Coherence or interest: Which is most important in online multimedia learning?

Derek A. Muller, Kester J. Lee and Manjula D. Sharma
The University of Sydney

The coherence principle states that all non-essential information in multimedia messages should be eliminated to minimise demands on cognitive resources. This assertion has been empirically verified in controlled laboratory studies with learners who have little prior knowledge and limited interest in the domain of instruction. It has not been investigated, however, whether the coherence principle generalises to real learning environments. In this study, 104 students from year 10, year 11, and first year university viewed either a concise or an extended online multimedia treatment on stellar spectra. The extended treatment included additional interesting information about the formation of black holes, galaxy collisions and the observation of dark matter. Following the multimedia, participants completed a retention and transfer test that covered only the material common to both treatments. Results showed students in both treatment groups achieved similar performance. This suggests that in authentic learning settings, interest may mitigate the effects of the coherence principle. Difficulties involved in measuring differences in learning within the constraints of a real learning environment are also addressed.

Introduction

There is a long-standing debate surrounding what should and should not be included in instruction. Some argue that interesting material is essential for learning, even if only of tangential relevance. Others contend that instruction should consist solely of information related to the learning outcomes.

Recent multimedia research has made a significant contribution to this debate in the form of the coherence principle:

> Students learn better when extraneous material is excluded rather than included. Extraneous material competes for cognitive resources in working memory and can divert attention from the important material, can disrupt the process of organising the material, and can prime the learner to organise the material around an inappropriate theme. (Mayer, 2001, p.113)

The opposing viewpoint claims interest plays a key role in allocating limited cognitive resources, making it essential for learning (Schank, 1979). For example, Mitchell (1993) proposed that a learner's interest could be caught and held during instruction to improve retention. This purpose could be well served by including highly interesting but unimportant information, often called 'seductive details', in instruction (Schraw & Lehman, 2001).

This paper reports a preliminary empirical investigation of these two viewpoints in an authentic online learning setting with students with three levels of prior knowledge.

Previous research

In multimedia research to date, the coherence principle has received the most empirical support. Multimedia on the formation of lightning storms produced better performance on transfer tests when seductive details were excluded rather than included (Mayer, Heiser & Lonn, 2001). In this case the irrelevant material involved a story about a football player struck by lightning on the playing field. However, even relevant equations have been shown to have a detrimental effect when the goal of instruction is a strong qualitative understanding of scientific processes (Mayer & Jackson, 2005).

Empirical support for multimedia principles in laboratory settings, however, does not guarantee applicability to real learning environments. Tabbers, Martens and van Merrienböer (2004) reported a study in which they investigated the modality and cueing effects in an authentic learning setting. The modality effect states that verbal information should be presented as narration rather than on screen text, to facilitate learning (Sweller et al., 1998). The cueing effect suggests that learning is enhanced when visual cues in an animation help link images to their associated narration (Kalyuga, 1999). Although these effects had been well established in previous studies, Tabbers et al. failed to replicate the results in a classroom environment. Their study found only a slight cueing effect and even a reverse modality effect. This suggests that the translation of multimedia principles into effective practice is not trivial.

One might expect the coherence principle to be particularly difficult to observe in authentic settings. In a laboratory environment, researchers oversee the experimental procedure, ensuring learners concentrate on the instructional message. In contrast, while viewing online multimedia from home, learners are under no obligation to pay attention to the instruction. Inspiring interest in the multimedia would therefore be more important in real learning settings than in controlled laboratories.

Research on the coherence principle in multimedia environments grew out of investigations of the 'seductive details effect' in text based learning (e.g. Harp & Mayer, 1998). Early studies found that the inclusion of seductive details in texts resulted in lower performance on retention tests (for a review, see Schraw & Lehman, 2001). However as Goetz & Sadoski (1995) argued, the findings were not as conclusive as was claimed. Methodological confounds and conflicting results signalled the need for further research to establish the effect of seductive details on learning.

