



Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

The role of interest and images in slideware presentations

Jason M. Tangen^{a,*}, Merryn D. Constable^a, Eric Durrant^b, Chris Teeter^b, Brett R. Beston^b, Joseph A. Kim^b

^a School of Psychology, University of Queensland, St Lucia, QLD, 4072, Australia

^b McMaster University, Canada

ARTICLE INFO

Article history:

Received 5 August 2010

Received in revised form

28 October 2010

Accepted 28 October 2010

Keywords:

Multimedia learning

Memory

Pedagogical issues

Media in education

PowerPoint

Interest

ABSTRACT

With the advent of technologies that allow lecturers to develop presentations using software such as Microsoft PowerPoint, Apple Keynote, and OpenOffice Impress (referred to generically here as “slideware”), lectures and meetings are beginning to resemble cinematic experiences rather than the text filled transactions that have been the norm for the last decade. Unfortunately, there has been little research on slideware use. Furthermore, literature on this topic indicates that lectures typically do little to encourage elaborative thinking. To address this gap in the literature, we tested 90 participants who each viewed a lecture on the visual system that consisted of three different styles of presentation: (1) Image Congruent: presenting images that were relevant to the target information, (2) Image Incongruent: presenting images that were relevant to the narration but not the target information and, (3) Text Based: presenting text summaries of the narration. Throughout each of these conditions, participants provided ratings of interest and then completed a quiz to measure recognition, recall and transfer. Our results revealed that participants were more interested in the image-rich slides than bullet points, and that accuracy (but not interest) depended on whether the images were relevant to the content of the lecture or not.

Crown Copyright © 2010 Published by Elsevier Ltd. All rights reserved.

1. Introduction

Technology has developed rapidly in education and training. Today, educators, researchers, and professionals have a broad selection of multimedia tools available for presenting and visualizing complex information. Searchable image databases (e.g., Google Images, Flickr, iStockPhoto, etc.) offer free or inexpensive high resolution photographs, illustrations and video; and slide software (or “slideware”) such as Microsoft PowerPoint, Apple Keynote, and OpenOffice Impress come bundled with predesigned templates, themes, builds, and special effects that produce presentations that range from simple bullet point summaries to rich multimedia expositions. In an educational context, these multimedia tools certainly provide an immersive experience that students find appealing (e.g., Szabo & Hastings, 2000; Apperson, Laws, & Scepanisky, 2002), but it is less clear whether interest in these visually rich presentations encourages learning compared to traditional, text based presentations. Every presenter has likely wondered whether to include a comic or an amusing photo in their presentation and what impact this would have on the audience, or whether to include, say, a photograph of some apparatus or simply describe it in a list bullet points. Yet, despite the widespread adoption of multimedia tools, very few experiments have been conducted on their pedagogical utility. To address this gap in evidence in an educational context, we examine the visual style component of multimedia using slideware as a medium.

Slideware presentations allow users to position text, graphics, images, video and other media on individual pages or “slides”, and are often displayed on-screen to supplement the verbal narration of the presenter. Most research on multimedia learning has focussed on relative encoding, processing and retrieval of verbal and pictorial information (e.g., Chandler & Sweller, 1991; Morett, Clegg, Blalock, & Mong, 2009; Paivio, 2007; Mayer, Griffith, Jurkowitz, & Rothman, 2008). Dual coding and dual channel theories propose that learners possess a verbal and a visual processing system in working memory, allowing them to encode information in long-term memory on two different levels (Clark & Paivio, 1991). This dual processing system would allow information presented using multimedia to be encoded on multiple levels, possibly leading to deeper and more meaningful processing. For example, according to dual coding theory, the verbal channel would

* Corresponding author.

E-mail address: jtangen@psy.uq.edu.au (J.M. Tangen).

process verbal narration and written text. On the other hand, both visual and verbal channels are implicated in processing images. Images are, therefore, more likely to be encoded on multiple levels than verbal stimuli (Paivio, 2007).

Cognitive load theory is also important for learning and memory as it is based on the assumption that humans have a limited capacity for processing information in working memory (Miller, 1956). If working memory's capacity is overloaded, then some information will not be encoded adequately; however, some cognitive load is intrinsic to learning (Muller, Sharma, & Reimann, 2008). Intrinsic cognitive load is the result of the inherent nature of the learned information. For example, learning a complex physical science concept versus learning a string of numbers would be comparatively higher in intrinsic cognitive load due to its complexity. Trying to alter intrinsic cognitive load would compromise the understanding of the material (Leahy & Sweller, 2008). Extraneous cognitive load results from the format of the instruction. For example, giving a poorly designed lesson that requires students to integrate information from a prior slide onto a diagram would create extraneous cognitive load by the participants' effort to engage in integration. Reducing this extraneous cognitive load provides more processing capacity for relevant information that is implicated in schema acquisition and automation. Due to the structure of long-term memory, "germane" cognitive load is integral to the effective learning of instructional material (Leahy & Sweller, 2008). Therefore, to promote superior retention of material, reducing extraneous cognitive load, thereby increasing the capacity for "germane" cognitive load, is critical.

