From Neo-behaviorism to Neuroscience: Perspectives on the Origins and Future Contributions of Cognitive Load Research[1]

Richard E. Clark and Vincent P. Clark

University of Southern California
Rossier School of Education and
Department of Surgery,
Keck School of Medicine

University of New Mexico
Departments of Psychology and
Neuroscience and the
Mind Research Network

Version: January 5, 2009

"*Those who cannot remember the past are condemned to repeat it"*
George Santayana "The life of reason" 1905

Previous chapters have defined Cognitive Load Theory (CLT), described evidence for different origins and types of cognitive load, the way that load interacts with individual differences and the impact of load on learning from instruction. In this chapter we will step back a bit and describe some of the history of cognitive load theory in order to better understand both its current focus and future directions. We will then describe some of the current and planned research trends in biology and neuroscience that may provide insights about a few of the thorny measurement and design problems that confront cognitive load researchers. The discussion focuses attention on an older area of neuroscience, the use of pupilometry and eye movement technology to measure cognitive load and identify the source of load. The chapter concludes with a suggestion for collaborative studies based on neuroscience insights about the measurement of cognitive load and ways to distinguish between germane and extraneous load for individuals during learning and performance.

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 until George Miller's classic paper on working memory capacity appeared a half century ago (Miller, 1956). At the beginning of the twentieth century and fifty years before Millers 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

[1] Clark, R.E. and Clark, V. P. (In press), From Neo-Behaviorism to Neuroscience: Perspectives on the Origins and Future Contributions of Cognitive Load Research, In, Plass, J. Moreno, R., and Brüken, R. (Eds.). *Cognitive Load: Theory and Application*. New York: Cambridge University Press.

processing. Judd, an American who was Wundt's student in Leipzig at the end of the 19th 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 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. Since Thorndike was a student of the powerful William James who supported his work, Judd's theory and evidence was largely ignored.

William James's support for Thorndike's view of transfer marked a turning point in psychology. James 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, p. 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 1), James gave a series of lectures at Johns Hopkins University where 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 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 why many of the prime movers in cognitive load theory 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 in order 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. While 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's (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 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 the education enterprise 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 where 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 is consistently less effective and efficient than worked examples that demonstrate how to perform a task or solve a problem (Mayer, 2004; Kirschner, Sweller and 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 relevant 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 and 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. Since 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 (relevant) 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 1970's, Lee Shulman famously complained that an obsessive emphasis on aptitude in learning theories had led to the situation where 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, 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 and 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 (Kayluga, Chapter X; Mayer and Marino, Chapter X) and providing demonstrations of how to perform tasks or solve problems in "worked examples" (Renkl, Chapter X). 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 (Sweller, Chapter 1). The explanation for the benefits of these CLT instructional methods help to explain the half century of research that demonstrates the failure of discovery, problem-based, inquiry and constructivist learning (Kirschner, Sweller and Clark, 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 where 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 a very unobtrusive and reliable way to measure cognitive load and 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 and Bailey, 2005). Brünken, Seufert and Paas (Chapter 4) discuss different solutions and conclude, "cognitive load measurement is still in its infancy" (p. xxx). 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 and 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 – 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 (Salomon, 1983; Lohman, 1989; Snow, 1996; Clark, Howard and Early, 2006; Clark and Elen, 2006). 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 4) suggest that a learners 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 in order 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.

Distinguishing between relevant and irrelevant cognitive load: A second urgent problem, related to the measurement of gross load and also addressed by Bruken, Seufert and Paas (Chapter 4) 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. Cognitive Load Theory 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.

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 for a single individual. The next section discusses the construct definition and measurement problems faced by CLT and possible ways to handle those problems with neuroscience research methods and theories.

Possible Neuroscience Contributions to Measuring Mental Effort in CLT

Recent neuroscience research has made significant advances towards a better understanding of brain function during learning and problem solving (Szucs and Goswami, 2007). During learning, all information is coded in the brain in the form of synaptic activity that mirrors 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 cognitive load theory) and lower level data concerning neuronal and biological functions in the brain and sensory systems (Szucs and 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. While 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 as a Measure of Mental Workload

A very promising neuroscience measure of cognitive load may be available in an established technology called pupilometrics (Megaw, 2005). Considerable evidence supports the claim that pupil dilation is highly correlated with mental effort during learning and problem solving (Kahneman and Beatty, 1966; Beatty and Wagoner, 1978; Recarte and Nunes, 2003; Iqbal, Adamczyk, Zheng and Bailey, 2004; Iqbal, Zheng and Bailey 2005). 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 reduced 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 reduced. 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.

