

Price, J. L., Catrambone, R., & Engle, R. (2007). When capacity matters: The role of working memory in problem solving. In D. H. Jonassen (Ed.), *Learning to Solve Complex Scientific Problems* (pp. 49-76). NY: Lawrence Erlbaum.

Chapter 3

When Capacity Matters: The Role of Working Memory in Problem Solving

Jodi Price, Richard Catrambone, and Randall W. Engle
Georgia Institute of Technology

Each day we are faced with problem-solving tasks, many of which are dependent on working memory (WM) capacity, which we broadly define as the capacity for controlled processing (Engle & Kane, 2004). Take, for example, the seemingly simple task of selecting new furniture. This problem-solving task would seem to have very little to do with WM capacity at first glance. However, visualizing and deciding whether the chosen furniture will fit in the allotted space taps WM capacity as it requires holding an image of the space in mind while mentally juggling and moving the selected furniture within the imaged space. Similarly, the act of mentally calculating whether one can afford the more expensive furniture again utilizes WM capacity as one must hold a series of numbers in mind while performing additional mathematical operations before reaching a decision. Thus it becomes apparent that even a task such as selecting new furniture uses WM capacity and fits within Baddeley and Logie's (1999) notion of WM, which they defined as the following:

... multiple specialized components of cognition that allow humans to comprehend and mentally represent their immediate environment, to retain information about their immediate past experience, to support the acquisition of new knowledge, to solve problems, and to formulate, relate, and act on current goals. (p. 28)

Tasks such as selecting new furniture rarely seem overly complex and yet other tasks that also rely on WM processing can be cognitively overwhelming. This raises the question of why two tasks, equally dependent on WM capacity, can seem disproportionately difficult. The purpose of this chapter is to address this issue by providing a brief history of WM research while detailing the real-world implications of WM capacity limitations and discussing what WM capacity is and is not related to as well as when WM capacity should be expected to matter. Finally, we discuss how this relates to problem solving and detail a line of problem-solving research specifically designed to address what we currently know about WM capacity limitations.

HISTORY OF WORKING MEMORY RESEARCH

To understand why WM capacity might be expected to play a role in some tasks but seem irrelevant in others, it is necessary to address how the WM system is presumed to operate. One of the most influential models of WM was put forth by Baddeley and Hitch (1974; see also Baddeley, 1986, 2001; Baddeley & Hitch, 1994). The Baddeley and Hitch (1974) model focused and expanded on the short-term store in Atkinson and Shiffrin's (1968) "modal" model and proposed that short-term memory (STM) acts as a WM system that is responsible for temporarily maintaining and manipulating a limited amount of information to support the performance of a variety of cognitive tasks (e.g., comprehension, learning, and reasoning; Baddeley, 1986).

The three components that comprise the WM system in the Baddeley and Hitch (1974) model are a supervisory system, called the central executive, and two specialized temporary memory slave systems, the visuospatial sketchpad and the articulatory loop. In its supervisory role, the central executive oversees and coordinates the two slave systems, switches the focus of attention, and activates information previously stored in long-term memory (LTM; Baddeley & Logie, 1999). In this sense, the central executive is very similar to Norman and Shallice's (1986) supervisory attentional system. The task of temporarily maintaining information is handled by the two slave systems which are believed to briefly store information by either creating images, in the case of the visuospatial sketchpad, or utilizing rehearsal processes, in the case of the articulatory loop (Baddeley & Logie, 1999).

OTHER WAYS OF CONCEPTUALIZING WORKING MEMORY

Other approaches to WM have been proposed since Baddeley and Hitch's (1974) model. Consistent with the Baddeley and Hitch model are multistore, distributed processing models which assume that different components handle different aspects or types of processing (e.g., Bayliss,

Jarrold, Gunn, & Baddeley, 2003; Carlson, Khoo, Yaure, & Schneider, 1990; Carlson, Sullivan, & Schneider, 1989). Such models are contrasted by those that conceptualize WM as a unitary construct (Colom & Shih, 2004; Engle & Kane, 2004; Kyllonen & Christal, 1990). For example, Engle and Kane (2004) viewed the WM system as a single store in which controlled attention and related processes keep a limited amount of information from LTM activated above threshold. Debates about the unitary versus multidimensional nature of WM thus reflect differences in opinion about how information is processed and maintained in WM. Such distinctions constitute only one way to differentiate among different models of WM.

Another basis for distinguishing among the different views of WM is whether WM capacity is conceptualized as being domain and task specific or something that may be generalized across many domains and tasks. Researchers such as Daneman and Carpenter (1980, 1983) suggested that WM capacity is domain specific and that WM capacity tasks will only have predictive validity when they tap specific skills necessary in the criterion task (Hambrick & Engle, 2003). Others such as Engle and colleagues (e.g., Hambrick & Engle, 2002, 2003; Hambrick, Kane, & Engle, 2005; Kane et al., 2004) argued that, although the coding formats are specific to language or visual and spatial domains, the supervisory attention aspect of WM is domain general. They suggested that measures of WM capacity tap information processing capabilities that are useful in many tasks, thus accounting for the ability of WM capacity to predict performance in a wide variety of domains and tasks.

ASSESSING WORKING MEMORY CAPACITY

WM capacity is typically assessed using measures that combine the storage component of STM tasks with an additional requirement of simultaneous processing of other information. For example, Daneman and Carpenter's (1980) reading span measure requires participants to read and comprehend a series of two to seven sentences before being asked to recall the final word of each sentence. Turner and Engle's (1989) operation span task also requires the maintenance of words or letters while solving a series of simple math problems. One's reading or operation span (i.e., capacity) is the number of words one can correctly recall while correctly answering questions about the sentences or solving the math problems, respectively. In general then, these tasks assess how much information one can maintain in an active state while processing other information. The processing component in WM capacity tasks is believed to tap the central executive or ability to control attention, and is what distinguishes STM from WM capacity tasks (Engle, Tuholski, Laughlin, & Conway, 1999). Therefore those with higher WM spans (i.e., those that can maintain more items in an active state) are as-

sumed to have more attentional resources or greater ability to control their attentional focus relative to those who score lower on these WM span measures (Feldman Barrett, Tugade, & Engle, 2004).

TO WHAT IS WORKING MEMORY CAPACITY RELATED?

WM capacity, as reflected by scores in operation span, reading span, and other measures of WM capacity, has been found to be related to performance in a variety of tasks (Hambrick & Engle, 2003) such as reading and language comprehension (Daneman & Carpenter, 1980; Daneman & Merikle, 1996; Engle, Cantor, & Carullo, 1992; Engle, Nations, & Cantor, 1990; Turner & Engle, 1989), learning to spell (Ormrod & Cochran, 1988), and learning a new vocabulary (Daneman & Green, 1986).

WM capacity tasks have been found to predict such cognitive tasks as taking lecture notes (Kiewra & Benton, 1988), storytelling (Pratt, Boyes, Robins, & Manchester, 1989), writing (Benton, Kraft, Glover, & Plake, 1984), logic learning (Kyllonen & Stephens, 1990), comprehending and following directions (Engle, Carullo, & Collins, 1991), as well as the ability to effectively navigate in a hypertext learning environment (Lee & Tedder, 2003).

