temporal dynamics and utility in different cognitive situations. In particular, the proactive control mode is one that is prospective or future-oriented, and involves sustained active maintenance of task goals. It is primarily engaged in an anticipatory fashion, when predictive cues in the environment signal upcoming high control demands, which can be most successfully met based on advanced preparation. Proactive control stands in stark contrast to the reactive control mode, which instead is present-oriented, involving the transient re-activation or retrieval of task goals (e.g., from long-term memory) based on either the detection of conflict/interference, or via associative (i.e., spreading activation) mechanisms triggered by features of the current situation. Reactive control is engaged as a "late correction", in a just-in-time fashion in conditions for which control demands are largely unpredictable, but are instead detected via moment-by-moment environmental (or internal) signals.

Prior work supports these distinctions between reactive and proactive control, while also providing information regarding associated neural mechanisms. For example, in the classic Stroop task, when interference is relatively rare, and further, when incongruent trials are clearly signaled in a trial-by-trial fashion by a relevant feature of the stimulus (a particular presentation color), then reactive control appears to be the preferred mode. In contrast, when Stroop interference is more frequent and expected, but not linked with any specific stimulus features, then proactive control appears to be preferred (Gonthier et al., 2016).

Likewise, in working memory tasks, a high expectation for interference when making target decisions (due to irrelevant familiarity of probes) was found to be associated with a shift from reactive to proactive control (Burgess and Braver, 2010). This control shift was accompanied by a shift in brain activity dynamics within lateral PFC from primarily transient patterns triggered by probe interference (familiarity) to one in which activity was increased during the encoding and working memory delay period, thus in advance and independently of probe type. In another study of working memory, increasing task motivation (through monetary incentives given for fast and accurate performance) was also associated with a similar temporal dynamic shift to proactive control within prefrontal and parietal cortex (Jimura et al., 2010). Moreover, those individuals showing the largest shifts in brain activity dynamics were the ones to show the greatest motivation-related changes in task performance.

The DMC framework is compatible with our account of RITL in that both highlight the role of computational trade-offs. Within the DMC framework, it is acknowledged that both control modes are flexibly utilized as a function of task demands, since each has associated costs and benefits. Reactive control is primarily utilized under

conditions in which control demands are either unpredictable or expected to be low. This is because it tends to be a weaker form of control, dependent upon rapid detection and activation of task goals to resolve behavioral uncertainty or conflicts as these arise. Consequently, reactive control is vulnerable to not being engaged as quickly or robustly as may be necessary to intervene effectively when the situation dictates.

In contrast, whereas proactive control tends to be a more robust form of control, re-configuring the system in advance to optimally deal with upcoming control demands, it is also a more costly and less efficient form of control. This is because it requires sustained active maintenance of task goals or rules in working memory, keeping them prepared and in an accessible form to be implemented when the need arises. In addition to the computational (and potentially metabolic) cost of proactive control related to sustained active maintenance,<sup>1</sup> another cost is that while the system is in this state, it may be less flexible in response to bottom-up signals that provide new information regarding the current situation. Consequently, proactive control is most typically engaged in conditions for which there is high motivation to perform the currently relevant task in an optimal manner, and for which advance signals, such as contextual or instructional cues, provide clear information regarding upcoming control demands and the task goals or rules needed to deal with these demands.

The key insight that links RITL to proactive control is that under RITL conditions, the instruction period provides both a clear indication of high upcoming control demands (given that the task is novel), while also signaling in advance the task goals or rules that will be relevant. Moreover, because the task is novel, there are only weak or nonexistent long-term memory representations of these task goals or rules. Thus, when RITL target stimuli are presented, such stimuli are unlikely to enable successful retrieval or reactivation of task goals and rules through either episodic/associative pathways in long-term memory or via conflict-based triggering.<sup>2</sup> Consequently, in order to ensure successful RITL task performance, proactive control, implemented via sustained active maintenance of task goals or rules from the instruction period, is necessary.

Nevertheless, it is important to point out that when proactive control is engaged in order to successfully implement RITL, reactive control might also need to be engaged to deal with stimulus-based interference that might arise. In particular, as we describe more fully below, in the NEXT paradigm it is precisely the engagement of proactive control that we argue is the cause of intention-based reflexivity interference from NEXT stimuli (Fig. 2B). Moreover, it is likely that reactive control is utilized in such situations as a late-correction mechanism to prevent errors from occurring. For example, the conflict that may occur from NEXT stimuli may trigger activation of auxiliary goal representations that detect the current stimulus as a NEXT rather than GO trial, and help bias responding in the appropriate direction to such stimuli. Interestingly, as we describe further in the final section, it may be that the brain tries to optimize for such situations through the use of hierarchical goal representations and goal-subgoal coordination involving anterior PFC regions.

Several accounts of reactive control involve active monitoring for conflict, which is used as a signal of the need for additional

top-down control (Botvinick and Cohen, 2014; Braver et al., 2007). This is important for addressing the “homunculus problem” (Hazy et al., 2007), in which the actions of a controller are accounted for via specific mechanisms rather than a vague “little man” metaphor. However, it is possible that such an explicit monitoring mechanism is not necessary for some aspects of decision making. In particular, if one of two competing responses is better practiced, then reaction time will likely slow when activating the less practiced one (e.g., due to lateral inhibition) without necessarily involving cognitive control. In contrast, much experimental and computational evidence has supported the idea that strategic shifts in the allocation of cognitive control across trials, and the consequent changes in reaction time, do likely require the involvement of cognitive control mechanisms (Botvinick et al., 2004; Jones et al., 2002).

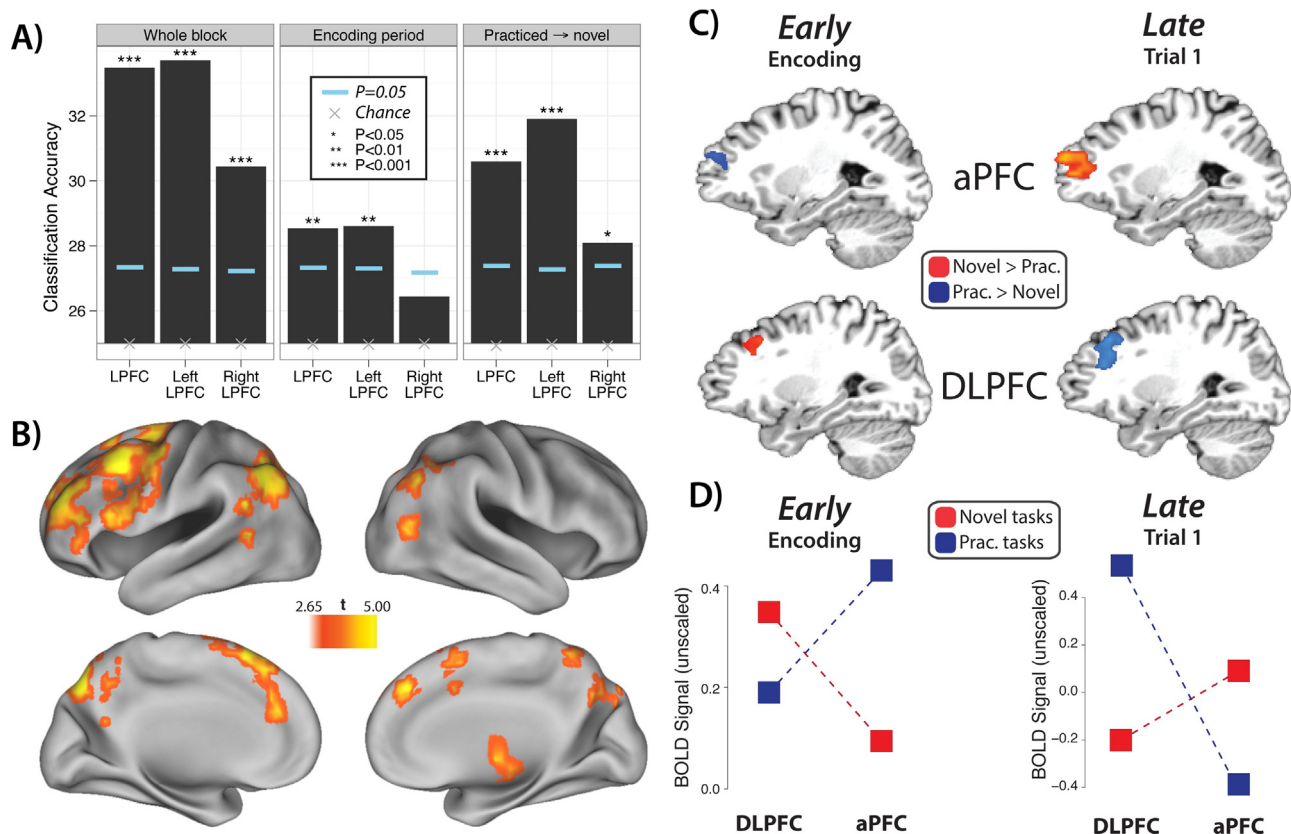
Our theoretical account of the role of proactive control in RITL makes a number of important predictions, some of which we have tested already. One of the most straightforward predictions relates to the differential importance of advance preparation in RITL situations relative to those involving well-practiced tasks. A surprising finding of earlier RITL studies involving the PRO paradigm (Fig. 2A) was that performance was only slightly worse when performing novel vs. practiced tasks (Cole et al., 2010a). Yet in the earlier paradigms a very long period was given to encode novel task instructions and prepare for the upcoming trial (5–9 s). Consequently, in ongoing follow-up work, we explicitly manipulated the duration of the encoding and preparatory period (Cole et al., under review). As expected, when this period was short (i.e., under 2 s), a strong “novelty cost” was observed, in terms of not only longer reaction times but also a significant (>6%) drop in accuracy. This was true even when occurring in a cued task-switching paradigm, which involved random alternation among a large number of both practiced and novel tasks. More importantly, in another variant of the paradigm, we allowed for the preparatory period to be self-paced (i.e., under participant volitional control). In this variant, we found that the preparatory time needed for novel tasks was over 300 msec longer than for practiced tasks, even when these tasks were randomly inter-mixed within a cued task-switching paradigm, as described above. This finding suggests that participants recognize the additional control demands associated with RITL situations and attempt to manage such demands via increased preparation (i.e., proactive control) prior to task performance.

