Using Time-Compression To Make Multimedia Learning More Efficient: Current Research and Practice

By Raymond Pastore, University of North Carolina Wilmington and Albert D. Ritzhaupt, University of Florida

Abstract

It is now common practice for instructional designers to incorporate digitally recorded lectures for Podcasts (e.g., iTunes University), voice-over presentations (e.g., PowerPoint), animated screen captures with narration (e.g., Camtasia), and other various learning objects with digital audio in the instructional method. As a result, learners are spending more time learning from audio-enhanced digital learning materials for both formal and informal purposes. In this paper, we present digital time-compression as a way to reduce the amount of time learners will spend on a learning task, while still maintaining acceptable intelligibility, pitch, and scores on important dependent measures (e.g., recall, recognition, comprehension, satisfaction). Research dating back to the 1950s is reviewed and framed in the context of multimedia learning environments. Recent research developments are reviewed and a discussion is provided emphasizing several design principles for this technology. Recommendations for future research are provided.

Keywords: time-compression, multimedia learning, research, design.

Introduction

The use of audio (e.g., narration, music, sounds) in learning environments comes with a serious design implication that is often ignored by designers. Audio, unlike its textual counterpart, is inherently time-dependent and sequential (Barron, 2004). That is, learners listening to a narrated production are required to move at the speed of the narrator. This creates a situation where text-based instruction takes significantly less time to complete than audio. What can we do to make audio-enhanced learning environments more efficient? Some have proposed the use of time-compression to reduce the amount of time learners spend listening to a narration. A lofty goal of an instructional designer is to maximize a learner's comprehension and satisfaction, while minimizing the amount of time a learner will spend on a learning task (Ritzhaupt, Gomes & Barron, 2008). This poses an interesting instructional design and research problem.

Some previous research shows that speech typically takes place at approximately 150 words per minute (wpm) (Benz, 1971; Nichols & Stevens, 1957). In speech, one is simultaneously listening and composing speech. Because one can speak at approximately 150 wpm, and the rate for speed reading is 250 to 300 wpm (Taylor, 1965) and the rate for silent reading is 275 to 300 wpm (Junor, 1992), it is reasonable to hypothesize that another 125 to 150 wpm of unused processing capacity might be available for listening to normal speech. This hypothesis has been studied and tested by researchers under a variety of conditions starting as early as the 1950s (e.g., Barabasz, 1968; Fairbanks, Guttman, & Miron, 1957; Goldhaber, 1970; Jester & Travers, 1967; Reid, 1968; Richaume, Steenkeste, Lecocq, & Moschetto, 1988).

Prior Research on Time-Compressed Speech

Fairbanks, Guttman, and Miron (1957) successfully executed one of the first major studies that investigated the effects of time-compressed speech. They used two technical messages on the subject of meteorology in their intervention. The passages of words were recorded at 141 wpm with compression levels of 30%, 50%, 60%, and 70%; the last produced speech at 470 wpm. Results showed significant differences with the largest gaps in comprehension after approximately 282 wpm.

Jesters and Travers (1967) designed and executed a study with speech speed, repetition and presentation patterns as the independent variables. Speech passages were recorded at varying speeds (200 to 350 wpm) of the same content. Presentation patterns refer to variations of sequencing the passages at different speeds. One condition progressively increased the rate from the slowest presentation to the fastest, the second decreased from the fastest to the slowest, and the third condition kept the speeds constant at approximately 263 wpm. At the end of four trials, there were significant main effects on speech speed, but the interaction effect between presentation pattern and speech speed was insignificant.

Foulke (1968) executed a study with 12 groups based on increasing 25 wpm increments from 125 to 400 wpm. After listening to the speech, participants were tested for comprehension by a multiple choice test. Comprehension did not seriously deteriorate by increasing word rate from 125 to 250 wpm, but it declined rapidly thereafter. Foulke (1968) suggests that time is required for the perception of words, and that as word rate is increased beyond a certain point, the perception time available to the listener becomes inadequate, and a rapid decline of listening comprehension commences after that point.

Barabasz (1968) conducted a study with 118 students in a human behavior and development class. Two lectures were used in a rotational research design to control for inter-group differences. The research investigated two different speeds and used both recall (administered after lecture) and retention (administered two weeks later) as dependent measures. The findings suggest that a lecture can be reduced to one-third the time without a significant difference in either recall or retention (Barabasz, 1968) or approximately 225 wpm.

