10

From Neo-Behaviorism to Neuroscience: Perspectives on the Origins and Future Contributions of Cognitive Load Research

RICHARD E. CLARK AND VINCENT P. CLARK

HISTORICAL PERSPECTIVES ON COGNITIVE LOAD RESEARCH AND THEORY

European and American psychology may have developed in a way that prevented or delayed the development of Cognitive Load Theory (CLT) until George Miller's (1956) classic paper on working memory capacity appeared a half century ago (Miller, 1956). At the beginning of the twentieth century and fifty years before Miller's paper kick-started the field of cognitive science, Charles Hubbard Judd (1908) lost an important argument with Edward Thorndike (1903) about the role of mental effort in the transfer of learning. The loss helped to sidetrack psychology into emphasizing behaviorism over cognitive processing. Judd, an American who was Wilhelm Wundt's student in Leipzig at the end of the nineteenth century, hypothesized that internal cognitive processes and external instructional strategies supported the mental work necessary to transfer knowledge between different problem contexts and settings. Judd had learned from Wundt to emphasize a version of scientific psychology that favored the study of consciousness, problem solving, thinking, and sensations. Judd's (1908) famous bow and arrow experiment demonstrated that effortful cognitive processes could support the generalization of a principle about the diffractive properties of water and so allow people to adjust their aim with the bow to hit an underwater target

that appeared to be somewhere else. Thorndike (1903) focused his research on animal maze learning and proposed an "identical elements" transfer theory, arguing that it was positive reinforcement that led to learning and transfer – and not cognitive processing. Because Thorndike was a student of the powerful William James, who supported his work, Judd's theory and evidence were largely ignored.

William James's support for Thorndike's view of transfer marked a turning point in psychology. James's earlier work had emphasized the role of mental effort in cognition when, for example, he described attention as "the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought. Focalization and concentration of consciousness are of its essence. It implies withdrawal from some things in order to deal effectively with others" (James, 1890, pp. 403–404). In an 1898 lecture that mirrors some of the arguments made recently about the possible evolutionary selection advantage offered by limitations on working memory by John Sweller (Chapter 2), James gave a series of lectures at Johns Hopkins University in which he claimed that consciousness had an evolutionary function or it would not have been naturally selected in humans. A few years later, James (1904) reversed himself and expressed strong misgivings in an article titled "Does 'Consciousness' Exist?" Judd (1910) later protested and argued for a selection bias for consciousness, but at the same time, Thorndike (1903) and others were more successfully arguing that learning was "not insightful" but instead was incremental. Thorndike's claim essentially denied any important role for consciousness or working memory in learning or problem solving.

A number of historians have proposed that the transition in psychology during James and Thorndike's era was due in large measure to an increasing interest by the American public in the development of the physical and biological sciences and a distrust of the introspective approach in philosophy and imprecise psychological research methods. This may have been the reason that American psychologists such as James, Thorndike, and others at that time were attracted to the learning research of 1904 Nobel Prize winner Ivan Pavlov and supported the use of animal experiments and the careful control of observable and measurable events favored in medical research. This exclusive focus on animal learning and connectionism was not reflected in European psychology, where researchers continued to be concerned with experimental work as well as introspection, Gestalt studies of consciousness, physiology, experimentation, and case study methods. The more flexible approach taken by European researchers may be the reason many of the prime movers in CLT have been trained in the

European psychological tradition. The irony is that behaviorism resulted in important advances in measurements, the specification of instructional method variables, and precise experimental methods while it discouraged hypotheses based on cognitive processing during learning and transfer. It also became increasingly obvious that behaviorism focused primarily on motivation to learn through reinforcement and emphasized very simple forms of learning. That recognition eventually made it possible for neo-behaviorists to hypothesize internal cognitive processes to explain complex learning.

One of the very early attempts to deal with complaints that behaviorism only focused on simple learning tasks was the neo-behaviorist research on complexity by Canadian psychologist Daniel Berline (1960). In the 1960's, information processing theory was developing, and Berline offered a model for representing cognitive stimulus and response bonds to describe the cognitive processing required for handling uncertainty and novelty. He proposed a method of measuring individual uncertainty about any stimulus and hypotheses that guided research on the relationship of problem uncertainty and learning. His internationalism and his neo-behaviorist theories made early attempts at cognitive science more acceptable to behaviorists in North America. During this time, cognitive science was developing slowly, forced to swim upstream against powerful behaviorists who resisted change. In addition to Miller's (1952) classic "Magical Number Seven" article, Ulric Neisser's (1967) book *Cognitive Psychology* also had a major impact on the development of CLT. Neisser proposed a computer processing metaphor for cognition and urged psychologists to study the function of working memory in daily activities. Although many cognitive psychologists now avoid the restrictive computer metaphor for cognition, educational psychology benefitted from the analogy during a formative stage. A decade after Neisser's book was published, the article by Schneider and Shiffren (1977) on controlled and automated processing had a huge impact on our view of complex learning, memory, and problem solving. With these events in the background, a decade later, John Sweller's (1988) article in *Cognitive Science* laid the groundwork for CLT.

An important lesson to be learned from the history of psychology is that education and psychology must permit more diversity in theoretical and methodological approaches. With a more interdisciplinary approach, we might have started to develop CLT a half-century earlier and so would have been considerably more advanced at this point. Yet, it may also be the case that one of the benefits of the historical delay caused by the dominance of behavioral theories was the development of a clear focus on pragmatic

instructional research. Behaviorists such as B. F. Skinner encouraged psychologists to conduct careful instructional research in schools. CLT researchers have retained the behavioral focus on instruction and as a result, CLT has made significant contributions to instructional design.

CLT CONTRIBUTIONS AND CHALLENGES TO INSTRUCTIONAL DESIGN

An emphasis on the application of research findings to instruction requires that we understand the conditions necessary for selecting and implementing the most efficient and effective instructional design for different learning tasks, learners, and delivery media. This decision has worked to the benefit of instructional design in at least two ways. First, we are no longer inclined to make quick inferences about how to support learning by reasoning from a descriptive theory of learning or from empirical studies unsupported by theoretical insights. Learning can accurately be described as a process in which people construct new knowledge by drawing on their prior experience and blending it with new information about a task (Mayer, 2004). We also have clear evidence that asking students to construct what they must learn without guidance is consistently less effective and efficient than worked examples that demonstrate how to perform a task or solve a problem (Mayer, 2004; Kirschner, Sweller, & Clark, 2006). CLT accurately predicts that ~~learning by being asked~~ to construct or discover how to solve problems or perform complex tasks overloads working memory and inhibits learning for students who have novice to intermediate levels of germane prior knowledge. Most of the chapters in this book and the research on the use of CLT for instructional design that preceded this book are clearly focused on helping those who design, develop, and present all types of instruction to learners at every age and level of expertise. Recent examples are Richard Mayer's (2001, 2005) edited handbooks on multimedia design, his book with Ruth Colvin Clark (Clark & Mayer, 2007) on designing e-learning instruction, and the systematic instructional design strategy for teaching complex knowledge published by Jeroen van Merriënboer and Paul Kirschner, (2007). These developments can be viewed as attempts to use CLT to identify the many ways that common instructional practices cause overload and suggest concrete and systematic ways to avoid them. Because many of the researchers who are committed to CLT development are also interested in instructional design, some of the most important educational contributions serve to define and clarify the role of instructional methods.