That WM capacity is predictive of performance in so many cognitive tasks has raised questions about the relation between WM capacity and intelligence, specifically fluid intelligence (Colom & Shih, 2004; Colom, Rebollo, Palacios, Juan-Espinosa, & Kyllonen, 2004; Conway, Cowan, Bunting, Theriault, & Minkoff, 2002; Mackintosh & Bennett, 2003; Schweizer & Moosbrugger, 2004), which Cattell (1963) proposed reflects the basic capacity to reason and solve novel problems. Engle, Tuholski, et al. (1999) used exploratory and confirmatory factor analyses as well as structural equation modeling to examine the nature of the constructs of STM, WM, and fluid intelligence and concluded that STM and WM are distinct, yet highly related constructs. However, despite the relatedness of STM and WM, only WM was found to relate to fluid intelligence. That WM capacity tasks involve an attention component that STM tasks do not led the authors to suggest that the link between WM capacity and fluid intelligence constructs is the ability to maintain an active representation, particularly in the face of interference or distraction (Engle, Tuholski, et al., 1999). Thus the ability to control attention is the component of WM capacity that is important to higher order functioning (Engle, Kane, & Tuholski, 1999).

WHEN DOES WORKING MEMORY CAPACITY MATTER (AND WHEN IS IT IRRELEVANT)?

The previous discussion of the many things to which WM capacity is related highlights the fact that WM capacity will matter in tasks requiring the simul-

taneous processing and storage of information. Engle, Kane, and Tuholski (1999) suggested that because WM capacity reflects the capability for controlled processing, only tasks or situations that encourage or demand controlled attention should yield individual differences in task performance. They specified seven contexts in which individual differences in WM capacity are likely to be observed:

1. When task goals may be lost unless actively maintained in WM.
2. When actions competing for responding or response preparation must be scheduled.
3. When conflict among actions must be resolved to prevent error.
4. When there is value in maintaining some task information in the face of distraction and interference.
5. When there is value in suppressing or inhibiting information irrelevant to the task.
6. When error monitoring and correction are controlled and effortful.
7. When controlled, planful search of memory is necessary or useful (p. 104).

Many of these contexts are present in tasks commonly used in psychology experiments and instructional settings. For example, category fluency tasks require individuals to say as many exemplars from a given category as possible without providing redundant answers. To do so individuals must maintain the category name (i.e., context 1) while conducting a controlled, planful search of memory (i.e., context 7), and keep a running list of items that have and have not been produced (i.e., context 4) so they can inhibit items already said (i.e., context 5) to prevent redundant responses (i.e., context 6). These same contexts are also apparent in classrooms when instructors discuss a concept and then ask students to provide examples of that concept (e.g., "What are some examples of sedimentary rocks?").

In another popular task, the Stroop (1935) task, individuals view colored bars and color words printed in opposing ink colors (e.g., the word *red* printed in green ink) and are asked to say the ink color (rather than the word) as quickly as possible. To successfully complete the Stroop task, individuals must maintain the goal of saying the ink color (i.e., context 1) while inhibiting the automatic tendency to read the word (i.e., contexts 2 and 5) to prevent erroneous responses (i.e., contexts 3 and 6).

Consistent with the Engle Kane, and Tuholski (1999) suggestion that these contexts can be expected to yield individual differences in WM capacity, Kane and Engle (2003) found that in the Stroop task, low span individuals were less able to inhibit the more automatic response (e.g., reading the word rather than saying the ink color of the word) than high span participants. Similarly, Kane, Bleckley, Conway, and Engle (2001) found that in

an antisaccade task, which required individuals to ignore a flashing light in the periphery to instead view stimuli that appeared in the opposite direction of the flashing light, low span individuals were less likely than high span participants to inhibit the automatic response (e.g., attending to the flashing light in the periphery) to achieve the goal-directed response of identifying a pattern masked letter that appeared in the opposite direction of the flashing light.

Cantor and Engle (1993) also found WM capacity-related differences in the ability to inhibit responses. They presented participants with unrelated sentences containing a subject paired with different predicates and then gave participants a speeded recognition test. Response times on the recognition test increased for both high and low WM span individuals as the number of predicates a single subject was paired with (i.e., FAN size) increased, but the increase in response times was much greater for the low than high span individuals. Cantor and Engle attributed the low spans' relatively higher response times to their having greater difficulty inhibiting previously associated subject-predicate pairs.

Similar differences in high and low WM capacity individuals have been found in tasks that require other types of controlled processing. Rosen and Engle (1997) examined the ability of their participants to use controlled versus automatic processing to overcome retroactive interference (i.e., the interference that occurs when newer material hinders memory for older items) in a category fluency task and found that high WM capacity individuals showed greater immunity to retroactive interference than the low WM span participants. Feldman Barrett et al. (2004) suggested that controlled processing depends on the central executive component of WM and "occurs when attention is applied in a goal-directed, top-down, or endogenous fashion" (p. 555). They further posited that the reason individual differences in WM capacity play a role in tasks that require the suppression or inhibition of automatically processed information is because controlled processing is necessary to suppress or inhibit this information.

Studies examining the relation between WM capacity and things such as stereotype threat, life stressors, and prejudice also provide support for the notion that WM capacity and controlled processing are necessary to suppress or inhibit task-irrelevant information. Schmader and Johns (2003) explicitly and implicitly activated stereotype threat, "the phenomenon whereby individuals perform more poorly on a task when a relevant stereotype or stigmatized social identity is made salient in the performance situation" (p. 440), and found that women and Latinos were more likely to experience a reduction in Operation span scores when placed in stereotype threat conditions, relative to control conditions. Schmader and Johns also found that WM capacity mediated the effects of stereotype threat on women's math performance and suggested that stereotype threat reduces

WM capacity because individuals utilize attentional resources, which would otherwise be devoted to task performance, to suppress the negative stereotypes. A similar explanation was offered by Klein and Boals (2001) for why life event stress was found to reduce functional WM capacity and result in lower Operation span scores. Klein and Boals attributed stressed individuals' lower scores to their devoting attentional resources to suppressing or inhibiting thoughts about the stressful event(s) rather than to their performance on the WM capacity task.

WM capacity and its role in inhibiting responses can also explain why prejudiced individuals showed reduced Stroop task performance after participating in interracial, but not same-race interactions. Richeson and Shelton (2003) measured implicit and explicit racial prejudice and then examined response modulation and behavioral control, two indicators of self-regulation believed to rely on attentional capacity, in participants asked to interact with interracial or same-race individuals before completing a Stroop task. The researchers hypothesized and found that prejudiced individuals exercised greater self-regulation during interracial interactions, relative to those in same-race interactions or less prejudiced individuals, which in turn hurt their performance on the Stroop task. Richeson and Shelton explained these findings in terms of WM capacity and executive control by suggesting that because self-regulation in interracial interactions and performance on the Stroop task involved the same attentional resources, prejudiced individuals exercising greater self-regulation were more likely to exhaust WM capacity and have fewer attentional resources left to devote to performance on the Stroop task.

