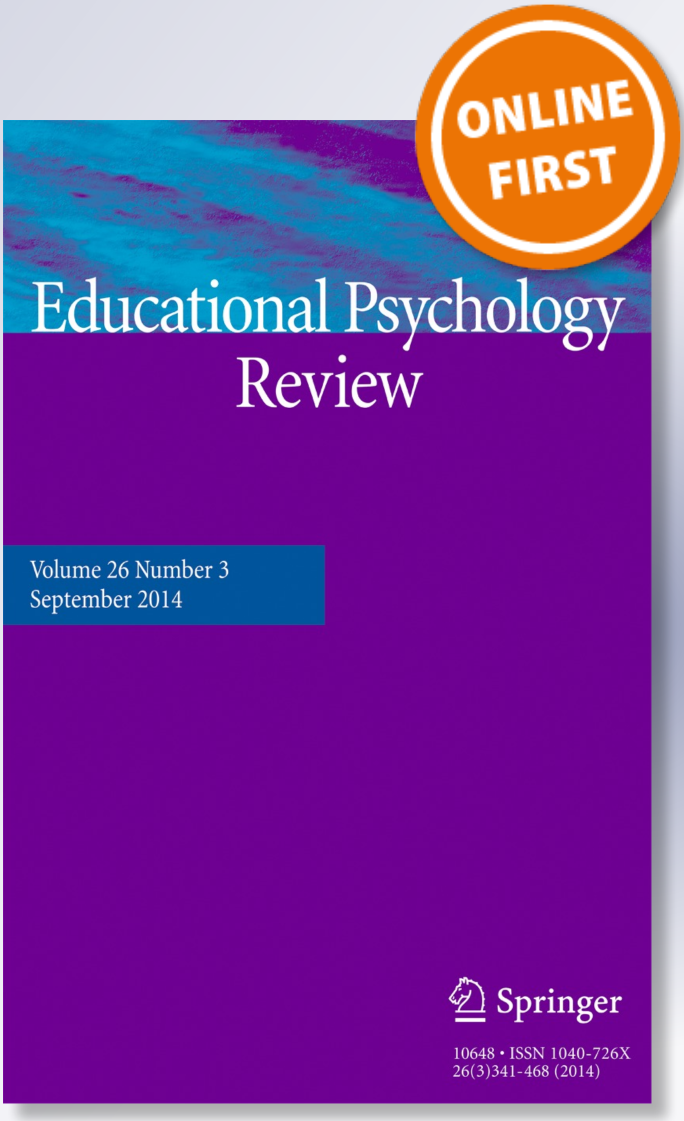
Reconceptualizing Working Memory in Educational Research

Barbara Fenesi, Faria Sana, Joseph A. Kim & David I. Shore

Educational Psychology Review

ISSN 1040-726X

Educ Psychol Rev
DOI 10.1007/s10648-014-9286-y



Educ Psychol Rev
DOI 10.1007/s10648-014-9286-y

REVIEW ARTICLE

Reconceptualizing Working Memory in Educational Research

Barbara Fenesi · Faria Sana · Joseph A. Kim · David I. Shore

Abstract In recent years, research from cognitive science has provided a solid theoretical framework to develop evidence-based interventions in education. In particular, research into reading, writing, language, mathematics and multimedia learning has been guided by the application of Baddeley's multicomponent model of working memory. However, an over-reliance on this single perspective has overlooked the theoretical diversity of contemporary research into working memory. We review the successes and shortcomings of applying Baddeley's model in accounting for a range of evidence and draw attention to alternative models that have been largely ignored within educational research. Specifically, we evaluate frameworks of working memory provided by Kane, Engle and colleagues (attentional control model) and Cowan (embedded process model). We conclude that these alternative views can support a reconceptualization of the contributions of working memory to academic learning that may not be afforded by interpretations of the prevailing multicomponent model.

Keywords Working memory · Learning · Education · Attention

Education and cognitive psychology show converging interest into the utility of working memory (WM) as a framework to understand how students think, learn and remember information. Expanded research in the domain of learning and memory highlights poor memory as a root cause of diverse learning problems. Indeed, WM capacity (WMC) may be the most relevant cognitive factor for both short-term and long-term learning (Dehn 2008; Kane and Engle 2000, 2002; Kyllonen 1996). Understanding this relation is critical to facilitate evidence-based interventions that address memory deficiencies. For example, students with a

B. Fenesi (✉) · F. Sana · J. A. Kim · D. I. Shore
Department of Psychology, Neuroscience, and Behaviour, McMaster University, 1280 Main Street West,
Hamilton, ON L8S 4K1, Canada
e-mail: fenesib@mcmaster.ca

F. Sana
e-mail: sanaf@mcmaster.ca

J. A. Kim
e-mail: kimjoe@mcmaster.ca

D. I. Shore
e-mail: dshore@mcmaster.ca

below-average WMC may suffer from poor academic performance that is misattributed as lack of motivation or reduced intelligence (Gathercole and Alloway 2004).

This paper provides an overview of the relation between WM and learning and discusses several WM frameworks that further our understanding of academic success and failure. Although the application of Baddeley's multicomponent model of WM has provided important insights into education, an over-reliance on this single perspective, which often overlooks the importance of attention and long-term memory, has ignored the theoretical diversity of contemporary research into WM. In particular, the attentional control model (Engle et al. 1999a; Kane et al. 2004) and embedded process model (Cowan 1988) each have the potential to offer important perspectives and interpretations into how WM influences academic skills and domains such as reading, writing, language learning, mathematics and multimedia learning. The goal of the current paper is to highlight how these contemporary models can guide future research to generate novel research questions leading to best practices in teaching and learning; specific research questions and programs are beyond the scope of this review.

What Is Working Memory and How Is It Measured?

WM is defined as the management, manipulation and transformation of information drawn from short- or long-term memory (Baddeley 2007; Cowan 2005; Engle and Kane 2004). WM is fundamental to thinking and learning because it allows individuals to string together thoughts and ideas currently active in memory and to link those ideas with information stored in long-term memory (LTM). A key characteristic of WM is its limited capacity to retain information while simultaneously processing additional information (Swanson 2000).

WMC has been measured using several valid and reliable tests, known as span tasks (e.g. Redick et al. 2013; Unsworth et al. 2005). Simple span tasks, such as recalling a list of presented items in correct serial order, only measure storage capacity (i.e. short-term memory (STM)). Complex span tasks, however, such as reading and verifying the truth of a series of sentences, as well as remembering the final word of each sentence for later recall, measure both storage and concurrent processing capacity (i.e. WMC). The WMC score is the total number of words recalled in correct serial order, with individuals who score in the top or bottom quartile identified as having high WMC or low WMC, respectively. Other complex span tasks preserve the idea of a dual task but use different stimuli. Importantly, complex span tasks have been repeatedly linked to higher-order cognitive abilities such as reading, writing, language learning and general academic achievement (e.g. Engle et al. 1999a).

How Is Working Memory Related to Learning?

