Metacognition: A Closed-Loop Model of Biased Competition

Evidence from Neuroscience, Cognition, and Instructional Research

Neil H. Schwartz
California State University, Chico, USA

Brianna M. Scott
University of Indianapolis, USA

Doris Förster
University of Frankfurt, GERMANY

Learning without thought is labor lost.

~ Confucius
Learning and thinking are synergistic actions of the way people develop knowledge to adapt to the world. The actions are collateral cognitive operations that share a unitary outcome of performance. And yet, it is not entirely clear how the operations actually take place—either at the neurological level of the brain, the metaphoric level of the mind, or the action-oriented level of behavior. In the present chapter, we will build a case for metacognition as an integral operator in learning and thought. We will put forth a position that thinking is best characterized by the metacognitive operations learners deploy when they attempt to learn—the planning, monitoring, and evaluating learners do to regulate their learning processes. We will support our case at the level of the neuro-anatomical structures of the brain, the metaphorical architecture of human cognition, and relevant features of instruction. We focus on these three levels because: 1) there is a rich literature on frontal lobe involvement specifically targeted to explain learners’ ability to think and learn; 2) decades of research on human cognitive architecture has been closely examined in the context of the neurological involvement of frontal lobe activation; and 3) learning and thinking are inextricably combined under the auspices of instruction. Finally, we will inventory the role of metacognition in some of our work, and selected works of others.

Differentiating Metacognition from Self-Regulated Learning

Metacognition and self-regulation are not synonymous terms. Individually, the concepts have a long, independent history with distinct theoretical bases (e.g., Flavell, 1979, Bandura, 1977); however, over the last few decades, the concepts have been blurred by inconsistent use and theoretical ambiguity, in addition to the necessary and inevitable revisions the concepts require to evolve theoretically over time. This led Dinsmore, Alexander, and Loughlin (2008) to review 255 articles published over the
last five years, asking the question: “should we expect to hold current generations to the conceptions first framed by Flavell, Bandura, and others, or is it assumed that alternative and contemporary conceptions are warranted?”

Dinsmore et al. (2008) conclude that metacognition is rooted in the theoretical foundation of Jean Piaget and centers around cognition and matters of the mind. Flavell, working from a Piagetian theoretical base, was responsible for conceptualizing metacognition as “thinking about thinking,” (Dinsmore et al., 2008); a definition that still stands 40 years later. Further, metacognition is conceptualized as being comprised of two factors: knowledge—what individuals know about their own cognition and cognition in general, and monitoring/regulation—the set of activities that help students control their learning (e.g., Flavell, 1979; Schraw & Moshman, 1995). Of most importance is the focus on endogenous characteristics (Moshman, 1982)—that is, metacognition is within the realm of the mind with much less concern over the human-environment interaction. Metacognition deals primarily with reflective abstraction of new or existing cognitive structures.

Self-regulation, on the other hand, originates from Bandura’s (1977) writings emphasizing the person-environment interaction, the importance of emotional and behavioral regulation, and the regulation of motivation. In short, Dinsmore et al. (2008) describe self-regulation as “the reciprocal determinism of the environment on the person, mediated through behavior. Person variables include the distinct self processes that interact with the environment through one’s actions” (p.393). Thus, self-regulation consists of the “higher order control of lower order processes responsible for the planning and execution of behavior”, —in addition to emotional control (Banfield, Wyland, Macrae, Munte, & Heatherton, 2004; Efklides, 2006).

We chose to start from Schraw, Crippen and Hartley’s (2006) definition of self-
regulation as consisting of three main components: cognition, metacognition, and motivation. It is within this framework that we examine the overlapping conceptual space between self-regulation and metacognition. We are most interested in the individual’s ability to monitor his or her own thinking, with or without environmental interaction. This monitoring action fits within the “multidimensional conceptual space of self-regulated action” that Kaplan (2008) put forward. This conceptual space, Kaplan (2008) contends, is the abstract “umbrella” under which metacognition and self-regulation stand. The commonalities between self-regulation and metacognition reveal an undeniable conceptual core binding the constructs; namely, that individuals make efforts to monitor their thoughts and actions and to act accordingly to gain some control over them. It is, in effect, a marriage between self-awareness and intention to act that aligns these bodies of work (Dinsmore, et al. 2008). In short, metacognition and self-regulation are not mutually exclusive; rather, they are “subtypes of the same general phenomenon of self-regulated action” (Kaplan, 2008).

The question becomes, then, not what is different between these concepts, but which subcomponent one is interested in studying. Thus, the distinction between self-regulation and metacognition is less important to the conceptual center of this chapter; one’s ability to monitor and control their thinking, regardless of theoretical roots, is of the utmost importance when examining the connection to cognitive architecture and the underlying neurological connections.