A series of experiments examining "choking under pressure" (Beilock & Carr, 2005; Beilock, Kulp, Holt, & Carr, 2004; Gimmig, Huguet, Caverni, & Cury, 2005) provides additional evidence that external sources of pressure (e.g., stereotype threat, life stressors, and interracial interactions) can strain WM capacity resources and result in reduced problem-solving performance. Beilock and colleagues (2004, 2005) presented modular arithmetic problems that varied in how much they taxed WM capacity, in both low and high pressure situations, to low and high WM span individuals and asked them to indicate the "truth" of the problems. To solve modular arithmetic problems (e.g., $62 \equiv 18 \pmod{4}$), one must subtract the middle number from the first (i.e., $62 - 18$), divide the difference by the last number (i.e., $44 \div 4$), and then declare the statement "true" if the resulting dividend is a whole number. Problems containing numbers larger than 20 or requiring borrow operations were classified as having higher WM demand relative to problems with smaller numbers that did not require borrowing (e.g., $5 \equiv 3 \pmod{2}$). Participants completed several practice problems and a low-pressure test, which participants viewed as more practice problems, before being given a scenario designed to create a high pressure environment

and completing the high-pressure test. Beilock and Carr (2005) found that high WM span participants outperformed their lower WM span counterparts on the high WM demand problems, but only in the low pressure test. The low spans' disadvantage disappeared in the high pressure test because their level of performance did not decrease under pressure whereas the high spans' performance did. Beilock and Carr suggested that this somewhat counterintuitive finding, that high pressure situations emphasizing WM capacity are more detrimental to high than low WM span individuals' performance, reflects the inability of high WM spans to use the WM-taxing strategies in high pressure situations that foster their performance in low pressure situations due to pressure tapping and reducing available WM capacity resources. Gimmig et al. (2005) offered a similar explanation for the choking under pressure they observed on the Raven's Standard Progressive Matrices task in which individuals are given increasingly difficult patterns with one missing piece and asked to decide which of eight pieces will complete the pattern. Consistent with the Beilock and Carr results, Gimmig et al. found that high pressure situations hurt high more than low WM spans' performance on the complex, but not the easier items. Engle and colleagues (Kane & Engle, 2000, 2002; Rosen & Engle, 1997) have also reported similar counterintuitive findings of dual task conditions hurting high more than low WM span individuals' performance (e.g., on verbal fluency and proactive interference tasks; i.e., tasks in which older items interfere with memory for newer items).

Together these studies highlight the role of controlled processing in task performance, as each fits one or more of the Engle, Kane, and Tuholski (1999) criteria of situations in which WM capacity should be important. Situations and tasks that do not fit these criteria are cases where WM capacity should be less important if not completely irrelevant. For example, WM capacity should be less important after extensive practice at a task because, with practice, performance of the task becomes more automatic and less dependent on controlled processing (i.e., automaticity develops). Reber and Kotovsky (1997) presented evidence to support the notion that practice and expertise serve to reduce the role of WM capacity in task performance. They found that WM load initially slowed the implicit learning of how to solve the Balls and Boxes puzzle task, but that after the task was learned, WM load no longer had an effect on problem-solving performance.

Also consistent with the notion that task experience influences how much controlled attention is necessary is another set of experiments conducted by Beilock and colleagues (Beilock & Carr, 2001; Beilock, Carr, MacMahon, & Starkes, 2002). Beilock et al. (2002) found that novel sensorimotor skills (e.g., golf putting and soccer ball dribbling) were performed better in skill-focused conditions in which participants had to attend to a particular component of how they were performing the sensorimotor task (e.g., not-

ing the completion of one's golf swing in putting; noting which side of the foot was being used to dribble a soccer ball), relative to dual-task conditions that required monitoring something unrelated to the sensorimotor skill (e.g., attending to and noting different auditory tones). Conversely, expert golfers and soccer players performed better under dual-task than in skill-focused conditions. Beilock et al. suggested that the skill-focused condition resulted in better performance for novices (and experts constrained to be novices in the soccer dribbling task by using a nondominant foot), because in the early stages of skill acquisition, greater attentional control is necessary and few attentional resources are left to devote to dual-task performance. However, dual-task conditions resulted in better performance in experts for whom it is counterproductive to attend to a component of a skill (as the skill-focused conditions required), because doing so takes proceduralized skills that occur essentially outside of WM capacity and breaks them back down into smaller, independent units that must each be processed in a step-by-step attention-demanding way, similar to the way processing occurs early in skill acquisition. These findings that WM capacity demands are reduced as skill level increased are consistent with Ackerman's (1988) theory of complex skill acquisition which suggests that early task performance will be influenced by general fluid intelligence (e.g., domain-specific perceptual speed) whereas later task performance will be driven by psychomotor abilities. Thus varying levels of task experience can be expected to require varying levels of WM capacity, regardless of whether one is dealing with cognitive or sensorimotor tasks.

By considering the various tasks and constructs to which WM capacity is related and the conditions under which WM capacity can be expected to exert its influence, it becomes apparent that there are many more situations in which WM capacity matters than in which it does not. Ericsson and Delaney (1999) argued that "WM is so central to human cognition that it is hard to find activities where it is not involved" (p. 259). For example, all novel problem-solving tasks or novel task components are likely to require controlled processing until practiced. Because WM capacity can be expected to influence initial task performance, it becomes necessary to address how WM capacity affects problem solving. In the sections that follow, we discuss what implications WM research has for problem solving and detail a line of research that has incorporated what we currently know about WM capacity limitations into instructional design.

IMPLICATIONS OF WORKING MEMORY RESEARCH FOR PROBLEM SOLVING

The discussion up to this point demonstrates that WM capacity can be expected to influence problem-solving performance for tasks or situations

that require controlled processing, particularly those involving interference from previous problem-solving situations. This suggests that both instructional methods and the design of instructional materials should consider the role of WM capacity in problem solving. Two major efforts have been seen in this regard, with the introduction of Sweller's (1989) cognitive load theory (CLT) and Mayer's (2001) cognitive theory of multimedia learning. Each of these theories is based on the goal of designing instructional materials in such a way so as to reduce the learner's WM load and thereby increase understanding.

COGNITIVE LOAD THEORY

In 1989, Sweller introduced CLT, which takes what we know about the structures and functions of the human cognitive architecture and incorporates this knowledge into a set of guidelines about how best to present information to maximize learning. These guidelines are based on assumptions about the roles of LTM, WM, and WM capacity in how people learn, as well as assumptions about different factors that serve to increase or decrease various types of cognitive load (i.e., the amount of mental capacity being used; Sweller, 1989).

The Human Cognitive Architecture

CLT assumes that humans have a very limited WM capacity, but a large LTM. CLT adopts Baddeley's (1986) multicomponent model of WM, with the central executive, visuospatial sketchpad, and phonological rehearsal loop, and suggests that under certain conditions, WM capacity might be increased by utilizing both slave systems (visual and verbal) simultaneously rather than relying on one or the other (Tindall-Ford, Chandler, & Sweller, 1997). Increasing WM capacity is important because Sweller (1989) argued that, contrary to Miller's (1956) notion that we can handle five to nine items in WM, in reality humans are only capable of dealing with two to three items simultaneously if the items must be processed rather than just held in WM. If the items that are being processed in WM interact with each other in any way then this will require additional WM capacity and will serve to reduce further the number of elements or items that can be dealt with at the same time. CLT therefore posits that instructional materials can compensate for WM capacity limitations by taking advantage of our large LTM (Sweller, van M rrienboer, & Paas, 1998).

LTM plays an important role in CLT by providing a way to overcome WM capacity limitations via the creation and storage of schemas. Schema formation allows many complex knowledge elements to be stored in LTM and worked on within WM as a single unit rather than many individual

pieces of knowledge, thus bypassing WM capacity limitations and allowing more processing to occur. Schemas help organize and store the information in LTM, but also play a major role in the development of skilled performance, as individuals combine several lower level schemas into one higher level schema to ultimately build increasing numbers of increasingly complex schemas (Sweller et al., 1998). The notion that the way in which information is encoded in WM and stored in LTM (e.g., in the form of schemas) can interact to influence the development of skilled performance is consistent with models of WM that emphasize the contribution of LTM to everyday skilled performance (e.g., Ericsson & Kintsch's, 1995, long-term working memory model; see also Ericsson & Delaney, 1999).