CLT and Instructional Methods

Another important advantage of the behaviorism that preceded the development of CLT may be CLT researchers' adaptation of the goal to provide specific, evidence-based operational definitions of "instructional methods" and welcome explanations of how different methods serve to maximize germane cognitive load and so lead to more learning. Most instructional design systems suggest that those who are developing instruction should "select appropriate instructional methods" without providing adequate guidance about the definition, design, or selection of effective methods.

When a young cognitive science was developing in the early 1970s, Lee Shulman famously complained that an obsessive emphasis on aptitude in learning theories had led to the situation in which instructional methods "are likely to remain an empty phrase as long as we measure aptitudes with micrometers and instructional methods with divining rods" (Shulman, 1970, p. 374). Cronbach and Snow (1977) reviewed all instructional research conducted for approximately four decades and recommended that we invest much more emphasis on understanding instructional methods.

Until CLT, our failure to focus adequate attention on the specification and presumed cognitive function of instructional methods continued to be one of the most embarrassing failures of instructional psychology. Instructional experiments typically employ treatments described as lectures, discussion, collaborative groups, graphic organizers, case studies, computer programs, and video and text materials. None of these descriptions (and often their accompanying operational definitions in research reports) are focused on the "active ingredients" in the instruction that may or may not have led to measured differences in outcomes (Clark & Estes, 1999; Clark, 2001). CLT's emphasis on elements of instructional methods that are germane and so contribute to learning and those that are extraneous and so distract and inhibit learning is a huge contribution to instructional psychology. Examples of methods suggested by CLT to support novice learners include formatting instructional content in focused, integrated pictorial and narrative presentations of topics (Chapters 3 and 7, this volume) and providing demonstrations of how to perform tasks or solve problems in "worked examples" (Chapter 5). CLT research provides strong indications that these methods maximize the processing time in working memory for task information that must be elaborated and stored in long-term memory while they minimize the extraneous cognitive effort required to support learning. CLT advocates also suggest that these methods provide effective

support for the limited executive learning functions available to learners with less prior knowledge (Chapter 2). The explanation for the benefits of these CLT instructional methods helps to explain the half-century of research that demonstrates the failure of discovery, problem-based, inquiry, and constructivist learning (Kirschner et al., 2006).

Challenges to CLT-Inspired Instructional Design

CLT has developed rapidly but like any theory, there are many unanswered questions and a number of areas in which current theoretical explanations and measures are inadequate. In the next section of this chapter, we review two urgent issues and examine the possible contributions we could expect from reconsidering the importance of biological, physiological, and neuroscience research. Two important problems that must be addressed before we can advance much further with CLT are that we have not yet found an unobtrusive and reliable way to measure cognitive load and we need to determine whether any specific source of cognitive load is productive for individual learners during instruction.

Measuring cognitive load during learning. Gross measures of mental workload, such as self-report and secondary tasks (Megaw, 2005), have been challenged (Gimino, 2000). Self-report measures appear to be confounded with personal judgments about the difficulty of a task rather than the amount of mental effort invested. Secondary measures capture the time required for individual learners to react to a random interruption during a task. These latency measures divert learners' attention from tasks and introduce a variety of messy confounds (see a review by Iqbal, Adamczyk, Zheng, & Bailey, 2005). Brünken, Seufert, and Paas (Chapter 9) discuss different solutions and conclude, "cognitive load measurement is still in its infancy" (p. xx). Past attempts to provide a definition of cognitive load in an educational context have focused either on the number of steps and/or interactions between steps required to perform a task – most often called "intrinsic" load (Sweller, 2006) – or on the mental workload experienced by individuals who are learning. One often-repeated example of the difference between low and high levels of intrinsic load is the difference between learning vocabulary in a foreign language and the presumably higher load required to learn to speak a foreign language (Sweller, 2006). Yet, there have been arguments that the construct of intrinsic load may be an unnecessary and distracting return to the behaviorist emphasis on the environment and the directly observable (Clark & Elen, 2006). Is load in the environment or is it a function of the amount of mental work necessary for any individual

learner to accomplish a task depending on individual differences in prior expertise – or some combination of the two factors?

Most definitions of cognitive load emphasize the non-automated cognitive operations that must be assembled by any given individual to complete the task (Clark & Elen, 2006; Clark, Howard, & Early, 2006; Lohman, 1989; Salomon, 1983; Snow, 1996). We could expect huge individual differences in cognitive load for any task depending on the amount of automated prior knowledge any one individual brings to the task. Brünken, Seufert, and Paas (Chapter 9, this volume) suggest that a learner's prior knowledge influences load and also that we do not have adequate measures of automated prior learning. We propose that more effort be invested in exploring physiological measures of mental workload to identify the amount of automated knowledge learners bring to instruction and to reliably quantify the mental effort they must invest to achieve a unit of learning.

Prior knowledge and germane cognitive load. From a cognitive perspective, the working load experienced during any task is determined in part (and perhaps entirely) by an individual's prior experience with the task (Chapter 2, this volume). The germane cognitive load necessary to succeed at a task is inversely related to the level of automation of necessary prior knowledge (Clark & Elen, 2006). Other things being equal, when we have less of the prior knowledge required when learning a new task, we must use more mental effort to construct new cognitive operations that support task performance. The more automated the prior knowledge, the less cognitive effort required to apply it during learning. For example, the amount of germane cognitive load required for children to learn division is lower if they have more automated addition and subtraction skills. Children who have learned addition and subtraction routines recently and have had less time to practice and automate would need to invest more mental effort at multiplication than those who have practiced longer (Clark & Elen, 2006). Conscious cognitive processing that serves to assemble and/or implement a productive approach to learning a task is the source of relevant cognitive load. Providing a worked example of a successful approach to a new task during instruction for students with highly automated prior knowledge reduces the necessary, relevant load to its lowest possible level (Kirschner, Sweller, & Clark, 2004).

Prior task experience fosters the development of implicit (automated, largely unconscious, procedural) task-relevant cognitive processes (e.g., Woltz, 2003) that are presumed to operate without consuming working memory space and so reduce the demand on working memory. Lohman (1989) described the problem of estimating the amount of cognitive load

from the prior experience measures of individuals on any task when he cautioned: "What is novel for one person may not be novel for another person or even for the same person at a different time . . . [thus] inferences about how subjects solve items that require higher level processing must be probabilistic, since the novelty of each [item] varies for each person" (words in brackets added, p. 348). Brünken, Seufert, and Paas (Chapter 9, this volume) discuss this problem and acknowledge that we have not yet found precise measures of cognitive load for individual learners. Kalyuga and Sweller (2005) have suggested that one way to measure implicit knowledge might be to provide students with a problem and some of the initial steps necessary to solve the problem, and then ask them to describe what must be done next. The difficulty with this approach is the evidence that people who have highly automated knowledge about a task can perform the task but cannot accurately or completely describe the steps they follow (see a review by Feldon & Clark, 2006). Variable levels of prior knowledge automation may account for some of the error reported in Kalyuga and Sweller's (2005) experiments. In general, the lack of a reliable, efficient measure of automated germane prior knowledge is a serious problem for CLT. A primary goal of CLT is to describe specific instructional methods that will maximize relevant and minimize irrelevant cognitive load for each learner at all stages of learning. Thus, when we have determined the total amount of cognitive load experienced by any individual in learning or problem-solving tasks, the next challenge is to break that total down into the proportion of germane (relevant) and extraneous (irrelevant) load being experienced. Yet, identifying the type and origin of mental workload is problematic because for any individual, the amount of load experienced during learning is influenced by the amount of prior knowledge he or she possesses and how automated that knowledge has become with use.