Prior knowledge

A rationale for the coherence principle is provided by cognitive load theory (Sweller, van Merrienböer & Pass, 1998). This theory considers the implications for instructional design of humans' limited capacity to process novel information in working memory (Miller, 1956). Three types of cognitive load are proposed: intrinsic, germane, and extraneous. It is difficult to alter intrinsic load because it is determined by the complexity of the material to be learned. Germane cognitive load refers to the mental effort required for schema formation, which is therefore essential for learning. In contrast, extraneous cognitive load is the additional mental effort required to deal with poorly designed instruction. From this perspective, extra interesting but irrelevant information constitutes extraneous cognitive load, which therefore should be eliminated from instruction.

To test this hypothesis, it was important to have students with a range of prior knowledge. Students with too much prior knowledge might not experience cognitive overload regardless of the amount of extraneous information. In contrast, those with too little prior knowledge might be overwhelmed simply by intrinsic cognitive load.

Objective

This study adds to the body of knowledge on multimedia learning in two ways. First, it extends previous research on the coherence principle by investigating the effect of seductive details in an authentic online learning environment. Second, the effect is explored using the same instructional materials with students with three different levels of prior knowledge.

Method

Participants

The participants were 26 year ten and 25 year eleven high school students, and 70 first year university astronomy students. High school students participated voluntarily and anonymously from three high schools and one tutoring college in Sydney. Their teachers felt the multimedia provided a worthwhile introduction to some of the fundamental astronomy curriculum topics. First year university astronomy students were encouraged to participate with a small mark incentive for visiting the study website. Students were not required, however, to complete the study in order to receive the mark.

Due to the voluntary nature of the study, some students at each level did not attempt the short answer questions on the post-test. They were therefore removed from the sample before the data was analysed. The remaining 104 participants consisted of 22 year 10s, 18 year 11s, and 64 first years. Of these, 55 received the concise treatment and 49 watched the extended multimedia.

Procedure

Students were asked to participate in this study voluntarily for homework. In their own time, they accessed the study website. An introduction informed participants that they could withdraw from the study at any time with no penalty. After signing in using an anonymous login, students were randomly assigned to view one of two multimedia treatments. The link to move on to the next part of the study was hidden until the completion of the multimedia. Participants then completed the post-test. Their answers and the times at which they were submitted were written to a *MySQL* database. This allowed for a determination of the time spent watching the multimedia and the time taken answering the post-test questions.

Allowing students to participate in this experiment from their home computers whenever it was convenient was a key feature of the methodology. It ensured that the results of the study are applicable to real learning environments. However it is important to note the substantial 'noise' this introduced into the system. Students may have consulted with their peers or other resources during the post-test. Or, they may not even have watched their prescribed multimedia treatment. It is possible that, as a result, some effects might be obscured by unforseen student behaviours, but this is a reality for any study that attempts to investigate learning in authentic settings.

There is a growing movement in education towards researching in both naturalistic and controlled laboratory settings (e.g. Brown, 1992; Klahr & Li, 2005). In what is often called 'design experimentation', researchers iteratively shift their work between the laboratory and the classroom. The goal is to ensure effects observed in the laboratory can be translated into improved classroom practice. In the present study, it was essential to uncover whether multimedia principles like the coherence effect make a discernable difference in online learning or if they are only observable in a controlled laboratory setting. Obviously if the latter, multimedia principles carry much less weight for the design of instruction.

Multimedia development

Two linear multimedia treatments were developed to explain the physics behind stellar spectra. One gave a concise summary of the physics involved in gathering and analysing electromagnetic radiation from stars, lasting about 7 minutes and 30 seconds. Table 1 shows a sample syllabus dot point, post-test question and corresponding script excerpt.

Table 1: Sample materials related to one syllabus dot point

Syllabus dot point	Post-test question	Script
Describe the technology needed to measure astronomical spectra	A spectrograph performs a similar function to a (n): A. Mirror B. Lens C. Prism D. Aperture E. Shutter	The Sun is the star we're most familiar with, but we can split light from other stars in a very similar way to find out more about them. The technology that we use to gather information about stars is just like the prism we were using. Things are a bit more complicated for other stars because they're far away – first, we need a telescope to focus the image, and then we use a spectrograph. A spectrograph splits up the light by its wavelength yielding a graph of light intensity versus wavelength.