Schema Construction and Working Memory Cognitive Load

Although schemas are stored in LTM, they are constructed based on controlled processing that occurs in WM (Sweller, 1989). CLT assumes that the effort one must exert to process the information in WM will depend both on the load that is imposed by the difficulty of the material itself, the *intrinsic* cognitive load, which is presumably unaffected by design manipulations, as well as the unnecessary or *extraneous* cognitive load imposed by poorly designed materials, which can be reduced by creating better instructional materials. *Germane* cognitive load can also be altered by design considerations, but reflects the effort that contributes to the construction of schemas. CLT therefore suggests that instructional materials should be designed to reduce extraneous cognitive load so as to free up capacity to apply toward germane cognitive load (Sweller et al., 1998).

The three types of cognitive load are additive in nature and combine to determine how difficult schema construction and ultimately learning will be for different types of instructional materials. Although both extraneous and germane cognitive load can be affected by design considerations, the intrinsic cognitive load imposed by materials depends on how many elements must be processed simultaneously in WM as well as how much element interactivity there is. When the elements can be learned or dealt with in isolation (e.g., solving a problem such as "Amy is shorter than Bobby. Bobby is shorter than Cory. Cory is shorter than Darren—Who is the shortest?"), there is low element interactivity and low cognitive load relative to materials or tasks that contain elements that must be processed simultaneously in WM to be understood (e.g., solving a problem such as "Amy, Bobby, Cory, and Darren are taking turns driving on a road trip. Each will drive one time and must drive in the following order—Amy must drive before Cory but after Darren. Darren must drive before Cory but after Bobby.—Who is the last to drive?"). Note that although both examples involve the same four people and are overly simplified examples of the an-

alytical reasoning problems one might see in graduate or law school admissions tests, there is low element interactivity in the first problem because each comparison can be processed in isolation, whereas high element interactivity exists in the second problem because the driving orders must be processed and compared at the same time in WM to determine the answer to the problem. Similarly, Sweller and Chandler (1994) noted that learning a new language involves both low element interactivity (e.g., learning individual vocabulary words) and high element interactivity (e.g., learning how to combine multiple words to form a syntactically correct sentence). Therefore, regardless of whether the learning domain is reasoning or language, as in our examples, or math, science, engineering, or technology, learners presented with materials low in element interactivity might be able to manage higher levels of extraneous cognitive load relative to those given materials high in element interactivity. The more element interactivity there is, the more important it is that extraneous cognitive load is reduced so that overall cognitive load is kept manageable for the learner (Sweller et al., 1998).

Determining what amount of cognitive load should be manageable for a learner and what constitutes too much is a difficult proposition, however. The difficulty centers on the fact that one cannot estimate the level of element interactivity in instructional materials without taking into account the learners because what constitutes a large number of interacting elements for one person might be a single element for someone with more expertise. Intrinsic cognitive load thus depends not only on the nature of the materials, but also the expertise of the learners (Sweller et al., 1998).

Schema Automation and Expertise

One aspect that contributes to experts being able to handle higher item interactivity than novices is the process of schema automation. As discussed earlier in this chapter in the context of when WM capacity is irrelevant (or at least less important), automaticity occurs after extensive practice and allows familiar components of tasks to be carried out with minimal cognitive effort. This serves to free WM capacity, making it possible to perform unfamiliar tasks at levels that otherwise might be impossible were conscious processing necessary for all of the task components. So what initially is a schema with high item interactivity might, with practice, become automatized to the point that the schema is processed with little to no load on the WM system, thus freeing the learner to focus on other aspects of the task. In keeping with this idea, Ericsson and Kintsch (1995) suggested that preexisting domain knowledge can ease encoding and processing demands on WM when dealing with domain-relevant information. Results from a study conducted by Hambrick and Engle (2002) are also consistent with the idea that domain knowledge

and WM capacity interact to determine how much element interactivity can be managed. They tested low and high WM capacity individuals' preexisting knowledge about baseball before presenting them with simulated radio broadcasts of baseball games such as the following excerpt:

Gabriel Garcia, the number seven hitter in the lineup, is next to bat ... Now Lawson delivers—and there goes the runner, and a groundball is hit into right field. That was perfect execution. The batter holds up at first base with a single, and here comes the throw to third base. The runner slides head first—and he's safe ... Sam Philipe is the next batter of the inning ...

The baseball task had high element interactivity as evinced by the number and types of things participants were instructed to keep track of while listening to the games (e.g., the number of outs, the score, which bases were occupied by which players), as well as the types of questions contained in the memory and comprehension tests (e.g., to correctly answer who struck out, which player was on third, and the score at the end of the inning required participants to simultaneously track, process, and update WM as changes occurred for multiple players in the game). Hambrick and Engle (2002) found that those with preexisting knowledge about baseball were better equipped to track changes in the games, as indicated by their test scores, but that WM capacity also influenced memory performance, regardless of the level of preexisting domain knowledge. These findings support the notion that schemas and domain knowledge (i.e., expertise) can serve to free up WM resources that are otherwise occupied in novices, thus enabling individuals with greater knowledge or greater WM capacity to handle more item interactivity (Sweller et al., 1998).

Although issues of expertise and schema automation create problems for calculating acceptable levels of cognitive load a priori, Paas and van Merriënboer (1993) have created an a posteriori computation method in which performance and cognitive load values are converted to *z* scores to allow comparison of the mental efficiency of different instructional conditions. High task performance combined with low mental effort yields scores indicative of high-instructional efficiency whereas low task performance with high mental effort indicates low-instructional efficiency. Paas, Tuovinen, Tabbers, and Van Gerven (2003) suggested that this computational method provides a way to meaningfully interpret participants' cognitive load ratings in terms of their actual performance and thus compare different instructional design methods across a variety of tasks. The usefulness of such a tool becomes apparent when one considers that the major goal of CLT is to provide guidelines for the development of instructional methods that are high in instructional efficiency (Sweller et al., 1998). This goal is also the basis for Mayer's (2001) cognitive theory of multimedia learning.

MAYER'S (2001) COGNITIVE THEORY OF MULTIMEDIA LEARNING

Mayer's (2001) cognitive theory of multimedia learning is based on the assumption that the human information processing system has two different channels, one for processing visual and pictorial information and the other for auditory and verbal processing, each of which has limited processing capabilities. The model assumes that active learning requires the learner to select and organize relevant words and images from text-narrations and illustrations before integrating them with prior knowledge. These assumptions of the cognitive theory of multimedia learning are the basis for several principles that Mayer argued should guide the design of multimedia instructional materials (i.e., those using both words [printed or spoken text] and pictures [e.g., static or dynamic graphics, illustrations, graphs, video, etc.]). These principles from Mayer's theory combine with suggestions derived from Sweller's (1989) CLT to yield a set of guidelines about how best to design and present instructional materials.

GUIDELINES FOR DESIGNING AND PRESENTING INSTRUCTIONAL MATERIALS

The guidelines that stem from Mayer's (2001) cognitive theory of multimedia learning and Sweller's (1989) CLT are empirically based suggestions about the content, format, and presentation methods one should use when designing and presenting materials to yield better learning. Following, we summarize these suggestions by stating each guideline and then detailing the supporting research.