Distinguishing between germane and extraneous cognitive load. A second urgent problem, related to the measurement of gross load and also addressed by Brünken, Seufert and Paas (Chapter 9, this volume) is that we have also not yet found a way to reliably determine whether mental work is being invested in productive or unproductive mental activity. CLT is based on the distinction between "extraneous" or irrelevant load (mental effort invested in activities that do not support learning goals) and "germane" or relevant load (mental effort that supports learning and problem solving). And yet, these two key constructs are only inferred post hoc from differences between learning scores that result from different treatments that are presumed to provide more of one than the other type of load. In a later section, we suggest that eye movement and gaze direction technology might be used

as an indicator of what is being processed cognitively and therefore an indicator of extraneous load.

It is likely that we will solve the measurement of gross mental workload before we are able to deal with the more difficult problem of distinguishing between different types of load being experienced by a single individual. The next section discusses the construct definition and measurement problems that exist with CLT and possible ways to handle those problems with neuroscience research methods.

POSSIBLE NEUROSCIENCE CONTRIBUTIONS TO MEASURING MENTAL EFFORT IN CLT

Recent neuroscience research has made significant advances toward a better understanding of brain function during learning and problem solving (Szucs & Goswami, 2007). During learning, all information is coded in the brain in the form of synaptic activity that underlies the symbolic representations hypothesized by cognitive psychologists. The combination of neuroscience and cognitive science permits the development of a common, integrated framework consisting of connections between higher-level cognitive representations (such as the hypothesized constructs and relationships in CLT) and lower-level data concerning neuronal and biological functions in the brain and sensory systems (Szucs & Goswami, 2007). The ultimate goal of this integration is to add to our ability to predict and explain how our brain function and biology give rise to our mental functioning during learning and problem solving. This integration would bring us full circle and perhaps redress some of the historical mistakes we made at the turn of the last century. Although neuroscience may not yet have much to offer instructional designers or teachers, researchers might benefit from its focus on precise measurements of brain and sensory processes. One exciting possibility can be found in neuroscience research on mental workload and pupil dilation.

Pupil Dilation and Vascular Constriction as Measures of Mental Workload

Promising neuroscience measures of cognitive load may be available in two established physiological measurement technologies called pupillometrics (Megaw, 2005) and peripheral vasoconstriction (Iani, Gopher, & Lavie, 2004; Marshall, 2007). In the case of pupillometrics, devices have been developed to measure the amount of pupil dilation along with the direction and

duration of a learner's gaze. Vasoconstriction measurement requires the wearing of a device on one finger that measures variations in blood flow to the finger.

Pupil dilation and mental effort. Considerable evidence supports the claim that pupil dilation is highly correlated with mental effort during learning and problem solving (Beatty, 1982; Beatty & Wagoner, 1978; Iqbal et al., 2005; Iqbal, Zheng, & Bailey 2005; Kahneman & Beatty, 1966; Recarte & Nunes, 2003). Kahneman and Beatty (1966) compared a variety of encoding, processing, and retrieval tasks, and found that pupil diameter increased proportionally with the mental workload required. In a digit storage and recall task, pupil width increased proportionally with the number of digits encoded and decreased as they were reported. In a separate experiment, digit encoding was compared with digit transformation for the same series of digits. Pupil width was larger when the numbers were added before encoding. They also found that pupil width decreased with task repetition over the course of the study, as task difficulty decreased. This work was extended by Beatty and Wagoner (1978), who examined pupil diameter for a series of letter comparison tasks that increased in complexity, from physical comparisons to comparisons by name, then by category. Again, pupil diameter increased with increasing task complexity. Beatty (1982) reviewed all experimental data on pupil dilation and effort, and concluded that the relationship survives alternative explanations.

One controversial aspect of these findings is that the neural circuitry thought to control pupil diameter, located in a variety of deep sub-cortical regions and in the brainstem, is not closely associated with the circuitry involved in working memory, located primarily in dorsolateral prefrontal cortex. A considerable amount of evidence seems to support the claim that increasing cognitive load affects pupil diameter indirectly through changes in affect-based arousal, perhaps caused by the need to perform mental work (Kahneman & Beatty, 1966; Iqbal et al., 2004; Recarte & Nunes, 2003). And yet, it must be noted that because pupil dilation is apparently mediated by affect-based arousal, we need to learn more about the nature of the relationship between arousal, working memory, and mental effort. If, as some neuroscientists have suggested (Iqbal et al., 2004), mental work is always accompanied by arousal, then pupil dilation might serve as a highly reliable measure of workload. If we find significant individual differences in arousal with prior task knowledge held constant, we would be less inclined to settle on pupil dilation as a measure of mental effort. Early studies of this issue seem to indicate that emotionality may not influence dilation as much as mental effort. A dissertation by Simpson (2006)

provided subjects with both abstract and concrete words that were either very pleasant or very unpleasant and found that, as expected, pupil dilation was greater for abstract words. However, dilation was not different for pleasant and unpleasant words. This question requires more research on individual and group differences in pupil dilation, but the uncertainty it raises does not eliminate the utility of pupil dilation as a measure of cognitive load.

Individual and group differences in pupil dilation. Studies have examined individual and group differences in pupil dilation with tasks held constant. For example, Van Gerven, Paas, van Merriënboer, & Schmidt (2004) found differences between the pupil dilation of younger and older subjects in a study that examined six levels of memory load based on the classic Sternberg memory task. They concluded that dilation might not always be a good measure of mental effort for older (senior) learners. In an age judgment task in which photographs of faces were either gazing directly at the observer or to the side, Gillian, Hood, Troscianko, and Macrae (2006) reported that pupil dilation was greater and more sustained in female than in male participants when analyzing directly gazing faces of both genders. The authors concluded that their female subjects invested more effort in processing socially relevant (direct-gaze) than socially irrelevant (deviated-gaze) faces regardless of the gender of the face. Heitz, Schrock, Payne, and Engle (2003) described two experiments in which groups of subjects with greater or lesser working memory spans engaged in a memory task. They reported that both groups demonstrated equal pupil dilation during tasks requiring similar mental effort, even though those with greater working memory span achieved higher scores. These data suggest that mental effort may not be a good explanation for differences in working memory but that dilation may be a good indicator of mental work.