Goldhaber (1970) studied the effects of compressed speech as a function of academic grade level. The study looked at speech delivered at 165 wpm and 330 wpm for students in junior high school (80) and college (80), with comprehension as the dependent measure. The narrative content was adjusted according to the Flesh Readability Formula (Flesh, 1949). The results showed main effects for speech and academic level, but no interaction effect was identified. This indicates individuals with varying levels of formal education perform differently (high school versus middle school), as one would anticipate.

Reid (1968) studied the effects of grammatical complexity and compressed speech on comprehension. He used a form of the Nelson-Denny Reading Test to make two difficulty levels of grammatical complexity and compressed speech at 175, 275, 325, and 375 wpm. Further, the Verbal Scholastic Aptitude Test was used as a covariate. Results suggest a significant main effect for both compressed speech and grammatical complexity and a significant interaction effect. Compressed speech was not statistically significant until 375 wpm level, which is more than double the speed of normal speech.

Short (1977, 1978) conducted an applied time-compression study in the context of a Food and Nutrition course with 90 students using a self-instructional method. The study compared students in groups that used recorded lectures on tapes with variable rate controlled speech (VRCS) compressors and the same tapes on normal speed (NS) tape recorders. Students who used VRCS compressors had an average time saving of 32 percent and an average grade increase of 4.2 points on posttest scores, indicating the group with the accelerated treatment actually performed better.

Richaume, Steenkeste, Lecocq, and Moschetto (1988) examined the effects of normal and compressed speech at 135, 202, 270, and 300 wpm on intelligibility and comprehension. Combining the results from three experiments, their findings suggest that intelligibility and comprehension do not decay until approximately 300 wpm is reached. The study also considered the complexity of the narrated stories. Their findings suggest that the poorest scores resulted from difficult stories and highest scores from the concrete and redundant stories, an indication that type of content moderates the effects.

The findings of these various research studies suggest that speech speeds somewhere near 275 wpm or more begin to negatively influence the dependent measures of interest (e.g., comprehension, recall, etc.) (Fairbanks, Guttman, & Miron, 1957; Foulke, 1968; Reid,1968). These studies also underscore control variables that may influence the dependent measures of interest, such as academic level (Goldhaber, 1970), grammatical complexity (Reid, 1968), or repetition (Jester & Travers, 1967). However, these previous research studies did not study the effects of time-compressed speech in the context of multimedia (with both pictures and words) learning environments or complex learning. This is a major limitation of the previous research and has led to a new set of studies in the context of multimedia learning environments.

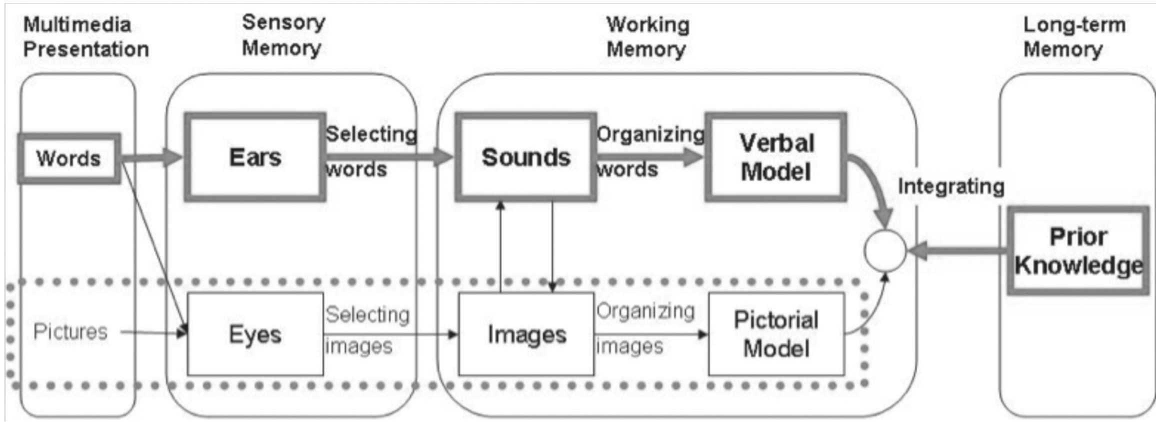

Figure 1. Modified cognitive model for multimedia learning representing previous research in time-compression (adapted from Ritzhaupt & Barron, 2008; Mayer, 2001).

