The Individual Differences in Working Memory Capacity Principle in Multimedia Learning

Jennifer Wiley
University of Illinois at Chicago

Christopher A. Sanchez
Oregon State University

Allison J. Jaeger
University of Illinois at Chicago

Abstract

Effective use of the working memory system is critical for successful learning, and this assumption has motivated much of the work on multimedia instruction. Interestingly, the limited capacity of human working memory has been invoked as part of explanations for both advantages and disadvantages of multimedia learning in comparison with learning from text or pictures alone. This chapter reviews several lines of reasoning that have guided explorations of the role of working memory in multimedia learning, including approaches that have emphasized the modality-specific buffer system and the potential for overloading the limited resources that are available to learners; as well as a newer approach that considers working memory capacity as an individual differences variable representing attentional control.

What Is the Individual Differences in Working Memory Capacity Principle?

Learning from multimedia is a higher-order cognitive process that relies on many subprocesses to be successful. For the purposes of this volume, multimedia comprehension is defined as learning from a combination of words and images. As such, it requires all of the processes involved in developing comprehension from written text or spoken discourse, in addition to the processes needed to interpret and represent information from images, diagrams, graphs, animations, or other forms of visualizations, as well as the processes necessary to alternate between and integrate multiple representations. Since the first studies on learning from multimedia, the
Individual Differences in Working Memory Capacity

The Limited Capacity of Working Memory

Because multimedia learning requires processing from multiple sources, codes, channels, or modalities of information, one of the earliest concerns was its relation to the limited capacity of the human working memory system. In particular, there has been great deal of concern about how having to attend to multiple information sources may put the learner under load. Many studies on multimedia learning have found that when learners are
given more information, including additional information that should be helpful for their understanding, they actually learn less (not more). Many of these studies have been situated in the literature using cognitive load theory as a construct to explain performance. Cognitive load theory (Sweller, 1988) originally emerged from a line of research on mathematical problem solving. In this work, it was found that naive students were able to learn mathematical principles more effectively from worked examples than from less-structured problem-solving attempts. The explanation offered for these results was that the demands of engaging in acts of problem solving, while also learning from those acts, were overwhelming and required too many resources for naive learners to both perform the necessary operations and abstract the principles. In other words, the working memory system was overloaded by attempting to learn from problem solving. This overloading concern was formalized as cognitive load theory and has since been applied to many other educational contexts, including, most prominently, multimedia instruction (see Chapter 2 for a more complete description of this approach).

Given this research, there has been a long tradition of assuming that poor learning outcomes obtained in multimedia contexts are due to learners being placed under load by the materials, meaning that the amount of information they are given, as a function of how information is presented or included in particular multimedia contexts, overwhelms the capacity of their working memory system. In this literature, the amount of load is frequently discussed as a characteristic or a property of the learning materials themselves.

**Working Memory Capacity as an Individual Differences Variable**

An alternative approach examines working memory involvement by considering the effects of individual differences in working memory capacity (WMC). In this approach, WMC is generally considered to be a trait of individuals in relation to their ability to use their working memory system. In other words, the main construct of interest in this approach is the central executive, or the ability to control one’s attentional resources. Complex span tasks have been explicitly designed to assess the functioning of the central executive; that is, they test for the maintenance of information in immediate memory and retrieval from secondary memory in the face of interference from an ongoing processing task (Daneman & Carpenter, 1980; Unsworth & Engle, 2007). As opposed to simple span tasks that do not involve an intervening processing component, complex span tasks involve both a memory storage component and a processing component. For example, in the operation span task (OSpan), learners participate in trials consisting of verifying a math equation and then remembering a word or letter presented after the mathematical verification task. These trials are presented in sets of two to seven items. At the end of each set of trials, a participant is asked to recall the to-be-remembered items. The following is an example set of trials during the OSpan task:
While simple memory span tasks (such as remembering a string of words or digits or sequences of spatial locations) without an additional intervening processing task are considered to assess primary memory, or the capacity of the buffers in the working memory system, the presence of the processing task in complex span tasks (like the equation verification task in OSpan) renders it a measure of the central executive system, or the ability to allocate or control one’s attention. In complex span tasks, participants are required to maintain, update, and retrieve to-be-remembered information from secondary memory, engage in active processing in a secondary task, switch back and forth between different task sets, and avoid proactive interference from previous trials and from intervening stimuli.