The cognitive theory of multimedia learning proposed by Mayer and his colleagues (e.g., Mayer & Moreno, 1998) is based on dual coding and cognitive load theory and is an applied approach aimed at providing direction for the creation of multimedia. It centers on three assumptions: (1) we have separate visual and verbal processing channels (i.e., dual coding theory, Paivio, 2007); (2) these two channels have limited processing capacity (Chandler & Sweller, 1991); and (3) meaningful learning can only occur if the learner attends to the relevant stimuli, organizes and integrates it into long term memory (Mayer, Heiser, & Lonn, 2001).

A factor that stems from this third assumption is the seductive details hypothesis. That is, interesting details that are relevant to the topic but not relevant to the conceptual goal will hurt learning presumably because high interest items are more likely to be processed at a deeper level than low interest items (Pintrich & Schunk, 1996). As working memory has a limited capacity, highly interesting details (or interesting images) prevent more relevant information from being processed effectively (Mayer, 2008).

The effect of seductive details has been exhibited using video clips (Mayer et al., 2001), in PowerPoint presentations (Harp & Mayer, 1998), in text based environments with illustrations (Harp & Mayer, 1998) and in text passages (Garner, Brown, Sanders, & Menke, 1992). Furthermore, a number of researchers have found that readers will often remember the interesting adjuncts within a textual passage rather than the conceptually important ideas (Garner et al., 1992; Garner & Gillingham, 1991; Hidi & Anderson, 1992; Lehman, Schraw, McCrudden, & Hartley, 2007). Seductive details or interesting-but-incongruent information can make processing difficult as it does not contribute to the coherence of the material and often make it less concise (Mayer & Moreno, 2003). This makes it difficult for the learner to select the relevant content for further processing, leading to an overall reduction in quality of learning. Consequently, if the material is made to be more concise and coherent, the learner will be inclined to do more essential processing rather than incidental processing (Mayer & Moreno, 2003).

Another factor that stems from the cognitive theory of multimedia is redundancy. Educators often construct teaching materials using onscreen text and narration (Kalyuga, Chandler, & Sweller, 1999), but this potentially creates extra cognitive load. Several studies have shown that duplicating narrated text as visual text, leads to a decrease in qualitative learning (Kalyuga et al., 1999; Mayer et al., 2001; Mayer & Moreno, 1998). For example, according to the cognitive theory of multimedia, text on screen (verbal channel) and narration (verbal channel) causes extra cognitive load as the learner is attending to unnecessary information that could be used in schema acquisition. Kalyuga et al. (1999) propose that a dual mode (visual and verbal) presentation cannot eliminate extraneous load; it can only alleviate it by increasing working memory capacity. Duplicating narrated text as visual text, however, seems to cause redundancy that overrides any benefit gained from a dual modality presentation. This redundancy effect was associated with an increase in cognitive load (measured by participants' responses about their subjective mental load on a 7-point scale). While a dual mode presentation without the redundant textual information was associated with reduced cognitive load (Kalyuga et al., 1999). Similar effects have been found with instructional physics material where the goal is qualitative learning (Mayer & Jackson, 2005), with narrated animations on the formation of lightning (Mayer et al., 2001; Mayer & Moreno, 1998), car brakes (Mayer & Moreno, 1998), or static diagrams and audio (Yaghou, Low, & Sweller, 1995).

It remains unclear how interest and accuracy vary as a function of multimedia type. Theories of multimedia learning deal primarily with the retention of the content that results from the presentation medium, with less regard for how relevant the verbal or visual information is to the content. Similarly, very few experiments have been conducted to isolate the role of interest and retention in learning about various forms of multimedia. Are people more interested in bullet point or image rich presentations? Do the images need to be relevant to the topic? One goal of this experiment is to compare the effect of image based vs. text based slides on ratings of interest and the retention of content. Secondly, we will contrast the relevance of the slide images to the presentation content, and examine its effect on interest and accuracy. In doing so, our aim is to disentangle interest and the deep cognitive processing of content, which are often confounded in measures of teaching performance.

2. Method

2.1. Participants and design

Ninety undergraduates (69 females, 21 males) between the ages of 17 and 50 ($M = 20.32$, $SD = 5.4$) from The University of Queensland participated for credit in an introductory psychology course. Each participant was presented with a 15 min lecture on the human visual system. They learned about the structure of light waves, the anatomy and function of the eye and visual cortex, and how visual information is processed in the brain. This was a within-subjects experiment, whereby participants experienced three different slideware presentation conditions:

2.1.1. Image Congruent: Highly image based with one image or word per slide (roughly 60 slides in 5 min)

Each image was selected to support the key concept of the narration (e.g., Narration: "Light travels in waves that code for the scene we're viewing"; Slide: an image of light waves traveling to the eye).

2.1.2. Image Incongruent: Highly image based with one image or word per slide (roughly 60 slides in 5 min)

Each image was selected to be related to the narration, but not support the key concept (e.g., Narration: “Light travels in waves that code for the scene we’re viewing”; Slide: an image of computer code).