Guideline 1: Give Learners Problems That Do Not Specify a Goal State

This first guideline represents what Sweller et al. (1998) called the goal-free effect (also known as the no-goal effect or the reduced goal-specificity effect), which reflects the finding that giving learners problems that do not specify a goal state (i.e., goal-free problems) alters how the learners go about trying to solve them. Sweller et al. (1998) suggested that goal-free problems result in better schema acquisition and lower extraneous cognitive load because rather than trying to figure out and keep in mind differences between the current and goal problem states, as one might do using means-ends analysis in conventional problems, learners given goal-free problems tend to adopt a strategy of finding any operator that can be applied to the current problem state, and, in so doing, end up with the identical result obtained by those using the more cognitively overwhelming means-ends analysis. Sweller, Mawer, and Ward (1983) found that students

given goal-free problems were superior to students given conventional kinematics and geometry problems in terms of schema construction. Similar results favoring goal-free problems have also been found using biology (Vollmeyer, Burns, & Holyoak, 1996) and trigonometry materials (Owen & Sweller, 1985).

In keeping with the idea that means-ends analysis results in greater WM load than goal-free problems, Ayres (1993) found that students given conventional two-step geometry problems made more errors at the subgoal phase than at the goal phase and rated their WM load highest at the subgoal phase because of the need to consider multiple elements in the problem at that point.

At odds with Ayres's findings, however, are the results of a series of studies conducted by Catrambone (1996, 1998) in which he examined the effect of manipulations designed to increase subgoal formation on problem-solving performance. Catrambone consistently found that learners given worked examples incorporating manipulations (e.g., labels and captions) designed to elicit self-explanation and subgoal formation had superior (i.e., more accurate) problem-solving performance, as assessed by near and far transfer tests, relative to those given worked examples without labels or captions.

Recently, Catrambone (2004) extended this line of work to examine the effect of labeling subgoals on cognitive load ratings. Participants were given paper-based study materials in the domain of physics mechanics and asked to study two worked examples after studying brief reviews of Newton's second law and trigonometry. Catrambone manipulated between-participants whether the subgoals were labeled or not in the two physics worked examples and collected cognitive load ratings once during the study phase and three times during the test phase (after completion of the near, medium, and far transfer test problems, respectively) to determine if highlighting subgoals would result in better performance, but perhaps at the expense of greater cognitive load either during training or testing. Contrary to Ayres (1993), Catrambone found that using labels designed to aid subgoal learning resulted in superior test performance on the medium and far transfer problems (ceiling effects were observed in both conditions on the near transfer problems) and lower cognitive load ratings during training and testing. This discrepancy in findings might be due to Ayres using conventional problems versus Catrambone's use of worked examples (i.e., the worked example effect) rather than subgoals per se.

Guideline 2: Give Learners Worked Examples Rather Than Conventional Problems to Solve

The finding that studying worked examples can be more beneficial than actually solving conventional problems has been called the worked example

effect by Sweller et al. (1998). The advantage of worked examples seems to be due to the fact that worked examples help focus the learners' attention on the pertinent problem states and solution steps in the problem whereas conventional problems tend to result in the learner using means ends analysis (Sweller et al., 1998). Paas and van M rrienboer (1994) found that students who studied geometry worked examples had higher transfer performance, lower extraneous cognitive load, and better schema construction than those given conventional problems. Moreover, using their mental efficiency computation method, Paas and van M rrienboer found higher instructional efficiency in the worked examples condition than in the conventional problem condition. Sweller and Cooper (1985; see also Cooper & Sweller, 1987) found similar effects with algebra worked examples. Because worked examples have been found to be effective instructional tools, an ever-increasing body of research exists examining how to design good worked examples, a topic the next guideline addresses.

Guideline 3: Avoid Forcing the Learner to Integrate Separate Pieces of Information

This guideline is based on what Sweller et al. (1998) called the split-attention effect and is derived directly from the worked example effect. It is based on the finding that worked examples that force the learner to integrate separate pieces of information to understand the material can result in more cognitive load than studying well-integrated worked examples. The split attention effect occurs when the same worked example with physically integrated material results in superior performance and reduced cognitive load relative to studying nonintegrated worked examples (Sweller et al., 1998). Work by Tarmizi and Sweller (1988) provides an example of the split attention effect in that they initially failed to obtain the worked example effect and only observed it after switching from conventional (i.e., nonintegrated) geometry examples to examples in which the information was well integrated.

Similar in nature and rationale to the split attention effect (Sweller et al., 1998) are Mayer's (2001) spatial and temporal contiguity principles. The spatial contiguity principle suggests that students learn better and experience less cognitive load when corresponding words and pictures are presented near rather than far from each other on the page or screen (Moreno & Mayer, 1999). The temporal contiguity principle addresses the timing rather than the location of information presentation and states that students learn better and experience less cognitive load when corresponding words and pictures are presented simultaneously rather than sequentially because it obviates the need to hold multiple things in memory for processing. The split attention effect, spatial, and temporal contiguity principles

combine to suggest that instructional materials should be designed to prevent the learner from having to try to hold in memory and integrate multiple pieces of information. However, if complete integration of materials cannot be addressed by location or spacing methods, it might be possible to deal with nonintegrated materials using dual modality presentation methods, a topic addressed by the fourth guideline.

Guideline 4: Present Materials Using Both Visual and Auditory Methods

Sweller et al. (1998) suggested that presenting materials both visually and auditorially can help compensate for split attention conditions when, for example, two pieces of visual information that would normally need to be physically integrated for understanding are instead combined by using both visual and auditory WM. Sweller et al. (1998) called this the modality effect, which is said to have occurred if presenting some information visually and some information auditorially results in better performance than using either the visual or auditory channel alone. Mayer (2001) suggested that the reason individuals learn better when two modalities are used rather than only one (e.g., an animation plus narration as opposed to an animation with on screen text), is because combining animations with narration utilizes both the visual and the verbal channels, whereas animations combined with on-screen text must both rely on the visual channel, which increases the likelihood that the visual channel will become overloaded and ignores the processing capacity available in the auditory channel. Experiments conducted by Mousavi, Low, and Sweller (1995) and Tindall-Ford et al. (1997) provided support for the modality effect in that both found lower cognitive load ratings when both audio and visual instructions were used than when only visual instructions were used, but only when the materials were high in element interactivity. The fact that the modality effect all but disappeared with low element interactivity materials supports the Sweller et al. (1998) claim that reductions in extraneous cognitive load via better designed instructional materials are most crucial when materials already have high intrinsic cognitive load.

Mayer (2001) argued that dual modality presentations not only help prevent WM overload, but are also more likely to induce learners to create verbal and pictorial mental models that may then be integrated for increased learning. This integration of the verbal and pictorial mental models might be facilitated due to the reduction of the effective WM load, through the use of multiple channels as suggested by CLT (Sweller et al., 1998). Mayer has named the improved learning that occurs when materials are presented with words and pictures, as opposed to words alone, the multimedia principle and cited multiple experiments in which he and his colleagues (e.g.,

Mayer & Anderson, 1991, 1992; Mayer & Gallini, 1990) have found that words and pictures resulted in better learning than words alone. Together the modality and multimedia principles suggest that combining visual and auditory presentation methods can serve to reduce WM load and increase learning. However, the benefits of dual modality presentations only hold true if the information presented in the two channels is not redundant and if extraneous, unnecessary information is removed from the materials to allow learners to focus on the relevant information. This caveat is the basis of the fifth guideline.

