University of Memphis

University of Memphis Digital Commons

Electronic Theses and Dissertations

7-22-2015

Attentional Guidance and Media Presentation during Explicit Instruction

Whitney Layne Cade

Follow this and additional works at: https://digitalcommons.memphis.edu/etd

Recommended Citation

Cade, Whitney Layne, "Attentional Guidance and Media Presentation during Explicit Instruction" (2015). *Electronic Theses and Dissertations*. 1228.
https://digitalcommons.memphis.edu/etd/1228

This Dissertation is brought to you for free and open access by University of Memphis Digital Commons. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of University of Memphis Digital Commons. For more information, please contact khggerty@memphis.edu.

ATTENTIONAL GUIDANCE AND MEDIA PRESENTATION DURING EXPLICIT INSTRUCTION

by

Whitney Layne Cade

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Major: Psychology

The University of Memphis

August 2015

Acknowledgements

This dissertation would like to acknowledge the supportive environment offered by the Institute for Intelligent Systems at the University of Memphis. This work was also supported by the National Science Foundation (IIS 1352207).

I would like to acknowledge my advisor, Dr. Andrew Olney, for his dedication over the years. He always supported my work, took my ideas seriously, and shaped me into the professional I am today. The rest of my committee, Dr. Art Graesser, Dr. Mark Conley, and Dr. Natalie Person, served on nearly all of my milestone committees and always had insightful comments, an editorial eye, and good jokes, and for that I thank you.

I would also like to recognize my real family and my IIS family, who supported me through the process of getting my degrees. Without you, none of this would be possible and it certainly would have been less fun.

The biggest thanks goes to Patrick Hays, who kept me motivated until the end, listened to all of my data woes and triumphs, and stopped me from taking any of it too seriously. I made it through graduate school only because of you, and there are no words to express how much you have meant to me.

Abstract

Cade, Whitney Layne. Ph.D. The University of Memphis. August 2015. Attentional Guidance and Media Presentation during Explicit Instruction. Major Professor: Andrew Olney, Ph.D.

While much is known about how certain types of images influence learning in multimedia educational systems, comparatively little work has been done on how different image types compare to each other in terms of the types of knowledge conveyed and transfer of knowledge. Two popular types of media found in many multimedia environments, pictures and concept maps, are capable of blending verbal information (such as in picture labels or node/link labels) and visual information (such as structural information) into a single image, which may result in increased exposure to vocabulary (improving learning) or may create split attention (decreasing learning). Both types can also be presented using animation techniques, although questions remain as to whether animation always improves learning in different kinds of media. This study explores media differences and animation techniques in two experiments, both of which utilize Khan Academy lessons as the basis for the multimedia presentation. In the first experiment, a 2x2 between-subjects design was utilized to examine different media types (labeled pictures vs. concept maps) and animation (animated vs. static). The results of this study indicate that animation improves relational knowledge and free recall scores, but an animation x media type interaction indicates that animated pictures are not very effective for conveying conceptual knowledge. In Study 2, a 2x2 between-subjects experiment dove deeper into the function of "labels" by examining how animation (animated vs. static) and labels (present vs. absent) interact, as both may be attention directing devices. It was found that animation and prior knowledge both had consistent

effects on learning, where those with high prior knowledge did not gain as much from viewing an animated presentation as those with low prior knowledge did, but labels had minimal effects on learning. In all, research indicates that different media should be used depending on the educational goals, animation may be particularly helpful for low prior knowledge students, and labels are not necessarily helpful for learning when the same information is presented orally.

Preface

Chapter 2 (Study 1) of this work was published in the Intelligent Tutoring Systems conference proceedings in 2014. The dissertation author was the main author of this work.

Table of Contents

List of Figures

Attentional Guidance and Media Presentation during Explicit Instruction

As online learning expands and educational software like intelligent tutoring systems (ITSs) becomes more and more ubiquitous, the need for interesting, diverse, and effective multimedia has grown. Text alone is no longer considered sufficient for learning environments (Mayer, 2009), and even traditional textbook learning, replete with static pictures and diagrams, has been overshadowed by modern multimedia options, such as videos and 3D models. However, understanding of the mechanisms and processes of multimedia comprehension has lagged behind technological innovation. As a result, researchers know little about how features inherent to the media, learner, and lesson interact to result in learning. Thus, with an increasing reliance on visual-heavy technology for learning, the factors and parameters that impact learning in multimedia environments need further exploration.

Multimedia Learning

In order to understand why multimedia learning is so popular (as, theoretically, it has been present in human culture for as long as we have taught each other with both language and visual aids), we must first understand how it operates. One theory proposed by Mayer (2005), the cognitive theory of multimedia learning, has been used to explain how students understand static multimedia in the context of language. It addresses language and visuals being co-presented in some respect (whether it is text and a picture, spoken words and diagrams, etc.) and taken in by the sensory organs. Then, representations from each modality (verbal and visual) are formed, integrated with each other for successful learning, and also integrated into and/or influenced by prior knowledge (Mayer, 2005). This theory has three underlying assumptions: 1) visual and verbal information can be received and processed separately, 2) there is a limited capacity of information humans can handle in each channel, and 3) humans engage in learning by actively attending to, organizing, and integrating information.

These limitations provide clues as to how information should be presented to students in order to optimize learning. For instance, failure to acknowledge the second limitation with fast or dense information presentation can force the student to use up extra attentional stores in an attempt to compensate for their limited capacity, or may even result in missed information as students attend to other, distracting information. This extra processing is typically referred to as extraneous load, which can be avoided by designing material presentation in a way that acknowledges the limitations of human processing (Sweller & Chandler, 1994). Another way to avoid violating these assumptions is to purposefully circumvent the "split attention effect", which applies to the first and second limitations. Split attention can occur when two or more learning components are presented in a single modality (e.g., text and a picture, two streams of audio, several relevant pictures), thus causing the student to switch their attention back and forth between components within a channel in order to integrate the information contained in each component (Sweller, 1988). Not only is this another example of creating extraneous load, but it violates assumptions of the cognitive theory of multimedia learning by having learners process too much information in a single modality. This violation is therefore detrimental to learning. For this reason, multimedia learning using technology will more often use audio for the verbal channel, such as with spoken narration, and allow images to occupy the visual channel. Although the integration of written text and pictures can be advantageous to learning (Schnotz, 2005), it has been demonstrated that audio narration and images can produce more learning than text and pictures when working in a learning interface (Moreno & Mayer, 2002).

In everyday practice, the cognitive theory of multimedia learning can be difficult to apply because, while the theory proposes general guidelines to follow in order to maximize understanding, no strict boundaries exist to know when these assumptions have been violated. A student failing to learn is not necessarily evidence of a violated assumption, and even when underlying assumptions are violated, it is difficult to ascertain when the violation occurred. For instance, is it the case that any text that is present on an image, such as labels on pictures or concept maps, creates split attention? On the one hand, text and pictures have been found to be suboptimal when compared to narration and pictures (Moreno & Mayer, 2002), but others have found that concept maps (structures of labeled nodes and links) are an effective means of conveying information when combined with narration (Nesbit & Adesope, 2006). Labeled pictures are also a popular multimedia presentation tool in presenting science material, and can lead to learning (e.g., Cromley, Synder-Hogan, & Luciw-Dubas, 2010). Therefore, it may be the case that understanding split attention and the visual modality is not as simple as, "Are text and pictures copresent in the visual modality?", but instead, may be a question of degree of compresence.

In fact, it may be useful to think about visual stimuli in terms of various continua. For instance, one such continuum may plot how "textual" an image is, with one end being pure text (such as in a book or word cloud) and the other being pure picture (such as a photograph). Nearly all image types that could be used in a multimedia environment falls somewhere along this spectrum, and explicitly plotting how much text a multimedia image has may help keep track of where learners begin to encounter problems, such as with split attention. Undoubtedly, the “problem point” on the continuum changes under various conditions, such as student prior knowledge, working memory, and verbal ability, but thinking along the lines of a continuum may be a useful quantification in studies that wish to examine split attention, text-picture integration, labeled images/diagrams, and other word-picture structures. This may make cross-study comparisons more accurate, as even two studies that use the same type of media may fall on different points on the continuum. Figure 1 below demonstrates where a few types of media may generally be plotted, although specific media may move around on the spectrum.

A similar continuum may be constructed for another popular multimedia presentation technique: image animation. Although the cognitive theory of multimedia learning considers animation to be a type of visual input rather than a treatment of visual input (Mayer, 2005), nearly every image can have some level of animation, ranging from completely static to pixel-by-pixel changes (such as what occurs in animations that show the process of drawing). These two extremes could be conceived of as endpoints on a continuum, with various types of animated styles falling between these two points. Thinking of animation as this continuum may help researchers track when animation creates too much cognitive load by exceeding a student’s limited capacity of information or creating split attention. Just as with the text-picture continuum, however, that point is likely to change given various features of the domain and student, but also just as before, the continuum may also make it easier for researchers to compare studies using animation, as all animation techniques are not the same. See Figure 1 for a plotting of some basic animation styles.

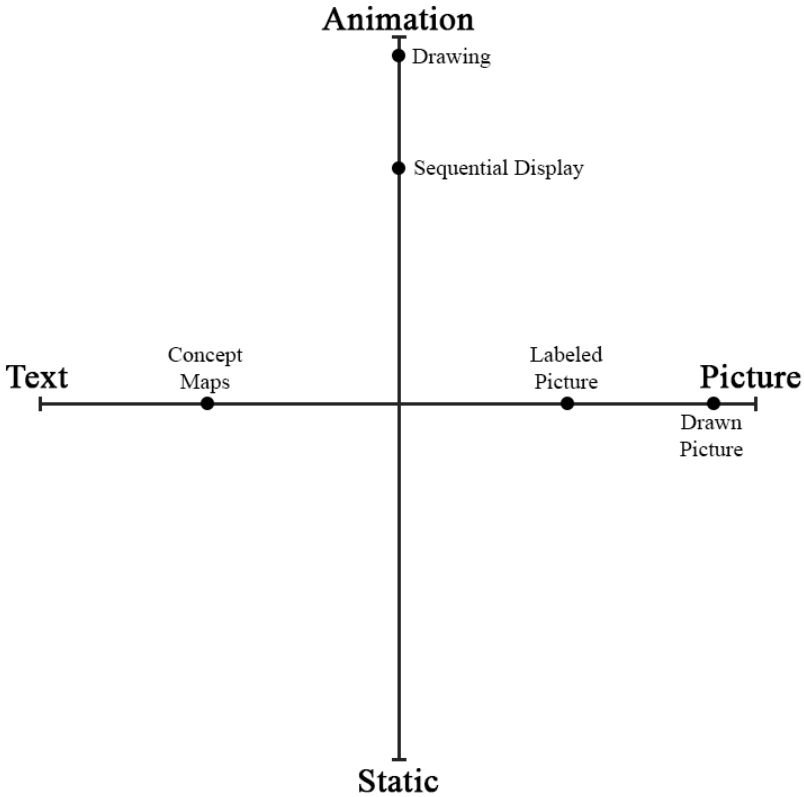


Figure 1. A plane depicting the text-picture and static-animation continua, with various examples of each type plotted. The combination of different examples would fall into the quadrants.

Although an infinite number of multimedia image properties exist which could be considered separate continua, only two have been presented here, as each represents some of the top concerns of the multimedia literature. Much work has been done on text-picture integration (Schnotz, 2005) and conversational text in image-heavy learning environments (Moreno & Mayer, 2002), but still left to consider is how images with text and pictures combined, such as in concept maps and labeled pictures/diagrams, function in a multimedia environment. While some studies have begun to investigate their effectiveness compared to more traditional media such as text (e.g., Ainsworth & Loizou, 2003; Nesbit & Adesope, 2006), little is known about how these forms of media compare to each other and other forms of “wordless” media. Therefore, one goal of this work is the consider how media that span various points of the text-picture continuum, such as wordless pictures, labeled pictures, and concept maps compare to each other in terms of effectiveness and/or distraction created from split attention. From this starting point, other researchers could compare their media to these and other points in the continuum so that a clearer picture arises of how text-picture images function in multimedia environments.

Another major concern of the multimedia literature has been when and how to animate these images. Although a fair amount of work has been dedicated to whether or not animation is effective (see Höffler & Leutner, 2007, a metanalysis on animation), investigations into what kinds of animation may be effective and what kinds may provide too much extraneous cognitive load are lagging behind. Therefore, it is the goal of this work to also investigate how a few lesser-studied animation styles, such as real-time drawing and sequential display (a block-by-block display method discussed more below), compare to their static counterparts. While there are a number of other factors that may affect the efficacy of animation, some of which are student, lesson, and domain features, this work attempts to start investigations into differing levels of animation, although these other factors will have to be considered in any between-study comparisons.

By combining these two orthogonal continua, a plane of potential research questions arises. Because there are an infinite number of points along each continuum, only a selection of popular media types and animation styles will be under investigation here. This work is meant to represent a starting point to such types of experiments (which may or may not also consider other continua), which is designed to guide future research towards parts of the space where interesting effects may be found. Therefore this work will be examining how various types of text-picture media, such as wordless pictures, labeled pictures, and concept maps, function in conjunction with the presence of absence of certain levels of animation. By considering these two factors together, a more nuanced view of how and when to use certain multimedia effectively will arise. The following sections address some of the issues germane to each of these topics in multimedia, before talking more about the specific investigation here.

Animations

One question that has received much attention when examining multimedia presentation for learning environments is this: should the image include *animation*? Lowe and Schnotz (2014) define animation as the "product of deliberate construction processes such as drawing" (p. 516), distinguishing it from captured video. The motivation behind desiring animated images is twofold: first, the animation may contain information not easily conveyed through text or static pictures, such as motion, and second, it serves as an attention-directing device, which may be particularly useful if the animation is being accompanied by narration (Lowe & Schnotz, 2014; Mayer & Sims, 1994).

However, it is not so easy to determine if animation should be used or not, as using a rule such as "animate if motion is involved" may not be helpful. As Lowe and Schnotz (2014) point out, motion can be displayed on a static image through arrows or dotted lines demonstrating an "after" position, while even stationary objects with many parts can be animated by "revealing" each part. Additionally, animation may also have costs. While animation may have a directing function for attention, it may be at the cost of extended exposure time to onscreen elements depending on the animation style. For example, in a style that uses a “slow reveal” animation method (more technically called “sequential display”, which will be discussed in more depth later), elements are absent from the screen and then added incrementally as the lesson progresses. In a static image of the same content, however, students would be exposed to all of these elements for the entire lesson, giving them more time with each element. Likewise, animation, such as might be found in a video, may only display key moments for a few seconds, while a static image can be inspected numerous times when the student needs it (Lowe & Schnotz, 2014). This would indicate that there may be certain conditions under which animation may not be optimal.

While there have been many studies that seek to find if animation is beneficial for learning (Höffler & Leutner, 2007), the results have been inevitably varied. For instance, the subject matter domain can greatly change the outcome, as animations have been more successful for Chemistry (d = 0.75) and military applications (d = 1.21) than they have for Biology (d = 0.13) when compared to static pictures. Likewise, the type of knowledge being queried may also matter, as animation may boost declarative knowledge (d = 0.44) and procedural motor knowledge (d = 1.06), but does not seem to be as beneficial for problem solving (d = 0.24). This suggests that there are a number of parameters worth considering when choosing to include animation, and there is no simple rule about when to animate. Others argue that even more constraints exist, as there are a number of perceptual, cognitive, and knowledge-based factors about the learner that must also be taken into consideration (Lowe & Schnotz, 2014).

Although there are many types of media to select from, much of the animation literature has focused on how to animate *pictures*, which are a very common media selection choice to accompany text or narration. Visual aids such as pictures naturally occur in human-to-human tutoring sessions, where the tutor may use pictures out of a book or sketch one herself to demonstrate some concept for the student (Williams, Williams, Volgas, Yuan, & Person, 2010). In fact, many ITSs include pictorial references to help ground their lessons, such as AutoTutor (Graesser et al., 2004), Guru (Olney et al., 2012), and MetaTutor (Azevedo & Witherspoon, 2009). These systems in particular may rely on pictures more than other systems due to their domain, as they are more science-centered and therefore often deal with abstract, microscopic, or complicated subjects that are best explained pictorially rather than verbally, such as the number of edges and faces of a cube. When dealing with such complex pictures, it has been found that highlighting or directing attention to relevant parts of the picture can be critical to understanding (Jeung, Chandler, & Sweller, 1997). Many researchers, therefore, find themselves following the “attention-guiding principle”, where, in order to optimize student understanding, attention must be visually guided to significant parts of the picture at the correct time (Bétrancourt, 2005).

Consequently, much work has been done on how to guide student attention to relevant parts of a picture, although the particular technique that should be used is still debated. These techniques fall all along the static-animation spectrum proposed in Figure 1, meaning that they have varying levels of “action”. Some researchers have explored pointing using an embodied pedagogical agent, which could be considered a method of animating a picture without changing the appearance of the picture. For instance, Craig, Gholson, and Driscoll (2002) used agent pointing combined with picture animation to explore the knowledge transfer after learning about lightning. However, they did not find that pointing enhanced student knowledge, although results of a later experiment demonstrated that pointing can be beneficial when the agent is more humanlike in appearance and the picture is not also being animated (Twyford & Craig, 2013). Others have used an animation style that Höffler and Leutner (2007) call “signaling cues”, where arrows or highlighting are employed to draw attention to an otherwise static image, a fairly low amount of action on the static-animation spectrum. Although their metanalysis found an overall disadvantage for signaling cues when compared to no cues ($d = 0.33$ vs. $d = 0.47$), some individual studies have found that cues such as flashing, highlighting, and pointing with hands and arrows were helpful for learning (Atkinson, Lin, & Harrison, 2009; Jeung, Chandler, & Sweller, 1997).

One of the more significant animation methods for pictures that has been discussed in the last decade is sequential display. In sequential display, a static image is hidden at the beginning of the lesson, and then parts of the image are suddenly revealed as they become relevant to the topic (usually being discussed through audio narration). In some ways, sequential display mimics the act of drawing or sketching a picture in a tutoring session without the fluid motion of line drawing; as the tutor wishes to discuss some component, it is added to the drawing. This may focus the student’s attention on the newly added part of the image while removing the distractions of future parts of the picture which may cause attentional interference in the visual channel. Along the static-animation spectrum, sequential display would fall short of drawing a picture while still having a large impact on the way the picture appears at any given moment. Lowe and Schnotz (2014) refer to this as a “build animation”, which is more appropriate when the instructional purpose of the lesson is to convey structure rather than process.

While sequential display has been used in at least one ITS (Guru; Olney et al., 2012), its roots appear in the document explanation literature. In one study, learners were shown the human brain and taught its parts and functions. Those who saw a sequentially displayed brain while listening to narration showed greater retention, particularly for function information, than those who experienced a static picture and narration (Jamet, Gavota, & Quaireau, 2008). Bétrancourt, Dillenbourg, and Montarnal (2003) found that when students were given a sequentially displayed presentation about finances, they outperformed those who saw a static presentation. More interestingly, however, those who were allowed to choose how the sequential display progressed trailed those who could not choose how information was organized or displayed (although often not by a large enough margin to be considered significant), suggesting that sequential display does not need to be under the control of the student to be beneficial, and in fact, may be slightly detrimental to their learning and knowledge transfer abilities. Sequential display may even affect how a student recalls information; students shown sequentially displayed town maps which used either a spatial or thematic organization recalled the maps in the order they were displayed (Bétrancourt, Bisseret, & Faure, 2001).

Sequential display may also be a good choice of animation style because of its ability to be closely aligned to narration. Because any attention-directing gesture or motion may be considered vague without the support of language to specify what the listener should be noticing (Wittgenstein, 1971), a good unity must exist between the moment of animation and the narration. Sequential display involves sudden changes to the picture due to objects being added, and can therefore be precisely timed to fit with what is being said. This may produce improved learning due to what is called the “contiguity effect” (Mayer & Sims, 1994), where the picture and verbal information (narration or text) must be aligned in time in order to facilitate the integration of the visual and verbal representations created by the student. If this temporal link is broken, it may be difficult for students to integrate what they have heard in a narration with a picture (even an animated one) they see later on.

In sum, while there is an extensive amount of research on how to animate a picture and whether these animations add to learning, more work needs to be done on when the context calls for animation, such as when the student has low prior knowledge, the content affords the demonstration of motion, and the domain of the content. The majority of previous research has focused on multimedia to the exclusion of properties of the content and the learner. Knowing more about when certain animations work and when they do not may be the next phase of picture animation research.

However, pictures have not been the only type of media of interest to the educational media community. These systems have included other forms of multimedia, such as concept maps, in order to enhance student learning. Concept maps, sometimes called semantic maps (Heimlach & Pittelman, 1986), graphic organizers (Stull & Mayer, 2007), or knowledge maps (O’Donnell, Dansereau, & Hall, 2002), are structures which consist of nodes and links, where nodes express some concept (e.g., “sun” and “star”) and links run between the nodes and are labeled to specify the relationship between the nodes (e.g., “is a”). Because their pictorial aspects are limited to their informational structure, they fall closer to the text side of the text-picture spectrum than a labeled image.

Concept maps have been proposed as an effective educational tool for classroom use (Novak, 1991, 1998; Novak & Cañas, 2008; Novak & Musanda, 1991), and have been incorporated into educational software (Cañas et al., 2001; Stull & Mayer, 2007). Two ITSs, Betty's Brain (Leelawong & Biswas, 2008) and Guru (Olney et al., 2012), have used interactive concept maps as learning activities. In Betty's Brain, students construct concept maps based on text readings in order to create a knowledge structure for Betty, a virtual student whose knowledge (the concept map structure) is evaluated through test taking. In Guru, partially-completed concept maps are filled in by the student by selecting labels from among several options. Students are not allowed to continue if they have selected the wrong option in any link or node, forcing students to correct flawed knowledge.

While these systems employ concept maps as interactive media, it has been shown that concept maps can be helpful aids during direct instruction, utilized in much the same way as pictures are used as visual aids (Adesope & Nesbit, 2013; Blankenship & Dansereau, 2000; Nesbit & Adesope, 2006; Stull & Mayer, 2007). In a meta-analysis, studying completed concept maps was shown to produce an average learning effect size of $M = 0.373$ when compared to controls (usually, studying text; Nesbit & Adesope, 2006). In fact, some studies have found that it is actually more advantageous to show students completed concept maps rather than allow them to fill them out themselves (Stull & Mayer, 2007). The authors state that, though it violates constructionist theories of knowledge which require students to take a hands-on role in their learning, it may be the case that the increased activity of concept mapping creates cognitive load that distracts the students from absorbing the material. Clark and Mayer (2008) indicate that there are times when studying completed concept maps is more beneficial for learning than allowing students to create their own, such as when expert created maps will be better than what they can create, when lectures are as good as (or better than) group discussions, and when worked examples are better than actual practice.