2.1.3. Text based: Highly text based with two or three bullet points per slide (roughly 8 slides in five minutes)

The bullet points summarized (but were not identical to) the content of the narration (e.g., Narration: “Light travels in waves that code for the scene we’re viewing”; Slide: a bullet point that states *Visual information is transmitted by light waves. Light is reflected from the environment and received by our eyes*).

The order was completely counterbalanced and the 15 min audio track was identical for each participant. During the presentation, participants rated their level of interest from 1 (Very Dull) to 5 (Very Good). Afterward, we tested them on the content of the lecture using a series of recall, recognition, and transfer questions. Before the presentation, participants were told that they were going to view a lecture on the human visual system and were provided with instructions about how to toggle the keys on the keyboard to indicate their level of interest. They were then told that they would be given a quiz immediately afterward. None of the participants had experience with the content of the presentation beyond that expected of a first year undergraduate student.

2.2. Materials and procedure

2.2.1. Presentation

We prepared a 15 min presentation on the human visual system and edited an audio recording of a male narrator reading the script. We then constructed three separate slideware presentations using Apple’s Keynote software according to the three conditions described above. The Image Congruent and Image Incongruent presentations each consisted of 184 slides with one image or word per slide, and the Text Based presentation consisted of 23 slides, and was modeled after a traditional “bullet point” presentation (e.g., two or three bullet points per slide that summarize the key concepts in the narration). We obtained high resolution images from various stock photo websites including iStockPhoto and Stock.XCHNG. The slides in each of the three presentations were then synced with the same audio track to create three separate movies using Apple’s iMovie software. We then divided the three movies into three, five minute sections and spliced them together to create the six possible presentation combinations. From a participant’s perspective, the first five minutes of the presentation might begin with the highly visual Image Congruent condition followed by the next five minutes consisting of the Text Based condition (presenting the information using bullet points rather than a quick succession of images), and the final five minutes consisting of the Image Incongruent condition, which is highly visual, but the images are less relevant to the key concepts depicted in the narration.

We selected the images in the Image Congruent condition to be as highly relevant to the narration as possible, while the images in the Image Incongruent condition were only loosely related to the narration. We did this by splitting up the narration into small sections, for example, given the passage “It is not surprising that 75% of North Americans claim...”, the slide in the Image Congruent condition depicts 75% of the North American continent in the form of a three dimensional pie chart. The slide in the Image Incongruent condition depicts a stock image of a man with his mouth agape conveying surprise. “Surprise” is certainly related to the content of the narration, but it not central. We repeated this image selection process throughout the narration.

2.2.2. Interest ratings

The presentations were shown to participants using Runtime Revolution, which was also used to collect their interest ratings. We used an interactive feedback tool to measure participants’ interest in the lecture content, a technique commonly used in market research and in political debates to gauge an audience’s reaction to visual stimuli over time. These dynamic ratings were used simply as a measure of temporary interest aroused by specific features of the task they are engaged in (Eccles & Wigfield, 2002), rather than a more persistent measure characterized by an intrinsic desire to understand the content (Schraw & Lehman, 2001). Throughout the 15 min presentation, participants were asked to indicate any change in their level of interest by pressing the up and down arrows on the keyboard. Each keystroke toggled a 5-point scrollbar on the screen accompanied by labels that ranged from “Very Dull” (1), “Dull” (2), “Normal” (3), “Good” (4), and “Very Good” (5). The value of the scrollbar was sampled and recorded every second of the presentation.

2.2.3. Test

In order to test participants on the content of the lecture, we created nine different questions based on each of the three, five minute sections of the presentation (i.e., nine questions were based on the content from the first third of the lecture, nine were based on the content of the second third, and nine on the final third). Three of the nine questions for each section were designed to test recognition (e.g., “What two ways can light differ?” – select from four multiple choice options), three tested recall (e.g., “Describe why the iris makes the pupil expand and contract?” – provide a short answer), and three tested transfer (e.g., “If the optic chiasm were severed, what would happen to our visual scene?” – provide a short answer). The 27 questions were presented in random order to participants using Runtime Revolution, and they made their responses using the computer mouse and keyboard. They could take as long as they want to make a response. After responding to each question, they were asked to rate how confident they were in their answer by moving a scrollbar from 1 (Guessing) to 6 (Very Confident).

3. Hypotheses

Considering the previous research around cognitive load, dual channels, and the cognitive theory of multimedia, we propose the following hypotheses:

3.1. Hypothesis 1

The Image Congruent condition will lead to superior learning outcomes as measured by performance on recall, recognition and transfer questions compared to the Image Incongruent and Text Based conditions.

This hypothesis is based on the idea that images are encoded in multiple processing channels leading to more elaborate encoding (Snodgrass, Wasser, Finklestein, & Goldberg, 1974). Encoding in the Text Based condition should be less elaborate than the Image Congruent condition as it is only using the verbal channel. This overuse of the verbal channel should also create extra cognitive load that contributes to poorer encoding for the Text Based condition (Mayer, 2008). Although the Image Incongruent condition is prompting multiple sensory modalities, the misdirection associated with the images is likely to result in inappropriate encoding in memory and more incidental processing resulting in poorer performance.

3.2. Hypothesis 2

Participants will consider the two image based conditions more interesting than the Text Based condition.