In support of this idea, we found an additional trial-by-trial relationship between preparation time and higher accuracy on novel task trials. Critically, this pattern was selective, in that it was not present when performing practiced tasks (even when these were unpredictable and inter-mixed with novel task trials). Together, these findings suggest that proactive control might be uniquely important during RITL situations, and that when participants are given the opportunity to prepare they do so effectively. Nevertheless, there may be counteracting pressures that do not allow for consistent proactive control on all trials, such as cognitive effort costs or urgency pressures (Westbrook and Braver, 2015), which produces the observed trial-by-trial correlations.

These behavioral findings are paralleled by predictions involving neural activity dynamics and neural coding of task representations. With regard to the latter, a key component of our account is that novel tasks are represented in frontoparietal regions, by relying on existing representations that implement more familiar tasks (Cole et al., 2013a). Indeed, we observed that task rules being used during novel tasks could be decoded within a widely distributed region of lateral PFC (Cole et al., 2011) (Fig. 3A). Moreover, we found that when training a pattern classifier to decode task rules from lateral PFC based on practiced tasks, these same rules could be decoded (when in new combinations) under RITL

<sup>1</sup> It is likely computationally costly in the sense that only a small number of cognitive representations can be maintained or manipulated in this sustained fashion at a time, such that this maintenance process severely limits computational capacity during the maintenance period.

<sup>2</sup> It is likely that a long-term memory is formed during the initial trial of a novel task. This memory is likely then strengthened with additional practice, potentially becoming automatic with time. However, in the initial stages of task execution, when RITL dominates, these long-term memory representations are not yet sufficiently stable enough to rely upon.



**Fig. 3.** Brain processes and representations supporting complex RITL. A) Logical decision rules from the PRO paradigm were decoded from lateral PFC fMRI data. Decoding was above chance. Decoding was also possible when classifiers were trained on practiced tasks and tested on novel tasks, suggesting reuse of practiced representations during novel tasks. This representation transfer may support RITL by allowing positive transfer of skill into novel circumstances. Adapted from (Cole et al., 2011). B) A searchlight decoding map of PRO paradigm logic rules using the same data and analysis as panel A. Statistically significant t-values are shown ( $p < 0.05$ , uncorrected for multiple comparisons). A liberal threshold was used to indicate potential locations of task rule coding, though some false positives are likely present (this issue does not apply to panel A). C) Statistically significant novel vs. practiced task activations during the encoding and first trial of the PRO paradigm. D) There is an interaction between dorsolateral PFC (DLPFC) and anterior PFC (aPFC) during the encoding and first trial periods. Briefly, DLPFC appears to be involved in individual rule representation, while aPFC appears to be involved in integrated (whole-task-set) rule representation. Adapted from (Cole et al., 2010a).

conditions. This supports the idea of “flexible re-use” of representations – the rule representations developed during practiced tasks are re-used, in new combinations, during novel tasks. More specifically, the re-use account suggests that familiar representations referenced by instructions serve as building blocks in the novel task, and that what typically makes the task novel is the combination of elements rather than the elements themselves.

Moreover, with regard to the proactive control hypothesis, we also found that activity within dorsolateral PFC was increased during encoding and preparatory periods on novel relative to practiced trials (Cole et al., 2010a) (Fig. 3B). Interestingly, however, in anterior PFC regions a distinct pattern was found in which activity during encoding and preparation was reduced on novel relative to practiced trials, though the pattern flipped during task performance trials. Below, we discuss this pattern in terms of the role of anterior PFC in balancing proactive control demands with the increased potential for interference that this produces. However, the main take-home point is that the work to date supports the idea that when performing in RITL situations, the increased reliance on proactive control is mediated through the representation and preparatory/sustained activation of task rules within frontoparietal regions. Nevertheless, further work is necessary to more precisely isolate preparatory periods during RITL situations, and to demonstrate that both sustained activation dynamics along with the quality of task rule representations govern the success of task performance.

## 5. Evidence that RITL is implemented via prepared reflexes

The operating characteristics of RITL have become an important theme within the cognitive behavioral literature. Here, the emphasis is on the hypothesis that RITL-based performance is implemented as a prepared reflex. The prepared reflex metaphor has a long history in psychology (Hommel, 2000, p. 200). In essence, this metaphor suggests that actors who respond to stimuli by taking an action do not necessarily have direct control of this act at the moment it is executed. Instead, direct control exists only beforehand, when setting the mind to respond in a particular manner in the future, in this case, by associating the stimulus with the relevant motor plan. So, when instructed that the letter “T” is associated with a left key-press, actors form an association between “T” and {left response}. It is during this preparatory stage that their full (volitional) control is exercised. Afterwards, seeing “T” would be sufficient to retrieve and execute the motor plan. In more modern terms, this suggests that people are able to upload into working memory a *procedure* (as opposed to factual information), and that once this procedure is represented in working memory it gains a life of its own. Recently, Oberauer and colleagues have termed this aspect of working memory “procedural working memory” (Oberauer, 2009; Oberauer et al., 2013).

This prepared-reflex action mode has several noteworthy advantages. Perhaps the most important advantage in the case of RITL, is that the prepared-reflex mode makes it possible to separate the instruction phase (when learning how to act while receiving



instructions) from the action phase. For example, when receiving directions for how to reach a particular destination, the instructions may need to be encoded and prepared as a procedure before they can actually be executed, given the potential need to execute rapid actions while driving. Yet once these instructions are properly encoded as a procedure, they should not require further preparation to implement while driving.

A related advantage is that, if the action plan also specifies the relevant triggering stimuli, this information can serve to direct attention towards this goal-relevant information (Tibboel et al., 2015), which in turn shields the intention from goal-irrelevant distraction (Dreisbach, 2012). Thus, when the driving instructions are encoded and being implemented they can also direct attention to the relevant intermediate waypoints. Yet another advantage is the ability to use instructions to rapidly overcome strong habits (Theeuwes et al., 2014). For example, a child with a mosquito bite, hearing the command from a parent, “Don’t scratch because you’ll cause a scab”, can use this information to immediately overcome an otherwise very strong habitual urge.