Most modern versions of complex span tasks are based on the reading span task (RSpan) originally developed by Daneman and Carpenter (1980). Although in the initial version of the reading span task sentences served as stimuli for both the processing and the memory components of the task (participants were asked to verify sentences and to remember their last words), newer versions have improved upon the design of the task by making the processing and memory elements independent. Newer versions of complex span tasks also use many different processing tasks as the intervening stimuli other than just sentence processing (e.g., verifying math equations, making symmetry judgments, reordering lists; see Conway et al., 2005) so that they might represent a more generalized measure of cognitive ability. Performance on complex span tasks thus provides an individual differences measure that represents the central executive, or the ability to control one’s attention. This estimate of generalized ability is particularly robust when a composite performance measure is derived from performance on multiple complex span tasks.

The individual differences approach to WMC has become more popular over the past 20 years, a trend driven largely by the development of and advances in complex span tasks to tap this construct, as well as an increasing interest in the role of individual differences in executive functioning and attentional control in many cognitive tasks (Engle, 2002; Oberauer, 2009). Much research has connected individual differences in WMC to the processes involved in developing understanding from text, commonly referred to as constructing a situation model, following the work of Kintsch (1998). Kintsch has delineated multiple levels of representation as a message is processed. Initial representations are coded at a surface level, representing the

```
IS (9 / 3) – 2 = 2  NO  aunt
IS (8 / 4) – 1 = 1  YES  bush
IS (6 / 2) + 1 = 4  YES  corn
IS (6 x 3) – 2 = 11  NO  bear
```
exact stimuli or words that are encountered. A secondary level, the textbase, represents a refinement of the actual phrasing into gist-level or macro-structural propositions. Again, this level largely represents information that was presented, with perhaps some summarization, simplification, or generalization. The final level of representation is called the situation model, and it is at this level that connections between the ideas in the text are created, often by also connecting the information with ideas in long-term memory. This level can also be seen as providing a mental model of the phenomena described by the text. This is the level of representation that captures the meaning of the text and involves the generation of inferences or connections across concepts. In other words, this level determines comprehension (Kintsch, 1994; Wiley, Griffin & Thiede, 2005), so its relation to WMC is of particular interest for research on learning from multimedia.

Indeed, individual differences in WMC have been shown to be highly correlated with standardized global reading comprehension measures like the VSAT and the Nelson-Denny reading test (Cain, Oakhill, & Bryant, 2004; Daneman & Carpenter, 1980; Engle, Cantor, & Carullo, 1992; Masson & Miller, 1983). A meta-analysis by Daneman and Merikle (1996) showed that global and local reading comprehension processes were better predicted by complex span measures than by simple span measures. In terms of the actual comprehension processes that may be supported, consistent with the intuition that generating inferences requires keeping multiple ideas active simultaneously so that they might be integrated, higher WMC has been shown to facilitate the generation of both bridging and elaborative inferences (Linderholm & van den Broek, 2002; Singer & Ritchot, 1996; Whitney, Ritchie, & Clark, 1991). WMC has also been related to the ability to derive global themes and main points from text (e.g., Budd, Whitney, & Turley, 1995; Cantor & Engle, 1993; Daneman & Carpenter, 1980; Lee-Sammons & Whitney, 1991).

Because multimedia learning represents a particularly complex cognitive task, it is also likely to depend on effective attentional control in several respects that build on those required for more general text comprehension. During multimedia learning, the learner must (1) attend to and maintain the goal of the learning episode; (2) attend to the available information in all media; (3) select the information relevant to the learning goal from the available information; (4) organize and represent the presented information in memory based on the goal of the learning episode; and (5) maintain the learning goal and representations of the incoming information in primary memory while retrieving necessary information from secondary memory in order to develop an integrated representation of the presented information (Lusk, et al. 2009; Mayer, 2001). Because of the many ways in which multimedia contexts may place demands on attentional control processes, individual differences in WMC should theoretically be an important factor in multimedia learning. Yet the relation of these individual differences to
multimedia learning, especially in terms of mechanisms that are affected and how better executive functioning may benefit learning, are just beginning to be understood. The next section describes several example studies that are consistent with this controlled attention perspective and have begun to document how individual differences in WMC can affect learning from multimedia.