In comparison to many cognitive factors, WMC is most highly related to learning and the best predictor of academic success (Dehn 2008; Kane and Engle 2000, 2002; Kyllonen 1996). WMC plays a significant role in allowing learners to string together thoughts and ideas by drawing on the contents of memory storage systems (short term and long term). Areas of academic achievement such as reading, written language, second language learning and math all have medium to large correlations with WMC ($r_s = 0.55–0.92$) (Engle et al. 1999b). Classroom learning also relies heavily on WMC, as students must process critical information while retaining verbal instructions given by the teacher. The effectiveness of instructional interventions (e.g. multimedia instruction, providing feedback) is influenced by WM limitations as well (Dehn 2008). For example, not allowing learners to segment (i.e.

pause) multimedia presentations selectively impairs learning for students with lower WMC, as they have greater difficulty controlling attention and ignoring irrelevant information (Lusk et al. 2009). Additionally, while many daily activities are accomplished through automatic processes (e.g. driving to work on your typical route), WM is necessary to develop automatized processing and is critical when dealing with novel problems. Learning involves distinct processes: inhibiting irrelevant information, maintaining new information and consciously retrieving information from LTM. Importantly, the majority of what must be learned and remembered must pass through WM, and thus, the capacity and functioning of WM play a major role in determining the rate and extent of learning. As a result, successful learning and academic achievement in multiple domains rely heavily on the functioning of WM.

Several theoretical approaches aim to define WM and its functional architecture. However, when it comes to examining the relation between WM and academic achievement, research has relied predominately on multicomponent views in which storage and processing components function together to accomplish learning goals. Below, we provide a brief overview of the conceptual development of WM (for a more detailed review, see Cowan 2014), highlight the impact of the multicomponent view of WM on educational research and suggest several contemporary theoretical frameworks that can provide alternative perspectives on academic success and failure.

Theoretical Frameworks of Working Memory

The modern origin of the scientific study of memory is often attributed to Hermann Ebbinghaus (1885). He examined his own memory for new information by testing himself with nonsense syllables. With relation to the study of WM, he anecdotally noted a fleeting grasp of the series of syllables during moments of concentration, but that repetitions were required to stabilize the memory for later recall (as cited in Cowan 2008). William James (1890) built on these earlier findings to propose a distinction between primary memory—information currently in consciousness that only lasts a matter of seconds (also known as STM)—and secondary memory, which provides permanent storage of memories residing outside of consciousness but available if desired (also known as LTM) (Cowan 2008).

The term working memory was introduced as an alternative conceptualization of STM by Miller and colleagues (Miller et al. 1960), who viewed WM as a system used to plan and carry out behaviour. Specifically, they proposed that an individual's WM limitation reflects the ability to maintain 'seven plus or minus two' items in active memory (Miller 1956). However, there are several important reasons why this is an overly simplistic conceptualization of WM. First, the abilities to maintain attention on relevant task goals, inhibit interference from distracting information and coordinate with pre-existing long-term stores are all significant WM factors that go beyond a simple seven plus or minus two perspective. Second, some researchers propose that the capacity limit of WM has multiple mechanisms, each of which process and retain domain-specific (e.g. auditory-verbal, visuospatial) information; therefore, a seven plus or minus two perspective does not sufficiently represent the variance in capacity limitations of domain-specific mechanisms. And third, WMC varies among people, predicts individual differences in intellectual ability and changes across the life span (Cowan 2005, 2010). Thus, the traditional notion of WM as seven plus or minus two items in active memory does not adequately represent the multitude of factors that influence WM limitations.

Nonetheless, Miller and colleagues' theoretical proposition of WM spearheaded a series of experimental studies that attempted to refine and operationalize WM, distinguish it from STM and LTM and define its functional architecture. Some researchers argue that WM consists of multiple components functioning together (Baddeley and Hitch 1974; Oberauer 2002), while others argue that WM reflects the attention-related aspects of STM (Engle 2002), and others argue that there is only one memory repository with WM comprising readily accessible information, which is available by virtue of its activation (Cowan 2008). The following will provide a detailed account of more contemporary theoretical frameworks of WM and discuss the implications of each perspective for educational research. Table 1 provides a comparison matrix of the key features among the theoretical frameworks.

Multicomponent Model

Early research on WM simply viewed short-term stores as a form of WM with important connections to cognition. However, Baddeley and Hitch (1974) focused on distinguishing WM from simple short-term stores. In a critical experiment, participants performed a reasoning task (presented visually) and concurrently recalled digit sequences of varying lengths (presented aurally). Although response time increased linearly with concurrent digit load, the error rate

Table 1 Comparisons among key features of three theoretical frameworks: multicomponent model, attentional control model and embedded process model

Key features	Theoretical frameworks		
	Multicomponent model	Attentional control model	Embedded process model
Structure	Central executive, two sub-systems (visuospatial sketchpad and phonological loop), and, episodic buffer.	No suggested hierarchy—LTM traces active above threshold, processes for achieving and maintaining activation, limited capacity controlled attention.	Hierarchically embedded subsets of memory: activated portions of LTM in response to STM cues, and subset of STM in the focus of attention.
Capacity limitation	WMC reflects limit of domain-specific stores involved in processing. Episodic buffer has capacity restriction of ~ four chunks.	WMC reflects capacity for controlled, sustained attention during interference or distraction (not about storage).	WMC is limited but flexible, (not specific to any processing domain). Capacity is approximately four chunks.
Domain generality (attention)	Less emphasis placed on attention. Central executive coordinates resource allocation between two sub-systems and integrates information.	Attentional control drives active maintenance of goal relevant information, especially during interference.	Focus of attention drives efficiency of WM processing—activation and maintenance of mental representations, attention switching, inhibition
Domain specificity	Visuospatial sketchpad and phonological loop are domain-specific stores that retain their respective information.	Acknowledgement of domain-specific stores (visuospatial, phonological), but emphasizes domain-generality as critical rather than specialized buffers.	Acknowledgement of domain-specific stores, but WM reflects attentional focus on activated sets of representations in STM or LTM.
LTM integration	LTM is functionally distinct system, which integrates active representations via the episodic buffer.	Very close connection between WM and LTM.	WM is an activated subset of LTM (not separate system).

For a comprehensive overview of multicomponent model, attentional control model and embedded process model, see Baddeley (2012), Engle and Kane (2004) and Cowan (2005), respectively remained constant at approximately 5 %. These results implicated a system that can support digit span, and also simultaneously manipulate visual and auditory information. As a result, the single unitary store model was abandoned and replaced by a three-component system: a limited capacity system with verbal and visuospatial short-term components and a central executive that was involved in the attentional control of action and coordination between the two short-term components. Baddeley (1986) also introduced a central storage system, but later removed it for the sake of what he perceived as parsimony. Recently, however, Baddeley (2000) reintroduced this central storage system as a fourth limited capacity component. This episodic buffer is a temporary multidimensional store that binds together different sources of information to form integrated chunks and forms an interface between the other two short-term storage components (verbal and visuospatial), LTM and the central executive. This buffer was added to address the incomplete taxonomy of memory systems; without the buffer, the model could not offer insight into phenomena such as where spatial information about sound or touch is stored and processed (Baddeley 1986, 2000).

Each of the short-term storage components serves different purposes. The verbal short-term store combines *phonological* processing, which recognizes, segments and blends phonemes into whole words, and *auditory* processing, which perceives, analyzes, synthesizes and discriminates all auditory stimuli to store information in the phonological STM (Hitch et al. 2001). The visuospatial store combines *visual* processing, which perceives, analyzes, synthesizes, manipulates and transforms visual patterns and images, and *spatial* processing, which localizes objects (e.g. scanning a field) to store information in visuospatial STM (Hitch et al. 2001). These domain-specific stores temporarily link the processed information to related items in LTM (Cowan 1999; Engle and Kane 2004). The links decay over time, but domain-general executive control processes, which coordinate ongoing processing within the subsystems, can maintain the links in an active state.