3.3. Hypothesis 3

The Text Based condition will result in superior learning outcomes as measured by recall, recognition and transfer questions compared to the Image Incongruent condition.

This third hypothesis is driven by the fact that our incongruent images should be analogous to “seductive details” causing an inappropriate schema to be created in memory, leading to poorer retrieval of the target information. The incongruent images are likely to result in more incidental processing as a result of their ambiguity, leaving less processing available, which is implicated in schema acquisition (Leahy & Sweller, 2008). Although, redundancy in the text based condition would usually create extra cognitive load because of overuse of the verbal channel, the text based summaries are thought to act as a signaling device, thereby mitigating the effects of higher cognitive load (Mayer, 2008). As such, due to the difference in cognitive load, the Text Based condition should result in better performance than the Image Incongruent condition.

3.4. Hypothesis 4

Participants will perform better on the transfer questions in the Image Congruent condition compared to the Text Based condition.

The cognitive theory of multimedia supports the assertion that encoding using multiple sensory modalities leads to more concrete and reliable memory retrieval cues as well as a deeper level of processing. Consequently, performance in the Image Congruent and Text Based conditions are unlikely to differ for recall and recognition, as deeper levels of processing are not required. The Text Based condition, however, only uses the verbal channel, leading to a less elaborate schema within memory (and therefore worse performance) compared to the Image Congruent condition – on the transfer questions in particular.

It is unclear how participants' ratings of confidence will track their accuracy on the test questions as previous studies of self assessment in student learning has resulted in conflicting evidence. For example, several experiments on self assessment have concluded that the relationship between actual performance and self-rated performance is modest at best (see Dunning, Heath, & Suls, 2004, for review). On the other hand, others have demonstrated that students can accurately assess their performance on particular exam questions (e.g., Rosenthal et al., 2010).

4. Results

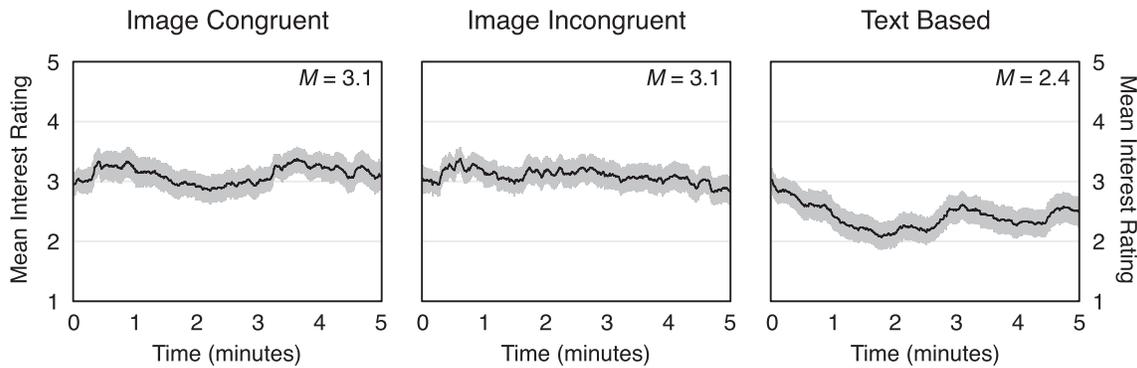
4.1. Interest ratings

During the presentation phase of the experiment, participants were asked to indicate their level of interest on a five point scale from 1 (Very Dull) to 5 (Very Good). The mean interest ratings across each of these five minute presentations are depicted in Fig. 1A. It is clear from this figure that ratings were higher for the two image based conditions compared to the Text Based condition, and that there was very little difference between the Image Congruent and Image Incongruent conditions. Indeed, the mean level of interest for the Image Congruent condition ($M = 3.11, SD = .77$) was virtually identical to the mean level of interest for the Image Incongruent condition ($M = 3.09, SD = .75$), and both were greater than the mean level of interest in the Text Based condition ($M = 2.41, SD = .84$). Since we were interested primarily in the difference between interest ratings across the three conditions (rather than the change in interest over time), we computed the mean level of interest for each participant for each of these conditions and conducted a one-way repeated measures ANOVA on these mean interest ratings. This analysis revealed a significant main effect of presentation type (Image Congruent, Image Incongruent, Text Based), $F(2, 178) = 29.74, p < .001$. Simple effects analyses revealed that the mean interest ratings in the Image Congruent presentation were not significantly different from those in the Image Incongruent presentation, $F(1, 178) < 1$, but they were significantly different from those in the Text Based presentation, $F(1, 178) = 46.02, p < .001$, and the mean interest ratings in the Image Incongruent presentation were significantly different from those in the Text Based presentation, $F(1, 178) = 43.16, p < .001$. As we expected, the two image based presentations were rated as more interesting than the Text Based presentation.