One controversial aspect to 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 dorsal-lateral 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 caused by the need to perform mental work (Kahneman and Beatty, 1966; Recarte and Nunes, 2003; Iqbal et al, 2004). And yet it must be noted that since 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. This question requires more research but the uncertainty it raises does not eliminate the utility of pupil dilation as a measure of cognitive load. Most important, in light of the obtrusiveness of alternative measures of load, is that a number of unobtrusive 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 and Nunes, 2003). Iqbal et al (2004) conclude, "…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 users mental workload and is less obtrusive than other physiological measures such as heart rate or EEG." (p. 1477). It is highly likely that pupil dilation is much less intrusive than the interruptions caused by secondary (latency) measures. Pupil size changes rapidly and is sensitive to subtle changes in mental effort and so allows a researcher to track changes in cognitive load over time in even very brief tasks. Pupilometry may improve our measurement of the amount of cognitive load but will not help us solve the problem of the relevancy of the load being experienced.

Measuring Proportion and Origin of Germane and Extraneous Cognitive Load

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 relevant (germane) and irrelevant (extraneous) load being experienced. Yet, identifying the type and origin of mental workload is problematic since for any individual, the amount of load experienced during learning is influenced by the amount of prior knowledge they possess and how automated that knowledge has become with use.

Prior knowledge and relevant cognitive load: From a cognitive perspective, the working load experienced during any task is determined in part by an individual's prior experience with the task (Sweller, Chapter 1). The relevant cognitive load necessary to succeed at a task is inversely related to the level of automation of necessary prior knowledge (Clark and 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 relevant cognitive load required for children to learn division is lower if they have more automated addition and subtraction skill. 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 and 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 and 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 that: "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 4) 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 difficulties 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 and 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, relevant prior knowledge is a serious problem for CLT and another area where pupilometry might be helpful.

Pupil Dilation as a measure of automated, implicit prior knowledge: The automaticity of task prior knowledge should be measurable using pupil dilation. 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 will 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. The less prior knowledge and the less automated that knowledge, the more it is necessary to provide instruction that eliminates all irrelevant 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) describe a study using pupilometry where the responses of different individuals to similar task conditions could be interpreted as different levels of prior automation. Yet we have not located studies where researchers set out to use pupilometrics as an indicator of knowledge automaticity. In most cases, researchers have assumed that inter individual differences in dilation during the performance of the same tasks are due to error rather than prior knowledge automation differences.

Assessing irrelevant load during learning: Most important to CLT researchers is developing reliable ways to measure the amount and origin of irrelevant load during learning. The instructional goal of this objective is to anticipate and eliminate all sources of irrelevant 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 irrelevant mental load experienced by drivers to attend to a "hands free" telephone conversation while driving and compared it to 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. It was also determined by eye movement tracking and behavioral observation that the irrelevant load imposed by the conversations resulted in a 30% reduction in their noticing of emergency cues during both the hands free and live conversations.