Attentional Control Model

Many researchers argue that WMC primarily reflects domain-general executive attention (i.e. the ability to use attention to maintain task goals and inhibit information in the face of interference; Engle 2002), analogous to Baddeley's central executive. Although cognizant of the domain-specific components of WM (i.e. short-term storage components that subserve a controlled executive; Brooks 1968; Kane et al. 2004; Oberauer et al. 2000), they argue that individual differences in executive attention are responsible for the correlations between WM span and complex cognitive measures (Engle et al. 1999a; Kane et al. 2004). This processing limitation is qualitatively different from a capacity limitation, which focuses on the number of items that can be stored (a representational limitation).

Evidence for this view comes from several studies using attention paradigms such as dichotic-listening tasks (Conway et al. 2001), anti-saccade tasks (Kane et al. 2001) and category fluency tests (Rosen and Engle 1997). For instance, in a dichotic-listening task, low-WMC subjects report hearing their name in the unattended channel while processing information in the attended channel more often than subjects with high WMC (Conway et al. 2001), suggesting that low-WMC subjects are less able to inhibit distractions. In another study (Kane and Engle 2000), when presented with a series of word lists and asked to recall words from the most recent list, low-WMC subjects are more heavily impacted by each new list compared to high-WMC subjects, suggesting that low-WMC subjects have difficulty blocking retrieval of irrelevant words. Moreover, increasing cognitive load (i.e. by performing a secondary task) adversely affected high-WMC subjects while leaving low-WMC subjects unaffected; high-WMC subjects combat interference using their attentional control, whereas low-WMC may not normally allocate attention, and thus are more affected by interfering stimuli.

Embedded Process Model

Cowan (1999) took a slightly different perspective from the multicomponent view and proposed an embedded process model in which WM is organized into two embedded levels. The first level consists of unlimited sets of LTM representations (i.e. sets of associated features or feature combinations), several of which can be temporarily activated at any given point. This activation is time-limited and subject to decay (Cowan 1988; Hebb 1949). Within these activated representations, there is a second embedded level, called the focus of attention. Those activated items within the focus of attention are even more salient. Only a few items (1–4) can reside in the focus of attention, while many more may be in a higher state of activation outside the focus of attention.

This model has two fundamental distinctions from the multicomponent view: (1) WM is viewed as a subset of LTM, rather than as a dedicated temporary storage system, and (2) STM for distinct types of stimuli (e.g. verbal, visuospatial) occurs within a common storage medium (i.e. LTM), and not in domain-specific maintenance subsystems. Thus, in the embedded process model, there is just one memory repository with WM comprising the subset of information readily accessible by virtue of its activation: Both information within the focus of attention and information in an activated state outside of attention comprise WM (Cowan 1999).

Dominance of Multicomponent Model of Working Memory in Educational Research

Despite the theoretical diversity in the WM literature, most researchers predominately use the multicomponent framework to suggest that the relation between WM and academic skills depends on domain-specific factors (i.e. auditory-verbal or visuospatial subsystems). There are clearly several empirical and practical advantages for using multicomponent theories of WM to examine educational domains. Since multicomponent models argue for highly compartmentalized WM processes, they offer evaluations and practical solutions geared towards specific sensory representations (i.e. visual vs auditory).

However, the attentional control model (Engle et al. 1999a; Kane et al. 2004) and embedded process model (Cowan 1988) are alternative theoretical approaches to WM that can further our understanding of how WM influences academic performance beyond the multicomponent model. The following section will highlight how the development and execution of educational skills such as reading, written expression, second language learning, mathematics and multimedia learning are traditionally explained from multicomponent views of WM, and how the attentional control model and embedded process model can expand the theoretical diversity concerning the role of WM in educational research.

Reading

Multicomponent Model

Reading skills heavily depend on several WM components to retain text information, which facilitates comprehension of subsequent sentences. For example, researchers employing the multicomponent view of WM predict that verbal components of WM (not so much visuospatial components) play a critical role in reading ability. Indeed, several studies demonstrate that verbal WM tasks (e.g. reading span task by Daneman and Carpenter 1980) and reading comprehension tasks both draw on verbal components of WM, including the phonological loop of WM (Swanson et al. 2006; Waters and Caplan 1996). Additional evidence corroborates the role for different components of WM in reading; verbal WM tasks are strongly related to reading tasks unlike visuospatial WM tasks, which are only moderately related to reading tasks (Daneman and Tardif 1987; Cornoldi et al. 1996). Moreover, skilled readers outperform less skilled readers on measures of verbal WM but not on measures of visuospatial WM (e.g. Cornoldi et al. 1996), and less skilled readers exhibit poor verbal STM spans (e.g. Nation et al. 1999). In fact, several literacy evaluations rely on the multicomponent view to diagnose individuals with learning difficulties using verbal complex span (reading and listening span) tasks, which target deficits in specific WM components (i.e. overall verbal WM or specifically the phonological loop). However, this multicomponent framework does not entirely account for a range of findings; it is not just domain-specific complex (reading) span tasks that predict reading comprehension, but domain-general (operation) span tasks also do, indicating that the verbal WM component in reading span does not fully account for WMC and reading comprehension associations (Daneman and Merikle 1996; Engle et al. 1992; Kane et al. 2004). Moreover, the multicomponent view does not elaborate on the role of LTM, which is arguably intricately tied to WM, especially during integration of semantic information with the current WM context.

Attentional Control Model

A theoretical perspective based on attentional control argues that WM functioning has more to do with the ability to control and allocate attention during complex cognitive tasks than with processing information specific to the phonological and visuospatial components. As a result, better WM functioning results from greater attentional control. For instance, the involvement of WM in reading comprehension is domain-general (i.e. dependent on the attentional/executive control component of WM tasks (e.g. Engle et al. 1999a; Turner and Engle 1989); therefore, proponents of the attentional control model would predict that reading proficiency is more heavily influenced by successful allocation of attention during reading. In fact, it has been reported that tasks requiring both maintenance and manipulation of information (i.e. those associated with executive function) correlate strongly with reading comprehension, regardless of task modality (e.g. Daneman and Merikle 1996). The strong link between attentional control and reading ability is also seen in individuals with attention deficit disorder (ADD), which is marked by difficulty controlling behaviour and paying attention. Individuals with ADD perform significantly worse than controls on attentional tasks involving impulse control and planning, as well as on standardized measures of reading comprehension (Willcutt et al. 2001). Poor reading comprehension is often mislabeled as dyslexia, and the focus is geared towards the reading problem as opposed to the attentional issue. Individuals with attention disorders are often easily distracted, citing the need to read things several times; they also tend to read slowly and spend their attention decoding information. Clearly, there is an important domain-general, attentional component influencing reading ability that extends beyond a multicomponent view of WM. There is undoubtedly ample evidence suggesting that domain-specific components (e.g. phonological loop) impact reading ability, but the strong link between domain-general attentional factors and reading ability suggests that it plays an equally large role in explaining the relation between reading and models of WM. It is important to note that the domain-general attentional mechanisms of the multicomponent model (i.e. episodic buffer and/or the central executive) can also be used to explain the link between attention and reading ability; however, the critical distinction here concerns the reliance on domain-specific stores, rather than attentional factors, as the key mechanisms driving the relation between WM and reading.