4.2. Accuracy

After viewing the 15 min presentation, we tested participants on the content. We were interested in their level of accuracy on the content presented during the three different presentation types (Image Congruent, Image Incongruent and Text Based). As indicated above, we also used three different types of questions (recall, recognition, transfer) to measure their depth of processing. Responses to each of the 27 questions were graded without knowledge of the presentation conditions and verified by a second marker, again blind to the order of the conditions. An open ended response was assigned .5 of a mark if a participant provided a response that was partially correct. For example, given the question “What is the matter that fills the main chamber of the eye?”, the correct response is “vitreous humor”, but it would be

A Interest Ratings



B Accuracy

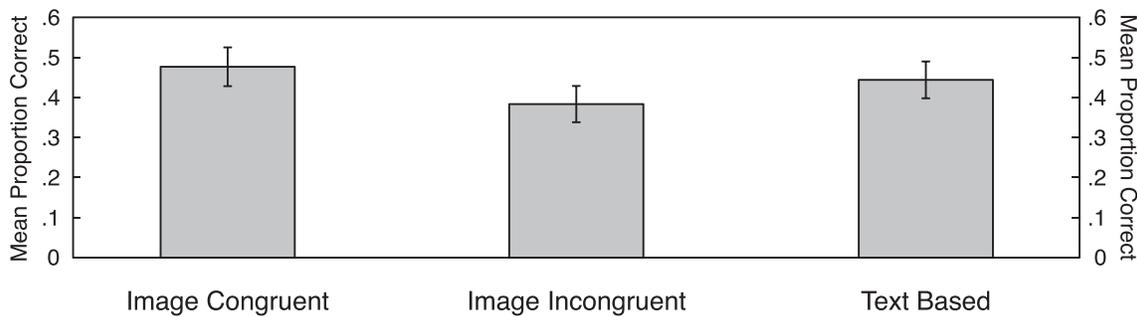


Fig. 1. The mean interest ratings sampled every second of the three, five minute presentations are presented in Fig. 1A, and the corresponding mean scores on a subsequent quiz (proportion correct) are presented in Fig. 1 B. Error bars represent 95% confidence intervals.

assigned .5 of a mark if they recognized that the chamber was comprised of liquid. The mean proportion correct for each presentation type is presented in Fig. 1B. It is evident from the results depicted in this figure that participants performed better on the Image Congruent and Text Based conditions compared to the Image Incongruent condition.

A 3 (presentation type: Image Congruent, Image Incongruent, Text Based) \times 3 (question type: recall, recognition, transfer) repeated measures ANOVA conducted on their accuracy scores confirmed this observation with a significant main effect of presentation type, $F(2, 178) = 5.13, p = .007$. Simple effects analyses revealed that accuracy rates were higher for the Image Congruent condition ($M = .48, SD = .24$) than the Image Incongruent condition ($M = .38, SD = .22, F(1, 178) = 9.97, p = .002$). The Image Incongruent condition also had significantly lower accuracy rates than Text Based condition ($M = .44, SD = .22, F(1, 178) = 4.2, p = .042$). The difference in accuracy between the Image Congruent and Text Based condition was not significant, $F(1, 178) = 1.23, p = .27$. The ANOVA also revealed a main effect of question type, $F(2, 178) = 49.93, p < .001$, where accuracy rates were higher for the recognition questions ($M = .56, SD = .19$) compared to the recall questions ($M = .36, SD = .23, F(1, 178) = 85.26, p < .001$), and transfer questions ($M = .38, SD = .20, F(1, 178) = 62.83, p < .001$), but the difference between the recall and transfer questions was not significant, $F(1, 178) = 1.71, p = .19$. The interaction between presentation type and question type was not significant, $F(4, 356) = 1.8, p = .13$.

In order to address Hypothesis 4 (i.e., participants will perform better on the transfer questions in the Image Congruent condition compared to the Text Based condition), we conducted simple effects analyses on each of the three question types comparing the Image Congruent and Text Based conditions. As predicted, there was no significant difference between the Image Congruent and Text Based conditions for the recall questions, $F(1, 356) < 1$, or recognition questions, $F(1, 356) < 1$, but there was a marginal difference for the transfer questions, $F(1, 356) = 3.14, p = .079$. That is, we found partial support for Hypothesis 4 in that participants were more accurate in the Image Congruent condition than the Text Based condition on the transfer questions, but not the recognition or recall questions. This should occur as a result of transfer requiring a “deeper” level of understanding (Mayer, 2008) and pictures prompting more elaborate processing due to dual coding.

4.3. Confidence

We conducted a 3 (presentation type: Image Congruent, Image Incongruent, Text Based) \times 3 (question type: recall, recognition, transfer) repeated measures ANOVA on participants' confidence ratings to determine how their confidence varied as a function of presentation and question type. This analysis revealed a significant main effect of presentation type, $F(2, 178) = 3.47, p = .03$, question type, $F(2, 178) = 51.1, p < .001$, as well as a significant interaction, $F(4, 356) = 2.74, p = .03$. As illustrated in Table 1, participants' confidence ratings followed a similar pattern to their accuracy ratings, such that participants were more confident with their answers in the Image Congruent condition compared to the incongruent condition, $F(1, 178) = 6.94, p = .009$, but their confidence in the Image Congruent condition did not differ

Table 1
Mean confidence ratings for each presentation and question type (standard deviations in parentheses).