An extended multimedia included excerpts from an interview with a professional astronomer interspersed throughout the concise treatment. The astronomer discussed exciting topics in the field like the formation of black holes, galaxy collisions, the observation of dark matter and new multi-billion dollar telescopes. These topics were irrelevant to the learning outcomes in the sense that none of the additional information presented could be used to answer the post-test questions. A sample interview segment is given below:

> One of the most violent events in the universe is a supernova - when a star collapses. These stars are very massive, much more massive than the sun. What happens is they crunch down into the centre and they squeeze that central mass, the core of the star, so much that it too collapses. And, it reaches a point where it cannot stop itself collapsing until it collapses into a point. This is now a very, very high mass object but all in a very small volume and what you've created is a black hole. What you have is a region where the gravitational pull is so strong that not even light can escape. So you have this object that is basically a little bit like a vacuum cleaner - if something gets a bit too close, it falls in and it can never get back out again… As material falls in, it crashes and bashes into other material and as it does that friction heats it all up and that can release huge amounts of radiation.

In total, 3 minutes and 15 seconds of extra material was added to make the extended multimedia.

Both multimedia treatments included live action demonstrations, animations, graphs, and 'talking head' narration as shown in Figure 1. As far as possible, instructional design adhered to research based principles (Mayer, 2005). Specifically, verbal information was presented as narration rather than on screen text, verbal and visual information were presented together to form a coherent message, and extraneous sounds and images were avoided. One obvious contravention of multimedia design principles was the inclusion of interesting but irrelevant information in the extended treatment. This was, of course, the intended variable for investigation.

a. live action demonstrations

b. animations

c. graphs
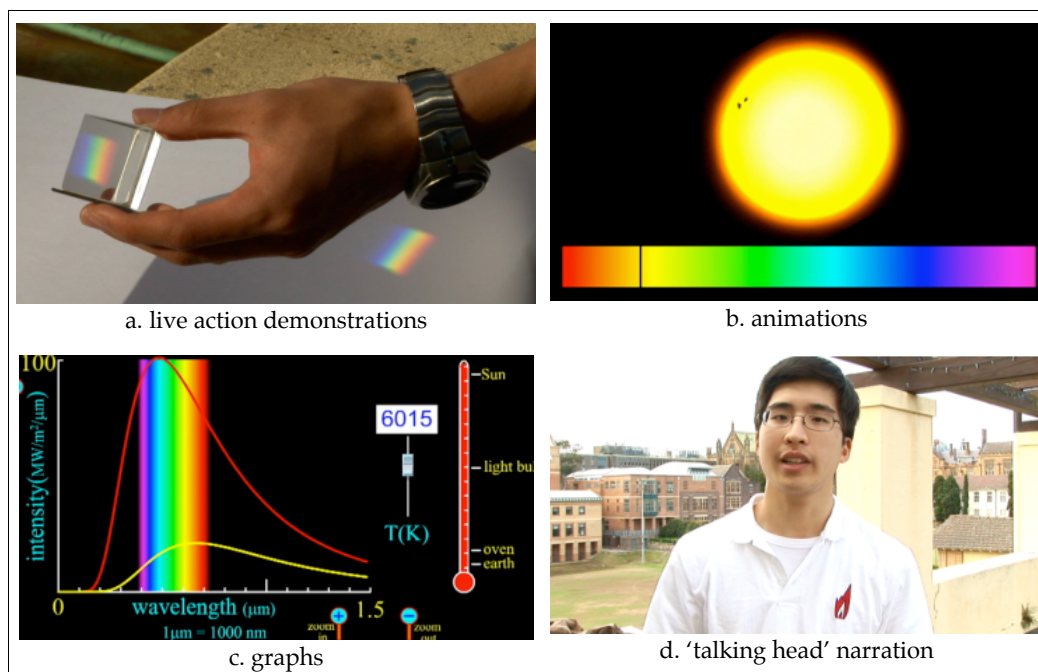
d. 'talking head' narration

Figure 1: Both multimedia included a. live action demonstrations, b. animations, c. graphs, and d. 'talking head' narration.