Embedded Process Model

As discussed earlier, reading skills and comprehension have been heavily linked with the coordination of several WM components, specifically the phonological loop of verbal WM (Swanson et al. 2006; Waters and Caplan 1996). However, Cowan's embedded process model, with its emphasis on activated LTM as the foundation for all memory representations, can highlight alternative links between reading ability and memory. Specifically, this model may predict that poor readers experience ineffectively activated LTM representations of text information and inadequate attentional focus during reading. Activated LTM elements can account for unique variance in reading comprehension, distinct from variance accounted for by traditional WM measures (Was and Woltz 2007). Poor readers may not necessarily have impairments in the function of verbal WM and the phonological loop, but rather, they may have difficulty activating LTM representations that facilitate appropriate links or references between concepts in the text. They may also have difficulty activating previously stored vocabulary items, making it challenging to understand the text's meaning, which further degrades the referential relation between sentences, and ultimately impairs comprehension. Clearly, the embedded process model and its emphasis on activated LTM elements during information acquisition go beyond a focus of domain-specific subsystems and make an important contribution to the understanding of how WM influences reading.

Written Expression

Multicomponent Model

Writing processes engage different WM components (Kellogg et al. 2007). The multicomponent model posits that writing involves both the visuospatial sketchpad and the phonological loop. Specifically, visualization of images and organization of diagrams require access to the visuospatial sketchpad (Dehn 2008), whereas translating ideas to writing and editing impose large demands on verbal WM, including the phonological loop (for a review, see Alamargot and Chanquoy 2001, chapter 5; Olive 2004). Storing phonological representations of words, or sentences under construction, taxes phonological WM, while producing language and recalling concrete nouns tax visual WM (Kellogg et al. 2007). Finally, the central executive coordinates the components necessary to successfully produce written expression (Alamargot and Chanquoy 2001; McCutchen 1996; Olive 2004). Similar to other educational skills, researchers have also used the multicomponent view to demonstrate that students with higher verbal WMC write more complex sentences than students with lower verbal WMC (Swanson and Siegel 2001). However, recent evidence suggests that problems with written output may be due to deficits in more general executive processes and broader memory problems (Kellogg 1999; Levy and Marek 1999). These findings are in line with the notion that written expression requires the active maintenance of multiple ideas, the retrieval of grammatical rules from LTM and the management of these simultaneous processes.

Attentional Control Model

Attentional models also propose that WM is crucial to written expression and predict that selective attention is needed to maintain focus and inhibit information that might interfere with successful writing (Engle and Kane 2004). Domain-general executive attention appears to be the only WM component that predicts planning, translating, revising and recalling vocabulary during the writing process (Vanderberg and Swanson 2007). Thus, writers with higher WMC tend to construct more organized representations of their expressions, since they are better able to exert attentional control and inhibit unwanted, scattered and intrusive thoughts. Conversely, writers with lower WMC may have difficulty sustaining attention on goal-relevant information and may be unable to inhibit distraction in order to stay on task during the writing process. Although writing is a multi-domain process (i.e. both verbal and visual), the importance of attentional control for successful written expression has largely been underestimated.

Embedded Process Model

The embedded process model goes beyond the domain-specific and domain-general processes and predicts that writing heavily depends on activated LTM. For instance, grammar and semantics can support more effective writing; for unrelated words, only five can be remembered and transcribed whereas for meaningful sentences, the capacity can be as much as 15 items (Baddeley 2010; Baddeley et al. 1987; Baddeley and Wilson 2002). In addition to such writing rules, LTM also holds historical knowledge and anecdotal experiences, all of which need to be effectively drawn upon during the writing process. If any of these sets of information are not activated, the quality of the written expression can be compromised. Several studies have demonstrated that LTM is critical for retrieval of grammar rules and recursive self-monitoring to promote high-level writing ability (Kellogg 1996, 1999; McCutchen 1996; Swanson and Berninger 1994). Researchers often overlook this theoretical framework of WM to explain such important academic skills with the justification that the multicomponent view has already been extended to work on language production and comprehension (Gathercole and Baddeley 1990, 1993).

Second Language

Multicomponent Model

Processes involved in reading and writing also apply to learning a second language. Overall, WM plays the critical role of constructing and integrating ideas from a stream of successive words (Just and Carpenter 1992). The multicomponent view of WM predicts that the phonological loop is a key component to understanding the acquisition of language (see Baddeley 2003 for an overview). Evidence for this came initially from the inability of a patient with a pure phonological loop deficit to acquire the vocabulary of a new language, despite otherwise normal verbal LTM (Baddeley et al. 1998). Further, factors disrupting the phonological loop, such as articulatory suppression, phonological similarity and word length, also disrupt the acquisition of foreign vocabulary (e.g. Gathercole and Baddeley 1990). Verbal WM is also necessary for the acquisition of new vocabulary, as it links the correct pronunciation with a semantic representation (Baddeley 1990; Crain et al. 1990). Furthermore, phonological STM capacity is a good predictor of the ability of children (Service 1992) and adults (Atkins and Baddeley 1998) to learn a second language (Baddeley 2003). However, these findings, while consistent with the multicomponent view of WM, underemphasize the critical roles for attention and LTM.

Attentional Control Model

The critical role for attention in successful second language learning comes through the use of feedback. In a small-scale study, Mackey et al. (2002) found a relation between noticing feedback on errors and WMC; learners with low WMC reported lower levels of noticing errors during task performance compared to learners with higher WMC who were more likely to have noticed the relevant feedback. Learners with greater WMC may also be better able to cope with the multifaceted demands of language learning, including incorporating feedback to facilitate language learning. Thus, perhaps attention control, at least with respect to second language acquisition, can be conceptualized as regulating attentional resources, maintaining access to relevant information, blocking access to irrelevant information and devoting attentional resources to competing cognitive demands of language learning (Kane et al. 2007).

Embedded Process Model

According to Cowan (1999), WM is not distinct from information in LTM but, rather, reflects information in LTM that either has greater activation or is in the focus of attention. Because verbal information in LTM is essentially knowledge of the language, the main difference between verbal WM and language knowledge lies in the degree to which the activity requires attentional control. Thus, the embedded process model postulates that WM reflects activated LTM in many ways with respect to second language acquisition. For example, immediate memory for a telephone number spoken in your native language is substantially better than that for a number spoken in a foreign language, reflecting the importance of long-term knowledge in WM. Similarly, understanding phrases and sentences requires access to representations in LTM and the integration of those representations with actively maintained information. Thus, language-processing impairments may not be due to subsystems operating inadequately, but rather the inability to activate related long-term stores. The embedded process model also represents larger issues in the domain of language processing. Given the nature of WM as dependent on LTM representations that are activated by ongoing cognition, rather than as a dedicated short-term store, it is clear why activity during verbal WM tasks overlaps neuroanatomically with activity in non-mnemonic language tasks, in that they likely utilize similar long-term representations (Chein et al. 2003).