Examples of the Individual Differences in Working Memory Capacity Principle

In one example of research on multimedia learning using an individual differences approach, WMC has been shown to predict learning from illustrated text (Sanchez & Wiley, 2006). In a series of experiments, participants read an expository text that was either not illustrated, illustrated with conceptually relevant images, or illustrated with conceptually irrelevant images, commonly referred to as seductive details (Harp & Mayer, 1997; see also Chapter 12 on the coherence principle). Seductive details are tangentially relevant text or imagery that, while interesting on an emotional level, do not share much conceptual connection with the to-be-learned material. As such, the addition of seductive imagery in a text often decreases understanding of the target material. It was previously speculated that seductive imagery can have negative effects by distracting, disrupting, or diverting attention away from more conceptually important information (Harp & Mayer, 1998). By thus interfering with the learner’s ability to appropriately attend to conceptually relevant concepts and facts, less relevant or even unrelated information would instead supplant conceptual information in the learner’s developing mental model. In other words, one could interpret the previously found seduction effect to be a direct result of the learner attending to information that is not beneficial for achieving the learning goal. Given that WMC is thought to represent the ability to manage executive control of attention, it was of interest to examine whether seduction effects might interact with individual differences in WMC. Specifically, it was hypothesized that high-WMC individuals might be less susceptible to the influence of interesting but ultimately irrelevant information, as high-WMC individuals had been shown previously to be better able to stay on task relative to goals, largely as a result of their heightened ability to suppress or ignore non-goal-relevant information (Kane & Engle, 2002).

To test this account, participants read a text about ice ages. Again, this text was either not illustrated, illustrated with conceptually relevant images, or illustrated with seductive details (see Figure 25.1). Participants were asked to read this text and then apply their knowledge to answer the question “What causes ice ages?” in a causal essay explaining why this phenomenon occurred. In addition, participants completed both OSpan and RSpan as measures of their WMC.
While performance was lower overall in the seductively illustrated condition, this effect was largely localized to lower-WMC individuals. Conversely, higher-WMC individuals were almost unaffected by the presence of this irrelevant imagery, and comprehension was equally good for these individuals regardless of what type of illustrations were included in the textual document. Thus, consistent with the general notion that seductive imagery interacts with attentional control, higher-WMC individuals seemed able to focus on relevant information, whereas lower-WMC individuals succumbed to the emotional appeal of the images. These learning results were further supported and expanded on by a second experiment that examined the eye movements of high- and low-WMC individuals as they read the seductively illustrated text. Again, consistent with the original intuitions of Harp and Mayer (1998), it was found that seductive details did in fact disrupt normal reading of the text and caused learners to attend to this irrelevant information in inappropriate ways and at inappropriate times. For example, seduced individuals often fixated on this irrelevant imagery midsentence, in essence stopping their efforts to comprehend the text midstream in order to attend to imagery that was conceptually irrelevant. It is perhaps not surprising, then, that these same individuals struggled to develop a coherent mental model of the text. It is important to note that this pattern of seduction was demonstrated only by lower-WMC individuals. Higher-WMC individuals effectively ignored the seductive imagery and instead focused their efforts on processing
the text. More recent work has extended these findings, showing that the presence of seductive images can also have a negative impact on metacomprehension accuracy (Jaeger & Wiley, 2010). Seductive but irrelevant information seems to make readers less able to gauge their comprehension when learning from expository science texts. While research on seduction effects demonstrates that WMC can aid in the development of understanding when irrelevant content is included in the learning environment, additional research has also demonstrated that WMC predicts multimedia comprehension even when no seductive information is present. In research by Sanchez and Wiley (2009), the way that illustrated text was presented, either in scrolling or paginated format, had significant interactions with WMC. In this set of experiments, participants who varied in WMC read a complex illustrated text either about ice ages (Experiment 1) or about the Irish potato famine (Experiment 2). Learning in both experiments was assessed by short-answer responses to the respective question “What causes Ice Ages?” or “What caused the Irish potato famine?” respectively. The critical manipulation in these studies was whether this information was presented in one long scrolling Web page or was instead segmented into smaller conceptual units that were presented on discrete pages on a Web site (see Figure 25.2). Note that the informational content (text and pictures) was consistent across presentations and that the only difference was whether the text was presented in multiple smaller pieces or as a whole.