The learning outcomes for which the multimedia was designed were based on content covered in senior high school physics in New South Wales, Australia (Board of Studies, 2001; Binnie, 2004). Astronomy is one of the compulsory modules under the heading of 'The Cosmic Engine' in year 11. 'Astrophysics' is an optional module in year 12. Thirty hours of teaching time is recommended for each module in its entirety.

Three syllabus dot points from the year 12 Astrophysics module were selected as the target learning outcomes for the multimedia:

• Describe the technology needed to measure astronomical spectra,
• Describe the key features of stellar spectra and describe how these are used to classify stars,
• Describe how spectra can provide information on surface temperature, rotational and translational velocity, density and chemical composition of stars.

Post-test

The post-test consisted of thirteen multiple choice and three short answer questions. Both types involved a mixture of retention and transfer questions. An example of a retention question is: A spectrograph performs a similar function to a(n): a) mirror b) lens **c) prism** d) aperture e) shutter. An example of a transfer question is: A tungsten filament in a light bulb has a temperature of about 3000 degrees Celsius. Evaluate the efficiency of light bulbs as light sources. A complete answer to this question involved a discussion of the black body radiation curve and its implications for the amount of infrared radiation given off at the temperature of the filament. Additional marks were awarded for stating the correct conclusion about the efficiency of light bulbs and suggesting how the devices could be made more efficient. One mark was awarded for each correct answer on the multiple choice questions. The short answer questions were each worth multiple marks based on a set marking scheme. The maximum possible score on the short answer questions was thirteen, bringing the maximum total score to 26.

The post-test was designed to thoroughly evaluate the syllabus dot points outlined above. A panel of physics educators including the first year astronomy lecturer reviewed all questions to check their validity.

Results and analysis

Answers to the short answer questions were marked independently by two of the researchers. These scores were then compared and discrepancies were resolved through discussion and clarification of the marking scheme.
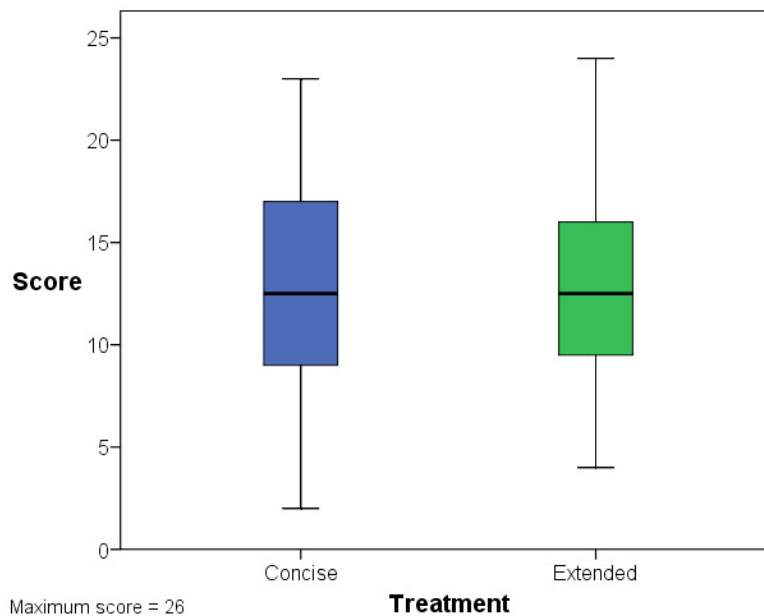

Figure 2: Post-test scores were not significantly different for the concise and extended treatments. Boxes extend from the first to third quartiles, enclosing the median. Whiskers show the extremes.

The distributions of total post-test scores for the concise and extended groups were each normally distributed over the entire sample population. A t-test revealed no significant difference between the two groups. In fact, performance on the post-test was almost identical for students in the concise and extended groups as shown in Figure 2.