	Recall	Recognition	Transfer	Total
Image congruent	3.14 (1.28)	3.93 (1.12)	3.37 (1.31)	3.48 (1.28)
Image incongruent	2.74 (1.27)	3.56 (1.11)	3.12 (1.16)	3.14 (1.23)
Text based	3.16 (1.40)	3.71 (1.22)	3.06 (1.20)	3.31 (1.30)

significantly from their confidence in the Text Based condition, $F(1, 178) = 1.7, p = .19$, and confidence ratings in the Image Incongruent and Text Based conditions were not quite significantly different, $F(1, 178) = 1.77, p = .18$, as they were for accuracy. There is no obvious explanation for the discrepancy between accuracy and confidence for these two conditions in particular, since the pattern of results for these two dependent measures are nearly identical in every other respect.

Regarding question type, participants were significantly more confident in their answers to recognition compared to transfer questions, $F(1, 178) = 54.78, p < .001$, and recall questions, $F(1, 178) = 93.39, p < .001$. They were also more confident in their answers to transfer questions compared to recall questions, $F(1, 178) = 5.12, p = .03$. As such, participants were the most confident in their responses to recognition questions, followed by transfer questions, and the least confident in their responses to recall questions.

4.4. Summary

Overall, our results indicate that participants found the two image based presentations more interesting than the Text Based presentation. Interest, however, does not seem to be directly associated with their learning outcomes as they performed significantly worse in the Image Incongruent condition than the Image Congruent condition – despite identical ratings of interest. Accuracy in the Image Incongruent condition was also significantly worse than the Text Based condition, despite participants rating it as much less interesting. Confidence ratings followed a similar pattern to the accuracy ratings, suggesting that participants knew when they were making a correct response.

5. Discussion

Despite the widespread use of slideware as a medium for teaching, we know very little about its effectiveness. As such, our aim was to investigate the role that text and images play in interest and accuracy. Our data suggest that image based presentations are more interesting than text based (bullet point) presentations. In terms of accuracy, interesting-but-irrelevant images resulted in worse performance than the text based condition that had no images at all. Indeed, if the purpose of the presentation is to capture the interest of the audience, then our data suggest that it does not matter whether the images are relevant to the topic or not. If, however, you want the audience to remember the content of the presentation, then images that are relevant to the topic will result in better memory for the content than irrelevant images.

Overall, our participants were less accurate in their responses to questions associated with the Image Incongruent condition compared to the Image Congruent or Text Based conditions. This partly supports our first hypothesis that the Image Congruent condition would lead to superior learning outcomes. The higher accuracy rate in the Image Congruent condition compared to the Image Incongruent condition is likely the result of more elaborate and appropriate processing within memory, consistent with a schema acquisition explanation. Both presentation conditions involved a large number of images that were matched semantically to the narration. However, only the photos in the Image Congruent condition depicted the key concepts. As such, the Image Congruent condition could be considered highly relevant leading to an appropriate schema in memory with reliable retrieval cues.

This finding may not be as intuitive as it seems. For example, research on memory for brands shows that incongruity can result in better memory (Houston, Childers, & Heckler, 1987). Incongruity has been shown to lead to greater amounts of tension, which in turn leads to an increase in arousal (Heider, 1958). Arousal has been linked to more processing in memory resulting in a greater likelihood of retrieval of incongruent stimuli such as: brands or advertisements (Goodstein, 1993; Houston et al., 1987); people (Hastie & Kumar, 1979; Srull, 1981; Srull, Lichtenstein, & Rothbart, 1985); and ethnicities (Baker & Petty, 1994). Furthermore, the degree of incongruity may play a role in the likelihood of retrieval (Meyers-Levy & Tybout, 1989; Peracchio & Tybout, 1996; Stayman, Alden, & Smith, 1992; Sujjan, 1985) such that high levels of incongruity will only lead to deeper processing if the person has the cognitive capacity to engage in the effortful processing required (Peracchio & Tybout, 1996). Incongruity, by its nature, requires deeper processing if one is to effectively integrate it into long-term memory (Lee & Schumann, 2004). Since our participants completed the quiz immediately after viewing the lecture, it is unlikely they had the resources available to engage in the effortful processing that would be prompted by incongruity. As such, it may be possible that the Image Incongruent condition could result in greater elaboration in memory if participants had the resources and motivation to engage in effortful processing.

In comparing the Image Congruent and Text Based conditions, there was no significant decrement in accuracy for the Text Based condition despite its lower mean. This result was inconsistent with our first hypothesis as we predicted that the Image Congruent condition would result in better performance compared to the Text Based condition. There are a number of possible explanations for these results. First, any cognitive load due to the redundancy of information in visual text and narration would likely be mitigated by the bullet point summaries that signaled the target information (Mayer, 2008). This signaling should be associated with appropriate schema formation within memory (Leahy & Sweller, 2008). It is possible that the signaling aided memory formation and retrieval insofar as encoding only in one processing channel (verbal) was not any worse than encoding in two channels (visual and verbal). In addition, the content of the presentation may not have been complex enough to require deeper processing. However, given that no participant scored at ceiling levels, this is unlikely. Second, it may be a question of a participant's cognitive resources, given that the quality of the schema acquisition process is somewhat governed by the amount of time and cognitive resources available to the learner (Lee & Schumann, 2004). In our experiment, participants had a limited amount of time to process the information during the presentation and they were tested immediately after. As such, the resources required for the elaborative processing in the Image Congruent condition were not available. This may be why we found

no significant difference between the Text Based condition and the two image based conditions: there was simply not enough time or resources for elaborative processing to occur. A third possibility is transfer appropriate processing (Morris, Bransford, & Franks, 1977). That is, the (mis)match between how information is encoded and retrieved. Each of the questions that tested our participants' knowledge of the presentation content was text based. This may have provided a particular processing advantage for the Text Based condition compared to the two image based conditions. It might be possible to reverse this difference if the test questions were image based.