Because presenting a student with a completed concept map has been shown to be an effective teaching tool, many researchers have explored using them to support lectures, or direct instruction. Just as pictures have formed the backdrop for narrated lessons, so too can concept maps. This means that concept maps, like pictures, can fall prey to the same multimedia learning pitfalls and pressures, particularly when it comes to student attention. Researchers have therefore looked into finding ways to animate concept maps in order to focus student attention on relevant portions of the map during the course of a narration. One of the earliest animated concept map studies examined whether animated or static concept maps or text improved learning (Blankenship & Dansereau, 2000). They found that animated concept maps improved memory for macrostructure recall (memory of large areas of the concept map rather than individual nodes and links) when tested 48 hours after viewing the map. Nesbit and Adesope (2011) also found that an animated concept map, when compared to text that does or does not align with the narration, is superior to both on free recall measures. However, it has also been found that animated concept maps are not better than static concept maps, although concept maps did outperform text (Adesope & Nesbit, 2013). The authors of this study suggest that narration can wash out the effects of animation for concept maps, although it may be the case that this is only true for narration that is closely tied to the text of the concept map. Despite a few studies that did not find a boost for learning from animated concept maps, a meta-analysis found a $g = 0.39$ effect size for studying a static concept map, while animated concept maps have a $g = 0.739$ effect size (Nesbit & Adesope, 2006), indicating that presenting students with an animated concept map as part of their multimedia learning experience can be highly beneficial for increasing their knowledge. As with animated images, animated concept maps produce inconsistent results across the literature, though they generally improve learning when present.

Despite studies examining the effectiveness of different kinds of media, there seem to be no studies which indicate how different types of media work to promote learning. No head-to-head comparisons of media types, animated or static, exist in the current literature. This is surprising given how easily they could be compared; pictures, as discussed above, can be animated using a number of techniques found along the static-animation spectrum proposed in Figure 1, including one called sequential display, which has been shown to be successful in improving student comprehension. Concept maps, on the other hand, are almost always described as being animated using sequential display. Nodes, links, or whole propositions (the node-link-node structure) are added to the display as the facts are mentioned in the narration. It is also clear from the animated concept map literature that, with some exceptions, animating concept maps in this way can be an effective way of promoting learning. However, given certain circumstances such as a Biology domain where animations historically produce weak effects, should one choose to include a picture or a concept map as their multimedia? This is one question this work seeks to address.

While animation has been the focus of much research, it is just one variable worth examining in the multimedia literature. As mentioned previously, concept maps and pictures may both have successful learning outcomes with animation as well as little effects from animation under certain conditions. Therefore, it may be worth examining how animation and text that is integrated into an image, such as in labeled images, may affect learning, a relationship proposed in Figure 1.

Labeled Images

While animated images are one way of directing and focusing student attention on a certain part of a multimedia image, there may be other ways to accomplish this. Animations, as mentioned previously, may have some downsides. According to Stull and Mayer (2007), animations may overwhelm the student, creating a cognitive load that cannot be handled, resulting in extraneous load and suboptimal learning. The animation literature is also plagued with questions of how and when one should animate an image given parameters such as the domain, lesson type, learner's prior knowledge, and structure of the lesson itself (Höffler & Leutner, 2007). Concept maps have less information to process due to their stripped down nature; they consist of two sets of shapes (typically an oval or rectangle for the nodes and lines or arrows for the links) and words, rather than shape and spatial information. In fact, while the action of animation may draw a student's attention to an area of the picture, labels may serve as a confirmation that the visual search for the relevant part of the picture can be terminated due to the potentially close match between what is being spoken and the words on the image.

Although this may be one attention-focusing asset of concept maps, pictures are also capable of containing word cues, such as in the case of a labeled picture; these kinds of images would fall between concept maps and unlabeled pictures in the text-picture continuum, as they contain both kinds of information, but typically would have fewer words visible than a concept map. Labeled pictures, also called a picture glossary

(Moline, 2011), are a type of diagram where parts of a complex picture are identified with a vocabulary term. Labeled pictures are typically depicted in a style more simplified than a photograph in order to remove extraneous detail. While labeling pictures may be a common classroom activity (Moline, 2011), researchers have also used labeled pictures (or, as is more commonly found in this literature, diagrams) to examine how pictures and text may improve learning (e.g., Berthold & Renkl, 2009; Cromley, Synder-Hogan, & Luciw-Dubas, 2010; Grosse & Renkl, 2006; Kragten, Admiraal, & Rijlaarsdam, 2012; Schnotz & Bannert, 2003). Several studies have found that diagrams are difficult for students to process, and thus, are sometimes not fully utilized. For instance, students may not process the diagram correctly (such as not following directional cues available in some diagrams), they may be attracted to unimportant details, and they may not understand basic diagram conventions (Cromley et al., 2013). In fact, the last issue may be the most critical to the success of labels – if students do not know that they should read all of the labels, and one study indicates that students may only read 36% of the labels on a picture (Cromley et al., 2010), then they may not be able to extract the full pedagogical benefit of labels. Other studies have indicated that having labels in diagrams may be distracting. The modality principle (Moreno & Mayer, 2002) suggests that printed text and pictures together may induce split attention. On the other hand, Sweller, Ayres, and Kalyuga (2011) state that text that is well-integrated into a picture may produce only germane load. Moreno and Valdez (2005) found that presenting a series of pictures without accompanying text lowered learning scores, even when compared to a text-alone condition. Additionally, much of the diagram literature, which does not explicitly study labels but often includes them, suggests that diagrams can provoke higher cognitive processes than mere text, such as drawing more inferences and using higher level strategies (Cromley et al., 2010).

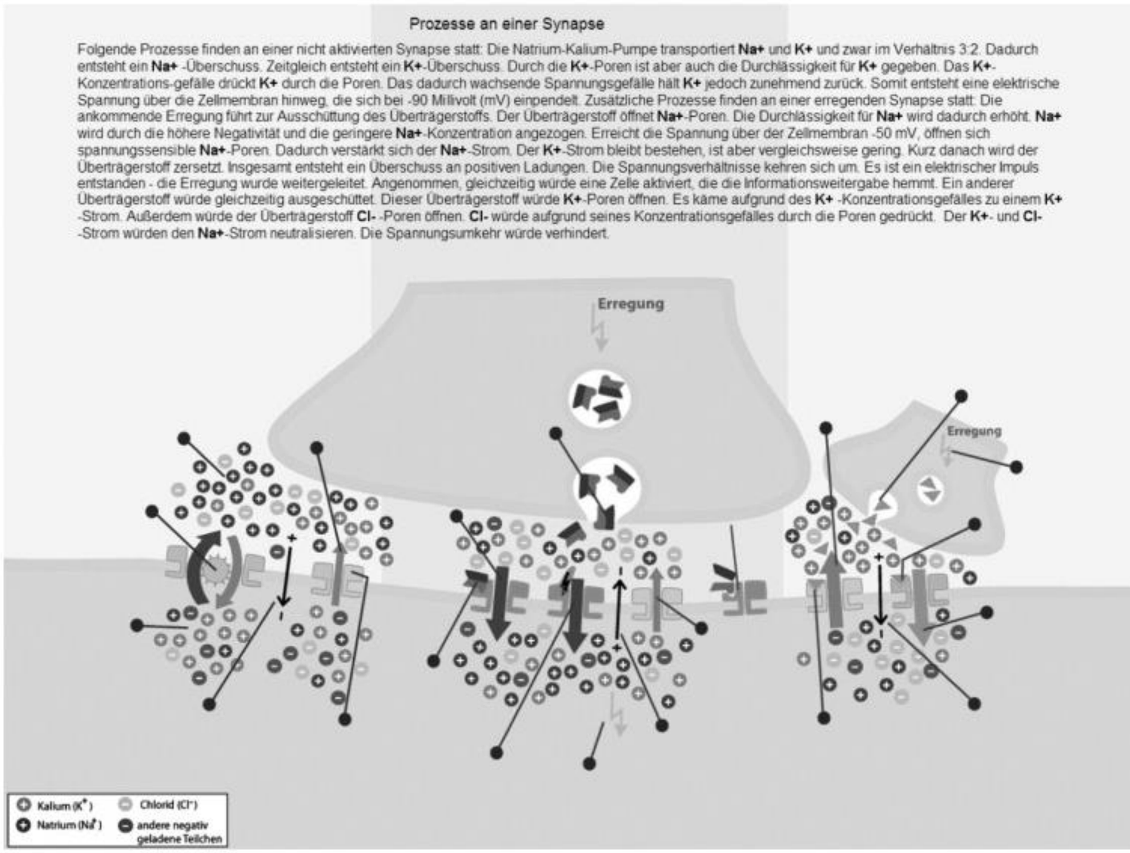


Figure 2. A labeled image with continuous text. Adapted from “What contributes to the split-attention effect? The role of text segmentation, picture labelling, and spatial proximity” by M. Florax and R. Ploetzner, 2010, *Learning and Instruction*, *10*, p. 221. Copyright 2010 by Elsevier. Reprinted with permission.

One study has specifically examined how the presence or absence of labels may affect learning outcomes. Pairing labeled and unlabeled images with continuous or segmented text, the authors hypothesized that labeling would boost comprehension and retention of information, but only found marginal impact for the presence of labels (Florax & Ploetzner, 2009). However, these results may be due to the way they chose to label their picture, which depicted the synapse of nerves cells and how information is processed and transmitted through the nervous system. Figures 2 and 3 demonstrate how the authors operationalized “labels”, which seem more aligned with arrow signaling cues (Höffler & Leutner, 2007) than the labels one might more commonly see in a textbook or classroom. The authors even cite signaling cues as the theoretical basis behind why labels may be effective, and may have visually adapted more traditional labels in order to fit this literature and their hypotheses concerning text segmentation (as they likely did not want the text from labels competing with the target, explanatory text). This conceptualization, with numbered and unnumbered arrows and larger segments of text, would put them much further rightward on the text-picture spectrum than would typically be the case in a traditional labeled picture, and so it is difficult to say if their effects would hold true for traditional, shorter labels as well. Indeed, it is suspect to generalize from this study that traditional labels may provide a minor contribution to learning. As labels sometime represent the only visible text, word labels are likely to be treated differently than numbered or unnumbered arrows. For instance, it is possible that word labels still induce some split attention between the figure, label, and narration. Although Florax and Ploetzner (2010) assert that labels may contribute to overcoming split attention between text and pictures, this may not be true in other instances. Therefore, more work needs to be done to determine how the presence of traditional labels affects learning.

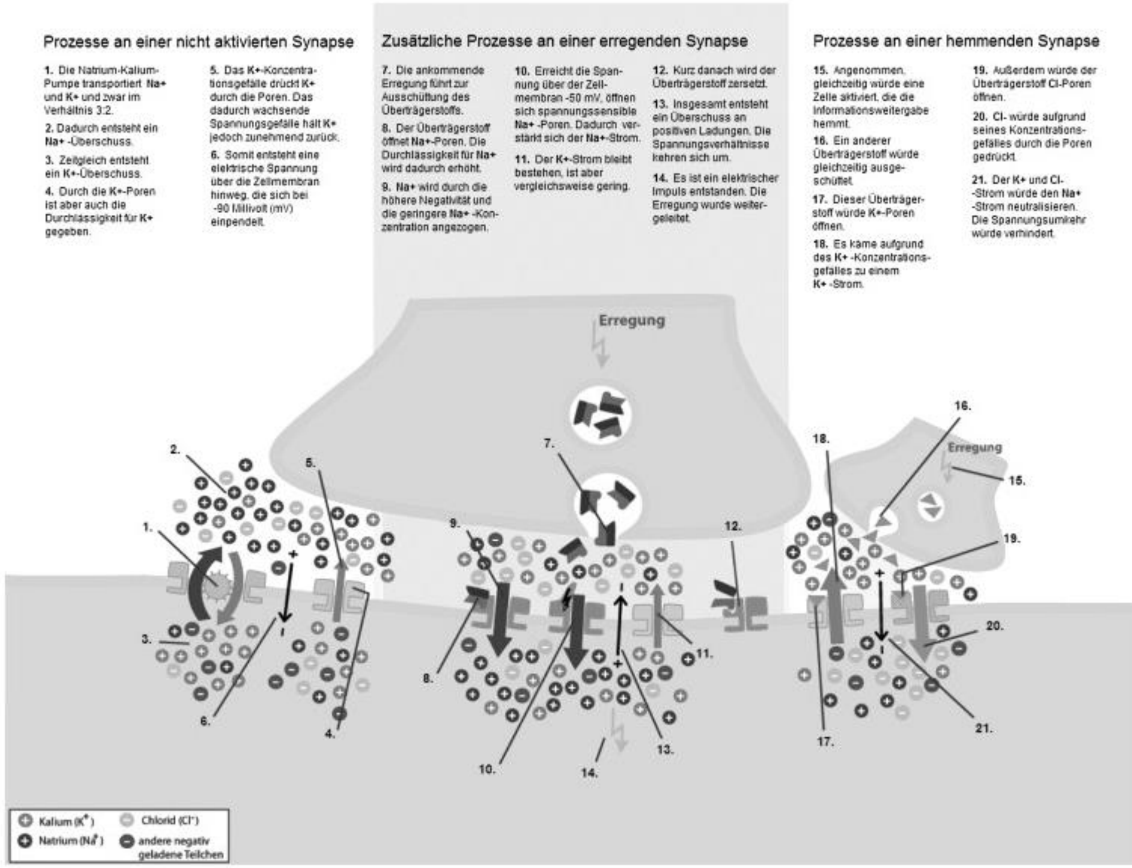


Figure 3. A labeled image with segmented text. Adapted from "What contributes to the split-attention effect? The role of text segmentation, picture labelling, and spatial proximity" by M. Florax and R. Ploetzner, 2010, *Learning and Instruction, 10,* p. 222. Copyright 2010 by Elsevier. Reprinted with permission.

Additionally, it may be worth examining how labels and picture animation interact. In perhaps the only other study to examine labeled versus unlabeled multimedia, Mayer and Gallini (1990) compared text presented alone, text with a static picture with labeled parts, text with a static picture with major actions labeled, and text with a dynamic picture that combined part and action labels (Figure 4). They found that only the dynamic picture with labels for processes and parts produced conceptual learning and creative problem solving. However, as seen in Figure 4, the "dynamic" picture in this study merely showed both the "on" and "off" states of the mechanism without including any actual animation. Because motion is only suggested through the comparisons between the two states, this provides only weak evidence that animation of some variety may be beneficial when combined with labels. Likewise, because this “on/off” state is also not displayed with the steps alone or parts alone labels, it is difficult to disambiguate the effects of the combined labels on learning from the implied-motion of the state comparisons.

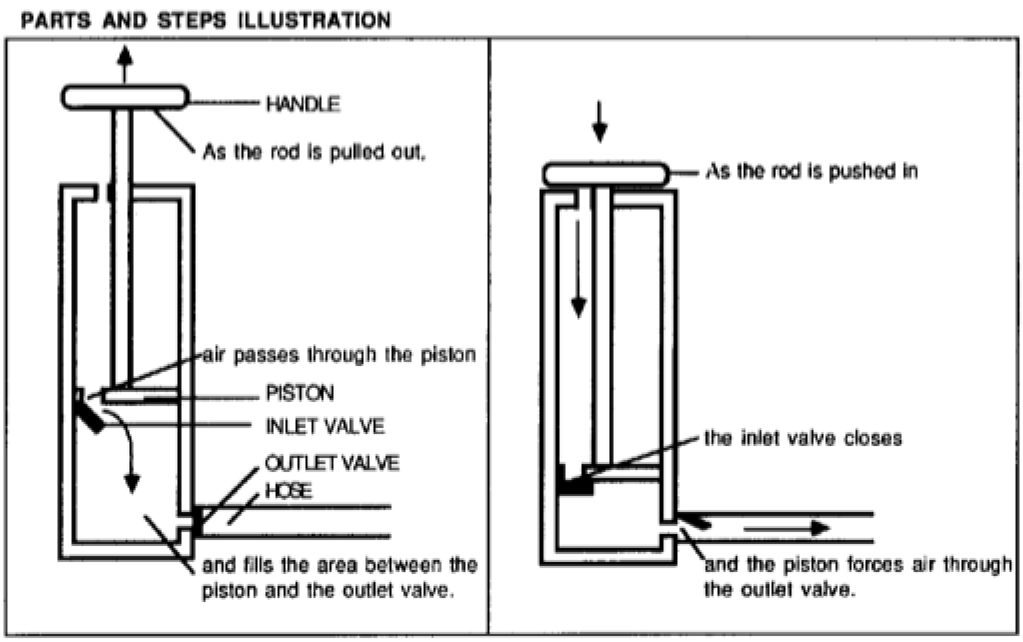


Figure 4. Illustration of the dynamic “parts and steps” labeled condition. Adapted from “When is an illustration worth ten thousand words?,” by R. E. Mayer and J. K. Gallini, 1990, *Journal of Educational Psychology, 82*(4), p. 717. Copyright 1990 by the American Psychological Association.

Based on this weak evidence, it still remains to be seen if animations and labels produce learning and if context may impact learning gains. Currently, no study exists that compares the effects of animation to the effects of labels. Additionally, if Mayer and Gallini (1990) are considered as showing evidence of how labels and animation may work together, then they only address one domain, mechanics, which depicts a single causal chain showing how the mechanism works. Other domains may have many parts which all act independently and serve different purposes, such as cell parts in Biology. This may change how easy it is to represent functions as labels and how difficult it may be for students to form a cohesive mental model from the lesson. The authors also indicate that the dynamic labeled condition benefitted only the low prior knowledge students, which could indicate that different multimedia presentations may affect learners differently, perhaps also in conjunction with other constraints, such as the domain.

It is also theoretically possible that labels are not helpful in multimedia learning. First and foremost, they may be highly redundant with the narration when combined with animation. For example, the narrator may say, “We have things called *mitochondria*” where *mitochondria* is the key term to be learned. At the same time, a mitochondrion may appear in the cell with the label “mitochondria” above it. There now exists redundancy between the spoken narration and the visual label. In previous studies, pairing text and narration in this way often produced less learning than having only narration and the picture, a phenomenon called the redundancy effect (Moreno & Mayer, 2002). They recommend leaving out text when narration and a multimedia image are present in order to reduce split attention between the image and the text, both of which would be occupying the visual modality (Chandler & Sweller, 1992).

Labels may also not be helpful if they are a redundant attention directing factor. Florax and Ploetzner (2009) regarded labels as signaling cues, a kind of attention directing device in the animating literature. However, Adesope and Nesbit (2013) found that animation may become redundant when there is a close pairing between the spoken narration and the words appearing on a screen in a concept map. Alternatively, it may be *labels* that are the redundant information stream rather than animation in the case of pictures, as signaling cues have been shown to be less effective than animation (Höffler & Leutner, 2007), and labels are less critical to a picture than they are to concept maps. Other studies have found that some signaling cues do not work when animation is present. For instance, Craig et al. (2002) found no effect for pointing when showing an animated picture of lightning. However, in a reversal of this, Cade (unpublished) found no effect for either pointing or animation in presentations on cellular functions, and hypothesized that the labels present on each picture may have provided sufficient attention directing such that the effects of pointing and animation were washed out. It is unclear whether labels are weak attention directing devices or if they have the power to wash out animation effects, as Adesope and Nesbit (2013) and Cade (unpublished) suggest.

On the other hand, labels may be beneficial to learning, as Florax and Ploetzner (2009) and the plethora of other multimedia studies that employ labels suggest. Labels may act as a visual way of segmenting a stream of narration. If a narrator is discussing mitochondria as one part of the cell and then switches and discusses another part, such as the Golgi apparatus, then there may exist some segmentation of the narration by topical switching, and labels may be a written but verbal representation of that switch. Mayer, Dow, and Mayer (2003) found that segmented text produced better learning than continuous text, and the same may be true for narration and labels. It remains to be seen if this audio segmentation happens in the absence of labels (if segmentation is occurring at all), which would mean that labels are not needed to flag when a new section is beginning. Labels paired with narration may also strengthen the contiguity effect (Mayer

& Sims, 1994) by providing a closer tie between the narration and the visual, since keywords may appear as labels in time with the narration. Therefore, contiguity would be formed by something changing on the visual in time with the narration, and also by echoing the key term of the narration on the visual. Additionally, the redundancy effect mentioned above may not be as impactful as it would seem on the surface, as the text being presented onscreen is a single word and not a sentence or paragraph, representing different levels of textual presence on the text-picture continuum, with labels possibly being so short that they may be treated almost symbolically. These signs would point to labels being beneficial for learning or less harmful than one may initially suspect.

Based on this review of the labeling literature, it is unclear if animation and labels separately contribute to directing the student's attention, if one is more powerful than the other, and how they work together when combined. To date, there has been no systematic study of how animations and labels interact, and so it is not known under what kind of conditions they may pair well together or become obsolete or harmful in the presence of the other. More carefully controlled studies of how labels function under certain constraints would contribute to the multimedia learning literature.

Present Research

While the image animation literature is fairly extensive, no unified rule of thumb concerning when to animate images has been found. Instead, researchers have discovered that the success of animations is impacted by certain conditions, such as the domain, type of knowledge being taught, and cognitive abilities of the student (Höffler & Leutner, 2007; Lowe & Schnotz, 2014). There may also be other parameters which are unaddressed by the literature. For instance, while work has been done separately on animating pictures and animating concept maps, no work has addressed how the two compare to each other. Additionally, while pictures have been a popular focus of the animated images literature, it is unclear how labels, a popular feature of pictures from science-based lessons, function in conjunction with animation, since both may serve an attention focusing purpose. The current research explores and compares various points on the plane created by the orthogonal intersection of two continua: the static-animation continuum and the text-picture continuum (Figure 1). Comparing a combination of points on this plane may help build a picture of how these two continua interact.

Study 1. Study 1 compares how animation affects two popular types of multimedia images: pictures and concept maps. While each type of media has its own literature, often comparing static and animated versions of the same image to text (Adesope & Nesbit, 2013; Blankenship & Dansereau, 2000; Höffler & Leutner, 2007), no comparisons exist between different types of media. Recalling the plane from Figure 1, this study looks at the more textual concept map, falling on the left side of the plane, and a labeled picture, falling to the right side of the plane. Additionally, each will be animated (using sequential display) and static, therefore falling into the top and bottom halves of the plane. This may give some sense of the strengths and weaknesses of these two "blended" forms of media, both of which consist of pictures and text but in varying degrees. Because pictures have a long-standing history in the multimedia literature, while concept maps are an up-and-coming alternative that have made their way into educational software already (e.g., Leelawong & Biswas, 2008; Olney et al., 2012), comparing these two forms of media may be a solid beginning to understanding how various media work given certain features of an educational environment.