Multimedia Learning Environments

Research on multimedia learning has evolved from simple media comparison studies to the basis of explaining the psychology of learning. Previous research in multimedia focused on the medium used for delivery rather than the instructional interventions that positively influence learning (Clark, 1983). This fundamental shift in research gave rise to cognitive theories in multimedia. Cognitive theories of multimedia learning share a few related theoretical underpinnings: sensory modality (input) and memory, working memory, limited-capacity and cognitive load, long-term memory, and dual-processing (Schnotz & Bannert, 1999; Mayer, 2001; Hede, 2002; Schnotz, 2005). Mayer (2001) provides the cognitive theory of multimedia learning, Schnotz and Bannert (1999, 2005) provide the integrated model of text and picture comprehension, and Hede (2002) outlines the integrated model of multimedia.

Based on the research literature, it would appear that Mayer's cognitive theory of multimedia learning has been most widely accepted and integrated model to explain the phenomena. Figure 1 visualizes the gap in research based on Mayer model. The perforated lines surrounding the visual/pictorial channel is not activated in prior research on time-compressed speech, while the auditory/verbal channel is often overloaded. Mayer's multimedia model is based on three tenets: dual channels, limited capacity, and knowledge construction. The first tenet, dual processing, suggests that humans have multiple separate channels for processing visual/pictorial and auditory/verbal information (Mayer, 2003). The second tenet suggests that humans' processors have a limited capacity to process information at any given instance in time. The third tenet is that humans are knowledge constructing processors that receive, organize, and connect incoming information with existing knowledge (Mayer, 2003).

"The process of meaningful learning from multimedia involves five cognitive processes: selecting words, selecting images, organizing words, organizing images, and integrating" (Mayer, 2003, p. 304). The model suggests that when a learner engages in a multimedia presentation, information is presented as either words or pictures. The next step in the model is sensory memory, in which the words, figures, animations, narration, and sounds impinge the eyes and ears of learners, who then selectively store the information in working memory. If the information is organized in working memory by the learner coherently representing sounds and images and connecting it with prior knowledge, an "integrated learning outcome" results (Mayer, 2003, p. 304).

Much of the time-compressed speech research predates the growth in multimedia learning research literature. From a theoretical perspective, speech or narration is effectively the same treatment as words communicated through an auditory channel. Rather, the tenants of multimedia learning provide a coherent framework and perspective with which to systematically investigate time-compressed speech. Research conducted in this manner can integrate knowledge and serve a multi-disciplinary audience.

Previous research has shown the combination of words and pictures leads to better learning than from words alone (Mayer & Gallini, 1990; Clark & Pavio, 1991; Pavio, 1986; Pavio, 1990). Further, it has been long established that a person's memory for pictures is better than memory for words alone (McDaneial & Pressley, 1987; Pavio, 1986; Standing, Conezio, Haber, 1970). This knowledge suggests that time-compressed speech should not be studied in isolation, but by including pictures as our previous research demonstrates doing so is a stronger instructional method. While researchers have known this information for more than 20 years, the combination of pictures and time-compressed audio has not been systematically studied until the past six years (e.g., Pastore, 2009; Pastore, 2012; Ritzhaupt, Gomes & Barron, 2008; Ritzhaupt & Barron, 2008; Ritzhaupt, Barron, & Kealy, 2011).

Under conditions of time-compressed audio, the presentation of a relevant picture may be able to represent verbal information and by doing so, provide the additional nonverbal memory representation that can be retrieved from memory if an individual's verbal information is inaccessible (Kullhavey, Lee & Caterino, 1985; Pavio, 1986). This can be explained by a referential process between the verbal and nonverbal information. Of particular importance is the strength of the relationship between the relevant picture and words used in multimedia materials. For instance, an individual listening to a speech about the history of the Chinese government while viewing a picture of a German flag is semantically incongruent, and according to theories of multimedia learning, may interfere with the learning process. Related information should be more easily accessible in memory than unrelated information (Kullhavey, Lee & Caterino, 1985).