**The Neuroanatomy of Executive Control**

Fernandez-Duque, Baird, and Posner (2000) suggested that metacognition could benefit from a cognitive neuroscience perspective where metacognitive regulation is examined in terms of the processes of executive control. The rationale for such a position is based on the work of Shimamura (2000) and others (cf. Bench, Frith,
Grasby & Griston, 1993; Rugg, Fletcher, Chua & Dolan, 1999) who have successfully mapped the concept of executive function onto specific mental operations, anchoring the operations within specific anatomical structures of the brain. Indeed, if Kaplan (2008) is correct that individuals monitor their thoughts and actions—exerting some control over them—then monitoring and control should be evident as neurological activity within known neuro-anatomic locations. Alternatively, describing the activity of specific brain locations implicated in metacognition helps delineate and define specific metacognitive functions.

Nelson and Narens (1990) suggested that metacognitive regulation is principally a coordinating activity made up of both bottom-up and top-down processes—cognitive monitoring, and cognitive control, respectively. Monitoring is responsible for such processes as error detection, attention, and source monitoring in memory retrieval; control is seen in conflict resolution, error correction, inhibitory control, planning, and resource allocation. The coordination is accomplished via a reciprocal influence at two levels of analysis—an object level and a meta-level. Metacognitive monitoring involves the flow of information from the object level to the meta-level where judgments of learning and feelings of knowing are evaluated by the learner; metacognitive control refers to the learner’s regulation of information processing where attention is monitored and cognitive strategies are deployed to manage learning performance. The point is that there is a strong relationship between metacognition in terms of monitoring and control and brain-based executive functions. In fact, there is now “incontrovertible evidence suggesting a trend toward a cognitive neuroscience perspective for many if not all aspects of human cognition” (Shimamura, 2000, p. 320) including metacognitive monitoring and control (Shimamura, 2008). That is, the spatial resolution of event-related fMRI has become
so precise that the ability to identify regions of brain activation has become extremely impressive, allowing for replicable patterns of activation to be observed across laboratories. This means that it is now possible to observe the metacognitive functions that were originally derived from theory, as in-vivo brain tissue activation in the context of behavioral activity within carefully controlled experiments of thinking and learning.

*In the neurosciences, the higher order processes of metacognitive monitoring and control are referred to as “executive control processes”*. We now know that separate, albeit interactive, frontal areas of the brain are critically involved in these processes (c.f. Mega & Cummings, 1994; Pannu, Kaszniaik & Rapcsak, 2005). In fact, recent evidence from neuroscience has lead to the conclusions that: 1) there is a strong correlation between indices of frontal lobe structural integrity and metamemory accuracy; and 2) the combination of frontal lobe dysfunction and poor memory severely restricts metamemory processes. The term metamemory is used here to note the synergistic effect of monitoring and control processes on successful memory functions. Specifically, patients with damage to the frontal lobe show impairments in metacognitive monitoring associated with feelings of knowing an answer before the answer is given, and evaluation—the kind of evaluation in which learners must evaluate contextual information such as remembering when or where some event occurred or who presented the information (Nolde, Johnson & D’Esposito, 1998; Rugg, Fletcher, Chua & Dolan, 1999).

Other judgments of feelings are also good indices of metacognitive monitoring; for example, feelings of knowing judgments, ease-of-learning, tip-of-the-tongue feelings and retrospective confidence judgments, and global predictions and post dictations. Indeed, all these indices have been used to provide evidence with
neurological patients that the prefrontal cortex is an essential region for performance, and reflects the *monitoring* function of metacognition. For example, in patients with Korsakoff’s syndrome, Moscovitch and Melo (1997) found that frontal lobe lesions or dysfunction results in a common occurrence of confabulation, because of a breakdown in search mechanisms and poor metacognitive monitoring. Schnyer et al. (2004) had learners with specific frontal lobe damage learn sentences and make judgments of feelings of knowing and retrospective confidence of the last word in each of several sentences. The learners performed poorer than normal controls on their feeling of knowing judgments. In fact, lesion analysis revealed an overlapping region of the right medial prefrontal cortex in the learners with frontal damage who performed the poorest on the task. Finally, Pannu, et al. (2005) examined differences in performance between patients with frontal lobe damage and healthy controls during a learning task in which the participants were asked to make feeling-of-knowing and retrospective confidence judgments in a face-name retrieval task. The two groups performed similarly when the faces were either extremely familiar or extremely unfamiliar, but quite different when the faces were of intermediate familiarity. Pannu et al. (2005) explained that the patients with damage to the right ventral medial prefrontal cortex monitored more poorly, suggesting that the “monitoring mechanism is engaged most critically when decisions are difficult” (p. 112).

By the same token, executive functions of the frontal cortex are involved in metacognitive *control* as well. Nagel (2009) found that high-level reasoning is an index of strategic behavior controlled by neural activity in the medial prefrontal cortex when learners believe they are controlling their cognition in the presence of a human rather than a machine. McGlynn and Kaszniak (1991) observed impairment in
metacognitive control process associated with the allocation of time when learners with Huntington’s disease had to search memory for answers to general information questions. Huntington’s Disease is an inherited degenerative disorder in which dysfunction exists in the frontal-subcortical circuits of the brain (Cummings, 1994).