One might argue that measuring participants' interest throughout the presentation had the potential to impose extraneous cognitive load, which may interact with their test performance. This is doubtful for three reasons: (1) participants were not required to constantly update their interest rating throughout the presentation, only when their interest changed; (2) mean interest ratings were virtually identical for the two image based conditions, but participants were more accurate in the Image Congruent condition (.48) compared to the Image Incongruent condition (.38). Because the images in the Image Incongruent condition were selected to be related to the content (but much less relevant than the images in the Image Congruent condition), it is very unlikely that participants in the Image Incongruent condition were working particularly hard to keep track of the lecture content; and (3) participants in the Text Based condition were indeed working hard, as we were taxing their verbal channel with both the narration and the text-based bullet points on the slide, yet they were still just as accurate (.44) as the Image Congruent condition (.48), even though they also provided ratings of interest throughout the condition. It is very unlikely, therefore, that our interest measure created an extra cognitive load that influenced the observed difference in accuracy.

We predicted that participants would consider the two image based conditions more interesting than the Text Based condition (Hypothesis 2), and that the Text Based condition would result in better learning than the Image Incongruent condition (Hypothesis 3). The basis for these predictions is the effect of the seductive details hypothesis. Participants performed better on questions related to the Text Based condition compared to the Image Incongruent condition, despite rating the image based conditions as more interesting. These data suggest that interest is not a sufficient condition for learning, which may seem counterintuitive given the large body of research which demonstrates that interest prompts elaborative processing (e.g., Pintrich & Schunk, 1996; Wade, Buxton, & Kelly, 1999). The difference in accuracy between the Image Incongruent and Text Based conditions is likely due to elaborate processing and integration of information within an inappropriate schema. In a sense, our incongruent images are signaling the information to be stored around unreliable or incoherent retrieval cues leading to inaccurate responses (see also Mayer, 2008). That is, seductive details prompt inappropriate encoding in memory thereby increasing the tendency for participants to use incidental processing.

The close correspondence between accuracy and confidence reported here stands in contrast to many of the findings in the self assessment literature (Dunning et al., 2004). Generally, the results from most studies cast doubt on our proficiency at self assessment, but many of these demonstrations have been based on people's generation of summative assessments of their own performance. In contrast, in the experiment reported here and in Rosenthal et al. (2010), participants rated their level of confidence in answering each question rather than assessing their overall performance. Our results are, therefore, consistent with the relatively accurate "micro-level" self-assessments obtained by Rosenthal et al. (2010).

6. Conclusions

The technology behind developing and broadcasting lectures through slideware is growing rapidly. Apple's iTunes University provides multimedia content on demand, and users can automatically synchronize this content with mobile devices or computers (Beldarrain, 2006). The trend in slideware development has moved away from the simple function of displaying text to more of a cinematic experience with sound effects, animated 3D builds and slide transitions, realistic shadow and reflection effects, and high definition aspect ratios. For students in the lecture hall of a tech savvy academic, the experience is not unlike an educational Pixar film. Although there are very little published data on the pedagogical benefit of these emerging technologies, it is clear that student satisfaction is favorable. Our results lend support to this finding as our participants rated the image rich conditions as more interesting than the text based condition. Indeed, it seems that images of any sort will capture people's interest, whether they are relevant or not. However, we also demonstrated that interest does not necessarily translate into better learning, as interesting-but-irrelevant images can actually lead to worse performance than no images at all.