Individual differences in the automaticity of task prior knowledge are an important issue in all studies of cognitive load (Clark & Elen, 2006). The more practice we experience with a task or critical components of a task, the less mental effort we require to learn related tasks or to assemble component tasks into a more complex set of skills. Cognitive load would presumably be less for a more experienced learner with more automated prior knowledge than one who has less prior knowledge. If we ask students to take pretests consisting of a sample of the types of tasks to be learned and/or tasks requiring the necessary prior knowledge for new learning, the amount of automated prior knowledge should be indicated by the correlation between the amount of pupil dilation during different pretest items and outcome measures, such as item solution speed and accuracy.

Individuals who dilate more and are slower and less accurate will most likely have less automated levels of prior knowledge or less access to relevant knowledge and therefore be required to invest more effort to succeed. The less prior knowledge and the less automated that knowledge is, the more it is necessary to provide instruction that eliminates all extraneous load and provide only the essential steps in a worked example of how to perform the task to be learned. Carswell (2005) used pupil dilation to assess mental workload when surgical residents were practicing with novel laparoscopic surgical technology. He was looking for novel ways to not only improve instruction but also to test alternative technologies for surgery. He tested surgeons with different prior experience levels with traditional technology and with laparoscopic technologies in order to reason about the relative contribution of prior knowledge and variations in the technology to mental workload. Recarte and Nunes (2003) described a study using pupillometry in which the responses of different individuals to similar task conditions could be interpreted as different levels of prior automation. Finally, van Gog, Paas, and van Merriënboer (2008) studied the eye movements of people with different levels of expertise at electrical trouble-shooting tasks. They reported that experts spent more time than novices looking at fault-related components of devices but did not measure pupil dilation. They also found an expertise reversal effect (Kalyuga, Ayres, Chandler, & Sweller, 2003) in which conceptual knowledge about troubleshooting interfered with the learning of experts, perhaps because it served as extraneous load for experts but was germane for novices. It would be interesting to replicate this study and others by van Gog and colleagues (e.g., Nievelstein, van Gog, Boshuizen & Prins, 2008) to collect data on the relative amount of mental effort invested by experts and novices during problem solving. In general, the combination of dilation, eye movement, and duration as measures of mental effort should be combined with subjects that differ in expertise and tasks that differ in complexity.

Pupil dilation as a method to assess extraneous load during learning. Most important to CLT researchers is developing reliable ways to measure the amount and origin of extraneous (irrelevant) load during learning. The instructional goal is to anticipate and eliminate all sources of extraneous load so that working memory processing is as efficient as possible. Recarte and Nunes (2003) designed a creative way to combine pupil dilation and eye movement technology to test the amount of extraneous (irrelevant) mental load experienced by drivers to attend to a "hands-free" telephone conversation while driving and compared it with the load experienced attending to the same conversation "live" with a person riding with them in a car. They

employed visual cues such as unexpected emergency road signals during the conversations to see if drivers noticed fewer of these important cues while engaging in conversations. It is interesting but not surprising to note that the amount of cognitive load was identical during both hands-free telephone and live conversations as measured by eye movement tracking and pupil dilation. It was also determined by eye movement tracking and behavioral observation that the extraneous load imposed by the conversations resulted in a 30% reduction in the drivers' noticing of emergency cues during both the hands-free and live conversations. In a very different study of extraneous load, Verney, Granholm, and Marshall (2004) used pupil dilation to examine the differences between college student performances on a backward masking task that required them to overcome distractions in order to solve target detection problems. Their analysis indicates that students with lower SAT scores invested more wasteful effort focusing on the distractions in the task, which were accounted for by socio-economic differences and prior target detection accuracy. Marshall (2007) describes three problem-solving studies in which pupil dilation reliably distinguished between rest and work; between germane or extraneous effort, and between rested and fatigued states.

Devices for measuring pupil dilation. A number of devices are currently available that will measure and analyze the pupil dilation for individuals during learning from computer displays or other fixed display technologies (Recarte & Nunes, 2003). Iqbal et al. (2004) concluded that "pupil size is the most promising single measure of mental workload because it does not disrupt a user's ongoing activities, provides real-time information about the user's mental workload and is less obtrusive than other physiological measures such as heart rate or EEG [electroencephalogram]" (p. 1477). In order to measure gaze and pupil dilation, it is often necessary to place a research subject's head in a vice-like frame (similar to those used during eye examinations) to prevent head movement. Recently, however, relatively light and unobtrusive equipment is beginning to be developed, such as a camera mounted on a light headband worn by subjects described recently by Marshall (2007). It is highly likely that pupil dilation measured by the headband technology is much less intrusive than the interruptions caused by head fixation devices or secondary (latency) measures.

Pupillometry may improve our measurement of the amount of cognitive load, and combining dilation with the direction and duration of gaze may also help to solve the problem of the relevancy of the load being experienced. Another less studied technology that also seems to offer the possibility of unobtrusive measurement of mental load is vascular constriction.

Vascular constriction and mental effort. Iani et al. (2004) reported that a measure of the constriction of blood vessels in the fingers is a measure of sympathetic nervous system activation and might serve as a reliable measure of mental effort. They conducted two experiments in which they varied task difficulty and the level of engagement of their subjects in the task, and reported that increased vascular constriction (reduced blood flow to the fingers) was highly correlated with performing tasks (constriction was greater when working than when resting) and was greater with more difficult than with less difficult tasks. They also reported a strong correlation between vasoconstriction and pupil dilation. Iani, Gopher, Grunwald, and Lavie (2007) examined the vascular constriction of pilot performance in a computer-based flight simulator in which the difficulty of the task could be manipulated. They found that constriction was greater with difficult than with easier tasks. In general, vasoconstriction seems to provide an alternative way to measure gross cognitive load, yet it does not seem to offer a way to determine the source of the load being measured. At this point, the most promising way to measure both mental load and the source of the load seems to be the use of technology that captures pupil dilation along with gaze direction and duration.

IMAGING METHODS FOR MONITORING CHANGES IN COGNITIVE LOAD

Whereas pupil dilation provides intriguing evidence regarding the changes in neuro-cognitive activity that underlie cognitive load, more direct measures of brain function are available. A large number of brain imaging studies have examined working memory. Working memory provides a temporary store that supports cognitive processing, and the capacity of working memory is commonly thought to be closely associated with cognitive load. The higher the cognitive load required to perform a task, the greater the demand on working memory. The imaging methods discussed in the following sections can be used to examine the neural activity that supports working memory and therefore can indicate how changes in cognitive load affect brain function.