Finally, metacognitive control has been observed in neuroimaging studies of the Stroop effect, where learners must resolve the conflict between the name of a color and the color in which the name is printed by inhibiting an incorrect response when the word and its color are incongruent. The neuroimaging data consistently reveal activation of the anterior cingulate within the prefrontal cortex (cf. Carter, Mintun, & Cohen, 1995).

In gist, Fernandez-Duque, Baird and Posner (2000) summed up the neurological evidence this way:

“Neuroimaging studies have shown activation of a network of frontal areas in tasks of executive control. The activated areas usually include the anterior cingulated and supplementary motor area, the orbitofrontal cortex, the dorsolateral prefrontal cortex, and portions of the basal ganglia and the thalamus. The tasks that activate these areas typically require subjects to deal with conflict, error, or emotion, therefore demanding effortful cognitive processing (Bush et al., 1998; Bush, Luu, & Posner, 2000). These mental abilities may be the building blocks that metacognitively-sophisticated thinkers use in their achievement of complex tasks, such as problem solving, strategy selection, and decision making.”

Based on the evidence above, we believe that knowledge of the neurological underpinnings of metacognition is important because it leads to testable hypotheses of
instruction. Consider recent work by Fugelsang and Dunbar (2005) on conceptual change. Fugelsang and Dunbar used fMRI to investigate the patterns of neurological activation when students were acquiring new scientific knowledge. The question was whether the students would change their relatively naïve understanding of scientific concepts when given new information either consistent or inconsistent with theory plausible to their previously held beliefs; the second question was whether different parts of the brain would be activated under the two plausibility conditions. What the researchers found bears directly on instruction. When given data consistent with their previous scientific understanding, the students showed activation of neural networks in the caudate and parahippocampal gyrus (C and PG)—networks well known to be involved in learning. However, when presented with data inconsistent with their previously held beliefs, activation was seen in the anterior cingulated cortex and the dorsolateral prefrontal cortex, with no activation of the C and PG. This suggests that when new information fits in well with information students already know, learning networks are activated, but when information does not make good sense in terms of students’ existing knowledge, students activate neural networks that actually inhibit the development of new learning. The finding has implications for our purposes here because it attests to the influence of students’ executive control on the ways teachers and instructional designers might approach their delivery of instruction.

**Working Memory: The Link Between Metacognition and Executive Control**

One of the best ways to make sense of neurologically based executive processes in terms of metacognitive monitoring and control is to examine both levels in the context of a single model. After all, a single model permits each to be explained relative to the other using a framework common to both. We chose Baddeley’s model of working memory (Baddeley, 2003) for such a purpose, because the model has had
a substantial influence in generating research of human cognitive processing. Indeed, Jonides, et al. (2008) pointed out that “between the years 1980 to 2006, of the 16,154 papers that cited ‘working memory’ in their titles or abstracts, fully 7339 included citations to Alan Baddeley.” (p. 195).

Baddeley’s model of working memory (Baddeley, 2003) is an extension of the tripartite model of human cognitive architecture originally proposed by Broadbent (1953) and later developed by Atkinson and Shiffrin (1968). Designed to explain the dynamic functions of in-vivo thinking, the model of working memory can be used to account for the executive control processes investigated within the neurosciences and the concept of metacognition evolving from studies of cognition and cognitive performance. *In short, we outline the model here because it is an effective framework with which to map the overlap of each and explain the importance of metacognition in instruction.*

Working memory, as described by Baddeley (2000), is a four-component model comprised of two slave systems—the visuospatial sketchpad and the phonological loop— an episodic buffer, and a central executive. The sketchpad is assumed to hold visuospatial information for further processing and is believed to be fractionable into separate visual, spatial, and possibly kinesthetic components. The phonological loop is assumed to hold verbal and acoustic information using a temporary store and an articulatory rehearsal system. The episodic buffer is postulated to be a limited capacity system providing temporary storage of information in the form of multimodal codes, and capable of binding information from the other components, and from long-term memory, into a unitary episodic representation. Finally, the central executive is conceived as the part of the model capable of “retrieving information from the episodic buffer in the form of conscious awareness,
metacognition: a biased competition model

reflecting on that information and, where necessary, manipulating and modifying it”, (Baddeley, 2000, p.420). Baddeley, Allen and Hitch (2010) contend that executive control is at the “heart of working memory” (p. 223).

The Role of the Central Executive in Working Memory

Baddeley’s (2000) concept of the central executive implicates the two hallmark features of metacognition; namely, monitoring and control. In effect, the central executive was postulated to be the component responsible for determining whether attention is necessary for deployment under conditions when a person is required to learn, solve a problem, or act in an unfamiliar way. Routine actions such as reciting the alphabet or driving a car are automatic and place only a light demand on attention; but, when routine action is impossible, a supervisory attention system (SAS) (Shallice, 1988) is probably deployed capable of reflecting on alternative plans of action and biasing behavior in the direction of the actions most likely to lead to a goal.