Time-Compression Technology

Early time-compression technology was based on playing back an audio recording at a faster speed than the original recording, which changed the tempo and pitch of the sound. This technique, though functional and easy to produce, resulted in the chipmunk effect, in which the vocal effect and intelligibility were adversely effected (Barron, 2004). Consequently, there was a desire to improve the quality of the time-compressed audio, while preserving the quality of the pitch and intelligibility to create a more enjoyable audio experience. The act of changing the tempo while preserving the pitch is referred to as the invariant timing hypothesis (Honing, 2006). This hypothesis states that one cannot distinguish between the quality of audio when there is a tempo change. This hypothesis makes time-compressed instruction a feasible method of instruction. The next iteration of analog time-compression technology involved removing small segments of the speech signal (Miller & Lichlinder, 1950). The Fairbanks method, for instance, would remove small portions of the signal at regular intervals (Barron, 2004), resulting in an audio recording requiring substantially less time to complete, but with reasonable quality.

Today, time-compression technology has evolved from analog format to one of a digital nature. More importantly, the technology is real-time: audio content can be manipulated by a learner while the audio is playing. This makes the technology much easier to use since the learners do not have to re-record the content at a faster or slower rate. The key digital technology that supports the increased or decreased playback of audio content involves time-compression algorithms. These sophisticated algorithms fall into two broad categories: linear and non-linear. Linear time-compression applies a consistent manipulative to the entire audio content, irrespective of the information in the audio recording. Short and fixed-length speech segments (called audio gaps) are discarded, and the retained segments are then abutted after cross-correlation (averaging the edges of audio frames before abutting) to diminish the effects of abrupt audible noises (He & Gupta, 2001). The result reduces the remaining audio segments by equal proportions.

Non-linear time-compression is more sophisticated than linear time-compression technology. Non-linear time-compression will first analyze the audio content, and compress based on the type of content recorded. Typically, non-linear time compression involves compressing redundancies in audio, including but not limited to pauses or elongated vowels in an audio stream (He & Gupta, 2001). Consequently, compression rates may vary from one point to another in the audio stream. Adaptive and hybrid algorithms including both techniques have been developed in more recent years, and have been successfully integrated into pervasive consumer products. Thankfully, digital time-compression algorithms have been built into common software packages, including tools like Audacity (open source audio manipulation utility) and Windows Media Player (common audio player on Windows operating systems). These tools can be used by designers and learners alike to minimize the amount of time spent engaging in multimedia materials. Other tools such as Sony Sound Forge also have built in time-compression algorithms. These tools should, however, be used with caution. Although there are many different methods and algorithms currently being used to compress speech, one has yet to stand out above the rest because they all reach a ceiling effect and produce equal quality (this refers specifically to the more recent methods using the software described) (Janse, Nooteboom, & Quene, 2001).

Time-Compression in Multimedia Learning Research

Since 2008, several studies have examined the use of time-compression in multimedia learning materials (Pastore, 2010; Pastore, 2012; Ritzhaupt, Gomes & Barron, 2008; Ritzhaupt & Barron, 2008; Ritzhaupt, Barron, & Kealy, 2011). These studies are framed by cognitive theories of multimedia learning and have used primarily experimental procedures to examine time-compression in the context of a multimedia learning treatments. As can be seen in Table 1, the results varied across these studies; however, some principles can be gleaned from this new body of knowledge. The remainder of this section will highlight the various studies in the past six years that have rigorously studied time-compression and other relevant variables. Across these studies, the authors expressed the time-compression using different metrics, which can be confusing, thus we provide Table 1 as a guide to interpret the time compression speeds in terms of acceleration speed, compression percentage, words per minute (wpm), and time for a 10-minute multimedia presentation.

In 2008, Ritzhaupt, Gomes, and Barron studied the relationship between verbal redundancy and time-compression in audio-enhanced multimedia learning materials. In the study, 183 undergraduates were randomly assigned to one of three audio-enhanced multimedia presentations that were recorded at three speeds (1.0, 1.4, and 1.8). The dependent measures in the study were comprehension and satisfaction. Time-compression was treated as a between-subjects effect, verbal redundancy was treated as a within-subjects effect. The results show no significant difference on performance across treatments (suggesting that one can accelerate the speed of narration without adversely influencing comprehension) and a significant difference on satisfaction in favor of 1.4 times the normal audio speed. The results also indicate statistical differences in favor of verbal redundancy, in which the same verbal information was presented on both an auditory and visual channel. One limitation of this study is that it did not isolate the effects of the pictorial information, which led to a sequence of two more studies.