Mathematics

Multicomponent Model

Both the verbal and the executive components of WM are implicated to varying degrees in mathematical calculations (Andersson and Lyxell 2007; Tronsky 2005). The effect of presentation variables (e.g. auditory versus visual, or horizontal versus vertical) and strategy use can be seen as influencing which components of WM are recruited. For instance, the phonological loop is involved in multi-digit arithmetic when problems are presented in either auditory or visual format, because individuals translate visually presented information into a phonological code for temporary storage (Noel et al. 2001). The visuospatial sketchpad is involved only when problems are presented visually (e.g. Logie et al. 1994). The central executive is implicated across operations as the number of carry or borrow operations increases and with higher values of the carry (Imbo et al. 2007, 2008). No doubt the multicomponent view significantly increased our understanding of the relation between WM and mathematics. However, WM is also limited in the degree of attentional control that can be exerted and may play a greater role than domain-specific processing in the early stages of math acquisition; greater attentional control may allow individuals to inhibit irrelevant information and draw on knowledge from LTM to develop accessible representations (e.g. number facts, counting procedures) and retrieve task-specific information. Indeed, individuals with arithmetic disabilities have difficulty inhibiting irrelevant information and retrieving appropriate task-specific information from LTM (Passolunghi and Siegel 2001; Rourke 1993).

Attentional Control Model

Although mathematical ability has been heavily linked to specific WM components, such as the phonological loop and the visuospatial sketchpad, successful allocation of attention and inhibition of distraction is also predicted to have an important role in mathematical ability. A key factor is the ability to focus attention on relevant parts of a problem, so that appropriate strategies and knowledge contained within LTM can be activated and applied to solve the problem. It is equally important to inhibit distraction from superficially relevant aspects of a math problem, which are embedded within math problems to distract learners from important elements requiring attention. This perspective is supported by research showing students with attentional problems perform better when given a computation task containing highlighted regions, which focuses their attention on important information (Kercood and Grskovic 2009). Poor allocation of attention to key features of a problem, either directly or indirectly, impairs math performance.

Embedded Process Model

In addition to considering domain-specific elements as influential in mathematical ability and learning, it is also important to consider a broader LTM representation. Successful mathematical problem solving requires the retrieval of appropriate problem-solving strategies and computational formulas. When faced with a math problem, whether purely computational or word-based, the embedded process model predicts that learners must activate appropriate strategies and rules previously learned and stored in LTM and apply that prior knowledge to generate a solution. Math difficulties might result from an inability to activate appropriate problem-solving strategies or an inability to recall specific rules that need to be applied to accurately solve a problem. Indeed, math difficulties share many common features with reading difficulties, including domain-general features such as LTM and attention to details (Das and Janzen 2004). However, it is important not to undermine the extensive research that argues for the significant role of domain-specific processes in math ability (e.g. verbal, visual-spatial components). The key difference between those studies and an embedded process approach is that perhaps the inability to retrieve appropriate strategies or rules from LTM might impair the functioning of more domain-specific components (e.g. verbal, visuospatial). Essentially, there is an interconnected system whereby ineffective LTM retrieval impairs adequate domain-specific functioning, so that the end result impaired mathematical ability.

Multimedia

Multicomponent Model

The two most dominant theories in multimedia design research (e.g. Mayer 1989; Sweller and Chandler 1994) rely heavily on a multicomponent view of WM. Cognitive load theory (CLT; Sweller 1994) of multimedia learning is part of a larger class of limited capacity theories, which posit that learners have a limited capacity WM system and an unlimited LTM store. Within these frameworks, impaired learning is a consequence of instructional materials overwhelming learners' limited WM resources. CLT is dictated primarily by multicomponent views of WM, because it states that WMC is distributed over two, partially independent and compartmentalized processors: visual and auditory. CLT uses dual-processing assumptions based in part on Paivio's (1986) dual-coding theory and Baddeley's (1998) theory of WM, both of which suggest that there are two separate channels for processing visual and auditory information. Importantly, WMC can be optimized by engaging both visual and auditory channels rather than relying on either processing channel alone. For example, presenting additional explanatory information as narration instead of adding it as text to an already complex visual display reduces WM burdens by allowing learners to 'off-load' processing from the visual channel to the auditory channel (Smith and Ragan 2005).

Cognitive theory of multimedia learning (CTML; Mayer 1989; Mayer 2009) also relies on a multicomponent framework of WM to drive instructional design research. Similar to CLT, CTML outlines two separate channels (auditory and visual) for processing information (i.e. dual-coding theory) and suggests that each channel has a limited capacity (similar to Sweller's notion of cognitive load). Learning is an active process of filtering, selecting, organizing and integrating information based on prior knowledge into mental representations. This theory not only compartmentalizes information processing into two separate channels (auditory and visual), but it also compartmentalizes specific memory stores involved during multimedia learning. Sensory memory receives stimuli and stores it temporarily (i.e. initial auditory and visual input from multimedia instruction), WM actively processes information to create mental constructs, and LTM stores all previously learned information.

Although the multicomponent view of WM has heavily influenced multimedia design research and has led to the development of design principles that promote effective multimedia instruction (for a comprehensive review of multimedia design principles, see Mayer 2009), the influence of attentional limitations and LTM representations also impacts multimedia design. Beyond the capacity limitations of dichotomous verbal and visual WM channels, a learner's ability to selectively attend to relevant images and sounds in a multimedia display (while simultaneously inhibiting attention to irrelevant information), as well as how a learner's pre-existing knowledge in LTM interacts with new incoming information, significantly impacts effective multimedia design and successful learning.

Attentional Control Model

The majority of multimedia principles outlined by CLT and CTML employ a multicomponent view of WM to predict learning from various combinations of words and images. The modality principle states that presenting information in separate modalities—visual and verbal—is better than presenting information in the same modality (Mayer 2009). Learners are able to use separate visual and auditory WM channels to process information and avoid overwhelming their visual channel with exclusively on-screen information. However, the detriment of multiple sources of visual information could also reflect an inability to effectively allocate attention to both information sources. This is reflected in the split-attention principle (Mayer 2009), which highlights how learners are often 'splitting' their attention between disparate sources of visual information, unable to reconcile information necessary for comprehension. Essentially, the attentional control model would predict that the requirements of attending to both the on-screen text and on-screen images exceed the learners' attentional abilities and hinders comprehension of presented material.

The verbal redundancy principle (Mayer 2009) also states that pairing verbatim on-screen text with narration impairs learning because it overwhelms the auditory/verbal WM channel. However, learners typically read words and sentences faster than their aurally presented counterparts. Since learner's eyes are always a few words ahead of what they are hearing in the auditory stream, proponents of the attentional control model might predict that visual attention may be constantly disrupted and redirected several words back to realign the on-screen and auditory verbal input. Although further empirical evidence is necessary to validate these claims, the redundancy principle may have a greater attentional control limitation than is currently emphasized.

Embedded Process Model

Cowan's embedded process model may acknowledge the contributions of visual and verbal WM components in multimedia learning but emphasizes that certain presentation styles make it more difficult to activate appropriate LTM representations. Importantly, this model suggests that long-term stores must be engaged to establish new information within a related context to achieve a stable mental model of the newly acquired information. Indeed, the embedded process model more readily integrates the influence of prior knowledge on multimedia learning (Schweppe and Rummer 2014) and directly stipulates how information is integrated between WM and LTM.

The modality principle is explained by CTML and CLT as effective multimedia design, because it allows separate subsystems of WM (visual and verbal) to simultaneously process information, rather than one subsystem engaging in all processing and becoming overloaded. However, the embedded process model could explain the modality effect in terms of modality-specific interference between LTM representations. As a result, if multimedia instruction engages the same modality, the prior knowledge representations that need to be activated to integrate with new incoming information might experience interference due to similar activation routes. Clearly, several multimedia principles can also be explained with a heavier emphasis on attentional limitations and LTM representations rather than capacity limitations of dichotomous verbal and visual WM channels.