However, there was a problem with the central executive as originally explained using the SAS. The central executive was conceived purely as an attentional allocation and deployment system, but evidence from a number of investigations suggested that this could not entirely be the case. Data from studies examining people’s capacity to focus attention, divide attention between two or more sources, switch attention between tasks, and link information between working- and long-term memory failed to entirely support the central executive in this capacity (Baddeley et al., 2010). Investigations did support the attentional focus function (c.f. Logie, Gilhooly & Wynn, 1994). Empirical results also supported the assumption that the central executive was likely responsible for dividing attention between two or more sources, (Logie, della Sala, Wynn & Baddeley, 2000). But, the other two
functions were not unequivocally supported by research. Specifically, task switching seemed to be better considered a result of a number of different processes rather than a single executive process (Mayer & Evans, 2001; Saeki & Saito, 2004); and, the linking function was probably better conceptualized in terms of a working memory component entirely different than the central executive. Thus, Baddeley et al. (2010) proposed an episodic buffer, the nature of which they described as “a buffer in the sense that it is a limited capacity temporary store that forms an interface between a range of systems all having different basic memory codes; having a multi-dimensional coding system; [and] episodic in the sense that it is capable of holding episodes, and integrating chunks of information that then became accessible to conscious awareness” (p.229).

And yet, there are serious questions as to how the episodic buffer functions in conjunction with the central executive, whether the episodic buffer and the central executive are clearly responsible for different cognitive functions, whether the two can be anchored in different or complementary neurological functions of the brain; and whether metacognition can be explained in terms of both components at both a neurological and cognitive level. It is certainly conceivable that the central executive may be responsible for metacognitive monitoring, and the episodic buffer may be responsible for metacognitive control. If this is true, then brain activation associated with the central executive might be expected to be principally attentional, and brain activation associated with the episodic buffer might be based on composite operations of specific brain systems acting to control the integration of information.

Implicating the Episodic Buffer in Metacognitive Control

Repovs & Baddeley (2006) postulated that the episodic buffer was the working memory component responsible for creating and manipulating novel
representations, creating a mental modeling space that enables the consideration of possible outcomes, and provides the basis for planning future action. Thus, the episodic buffer would seem to be the part of working memory responsible for more integrative processing during learning.

However, if the episodic buffer were the section where information is integrated (where “binding” takes place), two things would have to be evident. One, the central executive and the episodic buffer sections should have relatively independent actions on information during processing; and, two, the brain regions activated for the two working memory sections and their respective actions (e.g. attention and binding, respectively) should be different. Unfortunately, neither the first nor the second condition appears to be the case. The two components do not appear to have entirely independent actions, (Chein & Fiez, 2010), the actions are not contained in separate and unrelated regions of the brain (c.f. Baddeley, Allen & Hitch, 2010), and evidence from functional neuroimaging studies provide little, if any, support for the buffer’s binding function (Allen, Baddeley, & Hitch, 2006; Rossi-Arnaud, Pieroni & Baddeley, 2006). Instead, there is substantial evidence that the “concept of specialized buffers do not adequately map onto neural architecture at all. Findings appear more consistent with a system in which active maintenance involves the recruitment of the same circuitry that represents the information itself, with different circuits for different types of information” (D’Esposito, 2007, p. 764).

In addition, compartmentalization of working memory into components consisting of a central executive and an episodic buffer is based on research that is very complex and hotly debated; and, there is evidence that, as sovereign entities, there is no need for a central executive and episodic buffer to actually exist. As Rawley and Constantinidis (2009) explained, it is true that Baddeley’s working
memory components refer to functional rather than anatomic units, but there should be a functional neurology to the subsystems, and there does not appear to be. In fact, with regard to the central executive, the prefrontal cortex should be most actively involved when the central executive is activated, if the central executive is responsible for controlling and coordinating information through the slave systems and the episodic buffer. However, “physiologic evidence indicates that prefrontal neurons in area 46 represent spatial and object attributes of visual memoranda in correspondence to the visual-spatial sketchpad (Rao, Rainer & Miller, 1997) while at the same time providing neural correlates of executive functions such as rule execution and category classification (Freeman, Riesenhuber, Poggio & Miller, 2001; Wallis, Anderson & Miller, 2001)” (Rawley & Constantinidis, 2009, p. 133). The problem is apparently the same for the other subsystems. All of the subsystems appear to activate multiple brain areas, including both the prefrontal and anterior cingulate cortices (Smith and Jonides, 1999).

Thus, an episodic buffer is not a utilitarian concept with which to explain metacognitive monitoring or control.

**Implicating the Central Executive in Metacognitive Monitoring**

According to Repovs and Baddeley, (2006), the central executive has always been the “most important but least understood and least empirically studied component of the multi-component working memory model” (p. 12). However, based on a careful and exhaustive review of the evidence, Baddeley and his colleagues also contended that, “in complex cognitive abilities, the central executive seems to be mostly involved as a source of attentional control, enabling the focusing of attention, the division of attention between concurrent tasks, and as one component of attentional switching” (Repovs & Baddeley, 2006, p.14-15). Thus, the role of the
central executive seems to be the functional component of working memory principally responsible for the allocation, deployment, and maintenance of attention during learning.