Guideline 5: Avoid Redundant and Irrelevant Sources of Information When Designing Instructional Materials

Sweller et al. (1998) and Mayer (2001) suggested that presenting learners with redundant information or multiple sources of information that are self-contained and can be used without reference to each other can result in WM overload and reduced learning. For example, learners given an animation and narration have been found to learn more than those given an animation, narration, and text because the additional text in the latter case is redundant with the narration and runs the risk of overloading the visual channel (Mayer, 2001). The redundancy effect is said to have occurred when students not given redundant information perform better and report lower cognitive load than students presented redundant information. However, what is redundant information for one individual (e.g., an expert) might be necessary for another (e.g., a novice) to understand a diagram or worked example. For example, McNamara, Kintsch, Singer, and Kintsch (1996) found redundancy effects for experienced learners but not for novice learners, which aligns with the Beilock et al. (2002) observation that "experienced performers suffer more than novices from conditions that call their attention to individual task components or elicit step-by-step monitoring and control" (p. 14). Once again, this suggests that it is necessary to consider the learners' level of knowledge when designing instructional materials.

Although the redundancy effect reflects the need to avoid duplicate sources of information, Mayer's (2001) coherence principle suggests detrimental effects on WM load and learning can also occur if instructional materials include extraneous information (e.g., interesting but irrelevant words, pictures, music, and sounds). Mayer suggested that the inclusion of irrelevant pictures or sounds only serves to increase WM load as well as the likelihood that the learner will fail to notice the important aspects of the lesson because his or her attention has been drawn away by this other interesting, but irrelevant, information (Moreno & Mayer, 2000). A series of four experiments conducted by Mayer, Heiser, and Lonn (2001) yielded consistent support for redundancy and coherence effects.

Such findings detail the need to develop well-designed materials that highlight the important information without drawing attention to other, unimportant or redundant information that can reduce learning and increase cognitive load. Mayer's (2001) individual differences principle suggests, however, that some types of learners are more likely to benefit from well-designed materials than others. Mayer posited that it is more important for instructional materials to be designed well for low knowledge individuals who lack the knowledge base that would allow them to compensate for poorly designed materials, than for high knowledge individuals, who may draw on their larger body of knowledge to make sense of the lesson. On the other hand, the principle states that design effects will not benefit low spatial learners who will be using all their WM capacity to hold images in WM, leaving no additional WM capacity resources to integrate the visual and verbal representations, but will help high spatial ability learners who will have enough WM capacity to handle both maintenance and integration. For these reasons, well-designed visual and verbal materials are expected to benefit low knowledge, high spatial ability learners more than high knowledge, low spatial ability learners (Mayer, 2001). So again we see that knowledge and ability levels interact with design and presentation methods to influence overall levels of WM (cognitive) load and learning.

An additional factor that interacts with design and presentation methods to influence learning is type of practice. Consistency in practice can facilitate learning but may also increase the likelihood of proactive interference (Woltz, Gardner, & Gyll, 2000). For example, Luchins's (1942) classic water jug experiment demonstrated that individuals given a series of arithmetic computation problems requiring the same sequence of water jug manipulations to obtain the correct answer had difficulty switching to a new sequence of operations to solve transfer problems. Woltz, Gardner, and Bell (2000) also observed proactive interference, or what they called *einstellung* or negative near transfer effects, in a number reduction tasks, when learners were given consistent practice with possible rule sequences necessary to solve the problems. Functional fixedness, the inability to view or use common objects in new ways (Maier, 1930, 1931), and strong-but-wrong errors, which occur when learners incorrectly apply well-practiced skills (Reason, 1990), are other examples of proactive interference that may occur after consistent practice. These examples of proactive interference in problem solving all represent cases where consistent practice can have detrimental effects and WM capacity should be necessary to overcome the interference to produce accurate problem solving and transfer performance.

Sweller et al. (1998) suggested that one way to overcome these problems and enhance transfer performance is to give learners variability in practice. Variability in practice, whether the variability stems from the context in which the task is performed or how the task is presented, results in learners

being better able to transfer what they have learned to novel tasks. Such variability produces an apparent paradox, however, because variability results in higher cognitive load ratings during practice than if type of practice is held constant, yet the variability results in better transfer performance. Paas and van M rrienboer (1994) explained this paradox in terms of germane and extraneous cognitive load by hypothesizing and finding an interaction between the two types of cognitive load. They found that if extraneous cognitive load was high, then having variability during practice increased germane cognitive load to the point where learning and transfer performance were impaired. However, if extraneous cognitive load was low, then it was beneficial to transfer training to have variability during practice.

Together these five guidelines yield multiple ways to reduce WM load and increase learning by designing instructional materials that reduce extraneous cognitive load caused by poorly designed materials, thus freeing up WM to handle more germane cognitive load and schema construction. That each of these guidelines yields testable hypotheses has resulted in a large body of problem-solving research based wholly or in part on Sweller's (1989) CLT and Mayer's (2001) cognitive theory of multimedia learning. For example, aspects of the variability effect can be found in the research examining the effects of scaffolding on problem-solving ability. Renkl, Atkinson, Maier, and Staley's (2002; see also Atkinson, Renkl, & Merrill, 2003; Renkl & Atkinson, 2003) use of scaffolding to gradually move the learner from studying worked examples to eventually having the learner solve conventional problems represents one way of manipulating the variability of practice. By moving learners from worked examples to solving conventional problems after practice, the Renkl et al. (2002) scaffolding work also capitalizes on the worked example effect and the notion put forth by CLT (Sweller et al., 1998), that what is appropriate for novices might be inappropriate or redundant once expertise is achieved in a problem-solving domain. Thus multiple aspects of the Sweller et al. (1998) and Mayer (2001) theories play a role in scaffolding.

INTEGRATING WORKING MEMORY AND PROBLEM-SOLVING RESEARCH

Mayer's (2001) multimedia design principles, Sweller's (1989) CLT instructional design principles, and the body of research each has inspired are an important first step in integrating what we know about WM capacity limitations into instructional design methods. However, examination of the WM and problem-solving-instructional design literatures suggests several gaps and areas for future empirical research that should be addressed. For example, it seems that greater emphasis must be placed on explicating when

and how expertise can be expected to interact with WM capacity and the different instructional design methods to influence learning. CLT (Sweller, 1989) addresses the role of schemas and automaticity, both of which are components of expertise, in overcoming WM limitations and element interactivity, and Mayer's (2001) individual differences principle suggests that novices are more likely to benefit from well-designed instructional materials than experts. However, neither theory leads to direct predictions about how their principles are likely to interact with varying levels of knowledge or WM capacity to influence learning other than to suggest that instructional design methods that benefit novices might prove detrimental to experts (Kalyuga, Ayres, Chandler, & Sweller, 2003; see also Kalyuga, Chandler, & Sweller, 1998; Kalyuga & Sweller, 2004). This lack of specificity is problematic in light of Hambrick and Engle's (2002) finding that domain knowledge and WM capacity each accounted for unique and varying levels of variance in the ability to recall information about simulated baseball games. This suggests it is not sufficient to know how much an individual knows about a topic or whether an individual has low or high WM capacity because each can be expected to contribute to task performance in a different way. Therefore, a clear delineation of which principles are effective for different knowledge and WM capacity levels will ultimately be needed before the CLT (Sweller, 1989; Sweller et al., 1998) and multimedia design (Mayer, 2001) principles can be maximally effective and useful for students and instructors.