While this study can only represent a subset of the conditions that need to be examined, there are some existing parameter hypotheses that this research addresses. First, this study will be conducted in the domain of Biology. According to Höffler and Leutner (2007), Biology is one domain in which the typical gains seen when using animation vanish. This may be due to the fact that Biology is itself a diverse field and may employ many kinds of diagrams and pictures depending on the topic, some of which may or may not be improved with animation. This study works within the topic of “parts of a cell”, a common introductory Biology lesson, to see if the lack of animation effect is true of even basic Biology lessons.

Second, this study includes the same narration paired with each condition, which includes animated and static versions of a picture and concept map. Adesope and Nesbit (2013) hypothesized that they did not see a difference between the animated and static concept map because the narration washed out the effects of the animation by overscaffolding the concept map. In other words, they believed that the connection between the words on the concept map and the words in the narration was so strong that the animation did not provide any additional attentional directing. This study directly tests this hypothesis by including narration in all conditions, including the static and animated concept map conditions. If this hypothesis is true, then no differences should be seen between the static and animated concept map conditions.

Finally, this study examines what kinds of knowledge each media type impacts by testing the students’ conceptual and relational knowledge. Real differences may arise between the two types of media based not on general “learning”, but rather on the types of knowledge the student can acquire from each media type. For instance, concept maps may provide relational knowledge because they make relationships explicit through their structure, while pictures and concept maps may both be equally good at promoting conceptual knowledge because both may contain concept-based labels. Höffler and Leutner's (2007) meta-analysis indicates that animation also impacts the type of knowledge learned. Procedural motor and declarative knowledge respond fairly strongly to animation, while problem solving knowledge does not. Embodied cognition may also predict that animation could create greater memory for relational knowledge due to its heavy emphasis on verbs (Mahon & Caramazza, 2008).

The goal of this study is to begin establishing the conditions under which different types of media produce optimal learning. Although this experiment only considers a small number of these conditions (e.g., Biology, narration, pictures vs. concept maps, relational and conceptual knowledge), it may be a step towards obtaining a more nuanced view of how differing media facilitates learning within certain environments.

Study 2. The purpose of Study 2 is to examine a popular but rarely scrutinized feature of picture media, one of the most common media found in educational environments: labels. Labels are included in numerous studies that look at how multimedia impacts learning (e.g., Berthold & Renkl, 2009; Cromley et al., 2010; Grosse & Renkl, 2006; Kragten, Admiraal, & Rijlaarsdam, 2012; Schnotz & Bannert, 2003), but are rarely the focus of the study – they are usually included in a study with the presumption that they are adding something to the learning experience or reflecting the ecologically valid experience of a student viewing a picture. It is likely that, like all things in multimedia learning, labels may provide some benefits under certain conditions, although those conditions are not known. There are two studies that explicitly investigated the effects of labels on learning. In the study which most directly examined the impact of labels, Florax and Ploetzner (2009) found that labels provided only a marginal advantage in comprehending and retaining information extracted from text and a picture. However, their unconventional treatment of labels, where labels consisted of only numbered referents to other text, may have weakened these results. Mayer and Gallini (1990) also looked somewhat indirectly at labels in conjunction with a series of static images that implied motion, and found that only the condition with the static image series and labels for parts and functions produce learning gains for comprehension and retention. However, there was no real animation condition (only a series of images), and no condition which included the static image series without labels, and so it is difficult to discern how animation and labels may actually work together. Therefore, the goal of Study 2 is to examine how the presence of absence of labels interacts with animation to affect different kinds of learning.

Like Study 1, Study 2 also operates under a number of constraints, some of which are identical to Study 1. This study also works in the domain of Biology, which could produce weaker animation effects (Höffler & Leutner, 2007). It is also potentially a domain in which labels could do uncommonly well, as Biology often discusses systems, cycles, objects, and concepts with many parts which have specialized names and functions. Relational and conceptual knowledge are also tested in this study, as labels may produce especially strong conceptual knowledge due to the extra exposure to the written concept words that narration cannot provide. Additionally, other tests have been added to this study to see if labels produce better performance on certain tasks, such as a near transfer labeling task or a task where parts are identified based on their function. If exposure to labels does not produce more learning on these sorts of tasks, it may be the case that there are few advantages to including labels, and simplifying a picture by removing them may be preferable as a manner of producing less cognitive load.

This study is particularly focused on how labels and animation work together to impact learning. While Study 1 compares a labeled picture to a concept map, both of which have varying degrees of text and picture parts, this study examines how the total absence of words on the image compares to a labeled image when animated, shifting the focus of this study to the right side of the animation-picture plane in Figure 1. On the one hand, it may be the case that labels are redundant with the narration (Moreno & Mayer, 2002), thus producing unnecessary split attention between the image shapes and the labels (Sweller & Chandler, 1992). On the other hand, labels could provide extra exposure to the keywords the student will see on tests, and that extra "time on task" may help them more accurately recognize terms, rather than trying to recall them from the narration. Additionally, labels may produce superior learning gains on near transfer tests that directly relate to labeling, such as having to label a cell or identify the cell part based on its function, both of which are the kinds of tasks students are expected to perform on state tests. Therefore, it may be that labels are not generally "beneficial", but instead prepare students for certain kinds of tasks.

This study represents a step towards understanding how labels operate. This may be especially important if the results of Study 1 do not generally favor picture representations, as Study 1 will have labels in both picture conditions. Because labels are so ubiquitous in the multimedia learning literature, it is important to begin understanding how they function and when they are most optimally used.

Khan Academy. Both studies in this work use a single Khan Academy lesson as the basis for the multimedia interventions. Khan Academy is a site that creates and hosts freely available video lessons on a variety of topics in math, science, economics and finance, arts and humanities, computing, and test prep. Nearly all videos are created in the same manner: the tutor records him- or herself talking about some topic, which is synchronously aligned with a tablet-based screen capture of hand drawn pictures and worked problems. Therefore, each video natively has audio narration and a visual media representation (Khan Academy, 2015). See Figure 5 below for an example screen capture from the lesson that will be used in both experiments, *Parts of a Cell*.

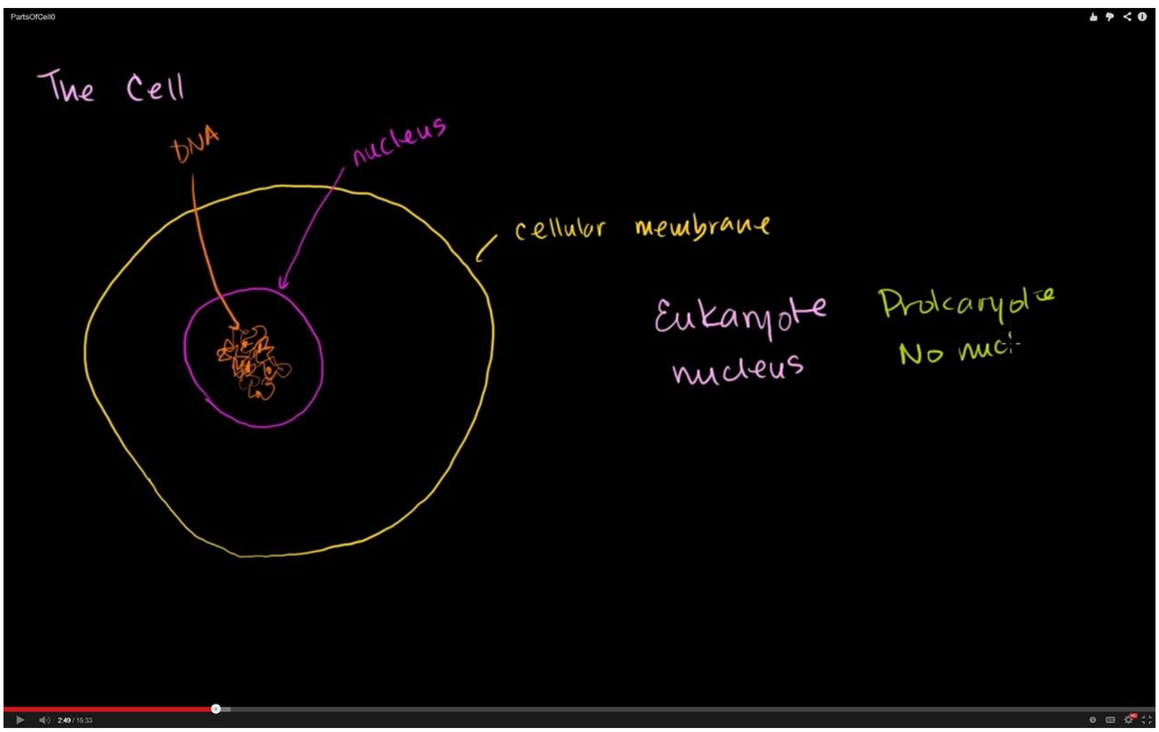


Figure 5. A screenshot taken from the Khan Academy lesson *Parts of a Cell*.

In these studies, the video from the Khan Academy lesson Parts of a Cell is used as the foundation for all manipulations. Changes were made to the visual portion of the video in accordance with the condition demands, such as removing labels from the video or replacing the hand drawn cell illustration with a concept map. The original narration, however, is preserved in all conditions in order to make the experience as close to the original lesson as possible.

The lessons created by Khan Academy are of interest to researchers for two reasons. First, it has become a highly popular form of freely available e-learning, and is therefore a good source of ecologically valid learning materials being used by real students every day. Second, the videos represent the kinds of media that researchers and software developers could easily produce for themselves with even a restrictive budget. All one needs to produce a video highly similar to a Khan Academy video is content knowledge, access to a tablet, and a modest ability to draw. Khan Academy has become highly successful from these highly simplistic videos, and it may be worth considering that more elaborate figures are not necessary to produce real learning.

2. Study 1: Animated Presentation of Pictorial and Concept Map Media in Biology

1 Introduction

As intelligent tutoring systems (ITSs) become more and more sophisticated, the types of media that can be included in such systems have become increasingly varied. In order to support the students' learning, ITSs have included static images (e.g., AutoTutor [1]), diagrams (e.g., Andes [2]), animated illustrations (e.g., Guru [3]), concept maps (e.g., Guru [3], Betty's Brain [4]), videos (e.g., Operation Aries! [5]), and other media. However, there are at this time very few rules in place to guide which media type to select and how to present it given a myriad of parameters such as the student's prior knowledge, student's spatial ability, and task demands [6]. More work is needed to understand what types of media work best under certain conditions.

Recently, the tension between static and animated images has been of particular interest. The literature on animated images demonstrates a strong division between results, where animations sometimes contribute significantly to students' learning and, at other times, they have no impact on learning whatsoever. For instance, in the document explanation literature, images animated using a technique called "sequential display" (where an image starts out blank and segments of an image appear when they become relevant in the narration) often result in better memory for the information on the image [7]. In a recent meta-analysis, animations were shown to have a $d = 0.37$ advantage over static images when it comes to learning [8]. However, the authors caution that this effect is not as strong under all conditions (for instance, animations had a weaker effect in Biology than they did in Chemistry), and may in fact disappear in some circumstances (such as when the animations are purely decorational). For instance, [9] found that students who viewed an ordered series of static images outperformed those who viewed an animated visual of the same dynamic process, which is one of the conditions under which animations are meant to operate best. Therefore, it seems that additional investigations must be done to discover the strengths and limitations of image animation.

However, it is not only animated pictures and illustrations that have been investigated for their efficacy. Researchers focused on concept maps, an educational device that is growing in popularity and has been incorporated into multiple ITSs, have also examined how animation can add to student learning. One of the limitations of concept maps is that they often contain no cues to guide specifically how they should be read. Eye tracking research bears this out, as gaze patterns vary largely between participants examining a concept map [10]. Therefore, animations are seen as a method of directing student attention and imposing a specified processing order. There have been two substantial investigations into concept map animation, but the results of these studies have been mixed, indicating that there may be conditions and best practice rules that guide the animation of concept maps as well. [11] found that animated concept maps resulted in better recall of the information 48 hours later over static maps or even animated text, but that animation had no effect on the ability to recall lower-level details. Recently, [12] also compared static and dynamic text and concept maps but found that animation provided no advantage for either text or concept maps. These opposing results may be due to at least one of two key differences in the experimental designs of the aforementioned studies: concept map complexity/size, where [12]'s map was more complex than [11]'s, and the use of accompanied narration, which [12] claimed counteracted the effects of animation in their study by providing too much scaffolded guidance.

While there seems to be indications that both animated concept maps and pictures can be advantageous to learning under the right conditions, very little is known about how they compare to each other. It seems intuitive to suppose that both have their own time and place in educational multimedia environments, but there are currently no rules to guide the selection of one over the other for ITS designers, and further still, there is no research to suggest whether the presence or absence of animation for either of these media forms should inform this selection decision. Currently, both the concept map animation literature and the picture animation literature focus primarily on how each media type stacks up to its own static version, as well as how it compares to and/or works alongside text (e.g., [11], [13]). How concept maps and pictures compare to each other in terms of learning, as well as how animation affects this comparison, is still an open question.

It may also be the case that it is not a simple matter of determining which media type is most effective, but rather, which type aids specific kinds of learning. For instance, one of the strengths of concept maps is that they explicitly model the relationships between concepts, which have been theoretically linked to creativity, understanding, and deep knowledge of the material [14,15]. However, both pictures and concept maps can convey conceptual knowledge, or information pertaining to the topic’s main concepts, such as through picture labels or labeled nodes. To date, none of the concept mapping literature has tried to differentiate between these different knowledge types; therefore, little is known about how concept maps, especially animated concept maps, may influence memory for these kinds of information. Picture animation research has revealed that animation can have an effect on memory for different types of knowledge. [8] found that animation had the largest effect on procedural motor knowledge, followed by declarative knowledge. Others have found that the method chosen for animation, such as displaying objects that are thematically related versus spatially related, can deeply impact how the information is later recalled [16]. It may be instructive to investigate how images, animated or not, impact conceptual and relational knowledge as well, as this would allow for a direct comparison between the performance of students exposed to either concept maps or pictures.

Likewise, there also remains an open question as to how narration impacts animated concept maps. Narration is the preferred mode of information delivery when pictures, animated or not, are available, so that the student's attention is not split between the text and the picture [17]. Narration presented with animated images is also not uncommon (e.g., [13]). However, questions have been raised about whether narration washes out the effects of animation in concept maps [12]. Narration may therefore be one parameter for deciding whether or not to use an animated image or concept map, but a replication of this "washing out" should be observed before deeply exploring this parameter.

In this study, we will look at how pictures and concept maps, both animated and static, effect students' relational and conceptual knowledge learning in Biology, as well as their free recall of information. This will allow for a direct comparison between pictures and concept map media types in terms of their learning efficacy, which may help guide selection principles for their inclusion in educational multimedia environments.

The visual in every condition will also be accompanied by spoken narration in order to further test [12]'s hypothesis that spoken narration removes the animation effect that had been observed by [11]. Although no advantage was found for animation in Biology visuals [8], this domain relies heavily on visual aids, and so discovering the best practices for displaying these visuals is to the advantage of both educators and ITS designers within the field of Biology.

This experiment used a Khan Academy Biology video as the basis for the educational intervention. Khan Academy is a popular online company dedicated to making short, freely available video lectures that students find easy to understand. Khan Academy videos always feature audio narration of a lesson played in synchrony with screen capture of the narrator drawing pictures or working out problems that support the lesson. Therefore, the videos produced by Khan Academy are ideal for this kind of investigation because they are ecologically valid learning videos that natively feature picture animation and spoken narration. Khan Academy is also at the forefront of online, self-paced education, and features the kind of media which could be in ITSs due to their low production costs. This experiment seeks to use and modify these materials, which already exist in the educational world, in order to compare the learning produced by animated pictures and concept maps.

2 Methods

A 2 x 2, between-subjects experiment was conducted in order to examine the interactive effects of media animation (animated vs. static) and media type (picture vs. concept map). Participants were randomly assigned to one of these four conditions.

Participants were recruited through Mechanical Turk (MTurk), an online service offered through Amazon. MTurk allows “requesters” to put up short tasks (“HITs”) to be completed by their “workers,” who are then paid a small wage for satisfactorily completing the task. Requesters can also place restrictions, called “qualifications,” on who can participate in their study. To ensure quality results, participants who wished to participate in the current study had to have previously completed 50 HITs and had to have at least 95% of those HITs approved by the requesters, meaning that they had done an adequate job on the task and had been paid for it. Additionally, participants in this study had to certify that they were above 18 years of age (an MTurk standard), were a native English speaker, were a United States or Canadian citizen (implemented to increase the odds of recruiting native English speakers and enforced via IP checks), had adequate hardware to complete the experiment, and did not have significant hearing impairments. Those who failed to meet these criteria were disqualified from proceeding to the experiment. Participants who completed the study were paid $1.00.

In this experiment, 214 participants completed the study, but six were disqualified due to their failure to meet the participation criteria. The average age of the participants was 35.91, with a minimum age of 18, a maximum of 72, and a median of 32.5. One hundred fourteen of the participants (54.8%) were female. Previous examinations of the Mechanical Turk workers found that workers are, on average, 31 years old, with ages ranging from 18 to 71, and 55% of workers are female [18], making our sample typical of the MTurk population with the exception that workers outside of the United States and Canada were excluded. Studies have shown that the MTurk population appears to function similarly (i.e., produce qualitatively and quantitatively similar results) to university populations and other online populations [19,20,21].

The materials for this study consisted of four edited videos which made up the stimuli, two interchangeable knowledge measures, and a brief demographics survey (portions of which are reported above). The interventions for this study are based on the "Parts of a Cell" video produced by Khan Academy. In Parts of a Cell, the narrator discusses various cellular components while drawing and labeling them on screen. The Parts of the Cell video was selected due to its straightforward nature and its popularity, as it is one of the most highly viewed videos from their Biology series. The original Parts of the Cell video was edited to shorten the overall video length from 21 minutes to 15 and to remove segments of the video where the narrator scrolls away from the main image to illustrate some point in an aside. This edited video comprised the animated picture stimulus. The animated concept map stimulus replaced the visual portion of the edited video with an animated concept map. In the concept map version, the nodes correspond to the same labeled and drawn cell parts that appeared in the pictorial version. The concept map is composed of 18 key propositions (facts in node-link-node format) arranged in a hierarchical layout, with much of the arrangement of the map determined by the order in which information is delivered in the narration. In the animated concept map, propositions are added to the map generally when the proposition has been stated for the first time. Once added to the map, propositions are not removed, and the map builds in complexity until it reaches its completed state near the end of the lesson. This is the traditional method of animating concept maps [11,12]. The static stimuli, both pictorial and concept map, were created by taking the final, complete version of the cell picture and concept map, respectively, and using that static image as the visual for the entire video while preserving the same audio narration.

While the "smooth drawing" of the picture and the chunked "sequential display" of the concept map are not visually equivalent forms of animation, both represent the ecologically valid and traditional display methods associated with their respective media types; concept maps have long been considered "animated" if displayed one proposition at a time, while pictures lend themselves to being drawn as a form of drawing attention to and elaborating certain areas of the image (as would be seen in, for instance, expert human tutoring [22]). This experiment considers both styles of animation as roughly functionally equivalent, as both are intended to guide the student's attention to specific parts of the media.

The knowledge measures were created by first extracting the propositional facts of the ensuing lesson (e.g., "Vesicles transport proteins"). These propositions were then made into multiple choice questions by removing either the equivalent of a proposition's node (e.g., "Vesicles transport ______") or its linking phrase ("Vesicles ______ proteins."). There were 18 key propositions in the Biology lesson videos, and therefore 18 node and 18 link questions were created for the knowledge measures. The questions were then randomly sorted into Form A and Form B such that each proposition is represented only once per form, resulting in 9 node question and 9 link questions per form. Participants experienced either Form A or Form B as their pretest, and received the opposite test for their posttest (counterbalanced).

To participate, MTurk workers had to first accept the assignment on MTurk, and were then transferred to the actual experiment, which took place in Qualtrics. Once the worker consented to participate and had made the necessary certifications, he or she first took a pretest to assess his/her prior knowledge on cell parts in Biology. After completing the pretest, participants then experienced one of the four conditions (animated picture video, animated concept map video, static picture video, static concept map video). Controls were removed from the video in order to help prevent starting and stopping the lesson, and participants were instructed merely to listen attentively while the video plays without taking notes. Once the video completed, participants performed a free recall task, where they were asked to write down as much information as they could remember from the material they just saw and heard. After the free recall task, participants took the posttest (the opposite test form from the pretest), and then filled out a brief demographics form. They were then given a password to enter into Mechanical Turk as proof of completion, for which they were then paid.

3 Results

This research seeks to investigate the effects of animation (animated versus static) and media type (picture versus concept map) on various types of learning, specifically conceptual learning, relational learning, and the general free recall of facts. This was accomplished by examining different types of questions: those questions querying the student's memory of node information (conceptual), link information (relational), and their free recall responses. Each of these research questions has been analyzed and considered separately below.

We first investigated how animation and the media type affected "link" questions, which tap into relationship knowledge. The nine multiple choice link questions from both the pre- and posttests were first scored for correctness, and then each participant's proportional learning gains score was calculated. Proportional learning gains, formulated as (Proportionalized Posttest – Proportionalized Pretest) / (1 – Proportionalized Pretest), are a useful learning gains metric because they control for prior knowledge. These were then analyzed using a 2 x 2 between-subjects analysis of variance (ANOVA). While there was not a significant main effect for media type (p = .39) or a significant animation x media type interaction (p = .645), there was a significant main effect for animation, F(1, 204) = 4.041, p = .046. We see that, when the media was animated (M = .542, SD = .377), participants scored significantly higher on the link questions than those in the static media conditions (M = .405, SD = .577; d = 0.281).

The analysis of the node questions was given a similar treatment; the scores from the nine node questions in the pre- and posttests were used to calculate a proportional learning gains score, which was then examined using a 2 x 2 between-subjects ANOVA. There was no significant main effect for animation (p = .741), but there was a marginally significant main effect for the media type, F(1, 204) = 3.402, p = .067, where those in the concept map condition (M = .427, SD = .39) scored higher on node questions than those in the picture condition (M = .319, SD = .452; d = 0.254). However, the results may be best explained by the significant animation x media type interaction, F(1,204) = 9.021, p = .003. When the media was animated, those in the concept map condition (M = .501, SD = .282) outperformed those in the picture condition (M = .222, SD = .537) on the conceptual node questions (d = 0.65). When the image was static, however, those in the picture condition (M = .414, SD = .347) learned more about concepts (nodes) than did those in the concept map condition (M = .347, SD = .468; d = 0.165).