Additional Theoretical Frameworks of Working Memory Less Directly Applicable to Educational Research

The previous section highlighted how the attentional control model and the embedded process model of WM provide additional insight into how WM impacts learning beyond the typical multicomponent approach. These alternative theoretical models can diversify our understanding of how WM influences different educational skills and domains and can potentially help identify mechanisms that alleviate WM deficits and improve learning. As suggested by Cowan (2014), there are several other WM frameworks with equal theoretical importance in explaining memory processes, such as models that specify components and parameters in mathematical terms. These frameworks have the potential to contribute to education by making the assumptions of the models more explicit. For instance, models such as time-based resource sharing (Camos and Barrouillet 2011; Gaillard et al. 2011; Oberauer and Lewandowsky 2011) and ACT-R (Anderson 2005; Anderson et al. 2004) have been applied to incorporate WM within an overall mathematical simulation of human information processing. In fact, Cowan et al. (2012) provided a mathematical version of the embedded process model that makes explicit the use of chunking with LTM for a putatively STM task. Although such modelling approaches are difficult to apply to educational practice at this point, they will present important opportunities for useful educational research in the near future.

Additionally, there are several other WM frameworks, such as the temporal context model and dual-component models, that provide important theoretical contributions to WM research, but are currently difficult to practically apply in the educational domains. The following section will examine the important theoretical contributions of these models to WM research and also highlight their drawbacks regarding educational applicability.

Temporal Context Model

The temporal context model suggests a purely unitary-store account of memory based on temporal discriminability and a gradually changing internal contextual state (Kahana et al. 2008; Nairne 2002; Sederberg et al. 2008; Usher et al. 2008). This is in stark contrast to multicomponent and attentional control models of WM. The main supporting evidence for the temporal context model comes from the long-term recency effect in continual distractor free recall, which involves a filled distractor interval not only between the last item and the recall test, but also in the interval between the study of each item. Interestingly, this task leads to the typical recency effect (end-of-list items are better recalled than early-list items). Dual-store models of memory implicating separable short-term and long-term stores cannot account for long-term recency if the end-of-list distractor in delayed free recall is sufficient to clear STM; then, it should also be sufficient to clear STM in continual distractor free recall. As a result, the temporal context model suggests that recall of an item depends not on its absolute recency, but on its relative recency to other list items. Importantly, when an item is recovered at test, it reinstates the temporal context active when that item was studied. Because this context overlaps with the encoding context of the items' neighbours, subsequent items from the list are recalled.

However, temporal context models that suggest similar STM and LTM stores are not universally accepted. Many researchers continue to experimentally demonstrate dissociations between short- and long-term recency (Davelaar et al. 2005). Also, while the temporal context model clearly highlights the potential importance of not compartmentalizing our theoretical conceptions of memory, current research using this model involves computational modelling and computer-simulated designs that involve learning of word lists and performance on serial and free recall tests. Although this method allows tight control of modelling parameters, most findings are not generalizable to an educational learning context, where learned material is greater in complexity and hierarchical in nature. Thus, the temporal context model makes it challenging to generate tractable research questions regarding the impact of its single store memory model on various educational skills and domains, which involve more complex learning processes beyond word list learning.

Dual-Component Model (Active Maintenance and Context-Based Retrieval)

Given the ongoing debate between unitary and dual stores of memory, some researchers have attempted to reconcile the competing models by combining notions of WM as active maintenance with work arguing for the importance of context-based retrieval processes (Unsworth and Engle 2007). Specifically, WM maintains new information in a heightened state of activity in the face of distraction and interference. Contextual cues then help to discriminate between relevant and irrelevant information in relation to the current task goal. In experiments using item-list recall, low-WMC subjects typically recall fewer items, emit more intrusions (particularly previous list intrusions) and are slower to recall items than high-WMC subjects. Low-WMC subjects are likely to use context cues that activate more irrelevant information (i.e. items from previous trials) than those used by high-WMC subjects.

Although this framework combines a flexible attentional component with a cue-dependent search mechanism, it encounters similar educational applicability issues as the temporal context model. Although experiments examining dual-component models require human participants (rather than computer simulations), the experimental materials also involve learning word pairs or items from lists. Again, this design allows tight control of modelling parameters but does not allow for reliable generalizability to an educational context, where learned content is greater in complexity and hierarchically organized. Effective educational research relies quite heavily on understanding the effect of various strategies and methods on memory function. As a result, research investigating optimal teaching and learning strategies must rely on theories of memory, such as multicomponent, embedded processes and attentional control models, which can be widely applied to educational research.

Conclusion

Educational researchers have generally relied on the multicomponent model to investigate the role of WM in an educational context and have typically interpreted their findings almost exclusively using this framework. We have described several theoretical models of WM that have the potential to make important contributions to evidence-based interventions in education. These models have been largely ignored and await rigorous testing to determine if their theoretical propositions are educationally relevant. We have demonstrated that many educational skills and domains are parsimoniously explained by attentional and activation-based models. For example, one study showed that students with ADD typically show poor reading comprehension skills due to problems in attentional control (Karatekin 2004) rather than deficits in phonological loop. In fact, Karatekin and colleagues found that children with ADD could maintain verbal information in the same manner as typically developing children; however, their impairment occurred in their ability to control attention. Thus, the attentional control model can be used to directly measure if reading impairments among this population are due to poor inhibitory control or goal maintenance.

Additionally, the embedded process model has already begun to formulate novel hypotheses for reduced multimedia learning when information is presented in similar modalities (auditory) as opposed to separate modalities (auditory and visual). This model posits that if multimedia instruction engages the same modality, the prior knowledge representations that require activation to integrate with new incoming information might experience interference due to similar activation routes. The critical role of LTM helps re-conceptualize the underlying mechanism for the detrimental impact of same-modality information presentation. Notably, these novel hypotheses are outside the focus of the multicomponent model, which concentrates on the function of discrete, domain-specific WM components.

We have argued that several frameworks of WM offer a rationale for expecting a demonstrable relation between WM and learning. By offering such alternative perspectives regarding the application of WM in educational research, these models may also speak to larger issues in the domain of applied cognition. Although researchers in education typically view WM as a dedicated short-term store, alternate theoretical frameworks encourage a more enriched view. For example, by conceptualizing WM as dependent on representations in LTM that are activated by ongoing cognition (i.e. embedded process model) or as a form of executive attention that controls attention under conditions of interference or distraction (i.e. attentional control model), researchers can explore the role of LTM or attention in WM, rather than the specific roles of verbal or visuospatial components. For instance, other theoretical models can provide further insight as to why a general attentional deficit can persist without deficits in verbal or visuospatial components of WM.

We conclude with a few words of caution. First, although we suggest that other WM models may provide a successful framework in which to interpret a wide range of educational findings, it is important to recognize that there are many additional models that could prove equally useful for understanding the theoretical basis of WM and its connection to learning. Our intent is not to suggest that the tenets of Cowan's or Engle and Kane's models should be accepted whole-cloth, as has largely been done with Baddeley's model, but to show that there is utility in considering the range of theoretical constructs that are offered by alternative models. Additionally, a deeper understanding of the central executive in the multicomponent model could also encourage researchers to consider the role of attentional control in WM and its impact on learning, as it is currently the domain-specific subsystems that receive the most empirical investigation. By demonstrating the usefulness of these alternative theories, our intent was also to inspire rigorous empirical testing to reliably determine the educational applicability of additional theoretical approaches. Second, it is important to acknowledge the difficulty of directly comparing theoretical frameworks, as often times the frameworks do not provide explicit enough predictions, and simply provide principles for guiding research in a more general way. Regardless, we are optimistic that by exploring several theoretical models, it will be possible to move beyond oversimplified notions regarding the nature of processing in WM and to reach a unified account that reconciles a wide range of seemingly incompatible or incomplete findings in the converging domains of research on WM and education.