Several intuitions can be invoked to predict that scrolling presentation should be less optimal. First, because all information is available at once, it may be distracting, particularly for readers with poor attentional control. Second, because the text appears as one long unit, readers might attempt to construct and maintain a surface or textbase representation of the entire text in memory as they read. In contrast, the explicit page boundaries in the paginated condition may help readers to represent and store more manageable units of text in memory. In order to maintain a surface-level representation of the entire text in the scrolling condition, participants may have to divert resources that would otherwise be available for more conceptual processing and the construction of a situation model representing what the text is about. Thus, fewer resources may be available for comprehension processes, which should in turn harm learning. Thus, persons with better attentional control, or higher WMC, may be better equipped to learn from scrolling interfaces, whereas segmented presentation may be particularly helpful for students with low WMC.

The results from both experiments corroborated these intuitions. While scrolling presentations reduced learning overall, this effect was again localized to individuals lower in WMC. In contrast, higher-WMC individuals learned equivalently across scrolling or paginated presentations, suggesting a resiliency to competing demands in the scrolling condition. The results of a third example study also support the observation that higher-WMC
individuals are better able to develop a coherent understanding when knowledge is distributed across multiple sources and learners must thus navigate a complex environment such as a Web site to gather and integrate this knowledge. In a study by Banas and Sanchez (2012), participants who varied in WMC, but who were equivalent in prior knowledge about the topic of the plant kingdom, read a wiki-like Web site where information about each unit within the plant hierarchy was presented on separate pages. Note that this format is consistent with most current online wiki repositories of knowledge.

Participants were asked to learn about the plant kingdom and were able to freely navigate the site in order to locate information necessary to answer
basic factual questions about the topic. However, at the end of this information search activity, participants were also evaluated as to how well they understood how the different units within the plant kingdom fit together within the overall conceptual hierarchy. It is important to note that the participants were never instructed to attend to this hierarchical information as they attempted to locate the information needed to answer the basic factual questions they were given. As such, an understanding of the hierarchical organization of the plant kingdom was not an explicit learning goal. The results indicated that while all participants were able to successfully find the answers to the search questions, only higher-WMC individuals constructed a more global understanding of the site’s structure (e.g., plant family taxonomy). Given that performance was equivalent on the search task, the divergence in learning the hierarchical relationships (which determined where to find the information) was attributed to higher-WMC students being more sensitive to potential connections and relations among concepts. Higher-WMC individuals had more capacity available to incorporate and situate the information they were learning within a broader conceptual framework.

In this section, we have provided several specific examples of studies that have established a direct relationship between individual differences in WMC and learning from multimedia presentations. A general principle emerging from these studies is that WMC is most critical when some of the information that is presented is irrelevant to the learning (i.e., seduction) or when the presentation context makes the learning process more difficult to manage (i.e., scrolling). In both cases, learners who are more susceptible to distraction may be harmed. WMC also is important for maintaining access to multiple representations so that information can be integrated and can allow for more global understanding of the information via construction of the situation model (generating inferences) or other higher-order representations, such as understanding the structure of the information (noticing connections or relations within a macro-structure). This aspect of WMC allows a reader to both process incoming information and do something with it or beyond it. On the basis of these studies it appears that the ability to control one’s attention becomes important particularly in the face of a
need to manage attentional demands at both a perceptual and a conceptual level. WMC thus allows a reader to select and focus on specific information, as well as to integrate and develop global understanding, when tasked with learning from multimedia presentations.

What Do We Know about How Individual Differences in Working Memory Capacity Relate to Multimedia Learning?