The free recall was scored automatically by comparing the responses to a list of keywords created from the transcript of the audio narration. One point was awarded for each of the keywords mentioned in the free recall response (although not for repeated mentions), and a coverage score for each person was then calculated by dividing the number of keywords mentioned by the total number of keywords on the list. This allowed us to examine their memory for technical vocabulary particular to the topic. The coverage scores were then analyzed using a 2 x 2 ANOVA to investigate the impact of animation and media type on the participants' memory for vocabulary. A covariate of the combined pretest scores for both link and node questions was also included in order to control for prior knowledge. There was no main effect for media type ($p = .374$), but there was a marginally significant main effect for animation, $F(1,202) = 3.524$, $p = .062$, where those who experienced an animated visual ($M = .349$, $SD = .2$) had better coverage of key vocabulary terms than those in the static visual conditions ($M = .318$, $SD = .19$; $d = 0.195$).

4 Discussion

In order to aid common ITS design decisions, this study sought to examine how animation, combined with picture representations and concept maps, affects memory for different types of information. The interpretation of the results is clearest when separately considering how relationships and concepts are best learned.

When it comes to knowledge of relationships, this experiment provides evidence that animation can contribute significantly to learning gains, indiscriminate of whether the image is a picture or a concept map. It seems that the action of animation, therefore, is better at guiding attention to the relationships between concepts, which included relationships such as part-of relations, properties, typology, and functional connections (“Vesicles – transport – proteins”). While this finding is not explicitly supported by the picture animation literature, there are some indications that it is in line with previous work. Animation has been shown to be somewhat effective in supporting declarative knowledge learning ($d = 0.44$), which would contain both concept and relationship knowledge, but it is especially effective in teaching procedural motor knowledge ($d = 1.06$; [8]). While procedural motor knowledge is undoubtedly also a combination of conceptual and relationship knowledge, it is mostly focused on the relational “how to” information. Therefore, it is somewhat expected that animations would aid more in teaching relationship knowledge. For concept maps, however, this is entirely new information; most recently, animation had been found to have no effect on learning [12], and there has not been an investigation on how animation would impact the learning of links or nodes. Therefore, the discovery that animation does in fact support learning with concept maps provides evidence that animated concept maps may need to be more deeply explored to understand the conditions under which they do or do not aid learners. Interestingly, although it seems intuitive that concept maps would be superior at teaching relational knowledge, no such link was found in this study, perhaps partially due to the topic (where many of the relationship are “part-of” relations, which are equivalently conveyed pictorially).

Conceptual learning is a more complex story. When the media is animated, concept maps provide superior support in teaching conceptual knowledge (operationalized by node questions). This is particularly interesting because it is not merely a case of concept maps explicitly spelling out the concepts while the picture merely represents them pictorially. The image on the picture drawn by the narrator is also labeled, and the labels of the picture and nodes of the concept map share a high overlap (93%, with the remainder being words jotted down on the picture in an aside). Therefore, the concepts are both equally visually represented in verbal form, but the concept map has the added advantage of removing extraneous detail, which may be the key to its success. Although animated pictures have been shown to aid in teaching declarative knowledge [8], which is at least partly conceptual, this study indicates that animated concept maps may be even better for creating gains in conceptual knowledge. For the static media, the picture fared slightly better than the concept map in terms of conceptual learning, although the difference is not great. This may be because, in the absence of animation, the more detailed picture has more unique cues to encode, and so more attention is paid to the labels and concepts. Further investigation is needed to determine if there is a true advantage of static pictures over static concept maps. However, both static conditions produced higher learning gains for concepts than did the animated picture, possibly due to its overwhelming volume of information and action.

The results from the free recall analysis show a more general (albeit slighter) trend, where animation affected participants' recall of technical vocabulary, which included both conceptual and relationship knowledge (e.g., terms such as "cytosol" and "transcribe"). This effect in and of itself is not surprising given that the literature shows that animation tends to improve learning [8], but what is interesting is the lack of effect for media type. Previous analyses of free recall responses in experiments with animated or static concept maps or text have found that concept maps produce better free recalls than text [11,12]. Here, when comparing two image-based media, this effect disappears;

it is possible then that animated image-based media may produce more recalled information than text, although additional research would need to be done to make this direct comparison. While the present free recall analysis is not as thorough as those typical of the concept map literature, where free recall responses are hand scored against a list of declarative knowledge statements, the free recall analysis done here does hint that animation may be useful in not just recognition of key terms, as may be demonstrated by the multiple choice questions, but in recall of information.

The pattern of results from this study implies that, generally speaking, there are conditions under which concept maps or pictures may be the preferred media, with animation being the main parameter considered in the present work. Animation in general seemed to contribute to relationship knowledge, while animated concept maps specifically were most efficacious in instilling conceptual knowledge. If animation is not an option, however, static pictures were more effective for conceptual knowledge. This underlines two general findings. First, different types of media seem to have their own contexts in which they are most effective in improving learning, and the learning environment and knowledge goals should be addressed in order to decide on the media type. Second, animation can have different effects on different types of media and learning, and further exploration of this little studied effect is in order. There are also some other interesting implications of this study. This study demonstrated that animated concept maps are not redundant with spoken narration, which would lead to a washing out of learning differences, as [12] suggested. While the parameters under which animation is not useful for concept maps is not yet known, narration does not seem to be one of those parameters. Additionally, it is interesting that these effects were found in the domain of Biology, which was one of the least successful domains in demonstrating differences between animation and static images. It may be the case that other domains would produce a stronger effect.

While this work fills gaps in our current knowledge of animated media, there are some limitations to this study. First, the results of this study do not take into account the effects of domain (in this case, Biology). It may be the case that certain domains or even certain properties of specific lessons are better represented with other types of media or other forms of animation. Likewise, this study also examines very specific kinds of knowledge measures, those that measure conceptual and relational knowledge, but it may be true that for other types of knowledge, such as general declarative knowledge, deep knowledge, or procedural knowledge, the results may vary. It is not the purpose of the present work to claim that one media type is superior to another in general, but rather, to relate that under the established conditions, animation and animated concept maps seem to produce larger learning gains in relational and conceptual knowledge, respectively. This work also does not explore every method of animating an image; there remains a breadth of animation methods in the existing literature to explore using this paradigm.

With the growing use of concept maps and other forms of media in ITSs, it is important that we continue to investigate the conditions under which they can be effective so that informed design decisions can be made. This will allow us to select the most effective media to use in our systems while avoiding investing in unnecessary “bells and whistles” that do not contribute to the student's experience. Future work which explores the limitations and advantages of different types of media in varying degrees of animation are necessary to contribute to the field's development.

References

1. Graesser, A.C., Lu, S., Jackson, G.T., Mitchell, H., Ventura, M., Olney, A., Louwerse, M.M.: AutoTutor: A Tutor with Dialogue in Natural Language. Behav. Res. Meth. Instum. Comput. 36, 180–193 (2004)
2. VanLehn, K., Lynch, C., Schultz, K., Shapiro, J.A., Shelby, R.H., Taylor, L., Treacy, D., Weinstein, A., Wintersgill, M.: The Andes Physics Tutoring System: Lessons Learned. IJAIED 15(3), 147–204 (2005)
3. Olney, A., D'Mello, S.K., Person, N., Cade, W., Hays, P., Williams, C., Lehman, B., Graesser, A.C.: Guru: A Computer Tutor that Models Expert Human Tutors. In: Cerri, S., Clancey, W., Papadourakis, G., Panourgia, K. (eds.) ITS 2012. LNCS, vol. 7315, pp. 256–261. Springer, Heidelberg (2012)
4. Leelawong, K., Biswas, G.: Designing Learning by Teaching Agents: The Betty's Brain System. IJAIED 18(3), 181–208 (2008)
5. Butler, H.A., Forsyth, C., Halpern, D.F., Graesser, A.C., Millis, K.: Secret Agents, Alien Spies, and a Quest to Save the World: Operation ARIES! Engages Students in Scientific Reasoning and Critical Thinking. In: Miller, R.L., Rycek, R.F., Amsel, E., Kowalski, B., Beins, B., Keith, K., Peden, B. (eds.) Vol 1: Programs, Techniques and Opportunities, pp. 286–291. Society for the Teaching of Psychology, Syracuse, NY (2011)
6. Schnotz, W., Cade, W.: Adaptive Multimedia Environments. In: Sottilare, R., Hu, X., Graesser A. (eds.) Design Recommendations for Adaptive Intelligent Tutoring Systems: Adaptive Instructional Strategies (Volume 2). Army Research Laboratory, Adelphi, MD (in press)

7. Bétrancourt, M., Dillenbourg, P., Montarnal, C.: Computer Technologies in Powerful Learning Environments: The Case of Using Animated and Interactive Graphics for Teaching Financial Concepts. In: De Corte, E., Verschaffel, L., Entwistle, N., van Merrienböer, J. (eds.) Powerful Learning Environment: Unravelling Basic Components and Dimensions, pp. 143–157. Elsevier, Oxford, UK (2003)
8. Höffler, T., Leutner, D.: Instructional Animation Versus Static Pictures: A Meta-analysis. Learn. Instr. 17, 722–738 (2007)
9. Lowe, R., Schnotz, W., Rasch, T.: Aligning Affordances of Graphics with Learning Task Requirements. Appl. Cognitive Psych. 25(3), 452–459 (2010)
10. Nesbit, J., Larios, H., Adesope, O.: How Students Read Concept Maps: A Study of Eye Movements. In: Montgomerie, C. Seale, J. (eds.) EDMEDIA 2007, pp. 3961–3970. AACE, Waynesville, NC (2007)
11. Blankenship, J., Dansereau, D.: The Effect of Animated Node-Link Displays on Information Recall. J. Exp. Educ. 68(4), 293–308 (2000)
12. Adesope, O. Nesbit, J.: Animated and Static Concept Maps Enhance Learning from Spoken Narration. Learn. Instr. 27, 1–10 (2013)
13. Mayer, R., Hegerty, M., Mayer, S., Campbell, J.: When Static Media Promotes Active Learning: Annotated Illustrations Versus Narrated Animations in Multimedia Instruction. J. Exp. Psychol.-Appl. 11(4), 256–265 (2005)
14. Bloom, B.S.: Taxonomy of Educational Objectives, Handbook I: The Cognitive Domain. David McKay Co. Inc., New York (1956)

15. Novak, J.D., Cañas, A.J.: The Theory Underlying Concept Maps and How to Construct and Use Them. Technical Report, Florida Institute for Human and Machine Cognition, (2008)
16. Bétrancourt, M., Bisseret, A., Faure, A.: Sequential Display of Pictures and its Effect on Mental Representations. In: Rouet, J.F., Levonen, J.J., Biardeau, A. (eds.) Multimedia Learning: Cognitive and Instructional Issues, pp. 112–118. Elsevier Science, Amsterdam (2001)
17. Sweller, J.: Cognitive Load During Problem Solving: Effects on Learning. Cognitive Sci. 12, 257–285 (1988)
18. Ross, J., Irani, I., Silberman, M.S., Zalvidar, A., Tomlinson, B.: Who are the Crowdworkers?: Shifting Demographics in Amazon Mechanical Turk. In: Mynatt, E.D., Schoner, D., Fitzpatrick, G., Hudson, S.E., Edwards, W.K., Rodden, T. (eds.) CHI EA 2010, pp. 2863–2872. ACM, New York (2010)
19. Mason, W., Suri, S.: Conducting Behavioral Research on Amazon's Mechanical Turk. Behav. Res. 44(1), 1–23 (2012)
20. Paolacci, G., Chandler, J., Ipeirotis, P.G.: Running Experiments on Amazon Mechanical Turk. Judgment and Decision Making 5, 411–419 (2010)
21. Suri, S., Watts, D.J.: Cooperation and Contagion in Web-based, Networked Public Goods Experiments. PLoS One 6(3), e16836 (2011)
22. Williams, B., Williams, C., Volgas, N., Yuan, B., Person, N.: Examining the role of gestures in expert tutoring. In: Alevan, V., Kay, J., Mostow, J. (eds.) ITS 2010. LNCS, vol. 6094, pp. 235–244. Springer, Heidelberg (2010)

3. Study 2: Labels as Attentional Guiding Devices in the Presence of Narration

The previous study suggests that pictures may not always function optimally as a visual medium during a narrated lesson. When conveying conceptual information, animated pictures performed below static and animated concept maps, as well as static pictures. This is particularly interesting because animated pictures are now a pervasive form of educational media (Höffler & Leutner, 2007; Lowe & Schnotz, 2014), but it may be the case that how and when they are presented matters in their educational efficacy.

One hypothesis put forward in the previous study about why animated pictures performed suboptimally is that, in this case, the picture is “busy”; there are many colors, shapes, and labels in the image itself, and when combined with animation, this may result in distraction rather than the directed attention that animation is supposed to add (Lowe & Schnotz, 2014; Stull & Mayer; 2007; Tversky, Bauer Morrison, & Bétrancourt, 2002). Perhaps more importantly, there may be too much information in the visual modality. Both shape information and the label are present in the drawing (as well as the act of animation itself, which may function to draw attention to that specific area of the picture), but the narration only has a single stream of informational input. The labels may be an unnecessary echo of the narration when a keyword is spoken, or they may be a distraction from the unique shape information presented in the picture. What is not known is, if a modality has multiple representations of the same information, such as shapes that represent a concept and a visual word which represents a concept, is the redundancy of information harmful, neutral, or in some respects, helpful?

As previous literature has demonstrated, competing visuals may be harmful to learning due to the split attention effect (Chandler & Sweller, 1992), where students may divide their attention between reading the label and examining the drawn picture. Although a label does not pose the same burden as a chunk of text, it may nonetheless present a distracting factor, especially when contiguity may already exist between the same narrated word and the shape information of the picture. Additionally, if labels are meant to direct students' attention to a specific area by matching key words in the narration (e.g., if the narrator says "Mitochondria" and the same word is written on a labeled picture, the student should know that that area is the one being discussed), the same task might already be accomplished by animation, which directs attention through movement or changes (Höffler & Leutner, 2007; Lowe & Schnotz, 2014). Labels, then, might be redundant as an attention directing factor.

The split attention effect with labels may account for the suboptimal performance of the animated picture when compared to the concept map in the previous study; even though both had the same conceptual words in their respective visuals, pictures also convey complex shape information, and when joined together, these two factors may create some level of split attention and distraction, especially when also combined with animation, another type of visual information.

On the other hand, a rival hypothesis exists which supports the importance of displaying both shapes and labels. Both potentially convey separate information which could be important for students; shape information may be useful in identification transfer tasks (e.g., a student who needs to identify on a diagram the part of the cell that produces energy) while exposure to a label may better ensure that a student can later identify the word in written questions, such as a multiple choice question. This may explain why participants who saw the static picture did better at the conceptual information task than those who saw an animated picture, as they had more exposure to the labels than the animation picture participants did. If this is the case, then animation is more to blame than labels that the animated picture condition poorly conveyed conceptual knowledge. Labels may also strengthen the contiguity effect (Mayer & Sims, 1994) because they exactly mirror key points in the narration, thus creating a stronger tie between the narration and the visual.

The focus of Study 2 will be on the *picture labels*, which are sometimes used in diagramming key parts of a picture (Berthold & Renkl, 2009; Cromley et al., 2010; Grosse & Renkl, 2006; Kragten et al., 2012; Schnotz & Bannert, 2003). Labeled images have many names (diagrams, picture glossary (Moline, 2011), etc.), but they all consistently refer to a picture which has its individual parts labeled with a key word or phrase. However, they are not a necessary picture component, and when paired with narration, may present either a helping hand or a stumbling block, as mentioned previously. This study will examine whether the presence or absence of labels in a picture visual representation will aid or dampen learning, particularly of the conceptual and relational information presented in the video. Additionally, this study will also look at how animation interacts with labels to test the hypothesis that animation and labels may be redundant if contiguity exists between the animation and the narration.

Methods

A 2 x 2, between-subject experiment was conducted in order to examine the interactive effects of labels (*unlabeled* vs. *labeled*) and image animation (*static* vs. *animated*). Participants were randomly assigned to one of four conditions upon agreeing to participate.

Participants

A total of 235 participants were collected through Amazon's Mechanical Turk in order to match the size of Study 1 (five were later removed due a failure to follow directions). Mechanical Turk workers were collected under similar qualifications to the previous study; participants were recruited from the United States (Canada was excluded from this study so that automatic filtering can take place rather than manual IP checks), they had to have completed 50 HITS (Turk tasks) previously, and had to have at least a 95% approval rating of the HITs they had completed on Mechanical Turk. Participants also certified that they were above 18, a native English speaker, had adequate hardware to play videos and hear audio, and were not significantly hearing impaired. Participants were paid $1.50 for successfully completing the HIT; the pay increase from Study 1 accounted for the additional time this HIT took due to added tests.

Mechanical Turk workers are, on average, 31 years old (median is 27, range: 18 - 71) and 55% female (Ross, Irani, Silberman, Zalvidar, & Tomlinson, 2010). The participants collected in this study were 36.2 years old (median is 33, range: 20 - 79) and 58.3% female. This sample was also demographically similar to the sample collected in Study 1.

Materials

While the testing materials for this study were identical to those of the previous study in many aspects, there were some major differences.

Video. The first major deviation from the previous study was in the main manipulation, the video; while this study used the same base video (*Parts of a Cell*), the visual was edited to focus on the animation and label variables. The base video from

Study 1, which had been cut for time from the original Khan Academy *Parts of a Cell* video, was used as a basis for all four videos, and the narration was not altered for any of them. For the two static picture conditions, a screen capture taken from the final moments of the video was displayed throughout the video and did not change. While this screen capture was not altered for the labeled static condition, the labels were erased so that only shape information remained for the unlabeled static condition.

The animation conditions were handled slightly differently than in Study 1. Study 1 previously had a fluid drawing animation for the picture condition and a sequential display animation for the concept map. Here, to eliminate this confound and ease image manipulation, both animated conditions used sequential display, where chunks of pictorial information were revealed at once rather than slowly watching a person draw the picture. This allowed for better experimental control of when objects appeared as well; previously, objects appeared at the content author's discretion with fluid drawing, which could be any time between first mention of a topic and the last.

For the labeled animation condition, upon first mention of a topic (e.g., "mitochondria"), the picture representation of that topic *and* the label appeared simultaneously. Rather than having the label and picture appear at separate times, this study controlled for timing to focus mainly on the issue of *copresence*. For the unlabeled animation condition, the timing of the pictorial shape information related to the topic remained the same as in the labeled animation condition, but in this condition, no label appeared to accompany it. Therefore, timing and shape information remained constant across each condition, and only the presence or absence of the label changed. See Appendix A for a visual representation of how these conditions differed.

Given these differences, the conditions have a high level of information equivalence. In all four conditions, the narration was the same and also carried all of the information critical to the tests. All four conditions also experienced the same "shape information" from the picture of the cell. The sequential display animation style did not carry with it extra information, as it does not convey motion – it simply revealed parts of a picture at key moments. Likewise, in the label conditions, the labels were included in the image, but they merely echoed the key term being spoken in the narration. In theory, a student could receive a near perfect score any of the tests that follow the video if they were to memorize and fully understand everything in the narration and shape information, which are constants in each condition. Therefore, what is being measured here is truly the focusing/enhancing effects of animation and labels.

Testing materials. The testing materials for this study closely resembled those of Study 1 but with a few additions.

First, the node and link questions from Study 1 were presented again in Study 2 (Appendix B). Although this study does not include a concept map representation, breaking key information into vocabulary/concept recognition (node questions) and relationship recognition (link questions) may prove useful in pinpointing where labels and animation strengthen referential connections in the understanding of fundamental facts. These questions were not altered from the previous study, as the content of the lesson itself had not changed. Additionally, students were asked to give a summary of what they learned, just as they did in Study 1, and this free recall task was scored using the same metrics; a list of key words and phrases that represent both conceptual and relational information was used to judge the completeness of each free recall task, both in terms of keyword coverage (Number of keywords used /Total number of words) and keyword density (Total number of words that are keywords /Total number of words).

This study also included Biology multiple choice questions related to the content presented (Appendix C). These multiple choice questions were culled in 2011 from all available state practice tests and sample items. Therefore, these questions are quite similar to those that would be seen on state Biology exams across the United States. Questions were selected from the pool of questions based on whether they related to any of the key facts used as a basis for the node and link questions. Although not every fact is represented by a question, every question from the pool that relates to the content was included, for a total of 26 questions.

Students also engaged in a fill-in-the-blank diagramming task in order to study how well information transferred. Students were shown a cell, seen in Appendix D, and asked to label the various parts that were presented in the lesson. This task was a free recall task, with no word bank given to participants. The unlabeled figure does not precisely reproduce the cell seen in the video lesson, but instead involves transfer past the superficial details of the original image to a new image. This reproduces a situation school-age students commonly come across when considering what they study, how they are tested, and how they are tested on state exams; while figures the student sees may resemble each other in essentials, they often have different orientations, arrangements, appearances, and even possibly different parts represented. The purpose of this task is to gauge how well students can label a figure given the different manipulations.

Students were also asked to visually identify which organelle has a particular function from among four choices on a new figure (see Appendix E). This task closely mirrors situations seen in state testing, where students are given the function of some part and asked to select that part from a diagram. This makes a connection between the function of a cell part, described in the narration, and the identification of the cell part, which is only described in the visual representation. Therefore, this task may be better at gauging the integration of verbal and visual representations in the students' mental models. Twelve questions were created for this task that mimic the language used in state tests, although no questions from the state tests could be used directly – questions typically query the name of an organelle indicated on a diagram or the name of an organelle that accomplishes a specified function. To remove the organelle names from the questions, four of the 11 questions were adapted from the state tests and seven questions were created in a similar style.

Motivation questionnaire. This study also included a motivation questionnaire in order to assess the learner's perceptions of their learning experience (Appendix F). It inquired about how they felt about the video, whether they were interested in the content, and how effortful/difficult they found the material to be. While there is not necessarily a strong connection between students' perceptions of their own learning experiences and their actual learning (Maki, 1998), it may be the case that student perceptions are ultimately tied to their attrition in a program of learning, where they may not return to the learning material if they perceive it to be boring or unhelpful.

Procedure

This study (HIT) was posted to Mechanical Turk, along with basic information about the type of task it was, requirements for participation, and an approximate study length (40 minutes to 1 hour). Participants who elected to participate were redirected to

Qualtrics through Mechanical Turk, where they first encountered the informed consent (Appendix G). Upon receiving their consent, participants were then taken to a cell labeling task pretest where a cell was presented with numbers in place of labels, and the participant had to fill in the coordinating numbered blank with the name of the cell part.