Nevertheless, there are also important disadvantages associated with the prepared-reflex mode. This most critical is the loss of online control, and the related unintentional triggering of the intention in inappropriate contexts – namely, automatic-like behavior. Meiran et al. (2015a, 2015b) described the example of a policeman accidentally shooting at an innocent civilian coming into their field of view while expecting a criminal as representing an unfortunate potential real-life consequence of the prepared-reflex strategy. A somewhat less dramatic example is when participants continue to execute a prospective-memory task even after it had been declared as no-longer relevant (Bugg and Scullin, 2013; Walser et al., 2012).

Recently, mounting evidence has suggested that (contrary to some prominent theories) the automatic retrieval and efficient implementation of actions is not restricted to highly over-trained skills (e.g., reading words, as seen in the Stroop effect). This evidence strongly supports the conclusion that novel action plans that have never been practiced beforehand also show this “automaticity” signature in terms of behavioral profile. Since these instances reflect a partly or even completely different underlying mechanism than skill-based automaticity, we have labeled them “intention-based reflexivity” (Meiran et al., 2012).

Intention-based reflexivity has been found in at least five different behavioral paradigms (Cohen-Kdoshay and Meiran, 2007; De Houwer et al., 2005; Liefooghe et al., 2012; Meiran et al., 2015b; Wenke et al., 2007). Intention-based reflexivity has also been demonstrated at the level of the retrieval of the action plan itself, as seen in the brain lateralized readiness potential (Everaert et al., 2014; Meiran et al., 2014). Specifically, Meiran et al. (2014) demonstrated that participants who received a novel task mapping two new stimuli to the right/left keys had elevated activation of event-related brain potentials in corresponding motor cortex when seeing the stimuli during a no-go phase that preceded the first task implementation. Together, these intention-based reflexivity phenomena provide strong support for the hypothesis that (at least relatively simple) action plans are implemented as prepared reflexes.

In our view (see Meiran et al., 2015a, 2015b), a clear demonstration of intention-based reflexivity must meet certain criteria. These include showing that action-activation occurs in spite of the instructions, while explicitly ruling out any potential involvement of practice-related long-term memory. We thus maintain that intention-based reflexivity can be seen most clearly and unambiguously in a paradigm we recently developed – the NEXT paradigm (Meiran et al., 2015a, 2015b) (Fig. 2B). In the NEXT paradigm performance is examined across a series of mini-blocks (55–110, in different experiments), which each involving a new task procedure that is instructed and implemented. The instruction phase

initiates each task block by visually presenting a few simple task rules (typically associating two new stimuli to the right/left button presses, e.g., if “T” press left, if “J” press right). After the instructions, the critical NEXT phase typically occurs, followed by a GO phase. In the NEXT phase, participants are told to withhold applying the instructions when the stimuli appear (typically in a distinct color), but instead to make a fixed response (e.g., right button press) to advance to the next screen. This phase ends after an unpredictable number of trials, wherein the very brief (typically, 2-trial) GO phase begins. The GO phase involves the new instructions being implemented for the first time. GO performance provides an assay of RITL, while the NEXT responses are used to assay intention-based reflexivity.

We have consistently found that NEXT (screen advancement) responses are slowed in incompatible conditions: the presented stimulus (e.g., “T”) is associated with the opposite response (left) than that used for NEXT responses (Meiran et al., 2015b). Importantly, this “NEXT compatibility effect” is most robust in the first NEXT response, thus meeting the pre-requisites we laid out: It represents reflexive behavior since participants are told not to apply the newly instructed rule during the NEXT phase, and it is not due to prior practice.<sup>3</sup>

This finding has several noteworthy implications. One is that, contrary to influential theorizing suggesting automaticity requires extensive task experience (Logan, 1988), the intention-based reflexivity phenomenon reveals that it can sometimes appear prior to the first instance of executing a new task. The underlying mechanism is quite likely different for skill-based automaticity and intention-based reflexivity, however. Evidence for this comes from findings suggesting that the two phenomena reflect opposite ends of the learning curve. Specifically, skill-based automaticity increases with practice (Schneider and Shiffrin, 1977), whereas intention-based reflexivity decreases with practice (Meiran et al., 2015b). Additionally, intention-based reflexivity can be eliminated by concurrent working memory load (Meiran and Cohen-Kdoshay, 2012a), whereas skill-based automaticity cannot (Kessler and Meiran, 2010). This means that although the two phenomena may appear superficially as similar, they are in fact likely to be quite different.

One thing that may be in common, however, is single-step retrieval (Logan, 1992). In the case of intention-based reflexivity, forming a strong association (in active working memory) between the relevant stimulus and response may allow the response to be retrieved in a single step, thus giving rise to the intention-based reflexivity phenomenon. Alternatively, there may be a neural substrate shared between intention-based reflexivity and skill-based automaticity, but with distinct means of forming this substrate. In particular, it may be the case that top-down influences from frontoparietal cortex implement novel task sets via temporary reconfiguration of the same neural circuits that would be more permanently reconfigured by extensive practice in the case of skill-based automaticity. Indeed, this is the account we favor, as described in detail in the subsequent section.

Recent research efforts have shed light on the intention-based reflexivity phenomenon and several conclusions can already be drawn. One is that for action plans to show this phenomenon, they must be uploaded into WM, such that the prepared-reflex strategy appears to be disabled under high working memory load (Liefooghe et al., 2013a; Meiran and Cohen-Kdoshay, 2012b). Specifically, Meiran and Cohen-Kdoshay found that when participants had to

<sup>3</sup> Given these results, it no longer seems absolutely essential to focus exclusively on the first NEXT trial, and thus any similar paradigm involving repeated presentations of tasks, such as Liefooghe et al.’s (2012) inducer-diagnostic paradigm, are appropriate for the examination of intention-based reflexivity.

keep additional unrelated rules in mind (working memory load), task performance dropped slightly, but more importantly, there was no longer evidence for intention-based reflexivity. This limitation also suggests that the prepared-reflex strategy is feasible as long as the action plans are not too complex, otherwise working memory would be taxed too heavily (see elaboration in Meiran et al., 2012). One reason is that complexity or high working memory loads might force retrieval to be done in multiple steps; since single-step retrieval may be a prerequisite for intention-based reflexivity, such reflexivity is eliminated in these conditions.

Automaticity in general is widely believed to reflect rigidity and poor ability to take the ever-changing context into account. A paradigmatic case of automaticity is the Stroop task, in that participants rigidly apply well-practiced word reading procedures to the stimuli and have difficulty in taking the more novel context (the instruction to name the ink color) into account. This difficulty is even more true for intention-based reflexivity, in that dozens of practice trials may be required for performance to become somewhat context-sensitive (Braem et al., 2016). Specifically, in the Braem et al. (2016) study it was shown that a word in a task-irrelevant location triggered intention-based reflexivity when the task was novel, but that this effect went away with extensive practice. Note that it is also possible that the task set was refined over time through practice, eventually resulting in a complex task set that more effectively filtered out distracting stimuli.

Finally, a recent study showed that individuals who are better able to prepare for a novel task show relatively little intention-based reflexivity (Meiran et al., 2016b). This finding is somewhat surprising given the aforementioned effects of working-memory load on intention-based reflexivity. Specifically, for the average person, lowering working-memory capacity by working memory load eliminates intention-based reflexivity. However, across individuals, those with poor working-memory capacity show enlarged intention-based reflexivity. Thus, this finding may suggest that the prepared-reflex strategy may be more costly for some people than for others. In the final section, we elaborate on this notion further, providing a possible neural account of such individual differences that involves balancing the demands of proactive and reactive control through engagement of the anterior PFC.