There is also evidence that the central executive is predominantly responsible for attentional processes. In the Time-Based-Resource-Sharing (TBRS) model proposed by Barrouillet and Valérie (2010), information in working memory is maintained by a rapid switching between brief processing and storage, allowing for memory traces to be constantly refreshed by attention. Raye, Johnson, Mitchell, Greene, and Johnson (2007) point out that this recursive refreshing involves the left dorsolateral prefrontal cortex. The same is true for the embedded processing model of working memory (Cowan, 1999). While structurally different from Baddeley and colleagues’ four-component model, and wherein a central executive is not postulated per se, the embedded processing model nevertheless does postulate the operation of a central controller. Most importantly, the controller is purported to supervise the preservation of information in working memory by iteratively subjecting it to a recursion of attentional focus (Cowan, 1999)—a reactivation strategy that Lewandowsky and Oberauer (2008) refer to as “attentional refreshing.” Finally, Chein and Fiez (2010) provide neuroimaging and corroborating behavioral evidence to support the central controller.

Thus, attention is manipulated by some sort of attention controller—a controller that is moderated by the individual, necessary for other cognitive processes to be deployed, and grounded in areas of the brain known to be involved in attentional focus. As Barrouillet and Valérie (2010) point out: “processing most often requires the selection, activation, and maintenance of goals and sub-goals, the selection of relevant information, the retrieval from long-term memory of related items of
knowledge, the planning and monitoring of adapted strategies, and response selection, *all activities known as requiring attention*” [italics added] (p. 356). Thus, the evidence suggests that the allocation, deployment, and maintenance of this attention are executive functions grounded in the neurological activity (principally, but not exclusively) of the prefrontal cortex. If attention can be assumed to be integral to monitoring per se, then the central controller is likely responsible for what is postulated to be the monitoring function of metacognition.

At the same time, most agree that attention directed top-down (e.g. internal representation to behavioral action) is based on information held in working memory (cf. Bundesen, 1990). Thus, from a top-down approach, using visual stimuli as an example, Lavie and colleagues (Forster & Lavie, 2007; Lavie, 2005) demonstrated that one’s ability to filter out irrelevant stimuli during selection of visual stimuli depends on the processing load in working memory. As the load increases, fewer resources are available to support the efficient selection of targets relative to the rejection of distracters. The net effect is an increase of the interference from distracters under conditions of high working memory load. On the other hand, there is a decrease of the interference effect of distracters even when the complexity of a visual display is high. This suggests—based on a biased competition model of attention—that the allocation of attention is selectively deployed (Desimone & Duncan, 1995). Stimuli compete for selection at multiple levels of representation, with the winner gaining control of both perceptual and response systems. Thus, working memory acts to bias the competition for attention to favor objects that fit the goals of the task. Soto, Hodsoll, Rotchtein and Humphreys (2008) suggest that prefrontal cells are implicated in this attention-biasing effect by being involved in prioritizing the relevant goals for tasks. Thus, “attentional refreshing”, “recursion of
attentional focus”, “biased competition”, and other ways of describing the allocation of attention is a voluntary process of cognitive engagement learners use to *monitor* the deployment of other cognitive processes.

**Implicating the Central Executive in Metacognitive Monitoring and Control**

The evidence above suggests that the prefrontal cortex is of critical importance in the monitoring function of the central executive. However, without an episodic buffer, it must have a controlling function as well—to be able to control “when behavior must be guided and controlled by internal states and intentions, when automatic responses have to be suppressed, and when tasks require the establishment of new or rapidly changing mappings between perception and action” (Wolters & Raffone, 2008, p. 2). Indeed, the prefrontal cortex is well positioned to coordinate processing in the rest of brain because it is strongly interconnected with reciprocal connections to virtually all other neocortical and subcortical brain regions. (Constantinidis & Procyk, 2004; Rawley & Constantinidis, 2009).

The evidence from neuroscience suggests that the prefrontal cortex *does* control behavior, but it does so by *modulating* rather than simply *transmitting* neural impulses.

According to Wolters and Raffone (2008) “simple adaptive behavior rests on a cycle of perception, action, and perception-of-action results,” but the prefrontal cortex allows an “internalization of this loop, freeing the organism of the restrictions of being aware of, or acting upon, physically present objects or situations only.” This means that the prefrontal cortex can orchestrate other brain regions in the manipulation of internal representations, independent of the present environment; it can maintain physically absent information in an active state by recurrent connections between itself and the rest of the cortex; and, it can redirect actions of monitoring,
attention and control by activating and or inhibiting particular motor programs. More importantly, its capacity for recurrent connections with memory systems, in addition to its mechanisms for combining information within neural loops, allows for the formation and updating of future goal states, and ways of achieving them. Taken together, the functions of the prefrontal cortex are clearly involved in cognitive control, relative to the regulation and influence of other brain regions.