A second issue that warrants further examination is how external and internal sources of pressure combine with WM capacity to influence performance, and whether any of the previously described design guidelines can be used to offset such sources of pressure. The WM research examining choking under pressure in math (Beilock & Carr, 2005; Beilock et al., 2004), life stressors (Klein & Boals, 2001), prejudice (Richeson & Shelton, 2003), stereotype threat (Schmader & Johns, 2003), and the problem-solving research investigating the effects of different aspects of instructional materials (e.g., number of elements and element interactivity) known to "pressure" and reduce WM capacity, all suggest that pressure can have deleterious effects on WM capacity or problem-solving performance. More disconcerting is the finding that pressure is most likely to negatively affect high WM capacity individuals who, under less stressful, less WM-demanding situations, would have superior problem-solving performance, relative to those with low WM capacity (Beilock & Carr, 2005). As Beilock and Carr (2005) noted, such findings have serious implications for the validity and interpretation of performance scores collected under highly stressful situations (e.g., Scholastic Assessment Test, Graduate Record Exam, Law School Admission Test, etc.) and raise questions about what scores on such measures represent (e.g., domain knowledge,

WM capacity, the effect of pressure on WM capacity). Questions such as these highlight the need for additional research into the various types of cognitive load (i.e., intrinsic, extraneous, and germane), the pressure each one induces (separately and together), and whether the multimedia design or CLT principles are in any way able to compensate for internal and external sources of pressure. For example, are there experimental manipulations that would "push" or alter intrinsic and germane cognitive load and how would such manipulations interact with WM capacity and situation-based pressure (see, e.g., Gerjets, Scheiter, & Catrambone, 2004)? Would training materials that emphasize learning subgoals increase germane cognitive load too much for a low WM span individual but be within acceptable WM load limits for a high span individual and would these acceptable limits vary as a function of internal and external sources of pressure (Catrambone, 2004)? In other words, do CLT (Sweller, 1989) and the cognitive theory of multimedia design (Mayer, 2001) need to include a "pressure principle" to account for the influence of pressure on WM capacity and problem-solving performance? Further research is necessary to see if the addition of such a principle is warranted.

Finally, although research has examined how CLT (Sweller, 1989) and multimedia design (Mayer, 2001) principles influence performance on cognitive tasks, it would seem worthwhile to also examine the usefulness of their application to the instruction of sensorimotor tasks. For example, would applying the principles stemming from the Sweller et al. (1998) modality and redundancy effects or Mayer's (2001) multimedia, spatial contiguity, and coherence principles to magazines, books, and videos designed to improve one's golf game result in one having a lower handicap? Although Ackerman's (1988) theory of complex skill acquisition and Beilock and colleagues' (Beilock & Carr, 2001; Beilock et al., 2002) findings of decreased reliance on controlled processing as sensorimotor skills develop suggest that the CLT (Sweller, 1989) and multimedia design (Mayer, 2001) principles might not apply once a skill has been acquired, it remains an empirical question whether the principles are useful in the "early" stages of sensorimotor skill acquisition.

These empirical questions will need to be addressed using a combination of methods, tasks, and manipulations commonly used in the WM and problem-solving literatures before we can definitively answer the question regarding when WM considerations can be expected to matter across different problem-solving tasks and learning environments. These literatures suggest a variety of factors that might influence WM capacity or problem-solving performance. However, until more research has been conducted to address the noted gaps, we are left with the speculation that, in problem solving and instructional design, WM capacity matters a lot, but research is needed to determine more precisely when it matters.

REFERENCES

- Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General*, 117, 288-318.
- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89-195). New York: Academic.
- Atkinson, R. K., Renkl, A., & Merrill, M. M. (2003). Transitioning from studying examples to solving problems: Effects of self-explanation prompts and fading worked-out steps. *Journal of Educational Psychology*, 95, 774-783.
- Ayres, P. (1993). Why goal-free problems can facilitate learning. *Contemporary Educational Psychology*, 18, 376-381.
- Baddeley, A. D. (1986). *Working memory*. Oxford, England: Oxford University Press.
- Baddeley, A. D. (2001). Is working memory still working? *American Psychologist*, 56, 849-864.
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47-89). New York: Academic.
- Baddeley, A. D., & Hitch, G. J. (1994). Developments in the concept of working memory. *Neuropsychology*, 8, 485-493.
- Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28-61). New York: Cambridge University Press.
- Bayliss, D. M., Jarrold, C., Gunn, D. M., & Baddeley, A. D. (2003). The complexities of complex span: Explaining individual differences in working memory in children and adults. *Journal of Experimental Psychology: General*, 132, 71-92.
- Beilock, S. L., & Carr, T. H. (2001). On the fragility of skilled performance: What governs choking under pressure? *Journal of Experimental Psychology: General*, 130, 701-725.
- Beilock, S. L., & Carr, T. H. (2005). When high-powered people fail: Working memory and "choking under pressure" in math. *Psychological Science*, 16, 101-105.
- Beilock, S. L., Carr, T. H., MacMahon, C., & Starkes, J. L. (2002). When paying attention becomes counterproductive: Impact of divided versus skill-focused attention on novice and experienced performance of sensorimotor skills. *Journal of Experimental Psychology: Applied*, 8, 6-16.
- Beilock, S. L., Kulp, C. A., Holt, L. E., & Carr, T. H. (2004). More on the fragility of performance: Choking under pressure in mathematical problem solving. *Journal of Experimental Psychology*, 133, 584-600.
- Benton, S. L., Kraft, R. G., Glover, J. A., & Plake, B. S. (1984). Cognitive capacity differences among writers. *Journal of Educational Psychology*, 76, 820-834.
- Cantor, J., & Engle, R. W. (1993). Working memory capacity as long-term memory activation: An individual differences approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 1101-1114.
- Carlson, R. A., Khoo, B. H., Yaure, R. G., & Schneider, W. (1990). Acquisition of a problem-solving skill: Levels of organization and use of working memory. *Journal of Experimental Psychology: General*, 119, 193-214.
- Carlson, R. A., Sullivan, M. A., & Schneider, W. (1989). Practice and working memory effects in building procedural skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 517-526.

- Catrambone, R. (1996). Generalizing solution procedures learned from examples. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 1020-1031.
- Catrambone, R. (1998). The subgoal learning model: Creating better examples so that students can solve novel problems. *Journal of Experimental Psychology: General*, 127, 355-376.
- Catrambone, R. (2004, November). *Subgoal learning reduces cognitive load during training and problem solving*. Paper presented at the 45th annual meeting of the Psychonomic Society, Minneapolis, MN.
- Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, 54, 1-22.
- Colom, R., Rebollo, I., Palacios, A., Juan-Espinosa, M., & Kyllonen, P. C. (2004). Working memory is (almost) perfectly predicted by *g*. *Intelligence*, 32, 277-296.
- Colom, R., & Shih, P. C. (2004). Is working memory fractionated onto different components of intelligence? A reply to Mackintosh and Bennett (2003). *Intelligence*, 32, 431-444.
- Conway, A. R. A., Cowan, N., Bunting, M. F., Theriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, 30, 163-183.
- Cooper, G., & Sweller, J. (1987). The effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology*, 79, 347-362.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 18, 450-466.
- Daneman, M., & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9, 561-584.
- Daneman, M., & Green, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language*, 25, 1-18.
- Daneman, M., & Merikle, P. M. (1996). Working memory and comprehension: A meta-analysis. *Psychonomic Bulletin and Review*, 3, 422-433.
- Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 972-992.
- Engle, R. W., Carullo, J. J., & Collins, K. W. (1991). Individual differences in working memory for comprehension and following directions. *Journal of Educational Research*, 84, 253-262.
- Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. Ross (Ed.), *The psychology of learning and motivation*, (Vol. 44, pp. 145-199). New York: Academic.
- Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102-134). New York: Cambridge University Press.
- Engle, R. W., Nations, J. K., & Cantor, J. (1990). Is "working memory capacity" just another name for word knowledge? *Journal of Educational Psychology*, 82, 799-804.
- Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, 128, 309-331.