Support for these roles for individual differences in WMC in multimedia learning can be found in several other studies that have used measures of WMC as part of their design. The importance of the ability to focus one’s attention on relevant information while learning from multimedia resonates with much recent work that has explored the benefits of segmenting multimedia presentations, or breaking presentations down into smaller units. Intuitively, it has been argued that learning should be improved by segmentation of a multimedia unit because it facilitates chunking and structures the instruction into meaningful parts. When text is not segmented into smaller content units, learners may be cognitively overloaded during multimedia instruction and their working memory capacity for maintaining information may be exceeded. Another related aspect of multimedia instruction is that it sometimes utilizes pre-set timing. Putting the timing of the presentation of each segment under learners’ control may also allow them to more effectively engage in cognitive processing. When the time allotted to each segment is extended, learners may not be as cognitively overloaded and their working memory capacity may be less likely to be exceeded (cf. Kurby & Zacks, 2008; Mayer, 2009; Schnotz & Lowe, 2008; see Chapter 13 on the segmenting principle).

Mayer, Dow, and Mayer (2003) investigated the possible benefits of segmentation using a multimedia unit on the working of an electric motor in segmented and nonsegmented versions. Students who were given the segmented presentation developed better conceptual understanding of the unit than students who experienced the nonsegmented version. Similarly, Mayer and Chandler (2001) also explored the possible benefits of segmentation with a multimedia presentation on the cause of lightning. In this study, students saw both segmented and nonsegmented versions, sequentially, with order counterbalanced across conditions. Mayer and Chandler found that the students who received the segmented presentation first developed a better understanding of lightning. Mayer and Chandler attributed the superior comprehension of the segmented-presentation-first group to the fact that these participants avoided cognitive overload and were able to build models of the component parts of the cause-and-effect relationships responsible for lightning formation during this first engagement. During the second
engagement, the participants were then able to construct a more global model. In a more recent study in which WMC measures were collected and that also entailed a segmentation manipulation, Lusk et al. (2009) examined learning from a multimedia historical inquiry unit. They found that segmentation of the unit specifically improved learning for low-WMC students. Taken together, these results provide evidence for the Sanchez and Wiley (2009) finding that segmentation (presenting multimedia in a paginated condition as opposed to a scrolling condition) improved understanding, particularly for low-WMC learners.

However, it is also interesting to note that simplifying presentations may have negative consequences for some learners. For high-WMC learners who can engage in multiple processes simultaneously, receiving the intact presentation and actively parsing a multimedia unit on their own, rather than receiving a segmented unit, may prove to be even more beneficial for these learners, as opposed to learners with lower WMC (Spanjers, Wouters, van Gog, & van Merriënboer, 2011). Similarly, directing learners’ attention during multimedia learning with a “spotlight” may be helpful for low-WMC but could harm high-WMC learners, who may be better off directing their own attention (Skuballa, Schwonke, & Renkl, 2012).

There is also support for the importance of WMC in helping multimedia learners to construct higher-order representations. Several researchers have found that superior attentional control aids in connecting and integrating information (Austin, 2009; Doolittle et al., 2009; Dutke & Rinck, 2006; Pazzaglia et al., 2008). In a study that was not on illustrated text comprehension but was technically a multimedia learning study because it employed words and images as stimuli, Dutke and Rinck (2006) used the RSpan and the spatial span task to investigate how individual differences in WMC affect learning of spatial arrangements of items (words or icons). Participants were exposed to pairs of stimuli. The dependent variable was the time to verify whether novel pairs or configurations depicted correct spatial relations. They found that WMC helped participants to integrate the learned pairs of objects into a complete arrangement, or alternatively WMC helped learners to create a global representation of all of the presented information into a single array.

In a study more directly related to developing an understanding from a multimedia unit, Pazzaglia et al. (2008) had Italian middle school students learn about Germany’s geography from a hypermedia system. After learning, the authors assessed the students’ awareness of the hypermedia structure as well as whether information was correctly integrated. Performance on complex span tasks predicted both of these learning outcomes, while simple span tasks did not. Similarly, in a study by Austin (2009, Exp. 2), participants learned about lightning in one of three groups: animation and written text, animation and spoken text, or animation, spoken, and written text. Regardless of condition, OSpan predicted understanding about lightning. In
a study by Doolittle et al. (2009), participants learned about pumps in one of three groups: animation and written text, animation and spoken text, or animation, spoken, and written text. Using an extreme groups design, OSpan predicted performance on both a recall test and a test of understanding. As in the previous study, no interactions were found between Ospan and the three multimedia conditions, indicating that WMC is important generally for multimedia comprehension and the construction of an integrated situation model or mental model of the content (for similar results see Doolittle & Mariano, 2008; Geiger & Litwiller, 2005; Lusk et al., 2009; Sanchez & Wiley, 2014).