## 6. RITL as proactively prepared reflexes

### 6.1. Flexible implementation of inflexible neural pathways

The previous sections reviewed literature supporting two seemingly contradictory findings: 1) the increased reflexivity observed during RITL in the context of the NEXT paradigm; and 2) examples of RITL instantiating some of the most flexible forms of human cognitive behavior. One way out of this seeming paradox is to consider the possibility that the increased global flexibility associated with RITL may come at the expense of (temporally) local decreases in flexibility. Further research will be necessary to directly test this, but there are already pieces of evidence in support of this possibility. For instance, this account is consistent with the flexible hub mechanism described above, in which functional pathways in the brain are temporarily altered according to task instructions. Such a reconfiguration pattern likely increases the flexibility of task implementation possibilities relative to stimulus-driven automatic responses. However, once functional pathways are configured according to task instructions, they may be somewhat resistant to further reconfiguration or updating (i.e., the task-set is somewhat “locked in”). One potential reason for this is that such a reconfiguration process takes time (Meiran, 1996; Monsell, 2003; Rogers and Monsell, 1995), leading to a degree of inflexibility, especially under time-pressured scenarios (e.g., responding

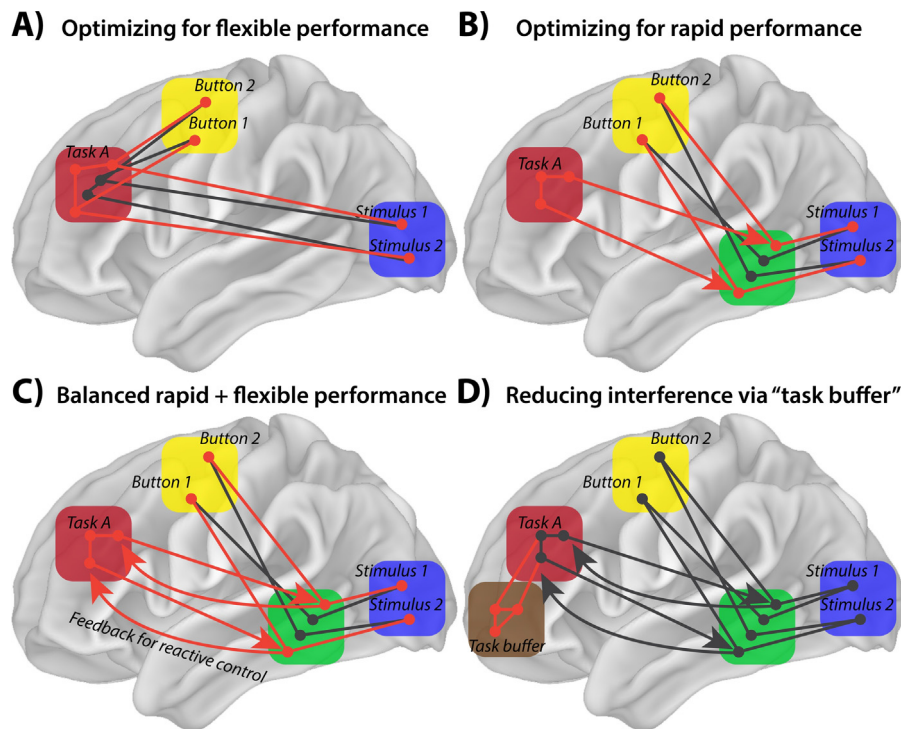
to NEXT stimuli). In other words, the temporary inflexibility that occurs during RITL scenarios may reflect the extent to which the brain's current task set cannot be immediately reconfigured by new instructions, once initially configured. This is consistent with the prepared reflex account described above, in which advanced preparation then temporarily limits subsequent flexibility.

The top-down reconfiguration of task-implementing pathways during RITL (e.g., visual-to-motor neural pathways for a visual-motor task) contrasts with an alternative possible mechanism, in which task-relevant information would pass through, and only be locally transformed within the frontoparietal control system during RITL implementation (Fig. 4A). This might be considered a particularly efficient neural mechanism, owing to the need for only a fixed set of sensory inputs and motor outputs with the control system. The control system would then just require internal “switch setting” to reroute sensory-motor information with rapid online control. However, inconsistent with this possibility, rapid reconfigurations of task-relevant functional connectivity pathways outside the control networks have been observed with fMRI (Cole et al., 2013b; Fuster et al., 1985; Gazzaley et al., 2007). The target brain regions for this sort of reconfiguration have tended to be in association cortex, typically in the posterior temporal lobe (Fuster et al., 1985; Gazzaley et al., 2007). These findings suggest flexible hubs may implement proactive control by rerouting/biasing existing (possibly automatic) pathways, momentarily restricting the natural dynamics of task-relevant neural representations to conform to the current task set (Fig. 4B).

One issue with this potential mechanism is the need to represent novel associations during RITL. Are these associations created on-the-fly in posterior cortex, or do they already exist there and are simply strengthened? It is unlikely that synaptic connections can be formed quickly enough to accommodate the speed of RITL trials (typically on the order of seconds). Instead, we postulate that there is extensive “latent connectivity” – connections that can be used when potentiated via top-down signals, but are unlikely to drive action potentials otherwise (Cole et al., 2013a). Such a mechanism would require excess connectivity that has not been (or is not often) used. Estimates of synaptic convergence on individual human cortical neurons range from 1000 to 10,000 connection inputs. More generally, there are estimated to be at least 100 trillion synaptic connections in the human brain. These large values indicate that the possibility of latent connectivity is at least plausible. For instance, consider the scenario in which a temporal lobe neuron receives inputs from triangle-representing neurons and sends output to right-button-press neurons. This could be considered a “triangle right-button task” neuron. With enough latent connectivity, it is plausible that the presence of triangle stimuli alone would not be sufficient to activate this neuron, but could do so when additional top-down signals from PFC bias its activity toward threshold. With regard to how instructions select the appropriate latent connections and neural units, one possibility is that instructions activate distributed neural representations of familiar task components (e.g., triangle and right button) that converge to optimally activate neurons that represent the conjunction of these components (i.e., triangle right-button task neurons). It will be important for future research to better establish the existence of latent connectivity and its role in RITL. One notable possibility is that individuals with more cortical neurons have greater latent connectivity, which may help explain higher fluid intelligence in individuals with larger cortical gray matter volume (Cole et al., 2012).

Other categories of evidence that support the role of top-down control in altering the natural dynamics of task-relevant (likely posterior, non-frontoparietal) neural representations. In particular, a number of relevant findings come from studies that have investigated conscious visual awareness – the ability to consciously detect near-threshold stimuli – through the use of visual masks.





**Fig. 4.** Alternative brain network mechanisms for RITL and reflexivity. Lateral PFC (red box) represents example Task A, with red circles and lines indicating activated neural populations and functional connections, respectively. A simple task set is depicted, with Stimulus 1 (e.g., the letter “T”) leading to Button 2 (e.g., pressing with the right index finger), and Stimulus 2 (e.g., the letter “S”) leading to Button 1 (e.g., pressing with the left index finger). The yellow box represents primary motor cortex, the blue box visual cortex, the green box posterior association cortex, and the brown box anterior PFC. **A)** A theoretical network architecture consistent with optimized flexible performance is depicted. However, this scenario would slow down processing due to requiring all task-related activity to flow through PFC, which is a long distance from the sensory input (requiring additional action potential propagation time). **B)** An alternative scenario involves preparatory top-down signals from lateral PFC to posterior association cortex (green box) closer to both motor and sensory regions. This would engender a faster flow of activity from the visual to motor regions. However, this would produce interference if a target stimulus was presented during a secondary or intermediate task (as in the NEXT paradigm), resulting in an increase of errors. **C)** We postulate the existence of an architecture in which multimodal associative activity is monitored for errors by lateral PFC (or anterior cingulate cortex). This would allow for online, reactive control of performance at the cost of slowing down performance. However, some slowing may be useful for preventing errors, based on a more complete representation of the full task context within lateral PFC (relative to posterior multimodal cortex). **D)** A potentially more effective way to reduce interference during, e.g., NEXT task performance is to temporarily store the Task A representation in a “task buffer”. This buffer is removed from immediate embodied representations, allowing for task representation maintenance without interference with ongoing task performance. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

For example, Woodman and Luck (2003) instructed individuals to press a button when they detected one of several visual targets. Using object-substitution masking they were able to titrate the stimuli such that they were detected approximately 50% of the time (according to verbal report by participants). Remarkably, event-related potentials centered on visual cortex were enhanced whenever instructed targets occurred, even when participants were not consciously aware of those targets. The lack of conscious awareness suggests the target information was not broadcast throughout the brain (e.g., to the frontoparietal network and language areas), such that the target information could have only been enhanced based on local processes within visual processing regions. Since the targets were arbitrarily set by conscious instructions, this result is consistent with top-down alteration of automatic pathways as described by the flexible hub theory (and related theories; Fig. 4B).