Other Imaging Methods for Monitoring Changes in Cognitive Load

While pupil dilation provides intriguing evidence as to the changes in neuro-cognitive activity that underlie cognitive load, there are also more direct measures of brain function. Working memory provides temporary store supporting other cognitive processing, and the capacity of working memory is commonly thought to be closely associated with cognitive load. The higher the cognitive load, the greater the demand on working memory. There are three basic forms of working memory that 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. Working memory follows three stages: encoding, maintenance, and retrieval. Each of these stages involves a different pattern of brain activity, and each can be affected differently by changes in load. These stages can also be distinguished by their differences in time; encoding must occur before retrieval. Isolation of these stages based on timing can be accomplished using event-related potentials (ERPs), which records 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 offers poor spatial resolution. By contrast, imaging methods that rely on hemodynamic measures, such as functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) typically measures changes in blood flow and/or oxygenation that are related to changes in brain function in brain function. These hemodynamic methods offer superior spatial resolution when compared with ERPs, which is necessary to unambiguously identify the anatomical location of brain networks supporting the different aspects and stages of working memory. However, because changes in blood flow are relatively slow, the methods of analysis tend to average over large periods of time (seconds to minutes) 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 et al. 1998, 2002) 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: Most working memory studies dealing with basic sensory processes have employed either a delayed response or N-back design. Delayed response tasks require subjects to maintain information in working memory for a period of time before a response is made. These tasks often evoke a characteristic sustained negative electrical potential over the scalp termed the contingent negative variation (CNV). The CNV is evoked during the maintenance of information stored in working memory (Tecce 1972). It is likely that it results from increased synaptic activity associated with maintaining information in the working memory store. Working memory tasks have been found to evoke a variety of other ERP components. Gevins et al. (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 with a preceding stimulus. All stimuli evoked the P200 component and the CNV, which varied in amplitude depending on task. Non-matching stimuli evoked a frontal P300 that was larger in the spatial WM task and a P450 potential over the left frontal cortex, while matching stimuli evoked a P390 potential. They concluded that WM 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 a 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, as variations in both attention and working memory can influence its production. This view is further supported by Luck et al. (2000) who suggest that many studies confound attention and working memory. They propose 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 therefore optimizing performance. Finally, Wager and Smith (2003) suggest that selective attention to features of a stimulus to be stored in WM leads to separate patterns of activation from WM storage. 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 can reveal much about working memory and cognitive load.

fMRI studies of cognitive load: Most fMRI studies of cognitive load effects 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 the additional cognitive load will increase the brain responses in a proportional way, otherwise known as the pure insertion hypothesis. Using these methods, a number of published studies have characterized brain networks supporting 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 series, in which subjects are asked to compare the current stimulus with stimuli presented one or more items earlier in a series. With an increasing delay between the first and second item to be compared, which increases the number of intervening items that must be maintained in working memory to perform the task, cognitive load increases. 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 and large regions of dorsolateral prefrontal cortex along with smaller regions of premotor cortex, superior parietal cortex and thalamus revealed changes in activity that followed an inverted U shape. 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. This finding reflects the inverted U findings in cognitive instructional psychology (e.g. Salomon, 1983; Clark & Elen, 2006) where prior knowledge is predicting mental effort under conditions where tasks become increasingly difficult.

These results demonstrated that a portion of the brain networks supporting working memory is sensitive to variations in cognitive load, while others do not appear to be as sensitive. Jaeggi et al. (2003) employed an n-back task with four levels of difficulty using auditory or visual material and did not find the inverted U result. Their 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 suggested that excessive processing demands in dual tasks are not necessarily accompanied by a reduction in brain activity. More recently, O'Hare et al. (2008) have examined the development of these brain networks using a Sternberg working memory task with three load levels. The brain networks were found to change with age, which ranged from 7 to 28 years in this study. 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 demonstrated that brain networks respond to increasing load differently from childhood through adulthood. As a result, we may find other developmental differences between the ways that younger children and adults handle cognitive load during learning.

Some of the differences observed across studies may be due to 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 and Nyberg, 2000). Area 6 activations are commonly found across tasks, including verbal, spatial, and problem-solving working memory, and so may be related to general working memory operations. 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 is found for verbal and numeric tasks compared with visuospatial tasks, which may be related to phonological processing. Activations in areas 9 and 46 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; Petrides, 1995). According to this model, 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. Other studies have found that object working memory engages ventral prefrontal regions while spatial-working memory engages dorsal prefrontal regions (Courtney, et al. 1998, 1998). However, other studies suggest that object working memory engages left-frontal regions while spatial working memory engages right frontal regions (Smith, Jonides, Koeppe, Awh, & et al., 1995; 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 frontal cortex, including parietal areas 7 and 40. In the case of verbal tasks, these activations tend to be larger on the left, which supports Baddeley's phonological loop model, where information is stored and rehersed in series (Awh, Jonides, Smith, Schumacher, & et al., 1996; Paulesu, Frith, & Frackowiak, 1993). Working memory tasks are also associated with altered activity in anterior cingulate, occipital, and cerebellar cortecies. 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 & D'Esposito (2001) who showed that activity in ventral extrastriate visual areas increased directly with load of a facial N-back working memory task. They concluded that both prefrontal and extrastriate areas worked together to meet the demands of increased cognitive load.

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 is a 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 that the solution to some of the thorny problems facing CLT requires that we step away from our century long dispute and open ourselves to the advances in neuroscience.