Although only a handful of studies have directly explored interactions between learning from multimedia under different conditions and WMC measures representing domain-free attentional control, the results found in the literature are consistent with the conclusions that WMC allows a student to effectively select and focus on information, as well as to integrate and develop global understanding, when tasked with learning from multimedia presentations.

**What Are the Implications for Cognitive Theory?**

This review has shown that there is evidence for two interrelated roles for individual differences in WMC when readers are tasked with learning from multimedia units. Both of these advantages of WMC are consistent with an attentional-control perspective of working memory capacity. In particular, individuals with greater WMC are less susceptible to distraction, which has been attributed to a heightened ability to direct and focus, or otherwise control, their attention. As a consequence of this heightened ability to control their attention and to focus only on relevant stimuli, which may lead to more efficient processing, individuals with greater WMC also are able to direct their remaining resources toward more integrative or global comprehension processes (Chiappe, Hasher, & Siegel, 2000; DeBeni & Palladino, 2000).

**What Are the Implications for Instructional Design?**

The obvious implication for instructional design is that designers need to be aware that individuals with low WMC may need support in both focusing their attention while learning from multimedia displays and in integrating information or constructing higher-order representations. As opposed to high-WMC learners, low-WMC learners may not engage in higher-order conceptual processes like generating inferences or abstracting a more global relational structure spontaneously (Banas & Sanchez, 2012;
Linderholm & van den Broek, 2002; Wiley & Myers, 2003). They may need explicit instructions to clarify that their goal when learning from a multimedia presentation is to attempt to construct a conceptual understanding of the material (Wiley, Ash, Sanchez & Jaeger, 2011; Wiley, Griffin, & Thiede, 2005; Wiley & Sanchez, 2010). Simplifying the perceptual demands during a first exposure also seems like a promising approach, as seen with the success of segmented presentations for low-WMC learners (Lusk et al., 2009; Sanchez & Wiley, 2009). In general, rereading manipulations, where readers are exposed to the same presentation twice, have also been shown to be particularly useful for low-WMC readers (Griffin, Wiley, & Thiede, 2008).

What Are the Limitations of Current Research?

One of the largest limitations on progress in this area is that there are still relatively few studies that measure individual differences in WMC and use them to predict or explain learning outcomes. State-of-the-art practices in WMC methodology suggest using a composite score derived from performance on a set of complex span tasks and using that score to predict learning outcomes by means of regressions. Since so few studies in the literature have taken this approach, even just replicating older work on illustrated text comprehension with updated methodology would be useful. Another impediment to progress is that although there are many more studies that have attempted to manipulate load and the demands on the working memory system by altering the presentation of materials, these studies rarely include measures of WMC. Although much work from the cognitive load perspective paints working memory demands as a function of the stimuli, and work from the controlled attention perspective by definition paints working memory demands as a function of the individual, the reality is, of course, likely to be found using a combination of both approaches and may be best explored in studies pursuing ATIs (aptitude-by-treatment interactions). Many interesting manipulations have been done with multimedia materials that have resulted in better learning outcomes, but they largely assume that improvements in learning are the result of reducing load on the learner. Collecting direct measures of load are one important step toward explaining any benefits obtained from these manipulations. Another approach would be demonstrating that manipulations that are intended to reduce load are specifically beneficial for low-WMC learners.

What Are Some Directions for Future Research?