After the labeling task, participants were randomly assigned to a condition and counterbalanced test order. Depending on their counterbalancing, students saw either Test A or Test B as a pretest. This test contained multiple choice questions from state tests and node and link questions, both mixed together and presented in a random order. As in Study 1, the link and node questions were split between tests such that the link and node question that coordinate with a specific key fact were sorted into different tests randomly, while still preserving an even split between link and node questions per test form.

Upon completion of the pretest, participants saw a screen warning them that they were about to watch a video, and to remove any distractions or take any necessary breaks so that they can concentrate on watching the video. They were also instructed not to take notes, but rather to listen and watch as the lesson plays. The video was one of four types, depending on condition assignment: unlabeled static video, unlabeled animated video, labeled static video, or labeled animated video. The video began as soon as the participant indicated that they were ready to watch the video, and controls for the video had been removed to prevent pausing or moving forward or backward through the video. Participants were not able to progress to the next stage of the experiment until an amount of time had passed that was equal to the length of video.

Immediately following the video lecture, students were asked to write a summary of everything they saw and heard in the video, which had to be a minimum of 140 characters to continue the study. Once they have finished this free recall task, they moved on to the cell labeling posttest task, which was in all ways identical to the labeling pretest task they completed at the start of the experiment. Participants then completed the posttest (Test A or B depending on the counterbalancing), followed by the organelle function identification task, where some organelle was described by a question and the participant must identify what organelle is being described by selecting the number of the matching organelle in a picture of the cell. Participants then filled out the motivation questionnaire and some basic demographic information. They were then given a password to fill in at the Mechanical Turk HIT site, where they received payment upon approval of a successful completion.

Results and Discussion

The goal of this research is to investigate the impact of animation and labels on different kinds of knowledge and learning tasks. To that end, the analyses have been split into two sections: in the first section, learning gains will be assessed on the different metrics proposed in the Methods (with a few additional analyses where gains are not able to be calculated), and in the second section, all analyses from the first section will be repeated to purposefully consider the effects and interactions of prior knowledge in conjunction with animation and labels. This allows for a more nuanced view of the data and the effects of prior knowledge, an important cognitive individual difference.

Learning

Each type of learning material presented in the pre- and posttests (state examination multiple choice questions, node questions, and link questions) was scored separately for correctness, and then proportionalized learning gains were applied, formulated according to the following equation: (Proportionalized Posttest – Proportionalized Pretest) /(1 – Proportionalized Pretest). For each analysis, individuals who fell greater than three standard deviations from the mean were excluded from that particular analysis, but could re-enter other analyses if they were not outliers in that analysis as well. All cell means are reported in Table 1.

A 2 x 2 between-subjects analysis of variance (ANOVA) for state examination multiple choice test questions revealed no significant main effects or interactions for animation or labels (all $p > .34$), indicating that for general topic knowledge found in state examination tests, animation and labels do not produce gains in learning. The same result was found in Study 1, although it was not reported at the time.

Table 1

Means and Standard Deviations for Administered Tests in Study 2

	Animation						**Static**					
	Label			***No Label***			***Label***			***No Label***		
	HPK	*LPK*	Total	*HPK*	*LPK*	Total	*HPK*	*LPK*	Total	*HPK*	*LPK*	Total
State Exam MC Qs	.75 (.18)	.55 (.21)	.32 (.37)	.75 (14)	.48 (.22)	.32 (.31)	.75 (.16)	.42 (.18)	.29 (.30)	.73 (.19)	.44 (.19)	.27 (.34)
Link Qs	.89 (.12)	.69 (.23)	.56 (.41)	.84 (.20)	.63 (.26)	.43 (.54)	.84 (.11)	.51 (.25)	.35 (.42)	.84 (.16)	.53 (.24)	.25 (.54)
Node Qs	.79 (.11)	.64 (.22)	.32 (.50)	.81 (.16)	.60 (.21)	.46 (.44)	.81 (.17)	.53 (.25)	.40 (.45)	.84 (.18)	.48 (.25)	.40 (.52)
Pictureless Label Qs	.96 (.08)	.76 (.25)	.68 (.37)	.92 (.17)	.71 (.35)	.63 (.43)	.89 (.21)	.61 (.32)	.52 (.52)	.90 (.15)	.51 (.33)	.42 (.52)
Coverage Score	.20 (.07)	.21 (.11)	.20 (.10)	.19 (.09)	.18 (.10)	.19 (.09)	.22 (.10)	.25 (.15)	.23 (.13)	.19 (.09)	.19 (.10)	.19 (.10)
Density Score	.28 (.07)	.29 (.13)	.29 (.10)	.29 (.08)	.27 (.12)	.28 (.10)	.33 (.11)	.32 (.15)	.33 (.13)	.28 (.07)	.27 (.12)	.28 (.10)
Cell Labeling Score	.75 (.20)	.38 (.21)	.46 (.29)	.72 (.21)	.32 (.25)	.42 (.34)	.72 (.19)	.28 (.20)	.42 (.30)	.67 (.24)	.18 (.17)	.28 (.31)
Function ID Score	.71 (.18)	.48 (.16)	.58 (.20)	.64 (.18)	.46 (.19)	.57 (.21)	.60 (.15)	.37 (.13)	.49 (.18)	.62 (.21)	.41 (.17)	.51 (.21)
Motivation Score	27.42 (5.82)	25.42 (7.24)	26.31 (5.82)	27.55 (5.70)	25.33 (6.97)	26.61 (6.30)	26.20 (6.66)	22.24 (7.11)	24.25 (7.11)	27.14 (6.71)	20.56 (5.64)	23.91 (6.99)

Note. HPK = High prior knowledge, LPK = Low prior knowledge
Items in Total column are means in that condition from the first section of analyses (gains and posttest-only, as indicated in the text); items in HPK or LPK columns are posttest-only scores from the second section of analyses
Standard deviations are in parentheses

2 x 2 between-subjects ANOVAs were also conducted for the link and node questions separately. For the link questions, which query relational knowledge, a statistically significant main effect was found for animation, $F(1, 224) = 9.73$, $p = .002$, and a marginally statistically significant effect was found for labels, $F(1, 244) = 3.40$, $p = .067$, but the interaction was not significant ($p = .817$). For the animation effect, being in an animated condition ($M = .50$, $SD = .48$) produced higher gains in relational knowledge than the static condition ($M = .30$, $SD = .48$; $d = 0.48$). For the main effect of labels, having labels present ($M = .46$, $SD = .42$) produced more gains than not having the labels present ($M = .34$, $SD = .54$; $d = 0.24$). The 2 x 2 ANOVA for node questions, which query conceptual knowledge, did not produce any significant effects (all $p > .25$).

Another metric was produced in order to examine the effects of knowledge that existed on the visual purely in the form of a label, with no accompanying picture. These "pictureless labels" represent a specific subset of cases where only a label would be present to represent some concept, and therefore these concepts have no pictorial grounding and would have no visual representation at all in the no-label conditions. To understand how animation and labels affected this class of topics (which primarily centered on the differences between prokaryotic and eukaryotic cells), a subset of questions on Test A (6 questions) and Test B (7 questions) were identified as querying these topics, resulting in a small mix of state questions and node/link questions. Proportionalized learning gains were then calculated. A 2 x 2 between-subjects ANOVA was then used to assess the pictureless label questions, resulting in a significant main effect for animation, $F(1, 224) = 8.88$, $p = .003$, but nonsignificant results for the label main effect ($p = .228$) and the animation x label interaction ($p = .775$). Those who saw an animated presentation (*M* = .66, *SD* = .43) did better than those who saw a static presentation (*M* = .47, *SD* = .52; *d* = 0.39), an interesting finding considering that even in an animated, labelless condition, no representation of these concepts would have been present on screen.

An additional metric was created in order to assess whether participants remembered the first or second part of the lesson better, and whether this interacted with the experimental manipulations to animation and labels. Therefore, pretest and posttest questions were split into groups based on whether they appeared in the first half or second half of the videos, proportionalized learning gains were calculated for each half of the video, and a 2 x 2 x 2 mixed between-within subjects ANOVA was conducted, with the half as a within-subjects variable, and animation and labels as between. There was a statistically significant main effect for half, Wilks' Lambda = .94, $F(1, 220) = 14.45$, $p < .001$, and for animation, $F(1, 220) = 7.89$, $p = .005$, but not for labels ($p = .555$). There was also a significant half x label interaction, Wilks' Lambda = .98, $F(1, 220) = 5.08$, $p = .025$, but not a significant interaction for half x animation, animation x label, or the half x animation x label interaction (all $p > .77$). Those in the animation conditions (*M* = .45, *SD* = .38) had greater gains than those in the static conditions (*M* = .33, *SD* = .41; *d* = 0.33). Although there was a significant main effect of half as well, this is best interpreted through the significant interaction between half and label, where those who saw labels did better on the first half questions (*M* = .50, *SD* = .33) than those in the labelless conditions (*M* = .40, *SD* = .36; *d* = 0.29), but those in the labelless conditions did better on the second half questions (*M* = .35, *SD* = .43) than those in the label conditions (*M* = .30, *SD* = .47; *d* = 0.11).

The free recall task was scored in two ways: by calculating the number of unique keywords mentioned in their response (a mixture of conceptual and relational vocabulary terms used in the lesson) and dividing that by the total number of words in their response (the keyword coverage score) and by taking the total number of keywords they mention regardless of repetition and dividing that by the total number of words in their response (the keyword density score). Once obtained, two 2 x 2 between-subjects ANOVAs were run, one for each free recall score, using the overall pretest score for the participant as a covariate to account for prior knowledge. For the keyword coverage score, a significant main effect was found for labels, $F(1, 225) = 5.18$, $p = .024$. Therefore, those who saw labels ($M = .22$, $SD = .11$) said proportionately more words on the "golden" list of keywords than those who did not see the labels ($M = .19$, $SD = .09$; $d = 0.27$). No statistically significant effects were found for animation or the animation x labels interaction (all $p > .24$). For the keyword density score, a similar pattern was found, where the animation main effect and animation x label interaction were not significant (all $p > .16$), but a statistically significant label main effect exists, $F(1, 225) = 4.33$, $p = .039$, where those participants who saw labels in their video ($M = .31$, $SD = .12$) used keywords more frequently in their free recall than those who did not see labels ($M = .28$, $SD = .10$; $d = 0.27$).

For the cell labeling task, the pretest and posttest labeling tasks were scored for simple percentage correct, which was then used in calculating proportionalized learning gains. A 2 x 2 between-subjects ANOVA was then run on the cell labeling gains, resulting in a statistically significant main effect for animation, $F(1, 225) = 4.70$, $p = .031$, and a significant main effect for labels, $F(1, 225) = 4.89$, $p = .028$. The interaction term was not significant, *p* = .262. Participants in the animated conditions (*M* = .44, *SD* = .32) made greater gains in the labeling task than those in the static conditions (*M* = .35, *SD* = .31; *d* = 0.28). Those participants in the labelled image conditions (*M* = .44, *SD* = .29) also made greater gains when labeling cell parts than those in the unlabeled conditions (*M* = .35, *SD* = .33; *d* = 0.28).

The function identification task was scored on percent correct since it does not have a pretest, and a 2 x 2 between-subjects ANOVA was then calculated using the participants' overall pretest score as a covariate to account for prior knowledge. A statistically significant main effect for animation was found, $F(1, 225) = 12.47$, $p = .001$, but the main effect for labels ($p = .864$) and the interaction term ($p = .272$) were not significant. Those participants who saw an animated presentation (*M* = .57, *SD* = .20) were better at linking the image of an organelle (cell part) to its function than those who saw a static presentation (*M* = .50, *SD* = .20; *d* = 0.37).

Finally, the Likert scale motivation responses were compiled into a single comprehensive score which indicated how enjoyable they found the video on the whole. Questions included items such as "I found the video to be boring", "I felt frustrated after watching the video", "I feel like I learned a lot from the video", "The video sparked my interest in learning more about Biology", etc. Negatively worded items, like the first two examples, were reverse coded such that higher scores indicated more enjoyability. This comprehensive score was analyzed using a 2 x 2 between-subjects ANOVA in order to detect how participants felt about the different conditions. While the main effect for labels ($p = .984$) and the animation x label interaction ($p = .715$) were not significant, there was a statistically significant main effect for animation, $F(1, 226) = 7.07$, $p = .008$, where those in the animated conditions (*M* = 26.46, *SD* = 6.47) thought more highly of the video than those in the static conditions (*M* = 24.09, *SD* = 7.03; *d* = 0.35).

These results represent an interesting and consistent pattern of outcomes from this investigation into the effects of animation and labels on different types of knowledge and tasks. Below, outcomes will be discussed first by those impacted by animation, and then those affected by labels.

Animation Discussion

In many previous studies, animation has proven to have a powerful impact on student learning, and for certain tests in this study, this effect has been replicated. Just as in Study 1, Study 2 found that animation had a significant effect on relational knowledge in the form of "link" questions. Link questions are recognition questions where learners select the correct relationship between two conceptual vocabulary terms, and here, those learners in the animation conditions produced higher learning gains in relational knowledge than those who saw static images, with a medium effect size of *d* = 0.48 (Cohen, 1988). As before, this is not entirely surprising given that previous work has found that animation aids both declarative knowledge learning (*d* = 0.44) and procedural motor knowledge (*d* = 1.06; Höffler & Leutner, 2007), both of which contain elements of conceptual and relational knowledge, with procedural motor knowledge representing more of that "how-to", action-oriented knowledge. When compared with Study 1, Study 2 demonstrates a much stronger effect in favor of animation (*d* = 0.48 vs. 0.28), perhaps indicating that animation may be especially useful in the case of pictures, which can become particularly complicated compared to a neatly laid out concept map (although, undoubtedly, complicated and messy concept maps also exist).

Interestingly, unlike in Study 1, animation did not aid conceptual knowledge or "node" questions and neither did the presence of labels. Although not significant, the pattern of the interaction term with node questions provides some intriguing hints as to why Study 1 found an interaction between media type and animation. In Study 1, concept maps and pictures were found to be similar when static, but animated pictures performed much lower on conceptual questions than animated concept maps. In Study 2, the same general pattern is observed with labels and animation. Here, the labeled and labelless image conditions are nearly identical when static (M = .404 vs. .402, respectively), but when animated, the labeled condition underperforms compared to the unlabeled condition (M = .319 vs. .462), although large standard deviations in these conditions may prevent a significant interaction.

Although this must be interpreted with extreme caution, it appears that the general trends between the two studies indicates that static images can cope with varying levels of visual complication for teaching conceptual knowledge, but when the image is animated, less may be more. Visual interest may become a visual burden (e.g., extraneous load; Sweller & Chandler, 1994) when animation is in the mix, and others have previously found that animation is not always a boon to learning (Höffler & Leutner, 2007; Tversky et al., 2002). This trend may have been more pronounced in Study 1 due to the different media forms so that it gained significance, but as Study 2 holds many visual elements steady (including color, timing of animations, etc.), this effect may have been dampened into nonsignificance. More evidence of this potential effect is needed before further supposition can be made.

These results also indicated that questions relating to labels that lack a picture component ("pictureless labels") produced a main effect for animation, where those who saw the animated multimedia display performed better on items with no pictorial reference over those who saw the static video. Here, it may be that animation itself draws enough interest to the presentation that the learner's heightened attention is responsible for the improved memory; this is somewhat borne out by the general motivation score, where people found the study more enjoyable and worthwhile when in the animation conditions, which may have then increased their attention and effort on the task. It is surprising that the presence of labels did not improve gains for pictureless label knowledge considering that there were essentially no visual elements for these items in the labelless conditions, animated or static. However, it seems that the narration, which was the same for all four conditions, provided sufficient information to answer the questions these topics pertained to, and it was animation that improved attention to this information, not the labels.

Similarly, the cell labeling task had main effects for both animation and labels with identical magnitudes of effects ($d = 0.28$). While the label effect will be discussed separately, it does seem that animation also improves attention enough that cell parts can be more correctly labelled than if they had not seen an animated presentation. Although improved attention may be one reason why students made greater gains on the cell labeling task in animated conditions, another mechanism at play here may be animation's improved contiguity with the narration. Animation may more closely pair audio and visual information via timing, thus creating referential connections between the audio and visual mental models according to the dual-coding theory (Mayers & Sims, 1994). These deeper mental models may then aid in recalling organelle names, an altogether more difficult task than recognizing them, as learners would do with the node questions.

This same mechanism may also be responsible for the main effect of animation on the function identification task, where those who saw animation could better pick out the unnamed organelle that performs the listed function compared to those who saw a static visual during the lesson. While it is likely that improved attention due to animation is one mechanistic pathway that increased gains, increased referential connections between mental models may be another. Additionally, there may be an embodied cognition explanation of these results, especially considering that link questions were also improved by animation. Animation may create better memory of “action” elements, such as relational information (like verbs; Mahon & Caramazza, 2008) and function information. To perform well on this task, function information needs to be tied to the pictorial information of organelle shape, and so an especially heavy emphasis was placed on the visual element, where animations operate. Animations may have therefore created a tie between organelle “actions” and the organelle shape through the action of animation itself.

Although previously mentioned, learners reported enjoying the study more when their visual was animated. This may have repercussions for long term use of multimedia; while this study and the previous were “one offs” for the participants, Khan Academy, ITSs, and other multimedia educational lines all depend upon returning users and students wanting to expand their knowledge outside of a single lesson. The ITS Guru, for instance, is designed to cover an entire year’s worth of Biology lessons (Olney et al., 2012). If users are excessively bored or unhappy in the course of a single multimedia display, they may not return to follow up on other lessons. Therefore, animations may be one key way of enriching their educational experience, to say nothing of the many ways it may improve learning.

Label Discussion

While animation may have a large backdrop of previous work, relatively little work has been done on the effects of labels. The results of this study indicate that, while they are not always enormously impactful, they make contributions to certain kinds of tasks and knowledge that may be valuable for learning.

First, while the animation effect on link questions was not surprising, it may be unexpected that labels also contribute to relational knowledge. Those who saw a labeled image presentation produced higher learning gains for relational knowledge than those who saw a static display. The reason this is so surprising is that, for the most part, the labels apply only to *conceptual* knowledge, naming cell parts or categories of cells. Very little relational knowledge is displayed through the labels. However, by having conceptual information written down and externally stored in the image, more time and energy can be spent on listening to the narration and comprehending the various organelle functions that are being discussed. In this way, labels may free up some resources so that participants can later recognize the correct relationship between concepts, although this effect may be limited, otherwise there would have been a label main effect for the function identification task as well. Because this effect is not entirely consistent with theory, replication is needed.

Labels, like animation, did not produce an effect on conceptual questions, which is perhaps more surprising than the lack of an animation main effect. With extra exposure to the written terminology that is later quizzed in the node questions, one would hypothesize that this added exposure would increase gains on node questions. However, it seems that this is not the case; it may be that this information is extracted from the narration with no difficulty, and so the labels do not add anything extra. Alternatively, labels were also hypothesized as the reason the animated picture condition underperformed on the node questions in Study 1 compared to all other conditions. Although this is not supported by a significant interaction, the pattern of the interaction does replicate those results, and so this may be a partial explanation of why the animated picture condition did poorly in Study 1.

Labels also improved the free recall scores, both in keyword "coverage" (or number of listed keywords mentioned) and "density" (amount of the free recall that was comprised of keywords). Learners used a wider range of keywords and a higher density of keywords when they had previously seen the labels in the multimedia presentation. Interestingly, the keywords used to score these free recall responses were a mixture of both conceptual ("mitochondria", "lysosome", "nucleus", etc.) and relational terms ("stores", "destroys", "transcribe"), although there are more conceptual vocabulary terms than specialized relational terms that could be considered a "keyword." However, having learners use more of those key vocabulary terms in their own free recall responses may be good starting place in getting students used to technical terminology and recalling it, rather than simply recognizing it. This also indicates that students are reading and absorbing at least some of the labels, a concern raised by Cromley et al., (2010), who found that students read only 36% of labels presented onscreen.

For the cell labeling task, there was a main effect for both animation and labeling, where learners who were exposed to labels produced higher gains in labeling a cell than those who did not see the labels. Again, it seems that students exposed to the labels are reading at least some of the labels presented to them, which in turn helps them recall the labels. Another marker indicating that labels are being read to some degree is an accidental spelling mistake. In the video, the narrator speaks frequently of the “cytosol”, but in writing the label, forgets to cross the “t”, writing “cylosol” instead. In this task, 8 different learners wrote “cylosol” rather than “cytosol” in their responses; this mistake occurs zero times in the pretest. However, finding an effect of labels on the cell labeling task represents almost the bare minimum contribution of labels; if exposure to labels had not improved students’ ability to perform a near transfer task and label a cell similar to the one they saw, the inclusion of labels in multimedia presentations would have been significantly called into question. It is true that the presence of labels has the same impact on learning as viewing an animated presentation, regardless of labels (d = 0.28), which does indicate that animation is just as impactful as labels when it comes to this “fill in the blank” labeling task.

There was also a significant interaction between labels and information retained between different halves of the video. Those who saw a labeled presentation did better on questions about the first half of the video over those who did not see labels, and those in the labelless conditions did better on information from the second half of the presentation. While there are more labeled items discussed in the first half of the video than the second, this does seem to suggest that interest in the multimedia presentation waned in the second half of the video for those who saw labels, but those who did not see labels kept a steadier level of attention throughout the video. This may be because those without labels needed to put in extra effort in order to link what was happening on screen to what was happening in the narration. Though their interest did flag a little bit over time, their extra effort kept them more alert than those who had labels to scaffold the connections between the visual and the audio information. This may have also more evenly distributed their referential connections, while those in the label conditions had more referential connections on the front end.

While these interpretations may give us some idea of how labels and animation may impact gains in types of knowledge, building mental models, and performance on certain tasks, the learner's prior knowledge may also deeply affect how they respond to both animation and labels (e.g., Höffler & Leutner, 2007; Lowe & Schnotz, 2014). Therefore, reinterpreting all of the previous analyses with the student's level of prior knowledge taken into account may create a more precise picture of how these parameters operate under the constraint of previous knowledge.

Prior Knowledge

A series of 2 x 2 x 2 between-subjects ANOVAs were run in a style similar to the previous analyses. The additional variable, prior knowledge, was calculated by performing a median split on the overall pretest score (which combined state examination multiple choice, link, and node questions). The dependent variables were then the posttest scores of the state examination multiple choice, link questions, node questions, pictureless label questions, and cell labeling task. Free recall coverage, density scores, and the function identification task have no pretest equivalent, so their ANOVAs are applied to their single outcome scores for each of those tests.

For the state examination multiple choice ANOVA, there was a significant main effect for animation, $F(1, 222) = 3.95$, $p = .048$, and prior knowledge, $F(1, 222) = 123.21$, $p < .001$, but not for labels, $p = .451$. When in the animated conditions ($M = .64$, $SD = .22$), participants scored higher on state test questions than those who saw a static presentation ($M = .59$, $SD = .24$; $d = 0.22$). Participants with high prior knowledge ($M = .74$, $SD = .17$) scored more highly on the state multiple choice questions than did those with low prior knowledge ($M = .48$, $SD = .20$; $d = 1.40$). There were no significant interactions (all $p > .14$).