Table 1. Acceleration speed, compression percentage, words per minute and time for a 10-minute multimedia production.

Acceleration	Compression	WPM	Time
1	0%	150	10
1.5	25%	225	6.67
2	50%	300	5
2.5	75%	375	4

In a follow up study, Ritzhaupt and Barron (2008) investigated the effects of time-compressed narration and representational adjunct images on a learner's ability to recall and recognize information as well as learner satisfaction. The experiment included 305 research participants in a 4 Audio Speeds (1.0 vs. 1.5 vs. 2.0 vs. 2.5) × Adjunct Image (Image Present vs. Image Absent) factorial design. The results showed statistically significant differences at 2.5 times the normal audio speed, in which performance on cued-recall and content recognition tasks was significantly lower than other audio speeds (Ritzhaupt & Barron, 2008). The presence of representational adjunct images had a significant positive effect on cued-recall, but not content recognition. The participants in the normal audio speed and image present groups were significantly more satisfied than those in other treatments.

Pastore (2010) examined the effects of time-compressed instruction and visual representations in a multimedia environment on recall, problem solving, and cognitive load. 216 university students were placed in Time Compression (0%, 25%, and 50%) x 2 Visual (Visual and No-Visual) treatments. Each treatment was presented in a multimedia environment with static visuals and narration on the human heart and its parts, which originally developed by Dwyer and Lamberski (1983). For the factual knowledge assessment, there was no significant difference in the 0% and 25% compression groups. These groups performed significantly better than the 50% compression group. Additionally, the with visual groups performed better than the with non-visual groups. The problem solving measure indicated that learning was not suppressed at 25%, when visuals were present, and was significantly lower at 50% regardless of presentation method. The cognitive load measure revealed that those presented 0% and 25% compression speeds indicated lower levels than those presented 50%. Additionally, learners presented visuals indicated lower levels than those presented no visuals. This study suggests that time-compression can be used to present complex material and retain problem solving knowledge as long as it is presented in a multimedia environment.

Ritzhaupt, Barron, and Kealy (2011) investigated why verbal recall of time-compressed narration is significantly enhanced when it is accompanied by a representational adjunct picture. They explored the potential of the Conjoint

Retention Hypothesis (CRH) as an explanation, which posits that mentally stored visual information can serve as a secondary retrieval cue that boosts recall of related verbal material. Four groups of participants ($N = 153$) listened to a compressed audio narration at different rates of speed accompanied by visuals, 50% of which were pictorially-related and 50% of which were pictorially-unrelated. Their results show the type of information significantly influenced the recall, but not the recognition performance. While CRH provides the most feasible explanation for the increased recall, the generative-recognize view best explains the differences between recognition and recall performance.

Pastore (2012) examined the effects of redundancy on learning (recall and problem solving) and learner's perceptions of cognitive load. 154 adult learners were given instruction on the human heart that was presented at speeds of 0%, 25%, or 50% and consisted of either narration only or narration and redundant text. The study did not find significant differences between the 0% and 25% conditions on factual, problem solving, or cognitive load measures. They were both (0% and 25%) significantly different from the 50% compression group. Participants presented only narration outperformed participants presented redundant narration and text. This study reaffirms the results from Pastore (2010) and reveals that presenting time-compressed instruction with 100% redundant text inhibits learning.

Most recently, Pastore (In review) sought to explore learners' perceptions of time-compressed speech in order to help determine what speeds they would prefer when listening for entertainment and learning purposes. Participants were presented with a 30-question survey concerning their preferences towards compressed speech. The questions asked how much they preferred speech compressed at 0%, 10%, 20%, 30%, 40%, and 50% both with and without visuals. Participants indicated that they preferred images to no images when speech was compressed. Surprisingly, participants consistently preferred the 10% compression speed. These results indicate that while learners can learn at higher speeds, they only prefer low compression speeds around 10%.