Prior knowledge was expected to play a significant role, however, in the applicability of the coherence principle. As mentioned previously, extraneous cognitive load may only hinder learning for novice learners whose cognitive resources are already burdened with novel information. Therefore, further analysis was carried out within each of the three prior knowledge levels to compare the performance of students in the concise and extended groups. Again, no significant difference was found in the performance on the post-tests. Results are illustrated in Figure 3.
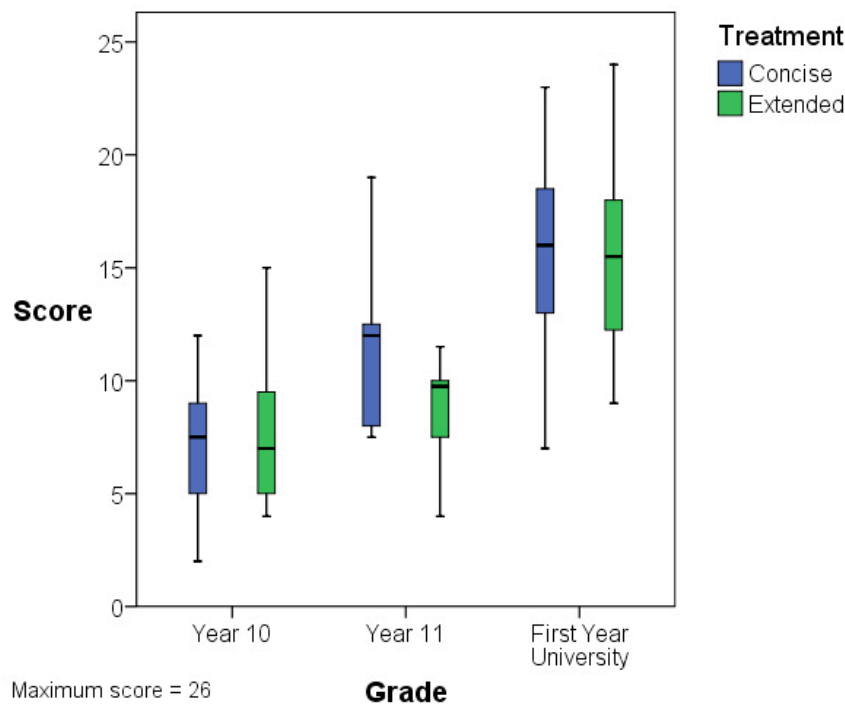

Figure 3: Grouping the sample by prior knowledge did not reveal any differences between treatments. Boxes extend from the first to third quartiles, enclosing the median. Whiskers show the extremes.

Only in comparisons of year 11 students' multiple choice scores did differences between the concise and extended groups approach significance ($t=1.78$, p(two-tailed)=0.096). Although this difference favoured the concise group, the sample size was small ($n = 9$ and 8 for the concise and extended groups, respectively), limiting the conclusions that can be drawn.

Discussion

This study failed to find evidence for the coherence principle in an authentic learning setting involving online multimedia. Students who watched a concise multimedia treatment on stellar spectra performed similarly on the post-test to those who watched an extended version. The findings suggest that, similar to the modality and cueing effects (Tabbers et al., 2004), the coherence principle does not generalise easily to authentic settings. In the following section we consider some possible reasons for this and discuss the difficulties in measuring learning differences in real learning environments.

One possible explanation for the lack of difference in post-test scores is that the extra material encouraged learners to pay attention to the extended treatment (Mitchell, 1993). As mentioned in the introduction, learner interest is likely a more important factor in a real learning environment than in a controlled laboratory. Detrimental effects of the increased cognitive load engendered by the extra material may have been balanced by heightened interest and attention. This could be further investigated through interviews or by running the same study again in a laboratory environment.