A more recent study came to similar conclusions using magnetoencephalography, which allowed for more precise spatial inferences regarding the involved brain regions. Marti et al. (2012) used rapid serial presentation of auditory followed by visual stimuli such that a fraction of stimuli were not consciously perceived (due to the attentional blink). Instructions were given to press a button at the presence of a target auditory stimulus and a target visual stimulus. The visual stimuli were of interest since only the visual trials involved attentional blink. They found that brain activity was consistent with target detection in visual regions during non-conscious

target trials, but not in frontoparietal regions. On conscious target trials frontoparietal regions were associated with a late response (~400 ms), but this target-related enhancement of activity was absent for non-conscious trials. In contrast, both early (~270 ms) and late (~400 ms) target-related enhancements of visual region signals were intact both with and without consciousness. These results suggest that bottom-up activation of frontoparietal regions are important for fully conscious target detection (possibly due to their hub status facilitating spreading of information), whereas top-down effects can alter visual processing even without consciousness. This is again consistent with frontoparietal flexible hubs altering visual processing pathways independently of target information entering the frontoparietal system (Fig. 4B).

The Marti et al. (2012) study raises another possibility typical of brain function – why use only one mechanism when (in a parallel processing architecture) multiple mechanisms will make for a more robust and flexible system? The Marti et al. (2012) result suggests that typical conscious instructed task performance involves activation not only of the modified automatic pathways in visual regions but also late activation of frontoparietal regions that more explicitly represent task instructions (Cole et al., 2011; Woolgar et al., 2011). This suggests that while top-down influences on automatic pathways temporarily reduce flexibility, a slower process of task information feeding back to the frontoparietal system may restore such flexibility (Fig. 4C). This suggests possible different interpretations of the individual difference findings observed with regard

to NEXT interference effects: individuals with the largest NEXT interference effects could either have especially strong top-down influence on automatic pathways or instead have poor reactive control – feedback that triggers the engagement of frontoparietal regions is either of low amplitude or slow speed. It will be important for future studies to test this possibility.

This account of frontoparietal feedback is consistent with reactive control, in which online monitoring of conflict between the task set and ongoing processing is used to calibrate top-down influences (Botvinick et al., 2001; Braver et al., 2007). It appears likely that a parallel process spreads from the task-involved visual pathways to frontoparietal regions to allow for online monitoring of task processing. This monitoring process can then lead to rapid reconfiguration of top-down signals in scenarios in which incorrect responses are likely to occur. Thus, we expect that such reactive control dynamics can compensate for the temporary inflexibility inherent in top-down influences of processing pathways. A potential target for future research will be to provide direct support for this claim, as well as to identify the particular brain regions (e.g., anterior cingulate cortex) involved in reactive control compensation for instruction-induced inflexibility.

One particular prediction this account makes is that there should be stronger task-evoked sensory-motor functional connectivity during RITL relative to non-RITL scenarios. This prediction is due to the additional proactive control needs present during RITL described above. We specifically expect that the strength of task-evoked sensory-motor functional connectivity (as opposed to resting-state functional connectivity independent of task performance) among posterior regions will be correlated with NEXT interference effects, on a trial-by-trial and individual differences basis. Demonstrating such an effect would provide strong support for the proposed frameworks shown in Figs. 3B and C.

One important caveat, however, is that we expect this effect to be much stronger for simple tasks than complex multi-step tasks. In particular, we would expect this effect to be stronger for the NEXT paradigm than the PRO paradigm (Fig. 2). This prediction is based on the assumption that NEXT task sets involve proactive-control-triggered activation of simple stimulus-motor associations within a single step, allowing for full potentiation of the posterior representations as depicted in Figs. 3B and C. In contrast, in the PRO paradigm (Cole et al., 2010a) the activation of stimulus-motor associations requires the sequential chaining together of three task rules (sensory, logic, and motor), and their associated neural pathways. We expect that the sensory rule (e.g., is a stimulus green) could be reflexively triggered in a manner similar to that hypothesized for NEXT paradigm rules. However, because activation of the appropriate motor rule is also dependent on integration with the relevant logic rule (i.e., to integrate the outcomes of the two sensory rule implementations), we predict that it will not be reflexively triggered within the same step. Rather, we expect that bi-directional interaction with the frontoparietal control network will be necessary to trigger the appropriate motor action. Testing this possibility will be important, since it would place a theoretical limit on the reflexive processes that occur during RITL, and potentially explain why novel multi-step procedures are carried out so slowly (typically hundreds of milliseconds per step) by the human brain.

Note that most RITL fMRI studies have been conducted with simple stimulus-response mappings (like the NEXT paradigm) and have primarily focused on task-evoked activations (Brass et al., 2009; Dumontheil et al., 2011; Ruge and Wolfensteller, 2010; Stocco et al., 2012). So far, only a small number (Cole et al., 2013b; Mohr et al., 2016) have investigated task-evoked functional connectivity effects, which are of primary interest here. Mohr et al. (2016) used a simple stimulus-response paradigm similar to the NEXT paradigm but did not focus on task-evoked func-

tional connectivity between control systems and their effects on stimulus-motor representations. Cole et al. (2013a, 2013b) focused on distributed representations between the frontoparietal control system and other systems, but used the complex PRO paradigm and did not isolate stimulus-motor representations. Thus, the predictions made here reveal an important gap in the RITL neuroscience literature.

## 6.2. Anterior PFC as a task representation buffer

As described above, a recent individual differences study revealed that individuals exhibiting minimal intention-based reflexivity effects on the NEXT paradigm also had better GO performance (Meiran et al., 2016b). This finding appears to be incompatible with the network architecture proposed in Fig. 4C. Specifically, the architecture in Fig. 4C suggests that a strong top-down effect from lateral PFC on posterior association cortex would enhance GO performance (implementation of the novel task) at the cost of increasing incompatible NEXT task performance interference effects. To accommodate the results reported in Meiran et al. (2016b) we suggest there might be an additional “task buffer” that is able to support task set representation in lateral PFC without directly triggering task-set implementation in posterior regions (Fig. 4D). Individual differences in the efficacy of this task buffer could potentially explain the counter-intuitive individual differences effect.

A likely location for this task buffer is anterior PFC. This hypothesis is based on a popular theory of PFC organization (Badre and D’Esposito, 2009; Fuster, 2001; Koechlin et al., 1999, 2003), which postulates that PFC forms a hierarchy with more anterior regions representing more abstract or temporally extended information. This account is also compatible with findings from the prospective memory literature. Neuroimaging studies of prospective memory have demonstrated that activity in anterior PFC increases in a sustained fashion when an intention is held online during performance of a secondary task (McDaniel et al., 2013; Reynolds et al., 2008). This pattern of findings suggests anterior PFC can act as a kind of task buffer, under conditions for which task implementation needs to be scheduled or deferred until after an intervening task is completed. By analogy, the NEXT paradigm can be considered as a kind of prospective memory (or branching (Koechlin et al., 1999)) task, with the novel instructions serving as a delayed intention. Thus, we predict that anterior PFC activity will be elevated during NEXT trials, and that this activity will correlate with individual differences in both NEXT trial and GO trial performance. Put differently, anterior PFC engagement in RITL contexts that involve deferred task implementation might be a means of balancing or optimizing the demands of proactive and reactive control. By utilizing anterior PFC as an actively sustained task buffer, the brain implements the form of proactive control needed for successful RITL, while at the same time minimizing the demands on reactive control, by reducing the likelihood of prematurely triggering task implementation (i.e., NEXT interference). Direct experimental confirmation of these predictions, as well as the functional connectivity hypotheses associated with them, would provide clear support for the conceptual model depicted in Fig. 4D.