Three basic processes of working memory have been identified: a brain network for the maintenance of auditory and verbal information, a separate network for the maintenance of visual and spatial information, and a central executive network for attentional control and manipulation of items in working memory (Baddeley, 1986), although evidence for the central executive is controversial (see the discussion by Sweller, 2004). Working

memory includes three processing activities that occur in sequence: encoding, maintenance, and retrieval. Each of these processes involves a different pattern of brain activity, and each can be affected differently by changes in load. Each can be distinguished by differences in time, for example encoding must occur before retrieval. Isolation of the brain processes supporting each of these stages based on timing can be accomplished using event-related potentials (ERPs), which record the small fluctuations in voltage at the scalp surface generated by neural activity in the brain. This can be used to infer the timing of events, but it offers poor spatial resolution and therefore inadequate information about where the processing is occurring in the brain. By contrast, imaging methods that rely on hemodynamic measures, such as functional magnetic resonance imaging (fMRI) and positron emission tomography typically measure changes in blood flow and/or oxygenation that are related to changes in brain function. These hemodynamic methods offer superior spatial resolution compared with ERPs, which is necessary to unambiguously identify the anatomical location of brain networks supporting the different processes of working memory. However, because changes in blood flow are relatively slow, these methods are usually unable to identify rapid changes in brain activity. Event-related fMRI is a method that can be used to achieve a balance between spatial and temporal resolution (Clark, 2002; Clark, Maisog, & Haxby, 1998) by focusing on the characterization of small changes in signal over short periods of time. These methods can distinguish changes in neural activity occurring on the order of a few hundred milliseconds apart, depending on how the data are acquired and analyzed.

ERP Studies of Working Memory

As described earlier, working memory has a fundamental role in supporting cognitive load. ERPs can be used to examine the neural activity that supports working memory, and therefore how changes in cognitive load affect brain function. Many neuroscience studies have employed delayed response tasks to study working memory. Delayed response tasks require subjects to maintain information in working memory for a period of time before a response is made. This might be a word, an object, a location in space, or some other sensory feature or groups of features that must be held in memory. Often, such tasks involve one or more items that must be held in memory, to be compared with additional items presented later in time before a response can be made. Delayed response tasks often evoke a characteristic sustained negative electrical potential over the scalp termed the contingent negative

variation (CNV). CNVs are evoked during the maintenance of information stored in working memory (Tecce, 1972). It is likely that the CNV results from increased synaptic activity associated with maintaining information in the working memory store. Working memory tasks have been found to evoke activity in a variety of brain regions. Gevins, Smith, and Le (1996) used high-resolution evoked potentials during verbal and spatial working memory tasks. In this study, verbal or spatial attributes were compared between each test stimulus and a preceding stimulus. All stimuli evoked the CNV and a number of other components, which varied in amplitude, depending on the specific requirements of the task. They concluded that working memory is a function of distributed neural systems with both task-specific and task-independent components and that these and other ERP components can be used to study working memory processes.

However, subsequent studies have shown that the interpretation of ERP components to study working memory can be more complex than is typically assumed. Kok (2001) found that the amplitude of positive components evoked from 300 to 500 msec post-stimulus reflected the activation of elements in an event-categorization brain network that is controlled by the joint operation of attention and working memory. This limits the use of these components as a measure of processing capacity or cognitive load because variations in both attention and working memory can influence their production. Luck, Woodman, and Vogel (2000) supported this view by suggesting that many studies confound attention and working memory. They proposed that attention may operate to adjust brain networks supporting working memory and other cognitive processes when brain systems are overloaded and therefore operates to adjust the brain's ability to process the extra information under conditions of higher cognitive load and thus optimize performance. Finally, Wager and Smith (2003) suggested that selective attention to features of a stimulus to be stored in working memory leads to separate patterns of activation from working memory storage. Selective attention is the process whereby specific objects or classes of stimuli are selected for further processing based on certain defining stimulus characteristics. Depending on the nature of these characteristics (e.g., color, shape, or spatial location), a different pattern of brain response is seen that is unique to those characteristics. Thus, the dynamic properties of these interrelated neural and cognitive systems make it difficult to use these measures to quantify specific features, such as cognitive load. Even with these limitations, carefully designed studies that take these and other issues into consideration can reveal much about how the brain deals with variations in cognitive load.

fMRI Studies of Cognitive Load

Most fMRI studies of cognitive load effects examine the identity of brain regions that support different aspects of working memory. These studies typically use parametric designs. These designs reveal the neural correlates of working memory load by identifying those regions in which activity changes as the level of cognitive load is changed across repeated measurements. This method assumes that additional cognitive load will increase the brain responses in a proportional way, otherwise known as the pure insertion hypothesis (Raichle, 1994). Using these methods, a number of published studies have characterized brain networks that support working memory and how these networks change with changes in cognitive load. N-back tasks are one such design that involves the presentation of stimuli in a series, in which subjects are asked to compare the current stimulus with stimuli presented one or more items earlier in a series. For a delay of one stimulus, the N-back task is similar to the delayed response task. However, for more than one stimulus delay, N-back tasks differ from delayed response tasks in the use of intervening stimuli presented between the two stimuli being compared. With an increasing delay between the first and second item to be compared, the number of intervening items that must be maintained in working memory to perform the task increases, and this increases cognitive load in turn. These tasks also differ in that two comparisons are made for most stimuli in an N-back task, first with the stimulus presented N stimuli before it and then with the stimulus presented N stimuli after. Callicott et al. (1999) used fMRI to identify characteristics of working memory capacity using a parametric N-back working memory task. In this study, as the number of items was increased, task performance decreased. As cognitive load was increased, some brain regions indicated changes in activity that followed an inverted U shape. Large regions of dorsolateral prefrontal cortex, along with smaller regions of premotor cortex, superior parietal cortex, and thalamus, revealed changes in activity. The authors concluded that this pattern was consistent with a capacity-constrained response. At lower levels of load, less activity was required to support the working memory processes. At middle levels, more activity was required to maintain the same level of performance. At very high levels of load, the performance of the network breaks down, resulting in both reduced activity and reduced performance. These results reflect the findings in cognitive instructional psychology (e.g., Clark, 1999; Clark & Elen, 2006; Gimino, 2000; Salomon, 1983) where prior knowledge predicts mental effort under conditions in which tasks become increasingly difficult.

These results demonstrated that a portion of the brain networks supporting working memory is sensitive to variations in cognitive load, whereas other portions do not appear to be as sensitive. Jaeggi et al. (2003) employed an N-back task with four levels of difficulty using auditory and visual material, and did not find the same inverted U-shape relationship. The participants' tasks were performed separately or simultaneously as dual tasks. When performed separately, activation in the prefrontal cortex increased continuously as a function of memory load. An increase of prefrontal activation was also observed in the dual tasks, even though cognitive load was excessive in the case of the most difficult condition, as indicated by reduced behavioral performance. These results suggest that excessive processing demands in dual tasks are not necessarily accompanied by a reduction in brain activity. More recently, O'Hare, Lu, Houston, Bookheimer, and Sowell (2008) examined the development of these brain networks using a Sternberg working memory task with three load levels. The Sternberg task involves asking subjects to encode a set of stimuli (e.g., "1," "3," and "9") and later presenting a series of stimuli and asking them to indicate which of these stimuli match the encoded set and which are new. The larger the size of the encoded stimulus set, the greater the cognitive load. The activated brain networks were found to depend on the participants' age, which ranged from 7 to 28 years. Adolescents and adults showed cognitive-load effects in frontal, parietal, and cerebellar regions, whereas younger children showed similar effects only in left ventral prefrontal cortex. These results demonstrate that increasing load produces different brain network responses from childhood through adulthood. As a result, we may find developmental differences between the ways that young children and adults handle cognitive load during learning.