References

- Apperson, J. M., Laws, E. L., & Scepanky, J. A. (2002). The impact of presentation graphics on students' experience in the classroom. *Computers & Education*, 47(1), 116–126.
- Baker, S. M., & Petty, R. E. (1994). Majority and minority influence: source-position imbalance as a determinant of message scrutiny. *Journal of Personality and Social Psychology*, 67(1), 5–19.
- Beldarrain, Y. (2006). Distance education trends: integrating new technologies to foster student interaction and collaboration. *Distance Education*, 27(2), 139–153.
- Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8(4), 293–332.
- Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, 3(3), 149–210.
- Dunning, D., Heath, C., & Suls, J. (2004). Flawed self-assessment: implications for health, education, and the workplace. *Psychological Science in the Public Interest*, 5(3), 69–106.
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53(1), 109–132.
- Garner, R., Brown, R., Sanders, S., & Menke, D. (1992). Seductive details and learning from text. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 239–254). Hillsdale, NJ: Lawrence Erlbaum.
- Garner, R., & Gillingham, M. G. (1991). Topic knowledge, cognitive interest, and text recall: a microanalysis. *The Journal of Experimental Education*, 59(4), 310–319.
- Goodstein, R. C. (1993). Category-based applications and extensions in advertising: motivating more extensive ad processing. *Journal of Consumer Research*, 20(June), 87–99.
- 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.
- Hastie, R., & Kumar, P. A. (1979). Person memory: personality traits as organizing principles in memory for behaviors. *Journal of Personality and Social Psychology*, 37(1), 25–38.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.
- Hidi, S., & Anderson, V. (1992). Situational interest and its impact on reading and expository writing. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 213–214). Hillsdale, NJ: Lawrence Erlbaum.
- Houston, M. J., Childers, T. L., & Heckler, S. E. (1987). Picture-word consistency and the elaborative processing of advertisements. *Journal of Marketing Research*, 24(4), 359–369.
- Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, 13(4), 351–371.
- Leahy, W., & Sweller, J. (2008). The imagination effect increases with an increased intrinsic cognitive load. *Applied Cognitive Psychology*, 22(2), 273–283.
- Lee, E. J., & Schumann, D. W. (2004). Explaining the special case of incongruity in advertising: combining classic theoretical approaches. *Marketing Theory*, 4(1), 59–90.
- Lehman, S., Schraw, G., McCrudden, M. T., & Hartley, K. (2007). Processing and recall of seductive details in scientific text. *Contemporary Educational Psychology*, 32(4), 569–587.
- Mayer, R. E. (2008). Applying the science of learning: evidence-based principles for the design of multimedia instruction. *American Psychologist*, 63(8), 760–769.

- Mayer, R. E., Griffith, E., Jurkowitz, I. T. N., & Rothman, D. (2008). Increased interestingness of extraneous details in a multimedia science presentation leads to decreased learning. *Journal of Experimental Psychology: Applied*, 14(4), 329–339.
- 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.
- 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., & Moreno, R. (1998). A split-attention effect in multimedia learning: evidence for dual processing systems in working memory. *Journal of Educational Psychology*, 90(2), 312–320.
- Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, 38(1), 43–52.
- Meyers-Levy, J., & Tybout, A. M. (1989). Schema congruity as a basis for product evaluation. *Journal of Consumer Research*, 16(June), 39–55.
- Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63(2), 81–97.
- Morett, L. M., Clegg, B. A., Blalock, L. D., & Mong, H. M. (2009). Applying multimedia learning theory to map learning and driving navigation. *Transportation Research Part F*, 12(1), 40–49.
- Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, 16(5), 519–533.
- Muller, D. A., Sharma, M. D., & Reimann, P. (2008). Raising cognitive load with linear multimedia to promote conceptual change. *Science Education*, 92(2), 278–295.
- Paivio, A. (2007). *Mind and its evolution: A dual coding theoretical approach*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Peracchio, L. A., & Tybout, A. M. (1996). The moderating role of prior knowledge in schema-based product evaluation. *Journal of Consumer Research*, 23(3), 177–191.
- Pintrich, P. R., & Schunk, D. H. (1996). *Motivation in education*. Englewood Cliffs, NJ: Merrill/Prentice Hall.
- Rosenthal, G. T., Barlow, S., McKnight, R. R., Price, A. W., Boudreaux, M., & Rachal, K. C. (2010). Do students know if they answered particular questions correctly on a psychology exam? *Journal of Instructional Psychology*, 37(1), 57–62.
- Schraw, G., & Lehman, S. (2001). Situational interest: a review of the literature and directions for future research. *Educational Psychology Review*, 13(1), 23–52.
- Snodgrass, J. G., Wasser, B., Finklestein, M., & Goldberg, L. B. (1974). On the fate of visual and verbal memory codes for pictures and words: evidence for a dual coding mechanism in recognition memory. *Journal of Verbal Learning and Verbal Behavior*, 13(1), 27–37.
- Srull, T. K. (1981). Person memory: some tests of associative storage and retrieval models. *Journal of Experimental Psychology: Applied*, 7(6), 440–463.
- Srull, T. K., Lichtenstein, M., & Rothbart, M. (1985). Associative storage and retrieval processes in person memory. *Journal of Experimental Psychology: Applied*, 11(2), 316–345.
- Stayman, D. M., Alden, D. L., & Smith, K. H. (1992). Some effects of schematic processing on consumer expectations and disconfirmation judgments. *Journal of Consumer Research*, 19(2), 240–255.
- Sujan, M. (1985). Consumer knowledge: effects on evaluation strategies mediating consumer judgments. *Journal of Consumer Research*, 12(June), 31–46.
- Szabo, A., & Hastings, N. (2000). Using IT in the undergraduate classroom. Should we replace the blackboard with PowerPoint? *Computers and Education*, 35(3), 175–187.
- Wade, S. E., Buxton, W. M., & Kelly, M. (1999). Using think-alouds to examine reader-text interest. *Reading Research Quarterly*, 34(2), 194–213.
- Yaghoub, M. S., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology*, 87(2), 319–334.