Aside from studies that include both manipulations of materials and measures of WMC, another frontier to push forward on is an exploration
of modality effects and how they might be related to a more general WMC construct or to other more specific abilities. Although there is a tradition of research that has been based on the assumption that we have distinct buffers for encoding the different kinds of stimuli that are part of a multimedia presentation (i.e., visuospatial and verbal codes) and that we have distinct working memory capacities based in those codes, a provocative question is whether there is a need to posit the existence of more than one mental workspace. An alternative tack is to assume that there is only one actual workspace but that other specific abilities (i.e., verbal and spatial abilities) might govern our ability to encode and pass information through that workspace, manipulate information in it, and retrieve relevant information from secondary memory. Instead of assigning information to a buffer based on whether it is graphic or verbal in its presentation format, we may find it more important to determine the kind of representation that we are trying to create from it. A generalized WMC construct, as well as individual differences in the ability to deal with the specific kind of representation that needs to be constructed, may then jointly determine the quality of processing in the mental workspace. This single workspace approach would be consistent with previous suggestions that spatial thinking may facilitate learning from multimedia (i.e., the ability-as-compensator hypothesis; Höfller & Leutner, 2011; Mayer & Sims, 1994; Moreno & Mayer, 2000). Alternatively, the act of translating between visual and verbal codes may be especially difficult, especially when the reader has difficulty with spatial thinking, which may be why learning spatial information from verbal text has been seen to depend on working memory capacity in some studies (Brunyé & Taylor, 2008; Kruley, Sciama & Glenberg, 1994).

More broadly, this line of reasoning suggests that simultaneously exploring the role of spatial thinking and WMC in learning from multimedia is another critical area for better understanding when students might learn best from multimedia and under what conditions. A recent study by Sanchez and Wiley (2014) provides a final example of learning from multimedia presentations that precisely targets this question. In this study, participants were asked to learn about the topic of plate tectonics from a multi-page Web site and write an essay about why Mt. St. Helens erupted. Pages of this web site were illustrated with conceptually relevant static images, were illustrated with animations, or were not illustrated at all. Participants were assessed on their WMC by means of both the RSpan and OSpan tasks, but they were also evaluated specifically on spatial visualization and dynamic spatial skills. Spatial visualization was measured using the paper folding task (VZ-2; French, Ekstrom, & Price, 1963) and involves the skills needed to mentally manipulate spatial information into cohesive new wholes. Dynamic spatial ability (DSA) involves the skills needed to integrate spatial information with temporal information to make discrete predictions about a future visuospatial state. DSA was measured using a modified version of the intercept task.
Individual Differences in Working Memory Capacity

(Law et al., 1993), in which participants predict when two moving objects that vary in speed will collide.

Because the topic of plate tectonics is itself highly visuospatial and dynamic (the tectonic system involves changes over time and space), the addition of imagery in a format that most closely matched the conceptual material (e.g., animations) was predicted to produce the best learning. Further, it was predicted that visuospatial skills would interact with the type and nature of imagery in a compensatory manner, such that more dynamic visualizations (e.g., animations) could potentially compensate for lower dynamic visuospatial skills, in essence eliminating the disparity found previously when lower-visuospatial learners interact with highly spatial content (Boucheix & Schneider, 2009; Hegarty & Sims, 1994; Hegarty & Steinhoff, 1997; Sanchez & Branaghan, 2009). Indeed, the animation condition did produce the best learning overall, for both low- and high-DSA individuals. In contrast, low-DSA individuals struggled to form an understanding of the topic in the nonanimation conditions. Without the benefit of animations, the other conditions forced learners to construct their own internal dynamic models. As lower-DSA individuals are less able to accomplish this, their learning suffered. This suggests that different visualization types can interact with individual differences in spatial skills to affect how well learn-ers may construct an understanding of the material. The benefit that can be realized from matching learners to visualization conditions is again a strong argument for undertaking future work that includes both manipulations of multimedia presentations and measures of individual differences in abilities and skills that may be relevant to the task.

Finally, an important question that is directly related to the focus of this chapter is whether independent effects might be seen for both visuospatial thinking and WMC on learning from multimedia. It may be important to recognize that WMC was significantly and positively correlated with both the spatial visualization and dynamic spatial ability measures obtained in the Sanchez and Wiley (2014) sample. This pattern has been observed before and suggests that these different measures may share overlapping variance. That means that it will be hard to draw conclusions about which particular skill or ability may be driving learning from any previous work that uses only one of these measures. However, because all three constructs were assessed in the present study, the unique effects of each could be analyzed. These analyses showed that WMC remained a significant overall predictor of comprehension even when variance due to spatial skills was accounted for. Further, there were no interactions between WMC and the multimedia conditions in this study. Individual differences in WMC were related to learning about plate tectonics from this series of Web pages regardless of whether the pages contained animations or illustrations. This result is consistent with several other studies that have found a relationship between WMC and learning across several different multimedia presentation conditions (Austin, 2009;
Doolittle & Mariano, 2008; Doolittle et al., 2009; Geiger & Litwiller, 2005; Lusk et al., 2009). This underlying relationship of WMC to multimedia learning suggests a broad influence of attentional control on comprehension processes, while the Sanchez and Wiley (2014) study shows that it is independent and significant beyond that of other relevant cognitive skills like spatial thinking.