But, the prefrontal cortex is responsible for cognitive control by virtue of three interdependent functions—maintenance, attentional control, and integration (Wolters & Raffone, 2008)—the same functions erroneously believed to be associated with an episodic buffer. Maintenance refers to the process of actively holding a limited amount of task-relevant information supplied by a preceding event; attentional control is the top-down selective activation of the representations of task-relevant stimuli and their corresponding responses; integration is the combination and reorganization of information from different sources in the service of controlling the execution of a task. Maintenance is the result of neurological patterns of activation borne from specific external inputs oscillating in a recurrent loop between multiple networks of prefrontal and other cortical cells in regions of the brain that are specialized for the nature of the input (Ranganath et al., 2004). Attentional control seems to operate in a biasing and competitive fashion where neuronal responses of the prefrontal cortex bias neuronal responses in posterior parts of the brain, creating a competition of activation and suppression for the task-relevant and task-irrelevant stimuli, respectively, required for task performance (Miller & Cohen, 2001). Integration appears to be a hierarchically arranged deployment of control, cascading down from super-ordinate prefrontal cortical modules specialized for large-scale integration, to subordinate modules that are relatively specialized for processing simple tasks.
(Koechlin, Ody & Kouneiher, 2003).

As D’Esposito (2007) explained it, there appear to be at least two types of these top-down signals—one that serves to enhance, and another that serves to suppress, task-relevant information. Both are important for our discussion here because enhancement and suppression mechanisms may actually exist to control both cognitive and metacognitive functions (Knight et al. 1999). After all, it is well documented that excitatory and inhibitory mechanisms are pervasively interleaved throughout the nervous system, in spinal reflexes, cerebellar outputs, and basal ganglia movement control networks, etc.—indeed, at multiple levels throughout the entire neuroaxis. That means “by generating contrast via both enhancements and suppressions… top-down signals bias the likelihood of successful representation of relevant information in a competitive system” (D’Esposito, 2007, p.768). In short, the top-down function and the biasing effect within the context of a competitive system could be a compelling way to think about a neurological explanation of metacognitive monitoring and control.

**Working Memory and Metacognitive Monitoring and Control**

Based on the evidence above, we conclude that it is not necessary to involve a central executive and episodic buffer as two distinct components of the working memory system to explain metacognition. Rather, it is necessary only to implicate a central executive controller of some kind that regulates attention and deploys operations of activation and suppression of internally stored and externally perceived input to reach a behavioral goal. In short, metacognition is certainly “in the brain,” but it is not in the central executive and episodic buffers of Baddeley and his colleagues’ working memory model.

So, just where would metacognition likely be?
We suggest that metacognition is manifest within the function of cognitive—and hence, neuroanatomic—activity of the brain best represented by the model of embedded processes (Cowan, 1999), the operations of which we have described in the evidence above (see Figure 1).

*Insert Figure 1 here*

To be specific, metacognitive monitoring and control is probably a reciprocal function of the same neurologic processes that excite and inhibit, in a recursive fashion, the regions of the brain responsible for two types of activities involved in learning—the activities involved in processing the information itself relative to the goals of a task—and, the activities involved in processing (evaluating and correcting) the original activities deployed to seek goal attainment—activities that are metacognitive. We believe that the monitoring function is probably principally attentional, and the control function is principally strategic. Thus, attention is probably allocated to evaluate the degree to which an individual is closer to the goal—a matching-to-sample function; the strategies are activated to change the person’s processing approach (and hence the corresponding brain activation) in meeting the goal. This alternating procedure is probably an interleaved activation of excitatory and inhibitory mechanisms based on two sources of information and two sources of goals, exchanged in a recursive fashion depending upon the degree to which the goal is being met. One source of information is composed of the stimuli that comprise the task in the context of the original task demands; the other is the information composed of the internal representation of the assessment of the correspondence between task demand and task success and the information about effective strategies for obtaining the success. In effect, we suggest that there may be no difference in the mechanisms operating between cognitive and metacognitive
processing when one considers activation of regions of the brain. Instead, it is the nature of the information being processed in the system that differentiates the two.

We believe the operations of active cognition and metacognitive monitoring and control probably look something like the patterns shown in Figure 2. That is,

*Insert Figure 2 here*

learners begin the process of learning by first directing their attention to two types of external information. One is the to-be-learned material; the other is the learning goal—in essence, the instructions with which the to-be-learned material is to be processed. This directed attention is an operation of the central controller where the learner seeks to differentiate between task-relevant and task-irrelevant stimuli in the external learning environment in order to find the stimuli having the highest probability of further processing utility. Once the differentiation is made, the central controller maintains attention on the task-relevant stimuli and the learning goal while concomitantly redirecting attention to the internally stored stimuli activated among the modules and cells distributed in the posterior cortex that are germane to the task. At the same time, while attention is being switched between directing and maintaining the two sources of stimuli—external and internal—the central controller is also engaged in the integration of these stimuli into new knowledge models—one model of the to-be-learned information, and the other model representing the learning goal.