References

Alamargot, D., & Chanquoy, L. (2001). *Through the models of writing*. Dordrecht: Kluwer Academic Publishers.

Anderson, J. R. (2005). *Cognitive psychology and its implications*. New York: Worth Publishers.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review, 111*(4), 1036.

Andersson, U., & Lyxell, B. (2007). Working memory deficit in children with mathematical difficulties: a general or specific deficit? *Journal of Experimental Child Psychology, 96*, 197–228.

Atkins, P. W. B., & Baddeley, A. D. (1998). Working memory and distributed vocabulary learning. *Applied Psycholinguistics, 19*, 537–552.

Baddeley, A. (1986). *Working memory. Oxford Psychology Series #11*. Oxford: Clarendon.

Baddeley, A. D. (1990). *Human memory: theory and practice*. London: Lawrence Erlbaum Associates.

Baddeley, A. D. (1998). Recent developments in working memory. *Current Opinion in Neurobiology, 8*, 234–238.

Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences, 4*, 417–423.

Baddeley, A. D. (2003). Working memory: looking back and looking forward. *Nature Reviews Neuroscience, 4*, 829–839.

Baddeley, A. D. (2007). *Working memory, thought and action*. Oxford: Oxford University Press.

Baddeley, A. (2010). Working memory. *Current Biology, 20*(4), R136–R140.

Baddeley, A. D. (2012). Working memory: theories, models, and controversies. *Annual Review of Psychology, 63*, 1–29.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–89). New York: Academic.

Baddeley, A. D., & Wilson, B. A. (2002). Prose recall and amnesia: implications for the structure of working memory. *Neuropsychologia, 40*, 1737–1743.

Baddeley, A. D., Vallar, G., & Wilson, B. A. (1987). Sentence comprehension and phonological memory: some neuropsychological evidence. In M. Coltheart (Ed.), *Attention and performance XII: the psychology of reading* (pp. 509–529). Hove: Erlbaum.

Baddeley, A. D., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*(1), 158–173.

Brooks, L. R. (1968). Spatial and verbal components of the act of recall. *Canadian Journal of Psychology, 22*, 349–368.

Camos, V., & Barrouillet, P. (2011). Developmental change in working memory strategies: from passive maintenance to active refreshing. *Developmental Psychology, 47*, 898–904.

Chein, J. M., Ravizza, S. M., & Fiez, J. A. (2003). Using neuroimaging to evaluate models of working memory and their implications for language processing. *Journal of Neurolinguistics, 16*(4), 315–339.

Conway, A. R. A., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: the importance of working memory capacity. *Psychonomic Bulletin & Review, 8*, 331–335.

Cornoldi, C., De Beni, R., & Pazzaglia, F. (1996). Profiles of reading comprehension difficulties: an analysis of single cases. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties: processes and intervention* (pp. 113–136). Mahwah: Lawrence Erlbaum Associates, Inc.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin, 104*, 163–191.

Cowan, N. (1999). An embedded-process model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: mechanisms of active maintenance and executive control* (pp. 62–101). New York: Cambridge University Press.

Cowan, N. (2005). *Working memory capacity*. Hove, East Sussex, England: Psychology Press.

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? In Sossin, W. S., Lacaille, J. C., Castellucci, V. F., & Belleville, S. (Eds.), *Progress in brain research, Vol 169*. Elsevier.

Cowan, N. (2010). The magical mystery four how is working memory capacity limited, and why? *Current Directions in Psychological Science, 19*(1), 51–57.

Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review, 26*, 197–223.

Cowan, N., Rouder, J. N., Blume, C. L., & Saults, J. S. (2012). Models of verbal working memory capacity: what does it take to make them work? *Psychological Review, 119*, 480–499.

Crain, S., Shankweiler, D., Macaruso, P., & Barshalom, E. (1990). Working memory and comprehension of spoken sentences: investigations of children with reading disorders. In G. Vallar & T. Shallice (Eds.), *Neuropsychological impairments of short-term memory* (pp. 477–508). Cambridge: Cambridge University Press.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*(4), 450–466.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: a meta-analysis. *Psychonomic Bulletin & Review, 3*, 422–433.

Daneman, M., & Tardif, T. (1987). Working memory and reading skill reexamined. In M. Coltheart (Ed.), *Attention and performance, XII: the psychology of reading* (pp. 491–508). Hillsdale: Lawrence Erlbaum Associates, Inc.

Das, J. P., & Janzen, C. (2004). Learning math: basic concepts, math difficulties, and suggestions for intervention. *Developmental Disabilities Bulletin, 32*(2), 191–205.

Davelaar, E. J., Goshen-Gottstein, Y., Ashkenazi, A., Haarmann, H. J., & Usher, M. (2005). The demise of short-term memory revisited: empirical and computational investigations of recency effects. *Psychological Review, 112*(1), 3.

Dehn, M. J. (2008). *Working memory and academic learning: assessment an intervention*. Hoboken: Wiley.

Ebbinghaus, H. (1885). *Uber das Gedachtnis*. New York: Dover.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*(1), 19–23.

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: Elsevier.

Engle, R. W., Cantor, J., & Carullo, 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., Kane, M. J., & Tuholski, S. W. (1999a). 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. 102–132). Cambridge University Press.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999b). Working memory, short-term memory and general fluid intelligence: a latent variable approach. *Journal of Experimental Psychology: General, 128*, 309–331.

Gaillard, V., Barrouillet, P., Jarrold, C., & Camos, V. (2011). Developmental differences in working memory: where do they come from? *Journal Of Experimental Child Psychology, 110*, 469–479.

Gathercole, S. E., & Alloway, T. P. (2004). Working memory and classroom learning. *Dyslexia Review, 15*, 4–9.

Gathercole, S. E., & Baddeley, A. D. (1990). The role of phonological memory in vocabulary acquisition: a study of young children learning arbitrary names of toys. *British Journal of Psychology, 81*, 439–454.

Gathercole, S. E., & Baddeley, A. D. (1993). Phonological working memory: A critical building block for reading development and vocabulary acquisition? *European Journal of the Psychology of Education, 8*, 259–272.

Hebb, D. O. (1949). *Organization of behavior*. New York: Wiley.

Hitch, G. J., Towse, J. N., & Hutton, U. (2001). What limits children's working memory span? Theoretical accounts and applications for scholastic development. *Journal of Experimental Psychology: General, 130*(2), 184–198.

Imbo, I., Vandierendonck, A., & De Rammelaere, S. (2007). The role of working memory in the carry operation of mental arithmetic: number and value of the carry. *Quarterly Journal of Experimental Psychology, 60*, 708–731.

Imbo, I., Vandierendonck, A., & Vergauwe, E. (2008). The role of working memory in carrying and borrowing. *Psychological Research, 71*, 467–483.

James, W. (1890). *The principles of psychology* (Vol. 1, pp. 403–404). New York: Henry Holt.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension. *Psychological Review, 99*, 122–149.

Kahana, M. J., Sederberg, P. B., & Howard, M. W. (2008). Putting short-term memory into context: reply to Usher, Davelaar, Haarmann, and Goshen-Gottstein (2008). *Psychological Review, 115*, 1119–1126.