Some of the differences observed across studies may result from variations in learning tasks. Using fMRI, working memory is often associated with increased activity in the prefrontal cortex, typically in Brodmann areas 6, 9, 44, and 46 (Cabeza & Nyberg, 2000). In area 6, located in the frontal cortex, activations are commonly found across tasks, including verbal, spatial, and problem-solving tasks, and thus may be related to general working memory operations that are not associated with other sensory or cognitive features of the task. By contrast, the exact pattern of activation in other brain areas is related to the specific nature of the task used. Increased activity in area 44, which lies next to area 6 in the lateral frontal cortex, is found for verbal and numeric tasks compared with visuospatial tasks, which may be related to phonological processing. Activations in areas 9 and 46, located on the frontal pole, are stronger for tasks that require manipulation of working memory contents compared with tasks that require only maintenance of

items in working memory (Owen, 1997; Petrides, 1994, 1995). Ventrolateral frontal regions (including areas 45 and 47) are involved in the selection and comparison of information held in working memory, whereas medial and anterior frontal regions (areas 9 and 46) are involved in the manipulation of multiple pieces of information. Some studies have shown that working memory for object information engages ventral prefrontal regions, whereas working memory for spatial locations engages dorsal prefrontal regions (Courtney, Ungerleider, Keil, & Haxby, 1996, 1997). However, other studies suggest that working memory for objects engages left frontal regions, whereas working memory for spatial information engages right frontal regions (Belger et al., 1998; Smith, Jonides, & Koeppe, 1996; Smith et al., 1995). Taken together, these studies suggest that the organization of frontal brain networks that support working memory still hold a number of secrets in terms of the cognitive basis around which they are organized.

Working memory studies also show activations in brain regions outside of the frontal cortex, including the parietal areas. In the case of verbal tasks, these activations tend to be larger on the left, which supports Baddeley's phonological loop model, which maintains that information is stored and rehearsed in series (Awh et al., 1996; Paulesu, Frith, & Frackowiak, 1993). Working memory tasks are also associated with altered activity in anterior cingulate, occipital, and cerebellar cortices. However, these tend to be more sensitive to stimulus characteristics and task demands, rather than cognitive load, suggesting that they perform operations that support working memory indirectly through their interaction with these other regions. One exception to this is the finding of Druzgal and D'Esposito (2001), who showed that activity in ventral extrastriate visual areas increased directly with load of an N-back working memory task using facial stimuli. They concluded that both prefrontal and extrastriate areas worked together to meet the demands of increased cognitive load.

Advanced methods of brain imaging offer many insights into the neural mechanisms that support working memory and the effects of changes in cognitive load on these mechanisms. Some progress has already been made in understanding the brain basis of processes related to cognitive load. Our ultimate goal is to achieve a unified theory that bridges the gap between cognitive psychology and neuroscience. We are beginning to see parallels across these two fields, as described earlier, but there is still much work to do. As brain imaging methods improve, and as cognitive psychologists are more willing to understand brain imaging technologies and to use this sort of information in forming hypotheses, a better understanding of cognitive load than could be achieved by either discipline alone can ultimately be achieved.

SUMMARY AND CONCLUSION

We have come full circle since Judd, an early cognitive psychologist, lost an argument to Thorndike, an early advocate of neurological and biological psychology. That lost argument serves as a cautionary metaphor for the bias that prevented American psychologists from focusing on cognitive questions for fifty years. It may also have produced a reaction whereby cognitive psychology is now experiencing a reverse bias against biological and neurological insights about learning and problem solving. The point of this review is to emphasize that the solution to some of the thorny problems facing CLT requires that we step away from our century-long dispute and become open to the insights offered by past and future advances in cognitive psychology and neuroscience. It seems reasonable to expect that neuroscience might aid the search for ways to reliably quantify cognitive load and to identify the sources of germane and extraneous load. We might also increase our understanding of how individual and group differences in prior knowledge, culture, and working memory span might influence brain function, resulting in quantifiable differences in activity recorded with brain imaging methods such as ERPs and fMRI, and how this affects our understanding of differences between various learning tasks and instructional methods.

We recommend a renewed commitment to exploring the use of pupil dilation accompanied by gaze direction and intensity studies to develop a more reliable and valid estimate of individual cognitive load and to help identify sources of germane and extraneous load. Pupil dilation could also be used to investigate how differences in the amount of prior expertise in a knowledge domain influence the type and amount of cognitive load experienced by learners. We expect that the more specific prior knowledge learners possess about the class of tasks being learned, the less load they will experience compared with learners who have less prior knowledge. It may also be possible that germane load for novices might become irrelevant load for experts and that this might be the source of the expertise reversal effect described by Kalyuga et al. (2003). We also suggest that fMRI methods of brain imaging offer many possible hypotheses based on evidence from neural mechanisms that support working memory and on the effects of changes in cognitive load on these mechanisms for different types of tasks and learners.

Neuroscience studies draw most often on Baddeley's (1986) model of working memory and search for evidence for three separate networks that maintain visual and spatial information, verbal information, and the control of attention and manipulation of items being held. In addition, neuroscience

looks for evidence for three processing activities that occur in sequence in each of the three networks during learning and task performance – encoding, maintenance, and retrieval. Although a number of technologies are used to identify and validate these processes, the most complete and accurate is event-related fMRI. In general, the brain regions associated with most of these processes have been identified, but complex ambiguities and arguments persist.

To this point, fMRI studies have provided additional evidence for the processes that occur in working memory and the brain structures that appear to support those processes. It also appears that working memory load consists of both task-specific and task-independent components. In addition, it appears that some experiments may confound working memory and attention processes. Claims have been made, for example, that when cognitive load increases, attention processes may be automatically evoked and serve to reduce load by forcing attention to more germane attributes of tasks (Wager & Smith, 2003). It is also possible that increases in load may evoke processes that focus attention on extraneous events (Clark, 1999). In addition, fMRI studies have provided evidence for the inverted-U hypothesis about the relationship between cognitive load and mental effort similar to the one suggested by Salomon (1983). When load is low, effort is also low, but as cognitive load increases, fMRI indicators of load also increase until it reaches a very high level in which the brain networks supporting working memory seem to fail, with accompanying decreases in mental effort and test performance. It also appears that some tasks may not produce the inverted U. At least one well-designed study (Jaeggi et al, 2003) identified dual-coding memory tasks in which increasing load (judged by both fMRI data and subject performance) did not yield decreasing effort.

In general, there appear to be a number of important interactions among variations in task types, working memory processes, and cognitive load. Some areas of the brain seem to be active during all working memory processing, and some areas seem to specialize in different types of processing. For example, separate areas have been associated with verbal and numeric tasks, whereas others seem to be active during visuospatial tasks. In addition, tasks that require manipulation of the contents of working memory (thought to be associated with executive functions) activate different areas than tasks that require maintenance of both visuospatial and verbal-numeric information in working memory. Other studies have found evidence to suggest different regions support spatial location and object information. fMRI studies have also provided strong evidence for age-related developmental differences in the operation of working memory. As load increases

in younger children, working memory activities appear in the left ventral prefrontal cortex, but in adolescents and adults, the same tasks produce cognitive-load activity in the frontal, parietal, and cerebellar regions. The reason for these differences and their consequence for instruction and/or learning are unknown.