In short, the direction of attention among relevant stimuli inside and out, the maintenance of that attention on the stimuli inside and out, and the integration of those stimuli inside and out is an oscillating process of active cognition which leads to the development of the new models of knowledge.

However, the function of active cognition in this oscillating process is improved when learners report competence in their general use of metacognitive skills
and when metacognitive activities are activated and supported within learners during the learning process. Thus, the question becomes how metacognition can possibly operate when active cognition seems to be sufficient for processing, but metacognition enhances performance beyond the outcomes of the tightly interleaved active cognition operations.

Our position is that metacognition must be comprised of the same processes as active cognition, but with attention and integration allocated to a different source of information. That source is no longer exclusively external per se, but rather originates from the internal cognitive environment instead—from the new model of the task-relevant information borne from the information intended to be learned, and the personalized model of the learning goal that was constructed during active cognition. This suggests that active cognition and metacognition form a closed loop, where metacognition is comprised of the same operations of the central controller in the prefrontal cortex and the posterior cortex’s activation of relevant models and cells, but at this point, oscillating to construct a knowledge model of learning goal attainment. That means that the cognitive system must negotiate, by attentional switching, three models of knowledge during the learning phase—one model that targets the to-be-learned material, one that targets the learning goal, and the third that monitors and controls whether the first two models are sufficiently formed to reach the learning goal.

Once the process begins, the system must balance attentional direction and maintenance between both sources of information along with an integration of each. Thus, active cognition and metacognition use the same resources, activate the same gross neuroanatomical regions, and balance the same operations to resolve the learning process until the final metacognitively constructed knowledge model of goal
attainment is complete.

**Metacognition and Learning**

If we are correct in our appraisal of the way metacognition works, then we should be able to interpret why manipulations of metacognition may, or may not, be instructionally successful. After all, we stated earlier that students’ executive control should exert an influence on the ways teachers and instructional designers might approach their delivery of instruction.

The literature on metacognition in learning and instruction is substantial. Thus, we sampled 28 investigations published between 2002 and 2009, with the intention of building a corpus of work from which to determine if our model is heuristically valuable. Half of the investigations addressed the degree to which learners actually deploy their metacognitive skills during the time in which they learn; the other half reported the results of conditions in which metacognitive operations were instructionally scaffolded. All of the investigations were situated in learning environments that were delivered on computers via arrangements of hypermedia.

What we generally discovered is that learners learn more when their metacognitive skills are well developed. The finding occurs with surprising regularity and is consistent across multiple types of manipulations of learning materials (c.f. Azevedo, 2005; Graesser, McNamara & VanLehn, 2005; Hartley & Bendixen, 2003; Schwartz, Anderson, Hong, Howard & McGee, 2004; Schwartz, Oppy & Gust, 1999; Scott & Schwartz, 2007; Veenman, Prins & Elshout, 2002). For example, Graesser et al (2005) noted that there are well-documented difficulties among learners when they do not possess adequate proficiencies in metacognitive skills; poor inquiry learning behavior and lower levels of comprehension characterize the difficulties. Azevedo (2005) reported that students who lack key metacognitive skills learn very little from
hypermedia when learning environments are open-ended. And, Hartley and Bendixen (2003) found that learners make better use of comprehension aids during learning within hypermedia environments, but only when the learners possess metacognitive skills that are high.

There are other supporting investigations as well. Veenman et al. (2002) could predict learners’ acquisition of high quality conceptual knowledge from the degree to which the learners had effective metacognitive skills; and, Scott and Schwartz (2007), and Schwartz et al. (1999; 2004) found that learners could navigate more effectively within, and learn more from, hypermedia environments when the learners’ metacognitive skills were high. Thus, learners with better-developed metacognitive skills do, in fact, learn better, and the evidence is apparent among multiple indices of performance.

And yet, not all learners have sufficiently well developed metacognitive skills (cf. Bannert, 2006), nor do all learners actually deploy those skills even if the skills are well developed (cf. Azevedo, Guthrie & Seibert, 2004). Manlove, Lazonder, and de Jong (2007), for example, found that learners typically show very few instances of metacognitive regulatory control operations during inquiry work in computer-based learning environments. Azevedo and Cromley, (2004) and Bannert, Hildebrand and Mengelkamp (2009) reported that learners rarely use metacognitive monitoring when negotiating complex hypermedia learning environments, in addition to failing to plan or activate their prior knowledge or use other effective knowledge acquisition strategies that would benefit their performance. Finally, Azevedo and Hadwin (2005) found that when learners learn about complex topics in computer-based learning environments without external metacognitive supports, their use of metacognitive control operations is very poor and they fail to gain a conceptual understanding of the
target instructional topics. Thus, the evidence is very clear that the failure to use metacognitive skills results in poor learning performance.