Considering anterior PFC functionality more broadly, it is likely that this region is also important for RITL for reasons other than reducing interference. In particular, prior work has shown that anterior PFC is involved in constructing novel multi-step task sets during RITL (PRO paradigm tasks) (Cole et al., 2010a) (Fig. 3B). Moreover, recent findings suggest that PRO rule representations within anterior PFC (fine-grained activity patterns that discriminate between task rules) are predictive of task performance on a trial-by-trial basis (Cole et al., 2016). In general, anterior PFC might be preferentially involved in complex multi-step RITL tasks like the

PRO paradigm precisely because task rule buffering is required in order to chain and integrate rules together in order to complete the stimulus-motor association pathway.

Together, these results suggest that anterior PFC might play a more generally important role in RITL situations – both during simple tasks as a means of optimizing control demands and minimizing interference, and in more complex tasks as a means of coordinating and chaining together multiple task steps. We look forward to future studies that can empirically test these possibilities.

## 7. Conclusions

Instructed task performance creates an apparent paradox: previously inflexible associations can be overcome via instruction, yet a new set of associations must (at least temporarily) be put in their place. We have reviewed evidence suggesting these new associations involve their own inflexibility. A recently developed theory involving flexible frontoparietal hubs that enable proactive reconfiguration of task-relevant functional connections leads to the somewhat counter-intuitive suggestion of how such inflexibility arises. However, we postulate that the proactive control process of transient but inflexible task-pathway reconfiguration can be balanced by reactive control, in which online monitoring of performance allows for rapid alteration of task-set parameters. We additionally suggest that anterior PFC can act as a “task buffer” to allow maintenance of task sets without immediate implementation, reducing the inflexibility inherent in task preparation, while optimizing the balance between proactive and reactive control demands. Investigations that examine the interplay between task preparation and prepared reflex-type interference in both simple (e.g., NEXT) and complex (e.g., PRO) RITL paradigms promise to provide much-needed insights into the cognitive and neural mechanisms of proactive and reactive control, as well as the role of anterior PFC in complex cognition. We hope that the theoretical potential of the ideas laid out here will encourage other researchers to join in these research efforts.

## Acknowledgements

We would like to acknowledge funding support from National Institutes of Health grants R37 MH066078 to TSB, K99-R00 MH096801 to MWC, and R01 MH109520 to MWC, as well as grant 2011246 from the USA-Israel Bi-national Science Foundation (to NM and TSB).

## References

Badre, D., D'Esposito, M., 2009. Is the rostro-caudal axis of the frontal lobe hierarchical? *Nat. Rev. Neurosci.* 10, 659–669, <http://dx.doi.org/10.1038/nrn2667>.

Bhandari, A., Duncan, J., 2014. Goal neglect and knowledge chunking in the construction of novel behaviour. *Cognition* 130, 11–30, <http://dx.doi.org/10.1016/j.cognition.2013.08.013>.

Botvinick, M.M., Cohen, J.D., 2014. The computational and neural basis of cognitive control: charted territory and new frontiers. *Cogn. Sci.* 38, 1249–1285, <http://dx.doi.org/10.1111/cogs.12126>, PMID: 25079472.

Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S., Cohen, J.D., 2001. Conflict monitoring and cognitive control. *Psychol. Rev.* 108, 624–652.

Botvinick, M., Cohen, J., Carter, C., 2004. Conflict monitoring and anterior cingulate cortex: an update. *Trends Cogn. Sci.* 8, 539–546.

Botvinick, M., Niv, Y., Barto, A., 2008. Hierarchically organized behavior and its neural foundations: a reinforcement learning perspective. *Cognition* 19, <http://dx.doi.org/10.1016/j.cognition.2008.08.011>.

Braem, S., Liefvooghe, B., De Houwer, J., Brass, M., Abrahamse, E.L., 2016. There Are Limits to the Effects of Task Instructions: Making the Automatic Effects of Task Instructions Context-Specific Takes Practice. *J. Exp. Psychol. Learn. Mem. Cogn.*, <http://dx.doi.org/10.1037/xlm0000310>, PMID: 27617776.

Brass, M., Wenke, D., Spengler, S., Waszak, F., 2009. Neural correlates of overcoming interference from instructed and implemented stimulus-response associations. *J. Neurosci. Off. J. Soc. Neurosci.* 29, 1766–1772, <http://dx.doi.org/10.1523/JNEUROSCI.5259-08.2009>.

Braver, T., Gray, J., Burgess, G., 2007. Explaining the many varieties of working memory variation: dual mechanisms of cognitive control. *Var. Work. Mem.*

Braver, T.S., 2012. The variable nature of cognitive control: a dual mechanisms framework. *Trends Cogn. Sci.* 16, 105–112, <http://dx.doi.org/10.1016/j.tics.2011.12.010>.

Bugg, J.M., Scullin, M.K., 2013. Controlling intentions: the surprising ease of stopping after going relative to stopping after never having gone. *Psychol. Sci.* 24, 2463–2471, <http://dx.doi.org/10.1177/0956797613494850>, PMID: 24091550.

Burgess, G.C., Braver, T.S., 2010. Neural mechanisms of interference control in working memory: effects of interference expectancy and fluid intelligence. *PLoS One* 5, <http://dx.doi.org/10.1371/journal.pone.0012861>.

Cohen, J.D., Dunbar, K., McClelland, J.L., 1990. On the control of automatic processes: a parallel distributed processing account of the Stroop effect. *Psychol. Rev.* 97, 332–361.

Cohen-Kdoshay, O., Meiran, N., 2007. The representation of instructions in working memory leads to autonomous response activation: evidence from the first trials in the flanker paradigm. *Q. J. Exp. Psychol.* 60, 1140–1154.

Cole, M.W., Schneider, W., 2007. The cognitive control network: integrated cortical regions with dissociable functions. *Neuroimage* 37, 343–360, <http://dx.doi.org/10.1016/j.neuroimage.2007.03.071>.

Cole, M.W., Bagic, A., Kass, R., Schneider, W., 2010a. Prefrontal dynamics underlying rapid instructed task learning reverse with practice. *J. Neurosci.* 30, 14245–14254, <http://dx.doi.org/10.1523/JNEUROSCI.1662-10.2010>.

Cole, M.W., Pathak, S., Schneider, W., 2010b. Identifying the brain's most globally connected regions. *Neuroimage* 49, 3132–3148, <http://dx.doi.org/10.1016/j.neuroimage.2009.11.001>.

Cole, M.W., Etzel, J.A., Zacks, J.M., Schneider, W., Braver, T.S., 2011. Rapid transfer of abstract rules to novel contexts in human lateral prefrontal cortex. *Front. Hum. Neurosci.* 5, 142, <http://dx.doi.org/10.3389/fnhum.2011.00142>.

Cole, M.W., Yarkoni, T., Repovs, G., Anticevic, A., Braver, T.S., 2012. Global connectivity of prefrontal cortex predicts cognitive control and intelligence. *J. Neurosci.* 32, 8988–8999, <http://dx.doi.org/10.1523/JNEUROSCI.0536-12.2012>.

Cole, M.W., Laurent, P., Stocco, A., 2013a. Rapid instructed task learning: a new window into the human brain's unique capacity for flexible cognitive control. *Cogn. Affect. Behav. Neurosci.* 13, 1–22, <http://dx.doi.org/10.3758/s13415-012-0125-7>.

Cole, M.W., Reynolds, J.R., Power, J.D., Repovs, G., Anticevic, A., Braver, T.S., 2013b. Multi-task connectivity reveals flexible hubs for adaptive task control. *Nat. Neurosci.* 16, 1348–1355, <http://dx.doi.org/10.1038/nn.3470>.

Cole, M.W., Repovs, G., Anticevic, A., 2014. The frontoparietal control system: a central role in mental health. *The Neuroscientist* 20, 652–664, <http://dx.doi.org/10.1177/1073858414525995>.

Cole, M.W., Ito, T., Braver, T.S., 2015. Lateral prefrontal cortex contributes to fluid intelligence through multinet connectivity. *Brain Connect.* 5, 497–504, <http://dx.doi.org/10.1089/brain.2015.0357>.

Cole, M.W., Ito, T., Braver, T.S., 2016. The behavioral relevance of task information in human prefrontal cortex. *Cereb. Cortex* 26, 2497–2505, <http://dx.doi.org/10.1093/cercor/bhv072>.