As neuroscience methods improve in spatial and temporal resolution and as new methods are developed, more precise information will be obtained. However, we know now that the cognitive sub-processes involved in performance of challenging learning and problem-solving tasks and the brain networks that support them interact in complex ways. In a single study, it is easy to confound the effects of changes in cognitive load on working memory with changes in attention as well as in perceptual and response processes, affect, and arousal, which all occur together in related ways. Therefore, it is vital that these methods are used carefully and alternative hypotheses be considered as we progress. Ultimately, though, we expect that these methods will lead to a better understanding of the neural and cognitive mechanisms that underlie cognitive load.

REFERENCES

Awh, E., Jonides, J., Smith, E. E., Schumacher, E. H., & Koeppe, R. A., Katz, S. (1996). Dissociation of storage and rehearsal in verbal working memory: Evidence from positron emission tomography. *Psychological Science*, 7(1), 25–31.

Baddeley, A. (1986). *Working memory*. New York: Oxford University Press.

Beatty, J. (1982). Task-evoked papillary responses, processing load and the structure of processing resources. *Psychological Bulletin*, 91(2), 276–292.

Beatty, J., & Wagoner, B. L. (1978). Pupillometric signs of brain activation vary with level of cognitive processing. *Science*, 199(4334), 1216–1218.

Belger, A., Puce, A., Krystal, J. H., Gore, J. C., Goldman-Rakic, P., & McCarthy, G. (1998). Dissociation of mnemonic and perceptual processes during spatial and working memory using fMRI. *Human Brain Mapping*, 6(1), 14–32.

Berline, D. (1960). *Conflict, arousal and curiosity*. New York: McGraw Hill.

Cabeza, R., & Nyberg, L. (2000). Imaging cognition II: An empirical review of 275 PET and fMRI studies. *Journal of Cognitive Neuroscience*, 12(1), 1–47.

Callicott, J. H., Mattay, V. S., Bertolino, A., Finn, K., Coppola, R., Frank, J. A., et al. (1999). Physiological characteristics of capacity constraints in working memory as revealed by functional MRI. *Cerebral Cortex*, 9, 20–26.

Carswell, C. M. (2005). Assessing mental workload during laparoscopic surgery. *Surgical Innovation*, 12(1), 80–90.

Clark, R. C., & Mayer, R. E. (2007). *E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning*. New York: Wiley.

Clark, R. E. (1999). Yin and yang cognitive motivational processes operating in multimedia learning environments. In van Merriënboer, J. (Ed.), *Cognition and multimedia design* (pp. 73–107). Herleen, Netherlands: Open University Press.

Clark, R. E. (2001). *Learning from media: Arguments, analysis and evidence.* Greenwich, CT: Information Age Publishers.

Clark, R. E., & Elen, J. (2006). When less is more: Research and theory insights about instruction for complex learning. In J. Elen & R. Clark (Eds.), *Handling complexity in learning environments: Research and theory* (pp. 283–297). Oxford, UK: Elsevier Science.

Clark, R. E., & Estes, F. (1999). The development of authentic educational technologies. *Educational Technology*, 37(2), 5–16.

Clark, R. E., Howard, K., & Early, S. (2006). Motivational challenges experienced in highly complex learning environments. In J. Elen, & R. E. Clark (Eds.), *Handling complexity in learning environments: Research and theory* (pp. 27–41. Oxford, UK: Elsevier Science.

Clark, V. P. (2002). Orthogonal polynomial regression for the detection of response variability in event-related fMRI. *NeuroImage*, 2002, 17, 344–363.

Clark, V. P., Maisog, J. M., & Haxby, J. V. (1998). An fMRI study of face perception and memory using random stimulus sequences. *Journal of Neurophysiology*, 79: 3257–3265.

Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods.* New York: Irvington Press.

Courtney, S. M., Ungerleider, L. G., Keil, K., & Haxby, J. V. (1996). Object and spatial visual-working memory activate separate neural systems in human cortex. *Cerebral Cortex*, 6(1), 39–49.

Courtney, S. M., Ungerleider, L. G., Keil, K., & Haxby, J. V. (1997). Transient and sustained activity in a distributed neural system for human working memory. *Nature*, 386(6625), 608–611.

Druzgal, T. J., & D'Esposito, M. (2001). Activity in fusiform face area modulated as a function of working memory load. *Cognition and Brain Research*, 10(3), 355–364.

Feldon, D. F., & Clark, R. E. (2006). Instructional implications of cognitive task analysis as a method for improving the accuracy of experts' self-report. In G. Clarebout & J. Elen (Eds.), *Avoiding simplicity, confronting complexity: Advances in studying and designing (computer-based) powerful learning environments* (pp. 109–116). Rotterdam, Netherlands: Sense Publishers.

Gevins, A., Smith, M. E., & Le, J. (1996). High resolution evoked potential imaging of the cortical dynamics of human working memory. *Electroencephalography and Clinical Neurophysiology*, 98, 327–348.

Gillian, P., Hood, B., Troscianko, T., & Macrae, C. N. (2006). Females, but not males, show greater papillary response to direct- than to diverted gaze faces. *Perception*, 35(8), 1129–1136.

Gillian, P., Troscianko, T., & Gilchrist, I. D. (2007) Effort during visual search and counting: Insights from pupilometry. *Quarterly Journal of Experimental Psychology*, 60(2), 211–229.

Gimino, A. E. (2000). Factors that influence students' investment of mental effort in academic tasks: A validation and exploratory study. Unpublished doctoral dissertation, University of Southern California, Los Angeles.

Heitz, R. P., Schrock, J. C., Payne, T. W., & Engle, R. W. (2003, May). *Working memory and mental effort: Data from pupillometry.* Paper presented at the Seventy-Fifth Annual Meeting of the Midwestern Psychological Association, Fort Wayne, Indiana.

Iani, C., Gopher, D., Grunwald, A. J., & Lavie, P. (2007). Peripheral arterial tone as an on-line measure of lead on a simulated flight task. *Ergonomics*, 50(7), 1026–1035.

Iani, C., Gopher, D., & Lavie, P. (2004). Effects of task difficulty and invested mental effort on peripheral vasoconstriction. *Psychophysiology*, 41, 789–798.

Iqbal, S. T., Adamczyk, P. D., Xheng, X. S., & Bailey, B. P. (2005, April). Task evoked pupillary response to mental workload in human-computer interaction. Paper presented at the 2005 Computer Human Interaction Conference, Portland Oregon.

Iqbal, S. T., Xheng, X. S., & Bailey, B. P. (2004, April). Towards an index of opportunity: Understanding changes in mental workload during task execution. Paper presented at the 2004 Computer Human Interaction Conference, Vienna, Austria.