Another possibility is that the materials were not targeted at the appropriate level of prior knowledge for the study participants. If the students already knew much of the material, the extraneous information would not have imposed an onerous cognitive load and learning would therefore be unaffected. In other words the learning material was below the comprehensive capacity of students in all groups. This is unlikely, however, as the post-tests for even the highest prior knowledge students did not appear to suffer from a ceiling effect. Furthermore, the test scores discriminated the three prior knowledge levels of students quite clearly. Within these groups, there was little evidence that the performance on the post-test depended on the type of multimedia treatment. We do note though, that a common misinterpretation of the coherence principle is that in fear of overloading students, learning material may be designed at too low a comprehension level.

Only the year 11 multiple choice scores showed a slight difference between the concise and extended groups. It is possible that the year 10s had such limited prior knowledge that they learned almost nothing from the multimedia and that first year university students had such well developed schemas that the extraneous information was not a burden. A greater sample size of year 11s would be required to investigate this matter.

Students' difficulties in constructing scientific concepts are very complicated with conceptual understanding playing a central role. Indeed, a dialogue between a tutor and a student involving discussions of alternative conceptions in short linear video interventions have shown higher gains than concise videos. (Muller, Sharma, Eklund & Reimann, 2007; Muller, Sharma & Reimann, 2007). Whether the discussion of alternative conceptions is germane load or not adds another facet to cognitive load theory.

One limitation of the study involved the interpretation of short answer responses. It was difficult to determine whether some students wrote brief responses because they possessed a limited understanding of the material, or because the study was voluntary. Short answer questions are commonly used to assess learning following multimedia (Mayer, 2005), and it could be argued that they are more sensitive to differences in

understanding than multiple choice questions. If students failed to demonstrate their understanding on these questions to save them time and effort, existing differences in learning may have been unobservable. It is difficult to circumvent this obstacle in authentic learning environments unless marks are awarded on the post-test to encourage earnest participation. This raises ethical issues, however, if in fact one treatment is superior to the other.

To improve the discrimination on the post-test, more questions could be used. Additionally, different types of questions, like two-tiered multiple choice items, would allow for deeper probing of student conceptions without requiring short answer responses (Treagust, 1987). Two-tiered questions involve a traditional multiple choice question followed by a multiple choice set of supporting explanations. Considering the two responses together helps identify random guessing and inconsistencies in reasoning. This technique has been used successfully in a number of other studies (e.g. Kearney & Treagust, 2001; Sharma, Khachan, Chan & O'Byrne 2005).

In future studies it would be essential to measure cognitive load directly for each treatment. Previous research has shown that self reports of mental effort provide a simple, reliable method of assessing cognitive load (Paas, Tuovinen, Tabbers & van Gerven, 2003). Equally important would be a direct measure of interest for each treatment. This information could help interpret the results of this study and determine which format of instruction is most effective.

There are many advantages of studying student learning with online materials. First, the logistical challenges of face to face experiments are easily overcome. For example, participants can partake anonymously and random assignments to different instructional treatments are straightforward. Second, online multimedia provides a transparent and repeatable means of comparing different forms of instruction. Third, the results of such studies not only yield recommendations for multimedia design, they also inform fundamental theories of the learning process. One of our recent studies has shown that misconception based additional material can dramatically improve learning and alter the way in which learners perceive the multimedia (Muller, Bewes, Sharma & Reimann, in press).

Summary

Adding approximately 50% extra interesting but irrelevant information to a multimedia treatment did not result in lower achievement on a post-test as would be predicted by the coherence principle. This finding appears robust across a range of age and prior knowledge levels. One possible explanation is that the additional interesting information served a useful function, maintaining attention in the authentic learning settings employed in the experiment. The results are far from conclusive, however. Further investigations can improve on this exploratory study by making simple alterations to the methodology. The post-test could be improved by including two-tiered questions or other additional items. Marks awarded for earnest participation would also ensure students demonstrate the full extent of their knowledge, especially on short answer questions. An experiment conducted in a controlled laboratory with the same materials could reveal whether the coherence principle is observable under any conditions with this subject matter and these participants. Finally, direct measures of the cognitive load and interest experienced by students in the different treatments would directly establish the effect of the extra material.