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 and Review, 9*, 637–671.

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., 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.

Kane, M. J., Conway, A. R. A., Hambrick, D. Z., & Engle, R. W. (2007). Variation in working memory capacity as variation in executive attention and control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp. 21–48). New York: Oxford University Press.

Karatekin, C. (2004). A test of the integrity of the components of Baddeley's model of working memory in attention-deficit/hyperactivity disorder (ADHD). *Journal of Child Psychological and Psychiatry, 45*, 912–926.

Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: theories, methods, individual differences, and applications* (pp. 57–71). Mahwah: Erlbaum.

Kellogg, R. T. (1999). Components of working memory in text production. In M. Torrance & G. Jeffery (Eds.), *The cognitive demands of writing* (pp. 143–161). Amsterdam: Amsterdam University Press.

Kellogg, R. T., Olive, T., & Piolat, A. (2007). Verbal, visual, and spatial working memory in written language production. *Acta Psychologica, 124*, 382–297.

Kercood, S., & Grskovic, J. A. (2009). The effects of highlighting on the math computation performance and off-task behavior of students with attention problems. *Education and Treatment of Children, 32*(2), 231–241.

Kyllonen, P. C. (1996). Is working memory capacity Spearman's g? In I. Dennis & P. Tapsfield (Eds.), *Human abilities: their nature and measurement* (pp. 49–76). Mahwah: Erlbaum.

Levy, C. M., & Marek, P. (1999). Testing components of Kellogg's multicomponent model of working memory in writing: the role of the phonological loop. In M. Torrance & G. Jeffery (Eds.), *The cognitive demands of writing* (pp. 25–41). Amsterdam: Amsterdam University Press.

Logie, R. H., Gilhooly, K. J., & Wynn, V. (1994). Counting on working memory in arithmetic problem solving. *Memory & Cognition, 22*(4), 395–410.

Lusk, D. L., Evans, A. D., Jeffrey, T. R., Palmer, K. R., Wikstrom, C. S., & Doolittle, P. E. (2009). Multimedia learning and individual differences: mediating the effects of working memory capacity with segmentation. *British Journal of Educational Technology, 40*(4), 636–651. doi:10.1111/j.1467-8535.2008.00848.x.

Mackey, A., Philip, J., Egi, T., Fujii, A., & Tatsumi, T. (2002). Individual differences in working memory, noticing of interactional feedback and L2 development. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 181–209). Amsterdam: John Benjamins Publishing.

Mayer, R. E. (1989). Systematic thinking fostered by illustrations in scientific text. *Journal of Educational Psychology, 81*, 240–246.

Mayer, R. E. (2009). *Multi-media learning*. New York: Cambridge University Press.

McCutchen, D. (1996). A capacity theory of writing: working memory in composition. *Educational Psychology Review, 8*, 299–325.

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.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Holt, Rinehart & Winston.

Nairne, J. S. (2002). Remembering over the short-term: the case against the standard model. *Annual Review of Psychology, 53*, 53–81.

Nation, K., Adams, J. W., Bowyer-Crane, C. A., & Snowling, M. J. (1999). Working memory impairments in poor comprehenders reflect underlying language impairments. *Journal of Experimental Child Psychology, 73*, 139–158.

Noel, M., Desert, M., Auburn, A., & Seron, X. (2001). Involvement of short-term memory in complex mental calculation. *Memory and Cognition, 29*, 34–42.

Oberauer, K. (2002). Access to information in working memory: exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 411–421.

Oberauer, K., & Lewandowsky, S. (2011). Modeling working memory: a computational implementation of the time-based resource-sharing theory. *Psychonomic Bulletin & Review, 18*, 10–45.

Oberauer, K., Sub, H. M., Schulze, R., Wilhelm, O., & Wittman, W. W. (2000). Working memory capacity—facets of a cognitive ability construct. *Personality and Individual Differences, 29*, 1017–1045.

Olive, T. (2004). Working memory in writing: empirical evidence from the dual-task technique. *European Psychologist, 9*, 32–42.

Paivio, A. (1986). *Mental representations: a dual coding approach*. Oxford: Oxford University Press.

Passolunghi, M. C., & Siegel, L. S. (2001). Short-term memory, working memory, and inhibitory control in children with difficulties in arithmetic problem-solving. *Journal of Experimental Child Psychology, 80*, 44–57.

Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., et al. (2013). No evidence of intelligence improvement after working memory training: a randomized, placebo-controlled study. *Journal of Experimental Psychology: General, 142*(2), 359–379.

Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology: General, 126*, 211–227.

Rourke, B. P. (1993). Arithematic disabilities, specific and otherwise: a neuropsychological perspective. *Journal of Learning Disabilities, 26*(4), 214–226.

Schweppe, J., & Rummer, R. (2014). Attention, working memory, and long-term memory in multimedia learning: an integrated perspective based on process models of working memory. *Educational Psychology Review, 1–22*.

Sederberg, P. B., Howard, M. W., & Kahana, M. J. (2008). A context-based theory of recency and contingunity in free recall. *Psychological Review, 115*, 893–912.

Service, E. (1992). Phonology, working memory, and foreign language learning. *The Quarterly Journal of Experimental Psychology, 45*(1), 21–50.

Smith, P. L., & Ragan, T. J. (2005). *Instructional design* (3rd ed.). Hoboken: Wiley.

Swanson, H. L. (2000). Are working memory deficits in readers with learning disabilities hard to change? *Journal of Learning Disabilities, 33*, 551–566.

Swanson, H. L., & Berninger, V. (1994). Individual differences in children's working memory and writing skill. *Journal of Experimental Child Psychology, 63*, 358–385.

Swanson, H. L., & Siegel, L. (2001). Learning disabilities as a working memory deficit. *Issues in Education: Contributions from Educational Psychology, 7*, 1–48.

Swanson, H. L., Howard, C. B., & Saez, L. (2006). Do different components of working memory underlie different subgroups of reading disabilities? *Journal of Learning Disabilities, 39*, 252–269.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*(4), 295–312.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction, 12*, 185–233.

Tronsky, L. N. (2005). Strategy use, the development of automaticity, and working memory involvement in complex multiplication. *Memory & Cognition, 33*, 927–940.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28*, 127–154.

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. *Psychological Review, 114*(1), 104.

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods, 37*(3), 498–505.

Usher, M., Davelaar, E. J., Haarmann, H. J., & Goshen-Gottstein, Y. (2008). Short-term memory after all: comment on Sederberg, Howard, and Kahana (2008). *Psychological Review, 115*, 1108–1118.

Vanderberg, R., & Swanson, H. L. (2007). Which components of working memory are important in the writing process? *Reading and Writing, 20*(7), 721–752.

Was, C. A., & Woltz, D. J. (2007). Reexamining the relationship between working memory and comprehension: the role of available long-term memory. *Journal of Memory and Language, 56*(1), 86–102.

Waters, G. S., & Caplan, D. (1996). The capacity theory of sentence comprehension: critique of Just and Carpenter (1992). *Psychological Review, 103*(4), 761–772.

Willcutt, E. G., Pennington, B. F., Boada, R., Ogline, J. S., Tunick, R. A., Chhabildas, N. A., & Olson, R. K. (2001). A comparison of the cognitive deficits in reading disability and attention-deficit/hyperactivity disorder. *Journal of Abnormal Psychology, 110*(1), 157.