But, the failure to deploy metacognitive skills is puzzling when it is clear that metacognition works to benefit learning performance. Thus, we questioned whether metacognitive operations scaffolded during instruction actually lead to better performance. Scaffolded instruction refers to the types of computer-based tools designed to detect, trace, monitor, and foster metacognitive skills (Azevedo, 2002); they are the human or non-human learning agents whose roles are designed to lead learners to strategic learning activities that result in better performance (Azevedo, Cromley & Seibert, 2004).

According to the preponderance of the empirical evidence, scaffolding does work (c.f. Azevedo & Cromley, 2004; Azevedo & Hadwin, 2005; Azevedo & Jacobson, 2008; Bannert, 2006, 2009; Graesser, et al. 2005; Manlove et al. 2007). When learners’ metacognitive processes are augmented via computer-based training systems and instructional strategies, metacomprehension accuracy and transfer task performance improves (Cuevas, Fiore, Bowers, & Salas, 2004). Also, learners learning with reflection prompts in computer-based learning environments show better performance on transfer tasks, and make navigation decisions that are more strategic (Bannert, 2006). Finally, Rouet and Le Bigot (2007) found evidence that training learners on meta-textual knowledge, hypertext navigation strategies, and methods to acquire problem-relevant information leads the learners to spend more time visiting relevant sections of the hypertext and write better essays containing more critical and more deeply processed information. The point is that teaching, prompting, and facilitating learners’ use of metacognitive skills results in improved learning performance.
So, why is it that some learners do not seem to develop metacognitive skills; why do some learners fail to deploy the skills on their own even when the skills have already been developed; and, why is it that scaffolds work to incur skill deployment?

We believe that the resources incurred to develop and deploy metacognitive skills are demanding. After all, learners must construct two knowledge models as we described above. Then, they must broker those models with the development of a third—the metacognitively constructed model built to negotiate learning goal attainment. This forces attentional direction and maintenance to be split across three large knowledge models of separate but related domains. It also forces integration within and between the three models—a heavy resource-consuming task of concomitant cognitive and metacognitive operations. Unless one or more of the models is well consolidated among the modules distributed in the posterior cortex, it is not at all surprising that learners economize their efforts in building any one of the three. Since the metacognitive model is, by definition, always secondary to the other two, it is quite likely that the model either does not initially get built, or more likely, is built incompletely. In the first case, it would necessarily fail to be deployed; in the second case, its partial construction would occlude and/or seriously compromise the construction of either or both of the primary cognitively constructed models. This would explain why learners: 1) learn very little from hypermedia when the environments are open-ended, 2) navigate within those environments more inefficiently, and 3) fail to plan or activate prior knowledge and other effective knowledge acquisition strategies that would benefit their performance. In short, the inadequacies become apparent because the monitoring and control processes comprising the learning-goal attainment model never get adequately constructed. If construction is attempted, on the other hand, the central controller would be expected
to be overtaxed and fatigued, accounting in part for the failure of learners to gain a conceptual understanding of target instructional topics.

This is exactly the reason why we believe metacognitive scaffolds actually work. When metacognitive scaffolds are available, learners make use of the scaffolded strategies. These scaffolds either incur the construction of a metacognitive model while relieving the central controller’s attention allocation and maintenance function of the controlling processes themselves, or they provide learners the learning control functions during primary knowledge model construction. In either case, both would lead to better performance on transfer tasks, better and more strategic navigation decisions in hypermedia, and more critically and deeply processed information following the learning phase.

Concluding Remarks

We began this chapter by stating that learning and thinking are synergistic actions of the way people develop knowledge to adapt to the world. We suggested that the actions are collateral cognitive operations sharing a unitary outcome of performance, with metacognition functioning as an integral operator. Assuming that metacognition is within the realm of the mind with much less concern over the human-environment interaction, we sought to find where metacognition might be operating at two levels—one, the metaphorical level of cognition, and two, the neuro-anatomical level of the brain. At the level of cognition, we discovered that metacognition does not fit well within the model of working memory described by Baddeley and his colleagues (e.g. Baddeley & Hitch, 2000). Instead, it is much better explained by the embedded processing model of Cowan (Cowan, 1999). At the level of neuroanatomy, based on an examination of the neurological processes forthcoming from fMRI research, we discovered that metacognition seems to be a reciprocal
function of the same neurologic processes that reciprocally excite and inhibit the regions of the brain responsible for the activities involved in processing the to-be-learned material relative to the goals of a task, and the activities involved in evaluating and correcting the original activities deployed to seek goal attainment. Taken together, both the embedded processing model and the neuroanatomical functions underlying it, lead to the conclusion that processes of metacognition and active cognition form a closed loop of operations occurring in the same areas of the brain. The construction of cognitive models (and hence neural processes) is derived from a biased competition of limited resources that lead to new learning.

References


Figure 1. Interdependent top-down function of cognitive control
Figure 2. Metacognitive monitoring and control: A closed neuro-cognitive loop