References

Binnie, A. (2004). Development of a senior physics syllabus in New South Wales. *Physics Education*, 39(6), 490-495.

Board of Studies (2001). *Physics Stage 6 Syllabus*. Sydney. http://www.boardofstudies.nsw.edu.au/syllabus_hsc

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *Journal of the Learning Sciences,* 2(2), 141-178.

Goetz, E. T. & Sadoski, M. (1995). Commentary: The perils of seduction: Distracting details or incomprehensible abstractions? *Reading Research Quarterly,* 30(3), 500-511.

Harp, S. F. & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology,* 90(3), 414-434.

Hidi, S. (2001). Interest, reading, and learning: Theoretical and practical considerations. *Educational Psychology Review,* 13(3), 191-209.

Kalyuga, S. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology,* 13(4), 351-371.

Kearney, M. & Treagust, D. F. (2001). Constructivism as a referent in the design and development of a computer program using interactive digital video to enhance learning in physics. *Australasian Journal of Educational Technology,* 17(1), 64-79. http://www.ascilite.org.au/ajet/ajet17/kearney.html

Klahr, D. & Li, J. (2005). Cognitive research and elementary science instruction: From the laboratory, to the classroom, and back. *Journal of Science Education and Technology,* 14(2), 217-238.

Mayer, R. E. (2001). *Multimedia learning*. Cambridge, UK: New York: Cambridge University Press.

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

Mayer, R. E, & Jackson, J. (2005). The case for coherence in scientific explanations: Quantitative details can hurt qualitative understanding. *Journal of Experimental Psychology Applied,* 11(1), 13-18.

Mayer, R. E., Heiser, J. & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology,* 93(1), 187-198.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *The Psychological Review,* 63, 81-97.

Mitchell, M. (1993). Situational interest: Its multifaceted structure in the secondary school mathematics classroom. *Journal of Educational Psychology,* 85, 424-436.

Muller, D. A., Bewes, J., Sharma, M. D. & Reimann, P. (2007). Saying the wrong thing: Improving learning with multimedia by including misconceptions. *Journal of Computer Assisted Learning.* DOI: 10.1111/j.1365-2729.2007.00248.x

Muller, D. A, Sharma, M. D. & Reimann, P. (in press). Raising cognitive load with linear multimedia to promote conceptual change. *Science Education. DOI 10.1002/sce.20244*

Muller, D. A., Sharma, M. D., Eklund, J. & Reimann, P. (2007). Conceptual change through vicarious learning in an authentic physics setting. *Instructional Science,* 35(6), 519-533.

Paas, F., Tuovinen, J. E., Tabbers, H. & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist,* 38(1), 63-71.

Schank, R. C. (1979). Interestingness: Controlling inferences. *Artificial Intelligence,* 12, 273-297.

Schraw, G. & Lehman, S. (2001). Situational interest: A review of the literature and directions for future research. *Educational Psychology Review,* 13(1), 23-52.

Sharma, M. D., Khachan, J., Chan, B. & O'Byrne, J. (2005). An investigation of the effectiveness of electronic classroom communication systems in large lecture classes. *Australasian Journal of Educational Technology,* 21(2), 137-154. http://www.ascilite.org.au/ajet/ajet21/sharma.html

Sweller, J., van Merrienboer, J. J. G. & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review,* 10(3), 251-296.

Tabbers, H. K., Martens, R. L. & van Merrienboer, J. J. G. (2004). Multimedia instruction and cognitive load theory: Effects of modality and cueing. *British Journal of Educational Psychology,* 74, 71-81.

Treagust, D. F. (1987). The development and use of diagnostic instruments to evaluate students' misconceptions in science. *International Journal of Science Education,* 10, 159-169.

Derek A. Muller, Kester Lee & Manjula D. Sharma
Sydney University Physics Education Research Group
School of Physics, The University of Sydney, NSW 2006, Australia
Author for correspondence: Derek A. Muller. Email: muller@physics.usyd.edu.au