The ANOVA on the link question posttest found significant main effects for animation, $F(1, 222) = 9.50$, $p = .002$, and prior knowledge, $F(1, 222) = 92.52$, $p < .001$, but there was no main effect for labels, $p = .34$. However, these main effects are best interpreted through the significant animation x prior knowledge interaction, $F(1, 222) = 4.97$, $p = .027$. When the learners had high prior knowledge, animated presentations ($M = .86$, $SD = .17$) produced only slightly higher link posttest scores than static presentations ($M = .84$, $SD = .14$; $d = 0.13$), while if the learners had lower prior knowledge, animated presentations ($M = .67$, $SD = .24$) produced much higher posttest scores than static presentations ($M = .52$, $SD = .24$; $d = 0.63$). There were no additional significant interactions (all $p > .23$).

For the node question ANOVA, a marginally significant main effect was found for animation, $F(1, 222) = 2.88$, $p = .091$, and a significant main effect was also found for prior knowledge, $F(1, 222) = 91.55$, $p < .001$. These effects are overridden by a significant animation x prior knowledge interaction, $F(1, 222) = 7.38$, $p = .007$, where high prior knowledge students perform slightly better when they had viewed a static presentation (*M* = .83, *SD* = .17) over an animated presentation (*M* = .80, *SD* = .14; *d* = 0.19), but when the learners had low prior knowledge, those who saw an animated presentation (*M* = .62, *SD* = .21;) outperformed those who saw a static presentation (*M* = .50, *SD* = .25; *d* = 0.52). There was no significant label main effect, $p = .717$, and no other interaction was significant (all $p > .21$).

The analysis of the "pictureless label" questions, or those questions that pertain to labels that did not have an accompanying pictorial representation, revealed a significant main effect for animation, $F(1, 222) = 11.18$, $p = .001$, and prior knowledge, $F(1, 222) = 67.69$, $p < .001$, but are perhaps best interpreted through the significant animation x prior knowledge interaction, $F(1, 222) = 4.21$, $p = .041$. Those with high prior knowledge performed better on these pictureless labeled questions when they viewed an animated presentation (*M* = .94, *SD* = .14) versus a static presentation (*M* = .90, *SD* = .18; *d* = 0.25), and low prior knowledge students also performed better when they saw an animated presentation (*M* = .74, *SD* = .29) over a static presentation (*M* = .56, *SD* = .33; *d* = 0.58), but the difference is more pronounced with low prior knowledge students. There was not a significant label main effect, $p = .148$, and no additional significant interactions (all $p > .35$).

A 2 x 2 x 2 x 2 mixed between-within subjects ANOVA was also conducted to determine the interactions between animation, labels, prior knowledge, and video half, where the questions from the posttest have been tagged as either pertaining to the first or second half of the video presentation. This ANOVA used only half scores derived from the participants' posttests as the dependent variable, keeping in line with the pattern of analyses conducted with the other ANOVAs. Here we find significant main effects for half, Wilks' Lambda = .88, $F(1, 222) = 30.23$, $p < .001$, animation, $F(1, 222) = 8.29$, $p = .004$, and prior knowledge, $F(1, 222) = 148.92$, $p < .001$, but no significant main effect of label, $p = .334$. There was a significant animation x prior knowledge interaction, $F(1, 222) = 5.07$, $p = .025$, and a marginally significant half x label interaction, $F(1, 222) = 2.90$, $p = .090$, where questions from the first half the of video are best answered by those in the label condition (*M* = .73, *SD* = .22) over the labelless condition (*M* = .70, *SD* = .23; *d* = 0.13), while questions from the second half of the video were slightly better answered by those in the labelless conditions (*M* = .66, *SD* = .25) over the labeled conditions (*M* = .64, *SD* = .25; *d* = 0.08). Additionally, there was a significant half x prior knowledge interaction, but all previous effects (minus the half x label interaction) may be best interpreted through the marginally significant half x animation x prior knowledge three-way interaction, Wilks' Lambda = .98, $F(1, 222) = 3.73$, $p = .055$. For the first half of the video, when viewing an animated presentation, high prior knowledge participants scored approximately the same when viewing an animated (*M* = .83, *SD* = .15) or static presentation (*M* = .83, *SD* = .15, *d* = 0), but low prior knowledge students benefited from the animated presentation (*M* = .67, *SD* = .21) over the static presentation (*M* = .51, *SD* = .22; *d* = 0.74). For the second half of the video, high prior knowledge students who saw an animated presentation (*M* = .82, *SD* = .18) slightly outscored those who saw a static presentation (*M* = .79, *SD* = .18; *d* = 0.16), while low prior knowledge students had a more pronounced effect, with those in the animated conditions (*M* = .54, *SD* = .22) outscoring those in the static conditions (*M* = .45, *SD* = .21; *d* = 0.42).

For the free recall responses, the ANOVA associated with the keyword coverage score demonstrated a main effect for labels, $F(1, 222) = 5.18$, $p = .024$, where those who saw labeled presentation (*M* = .22, *SD* = .11) outperformed those who saw no labels (*M* = .19, *SD* = .09; *d* = 0.30). No other main effects (all $p > .26$) or interactions (all $p > .24$) were significant. The same effect pattern was found for the keyword density analysis, $F(1, 222) = 4.14$, $p = .043$, where those in the labeled conditions (*M* = .31, *SD* = .12) outperformed those who saw the labelless conditions (*M* = .28, *SD* = .10; *d* = 0.27). No other main effects ($p > .19$) or interactions were significant ($p > .17$)

The ANOVA for the cell labeling task found a significant main effect for animation, $F(1, 222) = 8.32$, $p = .004$, label, $F(1, 222) = 4.49$, $p = .035$, and prior knowledge, $F(1, 222) = 233.64$, $p < .001$, but all interactions were nonsignificant (all $p > .16$). For the animation main effect, participants who saw animation (*M* = .55, *SD* = .29) were better at the cell labeling task than those who saw a static presentation (*M* = .47, *SD* = .31; *d* = 0.27), and for the label main effect, those who saw labels (*M* = .52, *SD* = .29) fared slightly better than those who did not (*M* = .49, *SD* = .32; *d* = 0.10). Additionally, those students with higher prior knowledge (*M* = .72, *SD* = .21) performed better on this task than those who had low prior knowledge (*M* = .29, *SD* = .22; *d* = 2.00).

For the function identification task, the ANOVA revealed a main effect for animation, $F(1, 222) = 11.05$, $p = .001$, and prior knowledge, $F(1, 222) = 89.27$, $p < .001$, although the label main effect was not significant, $p = .808$, and the interaction terms were also not significant (all $p > .133$). Participants were better able to match the organelle function with the organelle shape when they saw an animated presentation (*M* = .57, *SD* = .20) as opposed to a static presentation (*M* = .50, *SD* = .20; *d* = 0.35). Those with higher prior knowledge (*M* = .64, *SD* = .18) also outperformed those with lower prior knowledge on this task (*M* = .43, *SD* = .17; *d* = 1.20).

And finally, for the holistic measure of motivation (capturing how easy and enjoyable this experiment was), the ANOVA found a significant main effect for animation, $F(1, 222) = 7.69$, $p = .006$, and prior knowledge, $F(1, 222) = 18.23$, $p < .001$. The main effect for labels was not significant ($p = .837$). There was a marginally significant interaction between animation and prior knowledge, $F(1, 222) = 3.36$, $p = .068$. Those with high prior knowledge found the study slightly more enjoyable in the animation conditions ($M = 27.49$, $SD = 5.70$) versus the static conditions ($M = 26.66$, $SD = 6.65$; $d = 0.13$), while low prior knowledge students in the animation conditions ($M = 25.39$, $SD = 7.06$) found the experiment much more enjoyable than those in the static conditions ($M = 21.43$, $SD = 6.44$; $d = 0.59$).

Prior Knowledge Discussion

Because the results here are far more interdependent than in the previous section, this discussion will be integrated rather than broken into sections for the sake of clarity.

Overall, there was a general trend for both animation and prior knowledge affecting outcome scores. The state examination questions and function identification task both found main effects for animation and prior knowledge, with no interactions between the two. It seems that both are important components which affect how students perform at posttest, and interestingly, the addition of prior knowledge to the state examination question analysis produced an animation effect that was previously absent. This suggests that animation does help students remember basic facts and apply their knowledge so that correct inferences can be made (a hallmark of the state questions), but the effect is only produced when prior knowledge is also considered. Less surprising is that animation also produces higher scores in the function identification task, which was found in the previous analysis as well; because this task relies exclusively on building referential connections between the narration and the pictures, animation may be important in utilizing the contiguity effect in order to make these connections. Also noticeable in these results and in all other results with prior knowledge as a factor is that prior knowledge produces very large effect sizes, with the largest found here being a d = 2.00. This also indicates that prior knowledge is an important individual difference that can have a deep impact on students' learning experience.

The link and node questions saw a significant animation by prior knowledge interaction. For link questions, it seems that low prior knowledge students were most helped by an animated presentation when compared to a static presentation, while high prior knowledge students only saw small improvements to posttest scores with animation. This points to a recurring trend in these analyses, which is that low prior knowledge students benefit the most from animation, while high prior knowledge students, whether from near ceiling effects or otherwise, do not receive the same benefits. High prior knowledge students may not need the additional support of the attention directing (and to some extent, controlling) device that is animation, while low prior knowledge students may benefit from it as a scaffold for their thinking. This is similar to the results found by Schnotz and Rasch (2005), who found that using animation in a stepwise simulation of earth's rotation helped facilitate cognition (make cognitive processing easier) for those with low prior knowledge. This "expertise reversal effect" (Kalyuga, Ayres, Chandler, & Sweller, 2003) is a pattern found throughout these results. Additionally, the animation effect found again here reinforces an embodied cognition explanation for action aiding memory of "action-like" terms, which in this case were relational concepts. Also of note is that the previous effect of labels seen previously with link questions has vanished here. This effect was also not present in the Study 1 results, calling into question whether it was a real effect at all, or a spurious, marginal result.

On the other hand, the node (conceptual) questions had a more extreme interaction between animation and prior knowledge than the link question interaction. Here, high prior knowledge students actually did slightly better on conceptual questions when in the *static* conditions versus the animated conditions, while low prior knowledge learners did much better on the node posttest when in the animated conditions versus the static conditions. As Schnotz and Rasch (2005) point out, using certain kinds of animation may overscaffold knowledge for some students, where facilitating cognition may not actually benefit learning because cognitive processing they could have done on their own was instead done for them. In this case, the animation may have provided no help to high knowledge students, and instead, distracted them from learning the material. However, low prior knowledge students showed a clear benefit for animation, as it may help control the flow of information and scaffold knowledge for them, performing cognitive functions that they could not perform on their own such as chunking information, which is Phase 1 of the animation processing model (Lowe & Schnotz, 2014). This may have given low prior knowledge students the tools to learn more conceptual information than if they had seen a static presentation.

Like the link and node questions, the pictureless label questions also showed an animation by prior knowledge interaction. Here, both high and low prior knowledge learners had higher posttest scores if they had watched animated videos versus the static videos, but the difference between the two presentation types was much more pronounced with low prior knowledge students. Keeping with the pattern of the last two, animation seems to give structure and ease of processing to the presentation for low prior knowledge students that that high prior knowledge students do not necessarily need. This finding demonstrates this general increase in attention and processing even more than the last two analyses, because these questions query items that had no pictorial component and existed only as labels or only as audio (for those in the labelless conditions). Therefore, there was even less to look at or change onscreen, but nonetheless, animation aided these questions, indicating that it generally improves attention in these kinds of presentations.

The interaction between when an item is talked about in the video, labels, animation, and prior knowledge is a much more complicated story. Here there were two interactions, one between the video halves and labels, and the other, a three-way interaction between half, animation, and prior knowledge. The half by labels interaction, although marginal, revealed the same interaction seen previously: that those who saw labels did better on questions from the first half of the video, while those who did not see labels did better on the second half of the video. Like before, this may indicate that those who are not seeing labels may have to put more of their focus on the lecture, which pays off in the second half of the video when attention may wane and, in this case, the number of labels being referenced slightly decreases. The half, animation, and prior knowledge interaction indicates that, initially, high prior knowledge students do not benefit from animation while low prior knowledge students do, but in the latter half of the presentation, animation slightly improves higher knowledge students' scores while low prior knowledge students again see greater improvements due to animation (although not as large as in the first half). Like many of the other interactions presented here, it seems that high prior knowledge students are not receiving the same benefits of animation that low prior knowledge students are, although these benefits perhaps increase over the course of a video presentation as attention and cognitive resources become drained, with animations beginning to compensate for some of these losses. For low prior knowledge students, animation always means improvements, but these benefits slackened in the second half of the presentation. This may be because low prior knowledge students, in general, must use up more cognitive resources over time in order to learn and are not able to sustain the same level of learning throughout a longer presentation (here, roughly 15.5 minutes), and although animation certainly helps, it cannot overcome the effects of time on attention and learning. These results would lead one to believe that presentations aimed at lower prior knowledge students may want to include animation, frontload important information, and shorten presentation length, although further studies would be needed to test these hypotheses.

The story for writing free recall responses is much simpler and exactly mirrors the results found previously, even though this model includes prior knowledge. Here, only the presence of labels improves the number and density of keywords used in free recall responses. Like before, this is where the utility of labels is best observed. Regardless of prior knowledge or whether they saw an animated presentation, learners use more of those key vocabulary terms particular to this lesson if labels were present. While labels may not help in recognition tasks like multiple choice questions, they seem to be best at getting students to use the language of the lesson in recall tasks, although whether they also correctly convey the concepts of the lesson is another matter entirely that involves a much deeper, qualitative analysis of the free recall responses.

As in the previous analysis of the cell labeling task, a main effect for animation and labels was found, but this analysis also finds a main effect of prior knowledge. The main difference here is in the shift of effect sizes; although animation and the presence of labels before provided approximately the same effect on learning gains previously, here, animation impacts posttest scores more heavily than labels do ($d = 0.27$ vs. 0.10), indicating that labels have a much smaller effect than previously supposed. Prior knowledge also demonstrates an extremely large effect, $d = 2.00$, indicating that those who did well on the pretest were best at labeling the cell (with scores at about 72% at posttest) while low prior knowledge students score only 29% at posttest on average. This particular task seems to be one where learners either "get it" or they do not; a good use of further studies would be to detect and remediate students who are in danger of doing poorly on such tasks. While labeling tasks may be construed as fairly shallow, requiring only the recall of names of parts or processes and not necessarily a deep understanding of how something works, they are the kind of task students are expected to perform in class and on state tests.

Interestingly, the results of the motivation analysis follow most of the interactions here. The motivation analysis revealed that, while both low and high prior knowledge students preferred the animated conditions over the static conditions, it was the lower prior knowledge students who enjoyed the animated conditions the most ($d = 0.13$ vs. 0.59). This indicates that, not only do low prior knowledge students enjoy the animation more than high prior knowledge students, but they also learn more from it. High prior knowledge students seem able to take or leave the animation given that its effect on learning and motivation is so small. In short, high prior knowledge students may not have the same need for animation that low prior knowledge students have.

In total, these results demonstrate a consistent pattern, where prior knowledge and animation interactively affect many types of learning tests except written free recall responses about the lesson, which is affected primary by the presence of labels. While high prior knowledge students show little need for animation (which, in the case of node questions, actually negatively impacts their performance), low prior knowledge students are most helped by it.

General Discussion

This study revealed several key findings relating animation, labels, and prior knowledge. Overall, it appears that viewing an animated presentation is beneficial to students (particularly those with low prior knowledge of the subject) for improving a number of knowledge types, while labels primarily benefit students' ability to use vocabulary terms in written free recall responses and labeling tasks. Additionally, prior knowledge is deeply important in ascertaining if animation will improve students' knowledge, and could be crucial in determining the design of adaptive systems and other educational media. Let us consider each of these points in turn.

The first significant finding of this study is the degree to which animation affects the acquisition of different kinds of knowledge. Animation has been shown to be helpful in teaching relational knowledge, functional knowledge and its relationship to part identification, and the free recall of part names. Additionally, animation improved memory for items that existed only as labels in the visual and had no corresponding picture part, meaning that animation also sometimes helped students answer questions for items that were not on screen to be animated at all. When prior knowledge is considered in the model, animation also aids students in answering questions culled from national state tests.

But why is animation so effective? These results suggest a number of possibilities, some or all of which may be operating at once. First, animation may function as an attention directing device, focusing learners' attention through visual change on screen (Jeung, Chandler, & Sweller, 1997), which is most commonly called the attention-guiding principle (Bétrancourt, 2005). By guiding their attention to significant parts of the picture at the correct time, memory for those items is improved. However, above and beyond that, focus and interest in the lesson as a whole may rise, thereby improving memory for even those items that are not animated (as in the case of pictureless labels in the labelless condition). The motivation scores reported here, which query how enjoyable, easy, boring, etc. the task was demonstrate that animation may produce a more positive learning experience that is just interesting enough to produce learning without sinking the student into boredom and frustration, which are unproductive learning states (Baker, D'Mello, Rodrigo, & Graesser, 2010).

Second, as mentioned previously, animation may work to create referential connections between the visual and audio mental models by presenting them together in time. This contiguity effect then strengthens the bonds between these two models, providing a more cohesive and integrated mental model of the information which students then use to answer questions. There is strong evidence of this "contiguity effect" (Mayer & Sims, 1994) here due to the construction of this study; the animation is timed to “pop up” when the first mention of the cell part occurs, but exactly the same information is delivered to all conditions via the narration, which is also the source of all tested information. Additionally, the animation effect for the function identification task provides evidence that close ties are being made between the visual and audio information. While all conditions had access to the cell shape information and the narration, which were the two sources of information that had to be tied together to answer these questions, only those who experienced animation excelled at this task. Learners in these conditions clearly made links between the two information streams.

Finally, the animation may have worked here because of the style of animation used: sequential display. Previous studies have found that sequential display is an effective form of animation for teaching information from a variety of domains (Bétrancourt, Bisseret, & Faure, 2001; Bétrancourt, Dillenbourg, & Montarnal, 2003; Jamet, Gavota, & Quaireau, 2008). Sequential display differs from many other kinds of animation because it replicates some crucial elements of human drawing, and this simulation has a number of advantages. While it does serve to direct attention to new information, it also stops students from being overwhelmed by future information by withholding those visual components until they become relevant to the lesson. This reduces onscreen clutter, which may thus reduce extraneous load (Sweller & Chandler, 1994) and removes the cognitive theory of multimedia learning assumption being violated, where learners cannot process too much information in a single modality (Mayer, 2005). In this way, human drawing and sequential display are very similar, and in fact, when the animated picture condition from Study 1 (animated in the drawing style) is compared to the animated labels condition from Study 2 (which contains all the same visual information from Study 1 and is animated using sequential display), t-test comparisons reveal nonsignificant differences on the state examination multiple choice, link, and node questions (all $p > .333$), indicating that the two styles are very similar in learning functionality. Additionally, it cannot be said that there is added information in the animated conditions from Study 2 (an "information equivalence" argument), because sequential display does not convey additional information; it simply withholds parts of the picture and displays those parts when they are relevant. There is no additional "motion" information like other display types might convey, and all tested information is held in the narration, which is the same across conditions. Animation, therefore, has a number of cognitively-grounded reasons as to why it improves learning under certain conditions.

While animation effects were plentiful here, there were also a handful of effects for labels. Labels seem especially useful when asking students to perform recall tasks such as writing a summary (free recall) or labeling a cell. Learners were more likely to use key vocabulary terms from the lesson when labels were present in the display; this may have long-term value in educational settings, as teachers or tutors may want to get students comfortable using "jargon". However, the results here do not tell us if students are using these vocabulary terms correctly in their free recall responses (although they do for the labeling task), just that they are echoing more of those words they saw and heard if they had access to the labels during the lesson. Furthermore, there is evidence that at least some of the labels are being read given the learning effects in the free recall responses and labels, as well as the repeated spelling mistake mentioned before (mimicking the narrator's mistaken spelling of "cytosol" as "cylosol"), so the lack of effects on other knowledge tests cannot be blamed on students not reading the labels. It may be the case that learners (or perhaps simply adult learners) are good enough at parsing the audio stream for things like key vocabulary terms, and the labels are simply redundant with the narration. They did not provide any extra information, although having seen them written did prompt students to write them in their free recall responses and perform better on the labeling task.

It is also worth pointing out that there is no evidence for the hypothesis that labels create split attention. In no area measured here did labels decrease learning (other than when combined with animation in the nonsignificant node interaction), and so an extreme version of the modality principle, the idea that text on a picture induces split attention, may be too radical (Moreno & Mayer, 2002). On the other hand, Sweller et al., (2011) hypothesized that text that is well-integrated into a picture may aid learning, and although there is no measure for how "well-integrated" the labels were in the image from this study, it seems that this may also be too extreme in this case as well. Here, the modality effect occurred according to the extent that split attention was necessary (Schnotz & Cade, 2014), which in this case, was very little, as labels are so short to as to be nearly iconic and most likely were not consulted numerous times over the course of the presentation. Overall though, labels did not help or harm learning for most kinds of tasks, except in the free recall tasks where it improved scores with usually modest effects. The only potential hint of danger with labels may occur with long presentations, since this study found that labels were less helpful in the second half of the video presentation.

Finally, this study, like many others, found that the individual difference of prior knowledge is important to consider when designing multimedia. The interactions between prior knowledge and animation suggests that low prior knowledge students are helped the most by animation, while high prior knowledge students receive little to no improvements to learning when exposed to animation or could even be slightly hurt by it in the case of conceptual knowledge. It seems that low prior knowledge students need the animation as a way of scaffolding their learning and controlling the flow of the content, but high prior knowledge students have other cognitive tools to perform such tasks, such as previously existing mental models that information can be plugged into, and they therefore do not need all the accoutrements that animation brings. While prior knowledge is only one factor in determining the expertise of the learner (Schnotz & Cade, 2014), which is an important consideration when designing adaptive multimedia, it had a powerful impact on how students performed after viewing the multimedia presentation in this study, and may require strong consideration in determining the expertise of the learner.

Additionally, animation and prior knowledge were found to affect how the learners perceived the multimedia presentation; while high prior knowledge students only liked it slightly more if they had seen an animated presentation, low prior knowledge students reported being more positive about the presentation when compared to those that saw the static conditions. This likely has implications for a system's sustainability, as bored students may not return to a system for additional lessons, and so animation may be one way of staving off boredom and frustration for lower prior knowledge students. Additionally, it may simply make learning a less frustrating process.