**Training in WMC and Spatial Thinking**

Due to the fact that WMC is of central importance for a wide range of cognitive tasks as well as academic achievement, it is not surprising that attempts to improve WMC through training have a long history. From this perspective it is assumed that if WMC can be increased, then a host of other related skills and abilities will also benefit. However, the results of WMC training are mixed and have yet to systematically demonstrate that training can improve WMC and, more important, that improvements on transfer tasks can be attributed to increases in WMC. In a review of the WMC training literature, Shipstead, Redick, and Engle (2012) point out that, of the studies that do show increases in performance on WMC tasks after training, many obtain the effect using the task that participants were originally trained on. In these cases, improved performance may not represent training of WMC, but rather practice with a specific task. Shipstead et al. (2012) also point out that many of these training studies are conducted without appropriate control groups, making it impossible to rule out test-retest effects. In the best cases, no-contact control groups that participate in pre- and post-testing only are used. While these groups can be useful for addressing test-retest effects, they introduce other experimental confounds such as the Hawthorne effect, which suggests that people’s performance may be influenced by their level of involvement in a study. One of the few exceptions is the work of Chein and Morrison (2010), which has shown benefits of WMC training on reading comprehension, but results such as this have been rare.

Although there is only sparse evidence supporting the efficacy of WMC training, evidence from the spatial training literature is more encouraging. A recent meta-analysis of spatial training studies indicated that spatial skills do seem to be highly malleable, and that training effects on spatial thinking can be durable and transferable (Uttal et al., 2013). Further there has been some research indicating that training students in spatial visualization skills can also improve multimedia learning outcomes (e.g., Sanchez, 2012). However, there is much work to be done to understand what changes with training.

Because performance on spatial visualization tasks and WMC assessments generally share a moderate amount of overlap ($r = .40$), one possibility is that the training is changing the way students control their attention. Research looking at the effects of videogame training on spatial skills has
suggested enhancements in visual selective attention (Green & Bavelier, 2003, 2006, 2007; Spence & Feng, 2010), so it is possible that spatial training could be having its effects via relatively general changes in executive functioning or WMC. Another possibility is that spatial training may more specifically improve spatial information processing by helping students to develop strategies for understanding or representing spatial information. This notion of cognitive training offers an interesting avenue for future research to determine how WMC and visuospatial skills may interact during learning from multimedia presentations at a theoretical level, as well as how comprehension of multimedia information can be increased and appropriately supported in practice.

**Concluding Remarks**

This review has highlighted two main roles for individual differences in WMC on multimedia comprehension. Research supports the theory that individual differences in WMC allow readers to more effectively select and focus on information, as well as to integrate and develop global understanding, when tasked with learning from multimedia presentations. These results are interpreted from an attentional-control perspective, and opportunities to develop connections to studies motivated by cognitive load perspectives, modality-specific processing approaches, and the role of visuospatial skills in multimedia learning are discussed as important future directions that will help the field to develop a better understanding of when multimedia instruction may be most effective.

**Glossary**

*Attentional control*: An individual’s ability to selectively allocate attentional resources to some information while actively avoiding or ignoring other information.

*Seductive details*: Information that is interesting or enjoyable but only tangentially related to the main topic of a text.

*Complex Span tasks*: Tasks used to measure the capacity of an individual’s working memory. Complex span tasks differ from simple span tasks in that they require both the storage and processing of concurrent information.

*Visuospatial thinking*: Cognitive processes related to the understanding and conceptualization of both dynamic and static visual information and spatial relationships.

*Working memory capacity*: The ability to use the working memory system effectively and efficiently. This capacity varies across individuals.
Working memory system: A limited-capacity cognitive system that manages the retrieval and storage of information required to carry out complex cognitive tasks such as learning, reasoning, and comprehension.

References