Jaeggi, S. M., Seewer, R., Nirkko, A. C., Eckstein, D., Schroth, G., Groner, R., et al. (2003). Does excessive memory load attenuate activation in the prefrontal cortex? Load-dependent processing in single and dual tasks: functional magnetic resonance imaging study. *NeuroImage*, 19, 210–225.

James, W. (1890). *The principles of psychology* (Vol. 1). New York: Henry Holt.

James, W. (1904). Does 'consciousness' exist? *Journal of Philosophy, Psychology, and Scientific Methods*, 1, 477–491.

Judd, C. H. (1908). The relation of special training to intelligence. *Educational Review*, 36, 28–42.

Judd, C. H. (1910). Evolution and consciousness. *Psychological Review*, 17, 77–97.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, 154(3756), 1583–1585.

Kayluga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, 38(1), 23–31.

Kalyuga, S., & Sweller, J. (2005). Rapid dynamic assessment of expertise to improve the efficiency of adaptive e-learning. *Educational Technology Research and Development*, 53, 83–93.

Kirschner, P., Sweller, J., & Clark, R. E. (2006). Why minimally guided learning does not work: An analysis of the failure of discovery learning, problem-based learning, experiential learning and inquiry-based learning. *Educational Psychologist*, 41(2), 75–86.

Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology*, 38(3), 557–577.

Lohman, D. F. (1989). Human intelligence: An introduction to advances in theory and research. *Review of Educational Research*, 59(4), 333–373.

Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in Cognitive Sciences*, 4, 432–440.

Marshall, S. P. (2007). Identifying cognitive state from eye metrics. *Aviation, Space and Environmental Medicine*, 78(1), B165–B175.

Mayer, R. (2001). *Multi-media learning*. Cambridge, MA: Cambridge University Press.

Mayer, R. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *American Psychologist*, 59(1), 14–19.

Mayer, R. E. (2005). *The Cambridge handbook of multimedia learning*. New York: Cambridge University Press.

Megaw, T. (2005). The definition and measurement of mental workload. In J. R. Wilson & N. E. Corlett (Eds.), *Evaluation of human work: A practical ergonomics methodology* (pp. 525–551). London: CRC Press.

Miller, G. A. (1956). The magic number seven plus or minus two: Some limits on our capacity to process information. *Psychological Review*, 63, 81–97.

Neisser, U. (1967). *Cognitive Psychology*. New York: Appleton-Century-Crofts.

Nievelstein, F., van Gog, T., Boshuizen, H., & Prins, F. J. (2008). Expertise-related differences in conceptual and ontological knowledge in the legal domain. *European Journal of Cognitive Psychology*, 20(6), 1043–1064.

O'Hare, E. D., Lu, L. H., Houston, S. M., Bookheimer, S. Y., & Sowell, E. R. (2008). Evidence for developmental changes in verbal working memory load-dependency. *NeuroImage*, 42, 1678–1685.

Owen, A. M. (1997). The functional organization of working memory processes within human lateral-frontal cortex: The contribution of functional neuroimaging. *European Journal of Neuroscience*, 9(7), 1329–1339.

Paulesu, E., Frith, C. D., & Frackowiak, R. S. J. (1993). The neural correlates of the verbal component of working memory. *Nature*, 362, 342–345.

Perkins, D. (1995). *Outsmarting IQ: Learnable intelligence.* New York, The Free Press.

Petrides, M. (1994). Frontal lobes and working memory: Evidence from investigations of the effects of cortical excisions in nonhumans primates. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (pp. 9, 59–82). Amsterdam: Elsevier.

Petrides, M. (1995). Functional organization of the human frontal cortex for mnemonic processing: Evidence from neuroimaging studies. *Annals of the New York Academy of Sciences*, 769, 85–96.

Raichle, M. E. (1994). Images of the mind: Studies with modern imaging techniques. *Annual Review of Psychology*, 45, 333–356.

Recarte, M. A., & Nunes, L. M. (2003). Mental workload while driving: Effects on visual search, discrimination and decision making. *Journal of Experimental Psychology: Applied*, 9(2), 119–137.

Salomon, G. (1983). The differential investment of effort in learning from different sources. *Educational Psychologist*, 18(1), 42–50.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: 1. Detection, search, and attention. *Psychological Review*, 84, 1–66.

Shulman, L. J. (1970). Reconstruction of educational research. *Review of Educational Research*, 40, 371–393.

Simpson, H. M. (2006). The effects of word abstractness and pleasantness on pupil size during an imagery task. *Dissertations Abstract International Section B. The Sciences and Engineering*, 67(1-B), 584.

Smith, E. E., Jonides, J., & Koeppe, R. A. (1996). Dissociating verbal and spatial-working memory using PET. *Cerebral Cortex*, 6(1), 11–20.

Smith, E. E., Jonides, J., Koeppe, R. A., Awh, E., Schumacher, E.H., Minoshima, S. (1995). Spatial vs. object-working memory: PET investigations. *Journal of Cognitive Neuroscience*, 7(3), 337–356.

Snow, R. E. (1996). Aptitude development and education. *Psychology Public Policy and Law*, 2(3–4), 536–560.

Steele, C. M. (2003). Stereotype threat and African-American student achievement. In C. M. Steele (Ed.), *Young, gifted, and black* (pp. 109–130). Boston: Beacon Press.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(1), 257–285.

Sweller, J. (2004). Instructional design consequences of an analogy between evolution by natural selection and human cognitive architecture. *Instructional Science*, 3, 9–31.

Sweller, J. (2006). How the human cognitive system deals with complexity. In J. Elen & R. E. Clark (Eds.), *Handling complexity in learning environments: Research and theory* (pp. 13–26). Oxford: Elsevier Science.

Szucs, D., & Goswami, U. (2007). Educational neuroscience: Defining a new discipline for the study of mental representations. *Mind, Brain and Education*, 1(3), 114–127.

Tecce, J. J. 1972. Contingent negative variation (CNV) and psychological processes in man. *Psychological Bulletin*, 77, 73–108.

Thorndike, E. L. (1903). *Educational psychology.* New York: Lemcke & Buechner.

Van Gerven, P. W., Paas, F., van Merriënboer, J. J., & Schmidt, H. G. (2004). Memory load and the cognitive papillary response in aging. *Psychophysiology*, 412(2), 167–174.

van Gog, T., Pass, F., & van Merriënboer, J. J. G. (2008). Effects of studying sequences of process-oriented and product-oriented worked examples on troubleshooting transfer efficiency. *Learning and Instruction*, 18(3), 211–222.

van Merriënboer, J. J. G., & Kirschner, P. A. (2007). *Ten steps to complex learning: A systematic approach to four-component instructional design.* New York: Routledge.

Verney, S. P., Granholm, E., & Marshall, S. P. (2004). Pupillary responses on the visual backward masking task reflect general cognitive ability. *International Journal of Psychophysiology*, 52(1), 23–36.

Wager, T. D., & Smith, E. E. (2003). Neuroimaging studies of working memory: A meta-analysis. *Cognitive and Affective Behavioral Neuroscience*, 3(4), 255–274.

Woltz, D. J. (2003). Implicit cognitive processes as aptitudes for learning. *Educational Psychologist*, 38(2), 95–104.