On the whole, this study demonstrates that, while animation is a powerful attention-directing device that can improve learning in a number of arenas, especially for low prior knowledge students, labels seem not to function as attention directing devices, as they do not improve learning even in the absence of animation. Although they do enrich student free recall responses with more key vocabulary terms and they do improve students' ability to label parts better, labels generally do not help convey relational knowledge, conceptual knowledge, or answer state examination test questions that typically require some inferential abilities. Additionally, they have no effect on whether students view the multimedia presentation more favorably, and so an argument cannot be made to include labels just for the sake of improving students' perceptions of the lesson. On top of this, prior knowledge proved to be an important component in determining learning effects, with low prior knowledge students generally performing better when there was animation. While more work needs to be done concerning the extent to which the free recall responses convey correct knowledge, whether these results extend to other domains and lesson types, and whether picture complexity changes the pattern of results here, the goal of this work was to begin examining whether animation and labels, when combined, impacted learning.

4. General Discussion

The goal of this work is to expand our knowledge on the conditions under which some types of media and their presentation styles affect different kinds of learning. Of particular concern was how animation may interact with certain media types, as animation has become a predominant feature in more technologically advanced educational environments. In Study 1, animated and static versions of pictures and concept maps (two media types commonly when teaching science) were compared, and it was found that animation improved relational knowledge and marginally improved free recall, but students demonstrated better conceptual knowledge in the animated concept map condition, particularly compared to the animated picture condition. Study 2 expanded on Study 1 by delving further into the “picture problem”, examining the costs and benefits of a labeled picture. This study found that, while labels may aid students in improving their free recall scores and ability to label cells, animation and prior knowledge were much more important factors in answering state examination test questions, relational and conceptual knowledge questions, performing cell labeling tasks, the ability to pair a visual component with its function, and the student’s perception of the multimedia presentation.

Although many results have been discussed in this work, there are perhaps four main findings from these experiments. First, not all visual depictions of the same information are equivalent in terms of learning impacts. The type of media one is using, for instance, can have an impact on what kind of knowledge people learn. As we saw in Study 1, for conceptual information, concept maps produced much higher learning gains compared to pictures when both types of media were animated. While Study 2 hinted that the visual clutter of labels may have caused a dip in conceptual knowledge scores, it did not completely resolve why such a strong media type effect was found in Study 1.

Generally though, it does seem that concept maps may create a cleaner representation than do some types of pictures; their organization is fairly clear, and they consist only of conceptual and relational information. Their shape information is limited to ovals or boxes and labeled arrows, whereas pictures in their most base form are only shape information, with an option for labels. Both types of media also contain spatial information, but this information may differ between media; concept maps may be organized hierarchically, indicating subsumption or ordered chains of reasoning which are modeled for the student, while pictures may also contain subsumptive information, as well as information about shape, relative size, spatial relationships between elements that may not necessarily be conveyed in a concept map (e.g., in Study 1, the concept maps indicates that the nucleus is inside the cell, but not that it is in the center of the cell), and potentially other visual qualities that may not be distinctive enough to be mentioned (e.g., color or texture).

While concept maps and pictures each have their educational affordances, we still know relatively little about precisely which of these affordances affect certain kinds of learning. The studies conducted here have only established that there is in fact a discernable difference in how these two types of media function, but the particular mechanism within each type of media that impact learning are still relatively unknown. Previous literature has indicated that both concept maps and pictures accompanied by narration may be better than simply reading text (Adesope & Nesbit, 2013; Blankenship & Dansereau, 2000; Mayer, 2009), but nothing was said on how the two media types compared to each other, how they affected different kinds of learning, and what qualities from each media type affect knowledge acquisition. While the first two questions have been addressed in this dissertation, the last question has only been partly explored through Study 2's examination of the impact of labels in pictures. However, it is important to note that the media type effect was only true of the conceptual knowledge questions; other types of knowledge queried were not susceptible to this effect, which also leaves a question as to whether alterations to each type of media may produce or nullify effects for different kinds of knowledge. Clearly then, more work needs to be done on the specific elements that lead to different learning outcomes between concept maps and pictures in order to get a better idea of when pictures or concept maps are appropriate for certain learning contexts.

The second major finding of this research concerns animation. Both Study 1 and Study 2 found many powerful effects for animation on numerous knowledge tests. This is interesting for a number of reasons. First, Höffler and Leutner (2007) found that, while animation generally has fairly high learning gains on average, the effect size for animation in Biology was small, $d = 0.13$. However, effect sizes in the studies conducted here regularly ranged from medium to large effects. In spite of this, it is easy to see how Biology effect sizes could have great range, as it is a domain with large diversity in subject matter. It may be the case that animation is more necessary for certain kinds of Biology lessons than others or that certain animation styles interact with the instructional goals of the lesson, i.e. conveying process rather than structure (Lowe & Schnotz, 2014). A difference between animation styles (sequential display and fluid animation) was not observed here (see Study 2 Discussion), but both of these animation styles are functionally equivalent in that they withhold information until it needs to be discussed, allowing the image to build in complexity. However, other animation styles do not necessarily function in this manner, and this may have been one reason for the lower Biology scores in the animation meta-analysis. As mentioned before, the lesson type may also matter; the lesson used in these studies was largely taxonomic, while other lessons may be more process-oriented and therefore call for an animation style that emphasizes motion, which both drawing and sequential display do not.

While the fact that animation effects turned up in nearly every analysis here might lead one to believe that a general rule such as "animation is always better than a static presentation" is warranted, this would be an overgeneralization (Lowe & Schnotz, 2014). It is true that, under the correct circumstances, animation can produce large learning gains, but it is not true that this occurs across all animations, domains, knowledge types, and individual differences. This work indicates that for conceptual questions, animation may provide fewer gains than viewing static images when the animated image is a picture; this was reinforced in Study 2 with a similar pattern for animated labels, although this result was not significant. Additionally, Study 2 repeatedly found that, while animation was beneficial for those with lower prior knowledge, high prior knowledge students did not profit much from viewing animated presentations, and not simply because they performed at ceiling. It seems that animation may help model methods of thinking about the content, but higher prior knowledge students do not benefit from such modeling, as they may already have a knowledge structure in place which needs only to be updated. Therefore, if a lesson is being designed for students with some knowledge of the subject already, animation may not add much to their learning experience.

Animations may also function as a way of evoking embodied cognition. Animation does seem to improve relational knowledge and functional knowledge in particular, which in this case, involves largely remembering verb information. Because there has been a link found between verb information and action in embodied cognition (Mahon & Caramazza, 2008), it is possible that it is embodied cognition that is being tapped into here, improving students' memory for this "action" information. Given the evidence for this type of explanation that has been found here, further follow-ups which explicitly test for this effect may be warranted.

All in all, it does appear that animation can have powerful effects on learning when used with the appropriate audience and when the goal is to teach certain kinds of knowledge. As explained previously, animation may work to strengthen the contiguity effect, or the close connection between the narration and the image, by making well-timed visual changes that link the visual to a certain point in the audio stream (the source for all tested information in this study). This then creates referential connections between the visual and verbal mental models produced from each modality (Mayer, 2005; Mayers & Sims, 1994), leading to a more integrated mental model of the information. As mentioned previously though, those who have pre-existing (and correct) mental models may not benefit as much from animation because these connections have already been made. Additionally, animation paired with a visually complex image may sometimes cause learning gains to drop, as mentioned in the case of conceptual knowledge and animated pictures. It is therefore important to understand the context in which animation should be used as it does not have blanket positive effects.

The third major "take home" message from this work concerns the effect of labels. Because Study 1 found that animated pictures perform suboptimally when it comes to conceptual knowledge questions, Study 2 was designed to investigate whether labels, when combined with animation, created too much visual distraction which reduced the learning gains. While Study 2 found a potential suggestion that labels did adversely affect conceptual knowledge scores when combined with animation (although this effect was not significant), labels were largely unimportant to learning in other knowledge measures. Only the free recall vocabulary scores and the cell labeling task were affected by the presence of labels. The results they did bring could be construed as the bare minimum contribution to learning – people were more likely to insert those vocabulary words they had just seen into the free recall response over those who had simply heard the same vocabulary words. Likewise, labels should improve the students' ability to label another cell given that they had just had the behavior modeled for them, but animation was also found to improve cell labeling with the same degree of impact (and with a larger degree with prior knowledge was accounted for). Additionally, there is no evidence that labels even alter learner perceptions about the video, as enjoyability scores were unchanged by labels. In all, it appears that labels may only improve students' ability to label a cell (which animation also does) and insert more vocabulary into their writing, but they seem not to contribute much else.

Previous studies on labels had indicated that labels may have something to offer learning, but did not suggest that the effects would be strong. In Florax and Ploetzner's (2009) work, labels were found to boost comprehension and retention of information, but this effect was only marginally significant. They also operationalized "labels" as small segments of text, approximately one to two sentences in length, and so it was possible that the unconventionality of these labels were dampening effects. However, the studies in this work seem to partially replicate their results; while labels do cause significant effects on some learning tasks here, specifically free recall and labeling, the effects were not strong and were absent from many other learning tasks. This seems to indicate that labels are not a strong learning component in the visual of a multimedia display. Mayer and Gallini (1990) also found that labels only produced a learning effect when the labels 1) indicated both the part names and part functions and 2) were part of a “dynamic” display, or two pictures showing machinery in its on and off state to imply motion. Study 2 of this work specifically looked for label by animation effects using only part name labels, and found no significant interactions between the two. While this could ostensibly be due to the fact that the labels do not also indicate part functions, it is more likely the case that label effects are difficult to procure (as seen in the failure of their other label conditions), and only certain learning tasks and perhaps certain domains are sensitive to them. Previous work supports this conclusion, as it has been found that animation visual cues, or methods of enhancing an animation so certain parts draw more attention, are attended to but not understood better, as animation is too overpowering a force for the visual cues to compete with (de Koning, Tabbers, Rijkers, & Pass, 2010; Lowe & Boucheix, 2007).

Another hypothesis that could have been fulfilled by the labels has to do with split attention. Because labels are technically text, they do compete with the visual in some way, which could create split attention. However, there is very little evidence that labels do actual harm (other than the nonsignificant interaction with animation for conceptual questions). Because labels are so short, they are likely treated mostly symbolically, perhaps read once when they become relevant or first appear from the animation and then no longer attended to. It seems that humans may be able to simply parse the narration for their keywords, allowing them to answer multiple choice questions with these terms. Labels do boost the usage of these keywords in free recall tasks, but further analysis is needed to see if they also improve the correct usage of these keywords, which might make the inclusion of labels in multimedia more attractive. For now, they seem neither helpful nor harmful, on average, although they do take up valuable visual real estate, which may be one consideration for not including them in a more complex image.

The fourth major finding here concerns the effects of prior knowledge on learning with animation. Study 2 found several main effects of prior knowledge as well as interactions between prior knowledge and animation, which generally indicated that high prior knowledge students do not benefit from animation as much as low prior knowledge students. This has several implications. First, prior knowledge is an important consideration in animation effects. As others have pointed out (e.g., Lowe & Schnotz, 2014), prior knowledge influences both the bottom up and top down processing of animations since it is part of the cognitive schemata of the learner. Without knowing the prior knowledge of the learner, it is impossible to align the perceptual characteristics and goals of the multimedia presentation with the cognitive requirements of the learner since that information is unknown. This can be especially true of animations that enact some level of user control; user control should only be given to students who have enough prior knowledge to use the controls to strategically enhance their knowledge, but these controls should also be limited to the degree to which the student can perform such actions. While the pictorial display here did not allow for user control, such considerations must be made for user controlled displays, which may be a better way of improving high prior knowledge students' understanding over a no-controls presentation.

What is demonstrated here is a clear example of the expertise reversal effect (Kalyuga et al., 2003), where what works for low prior knowledge students will not necessarily work for those with high prior knowledge. This was most likely triggered by a number of things, such as the simple nature of the lesson and the fact that sequential display is a form of segmentation. Segmentation, where pieces of knowledge are chunked, has been shown to be helpful for low prior knowledge students but not necessarily for high prior knowledge students (Spanjers, Wouters, van Gog, & van Merrienboer, 2011), and by using sequential display, this may have overscaffolded the "chunking" process that high prior knowledge students are perfectly capable of without assistance (Lowe & Schnotz, 2014). While no differences in learning were found between sequential display and continuous fluid drawing, both have functional similarities, and so a different kind of animation altogether would be recommended for learning interventions attuned to high prior knowledge students.

The results of this work highlighted four main points: 1) the importance of choosing the displayed image that best enhances the knowledge goals of the lesson, 2) the power of animation for certain kinds of learning under certain conditions, 3) the general weakness of labels in improving learning outside of labeling tasks and vocabulary usage, and 4) the importance of prior knowledge when considering animation styles. While there were more findings from this study, the goal here was to begin understanding the constraints under which animation does and does not operate well, as well as how different types of images and their features affect knowledge acquisition.

Limitations and Future Directions

This work does have its limitations. First, it was entirely collected on Mechanical Turk, which, while being a more diverse population than university undergraduates, is still a particular population which may differ from others. Mechanical Turk users are people who have access to a computer and are inclined to participate in tasks that typically pay less than minimum wage either from necessity or desire. Therefore, their motivations may differ from the usual undergraduate participant or, most importantly, from a learner in a classroom. However, Turk's greater diversity in age and background is a step towards better generalization of results, and while they may not have the same motivations as a "traditional" student, they may be more motivated than the average subject pool student, allowing for a closer reflection of the learners who use educational systems in order to meet some personal goal.

The animation styles in this work are also a narrow selection of all animation methods that are available on the static-animation continuum. Sequential display is only one animation method, but was used as the sole animation method in Study 2. Study 1 used fluid drawing and sequential display for animation methods, but these animation styles were found not to significantly differ in terms of learning, and so they can be considered functionally equivalent. However, perhaps the most common animation, "objects working/in motion" (e.g., Mayer & Gallini, 1990) such as pumps and generators, is conspicuously absent from this work. Lowe and Schnotz (2014) suggest that lessons where the instructional goal is to convey structure should use "build" animations similar to the ones used here, but process-oriented lessons should process animations, which convey motion and change. Because the lesson was purely structural or taxonomic in nature, the appropriate type of animation was selected, but these results may not generalize to process animations in process lessons. This is an important example of how constraints may change the effectiveness of animation.

In a similar vein, the lesson used here represents only a small subset of Biology knowledge, the structure and function of cell parts. Even among taxonomic lessons, there is diversity in the subject matter and approach to the lesson. Because only one lesson was used here in order to make studies comparable and control for information, future work should consider using different pools of knowledge which require different pedagogical approaches in order to achieve more generalizability. This study also limited itself to content created by Khan Academy for its ecological validity, but other sources of content should also be considered. Likewise, expanding this line of work outside of the domain of Biology would also increase generalizability, as different domains have been found to be impacted differently by animations (Höffler & Leutner, 2007) and other domains are likely to have different image type options to expand out this work; pictures and concept maps are not the only kinds of displays used in educational settings.

Additionally, the free recall analyses here require further investigation, as they were analyzed using only cursory word counting tools rather than a deep, qualitative approach to grading. As such, those results can say nothing about the correctness of the knowledge written there, and can only make statements about the way vocabulary terms are used. Therefore, deeper analysis of these responses should be conducted in order to more fully understand if labels affect only the presence of vocabulary or if they truly contribute to student free recall.

There are many future directions this work can take. First, more exploration of the differences between pictures and concept maps can now be done given that differences between the two have arisen. Each has its own affordances, which could play into strengths and weaknesses of each when it comes to different kinds of learning and knowledge. For instance, using color information for gestalt chunking could be manipulated with both concept maps and pictures, and has already been done with concept maps alone (Nesbit & Adesope, 2011).

Additionally, it would be interesting to focus in on concept maps and examine how the arrangement of the concept map affects learning outcomes. The map used in Study 1 was a loosely hierarchical map (in that there were not discrete levels below the highest node), but other map arrangements are possible, such as a more structured hierarchy or a radial map. The layout of the concept map may affect how it is read (an interesting notion for eye tracking) as well as how it affects learning, particularly in comparison to a picture control.

To follow up on the expertise reversal effect, lessons with controls for high prior knowledge students may help reduce some of the lower gains they experienced. This could be done by allowing them to control the speed of the information, or by using a different kind of lesson where they could selectively observe processes (Lowe & Schnotz, 2014). With greater control, higher prior knowledge students may be able to target those areas where they have missing information without being subjected to a "railed" lesson where they already know the majority of the content.

Another interesting manipulation could be done to the complexity and/or abstraction level of pictures. While the image used here is fairly simple in its unlabeled form, another kind of cell picture could be used which is more "realistic" and visually complex. Additionally, Lowe and Schnotz (2014) discussed a certain reluctance for graphics designers to abstract animation in the way that static pictures are abstracted; everything else is more likely to become less visually realistic before the way something moves is simplified for a lesson. Tracking how abstracting the animation affects students' knowledge about how something works would be an interesting follow up to the discussion of animation styles and affordances.

While the label effects found here were weak, it may be the case that label effectiveness is constraint-based, just like image or animation efficacy. More work could be done to explore how labels function in a different kind of lesson, such as a process lesson, or in another domain where labels may be more crucial, such as diagramming forces in physics. It is too early to completely write off labels as effective visual information, although this work has called into question the blind inclusion of labels.

In total, this work represents only a handful of the conditions under which media may be tested, and so generalization outside of some of these parameters may be dangerous. The goal of this work was to begin investigating how media perform under certain constraints, not to discover which media is generally "best" since such generalizations would likely be wrong. As such, there are still a number of avenues to explore with different media types and how animation affects them, and this work only sought to begin this process.

References

Adesope, O., & Nesbit, J. (2013). Animated and static concept maps enhance learning from spoken narration. *Learning and Instruction, 27*, 1–10.

Ainsworth, S., & Loizou, A. T. (2003). The effects of self-explaining when learning with text or diagrams. *Cognitive Science, 27*, 669 – 681.

Atkinson, R., Lin, L., & Harrison, C. (2009). Comparing the efficacy of different signaling techniques. In G. Siemens & C. Fulford (Eds.), *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2009* (pp. 954–962). Chesapeake, VA: AACE.

Azevedo, R., & Witherspoon, A. M. (2009). Self-regulated learning with hypermedia. In D. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Handbook of metacognition in education* (pp. 319–339). Mahwah, NJ: Erlbaum.

Baker, R. S. J., D'Mello, S. K., Rodrigo, M. M. T., & Graesser, A. C. (2010). Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive-affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies, 68*(4), 223–241.

Berthold, K., & Renkl, A. (2009). Instructional aids to support a conceptual understanding of multiple representations. *Journal of Educational Psychology, 101*(1), 70 – 87.

Bétrancourt, M. (2005). The animation and interactivity principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 287–296). Cambridge, UK: Cambridge University Press.

Bétrancourt, M., Bisseret, A., & Faure, A. (2001). Sequential display of pictures and its effect on mental representations. In J. F. Rouet, J. J. Levonen, & A. Biardeau (Eds.), *Multimedia learning: Cognitive and instructional issues* (pp. 112–118). Amsterdam: Elsevier Science.

Bétrancourt, M., Dillenbourg, P., & Montarnal, C. (2003). Computer technologies in powerful learning environments: the case of using animated and interactive graphics for teaching financial concepts. In E. De Corte, L. Verschaffel, N. Entwistle, & J. van Merrienböer (Eds.), *Unrevealing basic components and dimensions of powerful learning environments* (pp. 143-157). Oxford, UK: Elsevier.

Cade, W. L. (2012). *Attentional guiding through embodied and image cues in an intelligent tutoring system*. Unpublished manuscript, Department of Psychology, University of Memphis, Memphis, TN.

Cañas, A. J., Ford, K. M., Novak, J. D., Hayes, P., Reichherzer, T., & Suri, N. (2001). Online concept maps: Enhancing collaborative learning by using technology with concept maps. *The Science Teacher, 68*(4), 49–51.

Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology*, *62*(2), 233–246.

Clark, R. C., & Mayer, R. E. (2008). Learning by viewing versus learning by doing: Evidence-based guidelines for principled learning environments. *Performance Improvement, 47*(9), 5–13.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Craig, S.D., Gholson, B., & Driscoll, D.M. (2002). Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features and redundancy. *Journal of Educational Psychology*, *94*(2), 428–434.

Cromley, J. G., Perez, T. C., Fitzhugh, S. L., Newcombe, N. S., Wills, T. W., & Tanaka, J. C. (2013). Improving students' diagram comprehension with classroom instruction. *Journal of Experimental Education, 81*(4), 511–537.

Cromley, J. G., Synder-Hogan, L. E., & Luciw-Dubas, U. A. (2010). Cognitive activities in complex science text and diagrams. *Contemporary Educational Psychology, 35*, 59 – 74.

de Koning, B. B., Tabbers H. K., Rikers, R. M. J. P., & Paas, F. (2010). Attention guidance in learning from a complex animation: Seeing is understanding? *Learning and Instruction, 201*, 111–122.

Florax, M., & Ploetzner, R. (2010). What contributes to the split-attention effect? The role of text segmentation, picture labelling, and spatial proximity. *Learning and Instruction, 10*, 216 – 224.

Graesser, A.C., Lu, S., Jackson, G.T., Mitchell, H., Ventura, M., Olney, A., & Louwerse, M.M. (2004). AutoTutor: A tutor with dialogue in natural language. *Behavioral Research Methods, Instruments and Computers, 36*, 180–193.

Grosse, C. S., & Renkl, A. (2006). Effects of multiple solution methods in mathematics learning. *Learning and Instruction, 16*(2), 122–138.

Heimlich, J., & Pittelman, S. (1986). *Semantic mapping: Classroom applications*. Newark, DE: International Reading Association.

Höffler, T., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction*, *17*, 722–738.

Jamet, E., Gavota, M., & Quaireau, C. (2008). Attention guiding in multimedia learning. *Learning and Instruction*, *18*(2), 135–145.

Jeung, H. J., Chandler, P., & Sweller, J. (1997). The role of visual indicators in dual sensory mode instruction. *Educational Psychology, 3,* 329–343.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38*, 23–31.

Khan Academy. (2015). Khan Academy. Retrieved from https://www.khanacademy.org/

Kragten, M., Admiraal, W., & Rijlaarsdam, G. (2012). Diagrammatic literacy in secondary science education. *Research in Science Education, 43*, 1785–1800.

Leelawong, K., & Biswas, G. (2008). Designing learning by teaching agents: The Betty's Brain system. *International Journal of Artificial Intelligence in Education, 18*(3), 181–208.

Lowe, R. K., & Boucheix, J. M. (2011). Cueing complex animations: Does direction of attention foster learning processes? *Learning and Instruction, 21*(5), 650-663.