Cole, M.W., Meiran, N., Braver, T., 2017. Proactive Control and Metacognitive Awareness During Rapid Instructed Task Learning (In preparation).

Cole, M.W., 2009. The Biological Basis of Rapid Instructed Task Learning. *Diss. Univ., Pittsburgh*.

De Houwer, J., Beckers, T., Vandorpe, S., Custers, R., 2005. Further evidence for the role of mode-independent short-term associations in spatial Simon effects. *Percept. Psychophys.* 67, 659–666.

De Jong, R., 2001. Adult age difference in goal activation and goal maintenance. *Eur. J. Cogn. Psychol.* 13, 71–89, <http://dx.doi.org/10.1080/09541440042000223>.

Desimone, R., Duncan, J., 1995. Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18, 193–222.

Desimone, R., Wessinger, M., Thomas, L., Schneider, W., 1990. Attentional control of visual perception: cortical and subcortical mechanisms. *Cold Spring Harb. Symp. Quant. Biol.* 55, 963–971, <http://dx.doi.org/10.1101/SQB.1990.055.01.090>.

Dreisbach, G., 2012. Mechanisms of cognitive control the functional role of task rules. *Curr. Dir. Psychol. Sci.* 21, 227–231, <http://dx.doi.org/10.1177/0963721412449830>.

Dumontheil, I., Thompson, R., Duncan, J., 2011. Assembly and use of new task rules in fronto-parietal cortex. *J. Cogn. Neurosci.* 23, 168–182, <http://dx.doi.org/10.1162/jocn.2010.21439>.

Duncan, J., Emslie, H., Williams, P., Johnson, R., Freer, C., 1996. Intelligence and the frontal lobe: the organization of goal-directed behavior. *Cognit. Psychol.* 30, 257–303.

Duncan, J., Parr, A., Woolgar, A., Thompson, R., Bright, P., Cox, S., Bishop, S., Nimmo-Smith, I., 2008. Goal neglect and Spearman's g: competing parts of a complex task. *J. Exp. Psychol. Gen.* 137, 131–148, <http://dx.doi.org/10.1037/0096-3445.137.1.131>.

Duncan, J., 2010. The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. *Trends Cogn. Sci.* 14, 172–179, <http://dx.doi.org/10.1016/j.tics.2010.01.004>.

Eichenbaum, H., Yonelinas, A.P., Ranganath, C., 2007. The medial temporal lobe and recognition memory. *Annu. Rev. Neurosci.* 30, 123–152, <http://dx.doi.org/10.1146/annurev.neuro.30.051606.094328>.

Everaert, T., Theeuwes, M., Liefvooghe, B., Houwer, J.D., 2014. Automatic motor activation by mere instruction. *Cogn. Affect. Behav. Neurosci.* 14, 1300–1309, <http://dx.doi.org/10.3758/s13415-014-0294-7>.