- Ericsson, K. A., & Delaney, P. F. (1999). Long-term working memory as an alternative to capacity models of working memory in everyday skilled performance. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 257-297). New York: Cambridge University Press.
- Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102, 211-245.
- Feldman Barrett, L., Tugade, M. M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological Bulletin*, 130, 553-573.
- Gejts, P., Scheiter, K., & Catrambone, R. (2004). Designing instructional examples to reduce intrinsic cognitive load: Molar versus modular presentation of solution procedures. *Instructional Science*, 32, 33-58.
- Gimmig, D., Huguet, P., Caverni, J. P., & Cury, F. (2005). *Choking under pressure and working memory capacity: When performance pressure reduces fluid intelligence (Gf)*. Manuscript submitted for publication.
- Hambrick, D. Z., & Engle, R. W. (2002). Effects of domain knowledge, working memory capacity, and age on cognitive performance: An investigation of the knowledge-is-power hypothesis. *Cognitive Psychology*, 44, 339-387.
- Hambrick, D. Z., & Engle, R. W. (2003). The role of working memory in problem solving. In J. E. Davidson & R. J. Sternberg (Eds.), *The psychology of problem solving* (pp. 176-206). London: Cambridge University Press.
- Hambrick, D. Z., Kane, M. J., & Engle, R. W. (2005). The role of working memory in higher-level cognition: Domain-specific versus domain-general perspectives. In R. J. Sternberg & J. E. Pretz (Eds.), *Cognitive intelligence: Identifying the mechanisms of the mind* (pp. 104-121). London: Cambridge University Press.
- Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, 38, 23-31.
- Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of expertise and instructional design. *Human Factors*, 40, 1-17.
- Kalyuga, S., & Sweller, J. (2004). Measuring knowledge to optimize cognitive load factors during instruction. *Journal of Educational Psychology*, 96, 558-568.
- Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, 130, 169-183.
- Kane, M. J., & Engle, R. W. (2000). Working-memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 336-358.
- Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review*, 9, 637-671.
- Kane, M. J., & Engle, R. W. (2003). Working memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, 132, 47-70.
- Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working-memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, 133, 189-217.
- Kiewra, K. A., & Benton, S. L. (1988). The relationship between information-processing ability and notetaking. *Contemporary Educational Psychology*, 13, 33-44.
- Klein, K., & Boals, A. (2001). The relationship of life event stress and working memory capacity. *Applied Cognitive Psychology*, 15, 565-579.

- Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity? *Intelligence*, 14, 389-433.
- Kyllonen, P. C., & Stephens, D. L. (1990). Cognitive abilities as determinants of success in acquiring logic skill. *Learning and Individual Differences*, 2, 129-160.
- Lee, M. J., & Tedder, M. C. (2003). The effects of three different computer texts on readers' recall: Based on working memory capacity. *Computers in Human Behavior*, 19, 767-783.
- Luchins, A. S. (1942). Mechanization in problem solving: The effect of *Einstellung*. *Psychological Monographs*, 54(6), 1-95.
- Mackintosh, N. J., & Bennett, E. S. (2003). The fractionation of working memory maps onto different components of intelligence. *Intelligence*, 31, 519-531.
- Maier, N. R. F. (1930). Reasoning in humans: I. On direction. *Journal of Comparative Physiological Psychology*, 10, 115-143.
- Maier, N. R. F. (1931). Reasoning in humans: II. The solution of a problem and its appearance in consciousness. *Journal of Comparative Physiological Psychology*, 12, 181-194.
- Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.
- Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: An experimental test of a dual-condition hypothesis. *Journal of Educational Psychology*, 83, 484-490.
- Mayer, R. E., & Anderson, R. B. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology*, 84, 444-452.
- Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology*, 82, 715-726.
- Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology*, 93, 187-198.
- McNamara, D., Kintsch, E., Singer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, 14, 1-43.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.
- Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology*, 91, 358-368.
- Moreno, R., & Mayer, R. E. (2000). A coherence effect in multimedia learning: The case for minimizing irrelevant sounds in the design of multimedia instructional messages. *Journal of Educational Psychology*, 92, 117-125.
- Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology*, 87, 319-334.
- Norman, D. A., & Shallice, T. (1986). *Attention to action: Willed and automatic control of behavior* (Report No. 8006). San Diego: University of California Center for Human Information Processing.
- Ormrod, J. E., & Cochran, K. F. (1988). Relationship of verbal ability and working memory to spelling achievement and learning to spell. *Reading Research & Instruction*, 28(1), 33-43.
- Owen, E., & Sweller, J. (1985). What do students learn while solving mathematics problems? *Journal of Educational Psychology*, 77, 272-284.
- Paas, F., Tuovinen, J. E., Tabbers, J., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, 38, 63-71.

- Paas, F., & van M rrienboer, J. J. G. (1993). The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Human Factors*, 35, 737-743.
- Paas, F., & van M rrienboer, J. J. G. (1994). Variability of worked examples and transfer of geometrical problem solving skills: A cognitive load approach. *Journal of Educational Psychology*, 86, 122-133.
- Pratt, M. W., Boyes, C., Robins, S., & Manchester, J. (1989). Telling tales: Aging, working memory, and the narrative cohesion of story retellings. *Developmental Psychology*, 25, 628-635.
- Reason, J. T. (1990). *Human error*. Cambridge, England: Cambridge University Press.
- Reber, P. J., & Kotovsky, K. (1997). Implicit learning in problem solving: The role of working memory capacity. *Journal of Experimental Psychology: General*, 126, 178-203.
- Renkl, A., & Atkinson, R. K. (2003). Structuring the transition from example study to problem solving in cognitive skill acquisition: A cognitive load perspective. *Educational Psychologist*, 38, 15-22.
- Renkl, A., Atkinson, R. K., Maier, U. H., & Staley, R. (2002). From example study to problem solving: Smooth transitions help learning. *The Journal of Experimental Education*, 70, 293-315.
- Richeson, J. A., & Shelton, J. N. (2003). When prejudice does not pay: Effects of interracial contact on executive function. *Psychological Science*, 14, 287-290.
- Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology: General*, 126, 211-227.
- Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology*, 85, 440-452.
- Schweizer, K., & Moosbrugger, H. (2004). Attention and working memory as predictors of intelligence. *Intelligence*, 32, 329-347.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643-662.
- Sweller, J. (1989). Cognitive technology: Some procedures for facilitating learning and problem solving in mathematics and science. *Journal of Educational Psychology*, 81, 457-466.
- Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction*, 12, 185-233.
- Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction*, 2, 59-89.
- Sweller, J., Mawer, R., & Ward, M. (1983). Development of expertise in mathematical problem solving. *Journal of Experimental Psychology: General*, 112, 634-656.
- Sweller, J., van M rrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10, 251-296.
- Tarmizi, R., & Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Educational Psychology*, 80, 424-436.
- Tindall-Ford, S., Chandler, P., & Sweller, J. (1997). When two sensory modes are better than one. *Journal of Experimental Psychology: Applied*, 3, 257-287.
- Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, 28, 127-154.
- Vollmeyer, R., Burns, B. D., & Holyoak, K. J. (1996). The impact of goal specificity on strategy use and the acquisition of problem structures. *Cognitive Science*, 20, 75-100.

- Woltz, D. J., Gardner, M. K., & Bell, B. G. (2000). Negative transfer errors in sequential cognitive skills: Strong-but-wrong sequence application. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 601-625.
- Woltz, D. J., Gardner, M. K., & Gyll, S. P. (2000). The role of attention processes in near transfer of cognitive skills. *Learning and Individual Differences*, 12, 209-252.