Lowe, R. K., & Schnotz, W. (2014). Animation principles in multimedia learning. In R. Mayer (Ed.), *Cambridge handbook of multimedia learning* (2nd ed.) (pp. 513–546). New York: Cambridge University Press.

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal. *Journal of Physiology - Paris, 102,* 59–70.

Maki, R. H. (1998). Test predictions over text material. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 117–145). Hillsdale, NJ: Erlbaum.

Mayer, R. (2005). Cognitive theory of multimedia learning. In R. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31–48). New York: Cambridge University Press.

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). New York, NY: Cambridge University Press.

Mayer, R. E., Dow, G., & Mayer, S. (2003). Multimedia learning in an interactive self-explaining environment: What works in the design of agent-based microworlds? *Journal of Educational Psychology, 95*, 806–813.

Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology, 82*(4), 715–726.

Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual coding theory of multimedia learning. *Journal of Educational Psychology, 86*(3), 389–401.

Moline, S. (2011). *I see what you mean* (2nd ed.). Portland, ME: Stenhouse Publishers.

Moreno, R., & Valdez, A. (2005). Cognitive load and learning effects of having students organize pictures and words in multimedia environments: The role of student interactivity and feedback. *Educational Technology Research and Development, 53*(3), 34–45.

Moreno, R. E., & Mayer, R. E. (2002). Learning science in virtual reality environments: Role of methods and media. *Journal of Educational Psychology, 94*, 598-610.

Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research*, *76*(3), 413–448.

Nesbit, J. C., & Adesope, O. O. (2011). Learning from animated concept maps with concurrent audio narration. *Journal of Experimental Education, 79,* 209–230.

Novak, J. D. (1991). Clarify with concept maps: A tool for students and teachers alike. *The Science Teacher, 58,* 45–49.

Novak, J. D. (1998). *Learning, creating, and using knowledge: Concept maps as facilitative tools in schools and corporations.* Mahwah, NJ: Lawrence Erlbaum Associates.

Novak, J. D., & Cañas, A. J. (2008). The theory underlying concept maps and how to construct and use them (Technical Report IHMC CmapTools 2006-01 Rev 2008-01). Pensacola, FL: Florida Institute for Human and Machine Cognition. Retrieved from http://cmap.ihmc.us/docs/theory-of-concept-maps

Novak, J. D., & Musonda, D. (1991). A twelve-year longitudinal study of science concept learning. *American Educational Research Journal*, *28*(1), 117–153.

O'Donnell, A. M., Dansereau, D. F., & Hall, R. H. (2002). Knowledge maps as scaffolds for cognitive processing. *Educational Psychology Review, 14,* 71–86.

Olney, A., D'Mello, S., Person, N., Cade, W. L., Hays, P., Williams, C., … & Graesser, A. (2012). Guru: A computer tutor that models expert human tutors. In S. Cerri, W. Clancey, G. Papadourakis, & K. Panourgia (Eds.), *Proceedings of the 11th International Conference on Intelligent Tutoring Systems* (pp. 256-261). Berlin: Springer-Verlag.

Ross, J., Irani, I., Silberman, M.S., Zalvidar, A., & Tomlinson, B. (2010). Who are the crowdworkers?: Shifting demographics in Amazon Mechanical Turk. In E. D. Mynatt, D. Schoner, G. Fitzpatrick, S. E. Hudson, W. K. Edwards, & T. Rodden (Eds.), *Computer-Human Interaction Extended Abstracts* (pp. 2863–2872). New York: ACM.

Schnotz, W. (2005). An integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 49-69). Cambridge, UK: Cambridge University Press.

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representations. *Learning and Instruction, 13*(2), 141–156.

Schnotz, W., & Cade, W. (2014). Adaptive multimedia environments. In R. Sottilare, A. Graesser, X. Hu, & B. Goldberg (Eds.), *Design recommendations for adaptive intelligent tutoring systems: Volume 2 - Adaptive instructional strategies*. Orlando, FL: U.S. Army Research Laboratory.

Schnotz, W., & Rasch, T. (2005). Enabling, facilitating, and inhibiting effects of animations in multimedia learning: Why reduction of cognitive load can have negative results on learning. *Educational Technology Research and Development, 53*(3), 47-58.

Spanjers, I.A.E., Wouters, P., van Gog, T., & van Merrienboer, J.J.G. (2011). An expertise reversal effect of segmentation in learning from animated worked-out examples. *Computers in Human Behavior, 27*, 46–52.

Stull, A. T., & Mayer, R. E. (2007). Learning by doing versus learning by viewing: Three experimental comparisons of learner-generated versus author-provided graphic organizers. *Journal of Educational Psychology, 99*, 808–820.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*, 257–285.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory: Explorations in the learning science, instructional systems, and performance technologies*. New York: Springer.

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

Tversky, B., Bauer Morrison, J., & Bétrancourt, M. (2002). Animation: can it facilitate? *International Journal on Human Computer Studies, 57*, 247–262.

Twyford, J., & Craig, S. (2013). Virtual humans and gesturing during multimedia learning: An investigation of predictions from the temporal contiguity effect. In T. Bastiaens & G. Marks (Eds.), *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2013* (pp. 2145-2149). Chesapeake, VA: Association for the Advancement of Computing in Education (AACE).

Williams, B., Williams, C., Volgas, N., Yuan, B., & Person, N. (2010). Examining the role of gestures in expert tutoring. In V. Alevan, J. Kay, & J. Mostow (Eds.), *Proceedings of the 10th International Conference on Intelligent Tutoring Systems*, (pp. 235–244). Berlin Heidelberg: Springer-Verlag.

Wittgenstein, L. (1971). *Philosophical investigations*. New York: MacMillan.

Appendix A

Conditions from Study 2

Appendix B

Node and Link Questions (Study 1 and 2)

Test A

LQ1) The cell is __________ a cellular membrane.
produced by
surrounded by
exterior to
turned into

LQ2) DNA is __________ mRNA.
a type of
produced by
destroyed by
transcribed into

LQ3) DNA is __________ the nucleus.
exterior to
transported by
surrounded by
on top of

LQ4) The nucleus is __________ eukaryotes.
present in
absent in
turned into
made of

LQ5) The nucleus __________ the nucleolus.
contains
is produced by
is part of
is contained within

LQ6) The nucleuolus __________ RNA.
destroys
produces
repackages
transports

LQ7) Proteins are __________ ribosomes.
made of
modified by
destroyed by
produced by

LQ8) Ribosomes are __________ "rough" ER.
part of
destroyed by
produced by
surrounded by

LQ9) Vacuoles __________ storage.
destroy
transport
function as
limit

NQ1) __________ dissolves organelles.
Cytoplasm (cytosol)
Golgi bodies
Lysosomes
Ribosomes

NQ2) DNA is turned into __________.
ribosomes
lysosomes
mRNA
proteins

NQ3) The nucleus is not present in __________.
eukaryotes
plants
fungi
prokaryotes

NQ4) Eukaryotes consist of __________, __________, and __________.
animals, plants, fungi
archaea, bacteria, viruses
bacteria, fungi, viruses
animals, mushrooms, fungi

NQ5) Prokaryotes consist of __________ and __________.
animals, plants
archaea, bacteria
plants, bacteria
archaea, protists

NQ6) __________ produce vesicles.
Golgi bodies
Ribosomes

Nucleoli
Vacuoles

NQ7) __________ transport proteins.
Vesicles
Mitochondria
Endoplasmic reticula
Cellular membranes

NQ8) Endoplasmic reticulum consists of __________ and __________.
normal ER, abnormal ER
smooth ER, rough ER
endoplasmic ER, nonendoplasmic ER
synthetic ER, authentic ER

NQ9) __________ produce energy/ATP.
Endoplasmic reticulum
Mitochondria
Golgi bodies
Proteins

Test B

LQ1) Lysosomes __________ organelles.
dissolve
strengthen
create
produce

LQ2) DNA is __________ proteins.
part of
created by
turned into
made of

LQ3) The nucleus is __________ prokaryotes.
part of
surrounded by
produced by
not present in

LQ4) Eukaryotes __________ animals, plants, and fungi.
mature into
evolved from
consist of
caused extinction in

LQ5) Prokaryotes ___________ Archaea and bacteria.
transport
are a type of
consist of
caused extinction in

LQ6) Golgi bodies ___________ vesicles.
are a part of
produce
dissolve
modify

LQ7) Vesicles ___________ proteins.
create
destroy
modify
transport

LQ8) Endoplasmic reticulum ___________ "rough" and "smooth" ER.
produced by
dissolved by
contained within
consists of

LQ9) Mitochondria ___________ energy/ATP.
produce
destroy
are a type of
repackage

NQ1) The cell is surrounded by ___________.
nucleolus
cellular membrane
nuclear membrane
cytoplasm (cytosol)

NQ2) DNA is transcribed into ___________ .
mRNA
proteins
tRNA
ribosomal RNA

NQ3) ___________ is surrounded by a nucleus.
DNA
mRNA

The cell
A vacuole

NQ4) The nucleus is present in ___________.
the nucleolus
eukaryotes
prokaryotes
all cells

NQ5) The ___________ contains a nucleolus.
nucleus
Golgi body
endoplasmic reticulum
prokaryote

NQ6) The nucleolus produces ___________.
the nucleus
DNA
RNA
ribosomes

NQ7) Proteins are produced by ___________.
lysosomes
Golgi bodies
mitochondria
ribosomes

NQ8) ___________ are part of the "rough" ER.
Lysosomes
Vacuoles
Ribosomes
Vesicles

NQ9) ___________ function as storage.
Lysosomes
Golgi bodies
Vacuoles
Nucleoli

Appendix C

State Examination Multiple Choice Questions

Test A

MC1) Hair is mostly protein. Which organelle would be much more abundant in an animal cell that produces hair than in an animal cell that stores fat?

Chloroplast
Mitochondrion
Nucleus
Ribosome

MC2) The long, thin, string-like molecules located primarily in the nucleus of eukaryotic cells are known as...

DNA
RNA
Genes
Chromosomes

MC3) Which structure is outside the nucleus of a cell and contains DNA?

Chromosome
Gene
Mitochondrion
Vacuole

MC4) The function of one cell organelle is to produce energy. What is the name of this organelle?

Golgi body
Mitchondrion
Nucleus
Ribosome

MC5) In the human body, the circulatory system transports and delivers substances. Within the cell, which organelle performs a similar function?

Nucleus
Golgi body
Mitochondrion
Endoplasmic reticulum

MC6) How is the prokaryotic bacterium different from a eukaryotic cell?

It has ribosomes to make proteins.
It stores its genetic information in DNA.
It has no membrane-bound nucleus.
It has a cell membrane.

MC7) Which of these is most responsible for carrying coded information from the nucleus?

Cell membrane
Ribosomes
mRNA
ATP

MC8) Which cellular organelle is responsible for packaging the proteins that the cell secretes?
Cytoskeleton
Cell membrane
Lysosome
Golgi body

MC9) Golgi body is to vacuole as packaging is to...
Protecting
Storing
Absorbing
Hydrating

MC10) Which of these is responsible for the "rough" • appearance of endoplasmic reticulum?
DNA
Enzymes
Lysosomes
Ribosomes

MC11) Specialized proteins control cell division in amoebas. Which cell part is responsible for making these proteins?
Mitochondrion
Nucleus
Pseudopod
Ribosome

MC12) A cell from heart muscle would probably have an unusually high proportion of...
Lysosomes
Mitochondria
mRNA
Golgi bodies

MC13) Which is the most accurate description of a eukaryotic cell?
Moves using cilia
Contains a nucleus
Produces food by photosynthesis
Reproduces only by binary fission

Test B
MC1) The building of proteins from amino acids occurs on the cell's...

Membrane
Ribosomes
Nucleus
Centriole

MC2) Which organelle produces proteins?
Nucleus
Lysosome
Ribosome
Golgi body

MC3) Which of the following organelles releases energy from sugars?
Ribosomes
Vacuoles
Chloroplasts
Mitochondria

MC4) The genetic information for making a protein must move from the nucleus to the cytoplasm. Which of these moves this information to the cytoplasm?
Ribosome
DNA
RNA
Amino acid

MC5) The part of a eukaryotic cell that allows it to remain separate from the outside environment is the...
Cell membrane
Ribosome
Cytoplasm
Golgi vesicles

MC6) What repackages proteins into forms the cell can use, expel, or keep stored?
Lysosome
Mitochondria
Golgi bodies
Centrioles

MC7) Proteins must enter the endoplasmic reticulum to be...
transported to other parts of the cell.
used in building new strands of RNA.
synthesize into new genetic codes.
excreted as waste material.

MC8) A particular toxin prevents cellular production of usable energy. Cells that are affected by this toxin are unable to carry out many of their normal functions. Which of these organelles would be most directly harmed by this toxin?

Ribosomes
Nucleus
Mitochondria
Vacuole

MC9) Compared to a skin cell, a muscle cell is likely to have more...
Golgi bodies
Mitochondria
Cell membranes
Chloroplasts

MC10) Under a microscope, a series of cells are observed that lack a membrane-bound nucleus. Which of these is the most likely cell type?
Plant cell
Animal cell
Eukaryotic cell
Prokaryotic cell

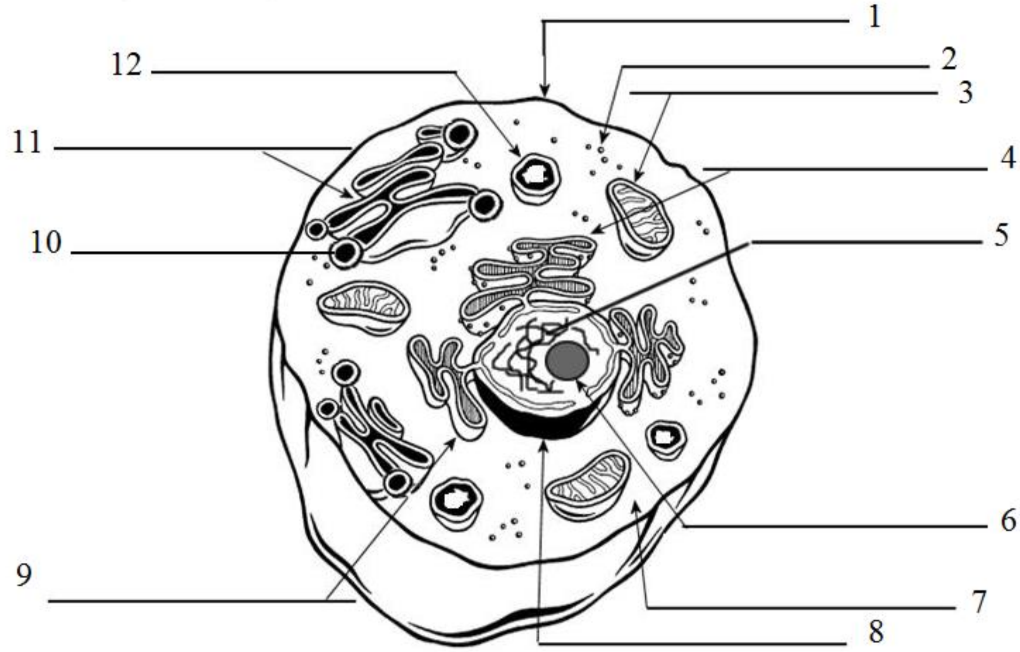
MC11) How is the storage of DNA in eukaryotic cells different from in prokaryotic cells?
Prokaryotic cells have a capsule around the DNA
Eukaryotic cells have DNA stored in the nucleus
Prokaryotic cells have DNA stored in a central vacuole
Eukaryotic cells have DNA free-floating in the cytoplasm of the cell

MC12) The outer surface of the endoplasmic reticulum may be smooth or rough. Which cell structures cause the outer surface of the endoplasmic reticulum to appear rough?
Ribosomes
Transport proteins
Mitochondria
Golgi bodies

MC13) Which best explains why a bacterial cell is classified as a prokaryote?
The presence of a nucleus
The absence of a nucleus
The presence of a cell wall
The absence of a cell wall

Appendix D

Picture Labeling Task



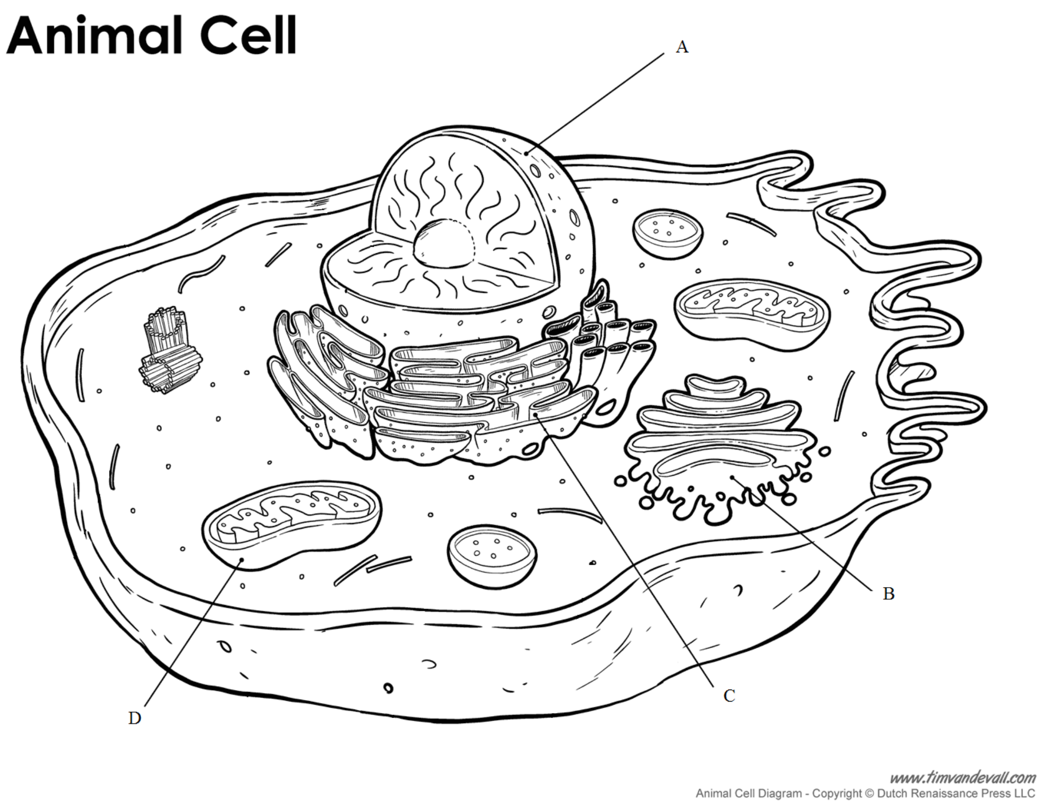
1. Cell Membrane
2. Ribosomes
3. Mitochondrion
4. Rough ER
5. DNA
6. Nucleolus
7. Cytosol
8. Nucleus
9. Smooth ER
10. Vesicle
11. Golgi Body/Apparatus
12. Lysosome

Image was adapted from BiologyCorner.com (http://www.biologycorner.com/anatomy/cell/chap3_notes.html). It is free to share under the Creative Commons Attribution-Noncommercial 3.0 Unported license (http://creativecommons.org/licenses/by-nc/3.0/deed.en_US).

Appendix E

Function Identification Task

Example Question



*1. Which organelle has the function of producing energy?

A
B
C
D*

Additional questions (using same picture but different answer choice lines)
* = question adapted from state tests

*2. Which structure stores most of the genetic information? A: nucleus C

3. Which organelle defines the cell as a compartment? A: cell membrane B

*4. Which cell structure directs cell activities? A: DNA A

*5. Which cell part is responsible for making proteins? A: ribosome B

6. What part of the cell produces the ribosomal RNA? A: nucleolus A

7. Which cell structure produces hormones and fatty compounds? A: smooth ER D

8. Which cell structure transports proteins? A: rough ER C

9. Which part of the cell moves proteins to the outside of the cell? A: vesicle B

10. What cell structure dissolves organelles? A: lysosome C

11. What organelle creates membranes to transport proteins through the cell? A: Golgi body A

Appendix F

Motivation Questions

All questions below are on a 6-point Likert scale, ranging from Strongly Disagree (score of 1) to Strongly Agree (score of 6). Items with an asterisk in front of them are reverse coded.

1. *I found the video to be boring.

2.* I feel frustrated after watching that video.

3. I feel like I learned a lot from the video.

4. The video sparked my interest in learning more about Biology.

5. I found the video to be easy to follow.

6. I put a lot of effort into this activity.

7. *I found the tests in this study difficult.

Appendix G

Informed consent

INFORMED CONSENT

The Mechanical Turk task you have selected is part of a research project that is described below. Please read carefully and feel free to ask questions. You may wish to print this consent form for your future reference.

Your participation in this research study is voluntary. You are also free to withdraw from this study at any time. In the event new information becomes available that may affect the risks or benefits associated with this research study or your willingness to participate in it, you will be notified so that you can make an informed decision whether or not to continue your participation in this study.

For additional information about giving consent or your rights as a participant in this study, please feel free to contact the IRB at 901-678-2533 or email irb@memphis.edu.

You are being asked to participate in this research study to help us investigate Biology education.

You will first be asked to demonstrate your current knowledge on a variety of topics. Then you will see a video lesson. You will then be asked to write about the video and take a short knowledge test. Afterwards, you will be asked to provide some demographic information.

This study should take no longer than 1 hour to complete. Beyond your time, effort, internet connection, and depreciation on your computer, there are no expected costs for participating in the study.

There are no anticipated sources of inconvenience or risk other than those associated with sitting in front of a computer. The U of M does not have a fund set aside for compensation in the case of study-related injury.

The potential benefit to you from this study is you might learn something about various academic topics from viewing these videos. Society will benefit as we use the judgments you make to create better educational systems.

If you should have any questions about this research study or possible injury, please feel free to contact Andrew Olney at aolney@memphis.edu. For questions regarding the research subjects' rights, the Chair of the Institutional Review Board for the Protection of Human Subjects should be contacted at 901-678-2533.

All efforts, within the limits allowed by law, will be made to keep the personal information in your research record private but total privacy cannot be promised. We will anonymize your data before storing it on our computers. However, your information may be shared with U of M or the government, such as the University of Memphis University Institutional Review Board, Federal Government Office for Human Research Protections, Institute of Education Sciences, Institute for Intelligent Systems, and the Department of Psychology, if you or someone else is in danger or if we are required to do so by law.

STATEMENT BY PERSON AGREEING TO PARTICIPATE IN THIS STUDY

I have read this informed consent document and the material contained in it has been explained to me verbally. I understand each part of the document, all my questions have been answered, and I freely and voluntarily choose to participate in this study.

By clicking this ">>" button, I agree to the above terms.