Specifically we recommend the use of pupilometry and eye movement studies to develop a much more reliable and valid estimate of individual cognitive load and to identify both relevant and irrelevant load. We also suggest that advanced methods of brain imaging offer many insights into the neural mechanisms that support working memory, and on the effects of changes in cognitive load on these mechanisms. For example, it appears that the relationship between load and mental effort may be an inverted U for some tasks and not for others. It also seems possible that we will find developmental differences in the processing of cognitive load by children and adults. As the 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 these 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 perceptual and response processes, affect and arousal, among other processes 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

Beatty, J. & Wagoner, B. L. (1978), *Science*. 199(4334), 1216-1218.

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

Brünken, Seufert and Paas (Chapter 4, this volume)

Callicott, J.H., Mattay, V.S., Bertolino, A., Finn, K., Coppola, R., Frank, J.A., Goldberg, T.E., Weinberger, D.R. (1999). *Cerebral Cortex*, 9, 20-26.

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

Clark, R. C. and 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. (2001). *Learning from Media: Arguments, Analysis and Evidence.* Greenwich, Conn: Information Age Publishers. ISBN 1-930608-77-2

Clark, R. E. and Elen, J. (2006). When Less is More: Research and Theory Insights about Instruction for Complex Learning, in Elen, J. and Clark, R. (Eds). *Handling Complexity in Learning Environments: Research and Theory*. Oxford: Elsevier Science Limited. 283-297.

Clark, R. E., Howard, K. and Early, (2006). Motivational challenges experienced in highly complex learning environments. In Elen, J. and Clark, R. E. (Eds.). *Handling complexity in learning environments: Research and Theory*. Oxford, G.B.: Elsevier

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.Ma., Haxby, J.V. (1998). An fMRI study of face perception and memory using random stimulus sequences. *Journal of Neurophysiology*, 79: 3257-3265.

Cowan, N. (2001). The magical number 4 in short term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24(1), 87-114.

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

Druzgal TJ, D'Esposito M. (2001). Activity in fusiform face area modulated as a function of working memory load.Brain Res Cogn Brain Res. 10(3):355-364.

Elen, J. and Clark, R. E. (Eds.) (2006). *Handling Complexity in Learning Environments: Research and Theory*. Oxford: Elsevier Science Limited.

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*, Rotterdam, The Netherlands: Sense Publishers. 109-116.

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

Gimino, A. E. (2000). Factors that influence students' investment of mental effort in academic tasks: A validation and exploratory study. Unpublished Ph.D. dissertation presented to the faculty of the Rossier School of Education, University of Southern California, Los Angeles.

Iqbal, S. T., Adamczyk, P. D., Xheng, X. S. and Bailey, B. P. (2005). Task evoked pupillary response to mental workload in human-computer interaction. Paper read at the 2005 Computer Human Interaction conference, Portland Oregon, April 2 – 7.

Iqbal, S. T., Xheng, X. S. and Bailey, B. P. (2004). Towards an index of opportunity: Understanding changes in mental workload during task execution. Paper read at the 2005 Computer Human Interaction conference, Vienna, Austria, April 24 – 29.

Jaeggi SM, Seewer R, Nirkko AC, Eckstein D, Schroth G, Groner R, Gutbrod K. (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.* New York: Henry Holt, Vol. 1, pp. 403-404.

James, W. (1904). Does 'consciousness' exist? The 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.

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. and 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. *Psychophysioloogy*, 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.

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 Wilson, J. R. & Corlett, E. N. (Eds.) *Evaluation of human work: A practical ergonomics methodology*. 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.

O'Hare ED, Lu LH, Houston SM, Bookheimer SY, Sowell ER (2008). Evidence for developmental changes in verbal working memory load-dependency. *NeuroImage*, 42, 1678-1685.

Paas, J., Kalyuga, S., & Leutner, D. (this volume). Individual differences and cognitive load theory.

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

Recarte, M. A. and Nunes, L. M. (2003). *Journal of Experimental Psychology: Applied. 9(2), 119-137. Available at http://www.apa.org/journals/releases/xap92119.pdf*

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.

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 Steele, C. M (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. (2006). How the human cognitive system deals with complexity. In Elen, J.

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

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

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

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

Wager TD, Smith EE. (2003). Neuroimaging studies of working memory: a meta-analysis. *Cogn Affect Behav Neurosci.*, 3(4):255-274.

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