- Fuster, J., Bauer, R., Jervey, J., 1985. Functional interactions between inferotemporal and prefrontal cortex in a cognitive task. *Brain Res.* 330, 299–307, [http://dx.doi.org/10.1016/0006-8993\(85\)90689-4](http://dx.doi.org/10.1016/0006-8993(85)90689-4).
- Fuster, J., 2001. The prefrontal cortex?an update: time is of the essence. *Neuron* 30, 319–333.
- Gazzaley, A., Rissman, J., Cooney, J., Rutman, A., Seibert, T., Clapp, W., D'Esposito, M., 2007. Functional interactions between prefrontal and visual association cortex contribute to top-down modulation of visual processing. *Cereb. Cortex* 17, i125–i135, <http://dx.doi.org/10.1093/cercor/bhm113>.
- Gick, M.L., Holyoak, K.J., 1980. Analogical problem solving. *Cognit. Psychol.* 12, 306–355.
- Gonthier, C., Braver, T.S., Bugg, J.M., 2016. Dissociating proactive and reactive control in the Stroop task. *Mem. Cognit.* 44, 778–788, <http://dx.doi.org/10.3758/s13421-016-0591-1>.
- Guimera, R., Mossa, S., Turtchi, A., Amaral, L.A.N., 2005. The worldwide air transportation network: anomalous centrality, community structure, and cities' global roles. *Proc. Natl. Acad. Sci. U. S. A.* 102, 7794.
- Hazy, T.E., Frank, M.J., O'Reilly, R.C., 2007. Towards an executive without a homunculus: computational models of the prefrontal cortex/basal ganglia system. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 362, 1601–1613, <http://dx.doi.org/10.1098/rstb.2007.0055>.
- Hommel, B., 2000. The prepared reflex: automaticity and control in stimulus-Response translation. In: Mosell, S., Driver, J. (Eds.), *Control of Cognitive Processes: Attention & Performance XVIII.*, pp. 247–273.
- Jimura, K., Locke, H.S., Braver, T.S., 2010. Prefrontal cortex mediation of cognitive enhancement in rewarding motivational contexts. *Proc. Natl. Acad. Sci.* 107, 8871–8876, <http://dx.doi.org/10.1073/pnas.1002007107>.
- Jones, A., Cho, R., Nystrom, L., Cohen, J., Braver, T., 2002. A computational model of anterior cingulate function in speeded response tasks: effects of frequency, sequence, and conflict. *Cogn. Affect. Behav. Neurosci.* 2, 300–317.
- Kessler, Y., Meiran, N., 2010. The reaction-time task-rule congruency effect is not affected by working memory load: further support for the activated long-term memory hypothesis. *Psychol. Res. PRPF* 74, 388–399, <http://dx.doi.org/10.1007/s00426-009-0261-z>.
- Koechlin, E., Basso, G., Pietrini, P., Panzer, S., Grafman, J., 1999. The role of the anterior prefrontal cortex in human cognition. *Nature* 399, 148–151, <http://dx.doi.org/10.1038/20178>.
- Koechlin, E., Ody, C., Kouneiher, F., 2003. The architecture of cognitive control in the human prefrontal cortex. *Science* 302, 1181–1185, <http://dx.doi.org/10.1126/science.1088545>, PMID: 14615530.
- Lee, S.W., O'Doherty, J.P., Shimojo, S., 2015. Neural computations mediating one-shot learning in the human brain. *PLoS Biol.* 13, <http://dx.doi.org/10.1371/journal.pbio.1002137> (e1002137–36).
- Lhermitte, F., 1983. Utilization behaviour and its relation to lesions of the frontal lobes. *Brain J. Neurol.* 106 (Pt 2), 237–255.
- Liefooghe, B., Wenke, D., De Houwer, J., 2012. Instruction-based task-rule congruency effects. *J. Exp. Psychol. Learn. Mem. Cogn.* 38, 1325–1335, <http://dx.doi.org/10.1037/a0028148>.
- Liefooghe, B., De Houwer, J., Wenke, D., 2013a. Instruction-based response activation depends on task preparation. *Psychon. Bull. Rev.* 20, 481–487, <http://dx.doi.org/10.3758/s13423-013-0374-7>.
- Liefooghe, B., Houwer, J., Wenke, D., 2013b. Instruction-based response activation depends on task preparation. *Psychon. Bull. Rev.* 20, 481–487, <http://dx.doi.org/10.3758/s13423-013-0374-7>.
- Logan, G.D., 1988. Toward an instance theory of automatization. *Psychol. Rev.* 95, 492–527, <http://dx.doi.org/10.1037/0033-295X.95.4.492>.
- Logan, G.D., 1992. Attention and preattention in theories of automaticity. *Am. J. Psychol.* 105, 317–339, <http://dx.doi.org/10.2307/1423031>.
- Luria, A., 1973a. The frontal lobes and the regulation of behavior. *Psychophysiol. Front. Lobes*, 3–28.
- Luria, A.R., 1973b. The frontal lobes and the regulation of behavior. In: Luria, A.R., Pribram, K.H. (Eds.), *Psychophysiology of the Frontal Lobes*. Academic Press, Oxford, England, p. 332.
- Marti, S., Sigman, M., Dehaene, S., 2012. A shared cortical bottleneck underlying attentional blink and psychological refractory period. *Neuroimage* 59, 2883–2898, <http://dx.doi.org/10.1016/j.neuroimage.2011.09.063>.
- McDaniel, M.A., LaMontagne, P., Beck, S.M., Scullin, M.K., Braver, T.S., 2013. Dissociable neural routes to successful prospective memory. *Psychol. Sci.* 24, 1791–1800, <http://dx.doi.org/10.1177/0956797613481233>.
- Meiran, N., Cohen-Kdoshay, O., 2012a. Working memory load but not multitasking eliminates the prepared reflex: further evidence from the adapted flanker paradigm. *Acta Psychol. (Amst.)* 139, 309–313, <http://dx.doi.org/10.1016/j.actpsy.2011.12.008>.
- Meiran, N., Cohen-Kdoshay, O., 2012b. Working memory load but not multitasking eliminates the prepared reflex: further evidence from the adapted flanker paradigm. *Acta Psychol. (Amst.)* 139, 309–313, <http://dx.doi.org/10.1016/j.actpsy.2011.12.008>.
- Meiran, N., Cole, M.W., Braver, T.S., 2012. When planning results in loss of control: intention-based reflexivity and working-memory. *Front. Hum. Neurosci.* 6, 104, <http://dx.doi.org/10.3389/fnhum.2012.00104>.
- Meiran, N., Pereg, M., Kessler, Y., Cole, M.W., Braver, T.S., 2014. Reflexive activation of newly instructed stimulus-response rules: evidence from lateralized readiness potentials in no-go trials. *Cogn. Affect. Behav. Neurosci.* 15, 365–373, <http://dx.doi.org/10.3758/s13415-014-0321-8>.
- Meiran, N., Pereg, M., Kessler, Y., Cole, M.W., Braver, T.S., 2015a. The power of instructions: proactive configuration of stimulus-response translation. *J. Exp. Psychol. Learn. Mem. Cogn.* 41, 768–786, <http://dx.doi.org/10.1037/xlm0000063>.
- Meiran, N., Pereg, M., Kessler, Y., Cole, M.W., Braver, T.S., 2015b. Reflexive activation of newly instructed stimulus-response rules: evidence from lateralized readiness potentials in no-go trials. *Cogn. Affect. Behav. Neurosci.* 15, 365–373, <http://dx.doi.org/10.3758/s13415-014-0321-8>.
- Meiran, N., Pereg, M., Givon, E., Danieli, G., Shahar, N., 2016a. The role of working memory in rapid instructed task learning and intention-based reflexivity: an individual differences examination. *Neuropsychologia* 90, 180–189, <http://dx.doi.org/10.1016/j.neuropsychologia.2016.06.037>.
- Meiran, N., Pereg, M., Givon, E., Danieli, G., Shahar, N., 2016b. The role of working memory in rapid instructed task learning and intention-based reflexivity: an individual differences examination. *Neuropsychologia*, Memory, Consciousness, and the Brain: A Special Issue in Honour of Morris Moscovitch 90, 180–189, <http://dx.doi.org/10.1016/j.neuropsychologia.2016.06.037>.
- Meiran, N., 1996. Reconfiguration of processing mode prior to task performance. *J. Exp. Psychol. Learn. Mem. Cogn.* 22, 1423.
- Miller, E., Cohen, J., 2001. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202.
- Mohr, H., Wolfensteller, U., Betzel, R.F., Mišić, B., Sporns, O., Richiardi, J., Ruge, H., 2016. Integration and segregation of large-scale brain networks during short-term task automatization. *Nat. Commun.* 7, 13217.
- Monsell, S., 2003. Task switching. *Trends Cogn. Sci. (Regul Ed)* 7, 134–140, PMID: 12639695.
- Muhle-Karbe, P.S., Duncan, J., De Baene, W., Mitchell, D.J., Brass, M., 2016. Neural coding for instruction-Based task sets in human frontoparietal and visual cortex. *Cereb. Cortex N. Y.* 1991, 032–115, <http://dx.doi.org/10.1093/cercor/bhw032>.
- Oberauer, K., Souza, A.S., Druey, M.D., Gade, M., 2013. Analogous mechanisms of selection and updating in declarative and procedural working memory: experiments and a computational model. *Cognit. Psychol.* 66, 157–211, <http://dx.doi.org/10.1016/j.cogpsych.2012.11.001>.
- Oberauer, K., 2009. Chapter 2 Design for a Working Memory, in: Motivation, B.-P. of L. and (Ed.), *The Psychology of Learning and Motivation*. Academic Press, pp. 45–100.
- Palmer, D.C., 2012. The role of atomic repertoires in complex behavior. *Behav. Anal.* 35, 59.
- Power, J.D., Cohen, A.L., Nelson, S.M., Wig, G.S., Barnes, K.A., Church, J.A., Vogel, A.C., Laumann, T.O., Miezin, F.M., Schlaggar, B.L., Petersen, S.E., 2011. Functional network organization of the human brain. *Neuron* 72, 665–678, <http://dx.doi.org/10.1016/j.neuron.2011.09.006>.
- Reynolds, J.R., West, R., Braver, T., 2008. Distinct neural circuits support transient and sustained processes in prospective memory and working memory. *Cereb. Cortex*, 1–14, <http://dx.doi.org/10.1093/cercor/bhn164>.
- Rogers, R.D., Monsell, S., 1995. Costs of a predictable switch between simple cognitive tasks. *J. Exp. Psychol. Gen.* 124, 207.
- Ruge, H., Wolfensteller, U., 2010. Rapid formation of pragmatic rule representations in the human brain during instruction-based learning. *Cereb. Cortex N. Y.* 1991 (20), 1656–1667, <http://dx.doi.org/10.1093/cercor/bhp228>.
- Schneider, W., Shiffrin, R., 1977. Controlled and automatic human information processing: i. Detection, Search, and Attention, 1–66.
- Sporns, O., 2012. From simple graphs to the connectome: networks in neuroimaging. *Neuroimage* 62, 881–886, <http://dx.doi.org/10.1016/j.neuroimage.2011.08.085>.
- Stocco, A., Lebiere, C., O'Reilly, R.C., Anderson, J.R., 2012. Distinct contributions of the caudate nucleus, rostral prefrontal cortex, and parietal cortex to the execution of instructed tasks. *Cogn. Affect. Behav. Neurosci.* 12, 611–628, <http://dx.doi.org/10.3758/s13415-012-0117-7>.
- Theeuwes, M., Liefooghe, B., De Houwer, J., 2014. Eliminating the Simon effect by instruction. *J. Exp. Psychol. Learn. Mem. Cogn.* 40, 1470–1480, <http://dx.doi.org/10.1037/a0036913>.
- Tibboel, H., Liefooghe, B., Houwer, J.D., 2015. Attention to future actions: the influence of instructed S-R versus S-S mappings on attentional control. *Psychol. Res.*, 1–7, <http://dx.doi.org/10.1007/s00426-015-0695-4>.
- Walser, M., Fischer, R., Goschke, T., 2012. The failure of deactivating intentions: aftereffects of completed intentions in the repeated prospective memory cue paradigm. *J. Exp. Psychol. Learn. Mem. Cogn.* 38, 1030.
- Wenke, D., Gaschler, R., Nattkemper, D., 2007. Instruction-induced feature binding. *Psychol. Res.* 71, 92–106, <http://dx.doi.org/10.1007/s00426-005-0038-y>.
- Westbrook, A., Braver, T.S., 2015. Cognitive effort: a neuroeconomic approach. *Cogn. Affect. Behav. Neurosci.* 15, 395–415, <http://dx.doi.org/10.3758/s13415-015-0334-y>.
- Wig, G.S., Schlaggar, B.L., Petersen, S.E., 2011. Concepts and principles in the analysis of brain networks. *Ann. N. Y. Acad. Sci.* 1224, 126–146, <http://dx.doi.org/10.1111/j.1749-6632.2010.05947.x>.
- Woodman, G.F., Luck, S.J., 2003. Dissociations among attention, perception, and awareness during object-substitution masking. *Psychol. Sci.* 14, 605–611.
- Woolgar, A., Thompson, R., Bor, D., Duncan, J., 2011. Multi-voxel coding of stimuli, rules, and responses in human frontoparietal cortex. *Neuroimage* 56, 744–752, <http://dx.doi.org/10.1016/j.neuroimage.2010.04.035>.