As was demonstrated in this literature review, the results of these studies vary but there are many common findings across studies. For compression speeds, it appears that audio can be compressed up to 2.5 times (Ritzhaupt and Barron, 2008) or 75% when presenting learners with declarative information. However, when presenting complex content (i.e., problem solving or high level knowledge) instruction can only be presented at about 1.5 times or 25% compression speeds (Pastore, 2009) and should be in a multimedia environment, otherwise learning will be inhibited and quickly diminish as compression increases. For satisfaction, studies show that learners do not prefer high speeds of compression. Ritzhaupt, Gomes, and Barron (2008) found that learners preferred speeds around 1.4 times (20% compression speed) and Pastore (in review) found that learners preferred 1.2 times (10%). As a result of these studies, a series of design principles, discussed in the next section of this paper, have been developed.

Design Principles: Time Compression in Multimedia Learning

A goal in writing this paper is for instructional designers to consider using digital time-compression to reduce the amount of time learners will spend on a learning task while not depressing important learning outcomes (e.g., recall, recognition, problem solving, cognitive load, satisfaction). Whether you believe it or not, this technology has already been widely adopted in advertising and most people have heard time-compressed audio as broadcasters compress audio clips to increase the number of advertisements they can air on their programs (e.g., Moore, Hausknecht, & Thamodaran, 1986). Thus, the implications of this technology and its use are large. The time savings alone is reason enough to consider compressing narration. Even a 5% compression of narration across a company training many people could lead to significant dollars saved. As a result, this paper aims at the practitioner in our field to consider using this technology in their products, but with caution. We have, therefore, developed a set of eight design principles to guide instructional designers on the use of this technology based on our previous research.

1.0 Use Visuals with Time-Compression

Across all of the recent studies on the use of time-compression in multimedia learning environments, the availability and use of a semantically-related visuals by learners had a positive effect on learning outcomes (Pastore, 2009; Pastore, 2012; Ritzhaupt, Gomes & Barron, 2008; Ritzhaupt & Barron, 2008; Ritzhaupt, Barron, & Kealy, 2011). This finding is unsurprising as we have a long history of observing the effects of meaningful visuals on learning. This finding attests to the durability and strength of the multimedia effect and modality effect on learning outcomes (Mayer, 2001; Mayer, 2003), even under time-compression conditions. First, the multimedia effect tells us that better learning occurs from consistent pictures and narration rather than narration alone. Second, the modality effect tells us that learners who received pictures and narration performed better on dependent measures than did learners who received pictures and onscreen text. See the recommendation *Activate a Referential Process* for further considerations on image use.

2.0 Limit the Compression Range

When designing a multimedia message with time-compression, stay within the range of 150 (normal speed) to a maximum of approximately 275 wpm, which is 1.8 times (40% compression) the speed of normal narration. We have seen from our research and the research dating back to the 1950s that learning outcomes begin to degrade when this ceiling is reached in learning materials (Fairbanks, Guttman, and Miron, 1957; Foulke, 1968; Reid, 1968). Ritzhaupt and Barron (2008) found that learners were most satisfied with 1.4 times (20% compression) the normal audio speed or approximately 210 wpm. This study included a normal speed of approximately 150 wpm, meaning learners were more satisfied with an accelerated treatment as oppose to the normal speed. Pastore (In review) surveyed learners and found that learners preferred a compression percentage of about 10%, which translates to about 165 wpm. Put simply, our results suggest learners actually prefer accelerated audio speeds in multimedia learning materials.

3.0 Consider the Intrinsic Cognitive Load

Consider the intrinsic cognitive load of the learning materials (based on learner prior knowledge) before employing time-compression in your design because the time-compression increases the extraneous cognitive load (Pastore, 2009; Ritzhaupt, Barron, & Kealy, 2011; Sweller, Van Merrienboer, & Paas, 1998). Ritzhaupt and colleagues (2008, 2011) used learning materials that might be considered low intrinsic cognitive load as one study used a discussion of the educational affordances of Podcasts and the other two studies used a non-fictitious story about various locations in Australia. Meanwhile, Pastore (2010, 2012) used a scientific explanation of how the human heart works, which might be considered a higher intrinsic cognitive load. There were differences between these two studies in terms of the degree of time-compression the learners could handle before diminishing the scores on dependent measures of interest. We believe the instructional designer must strike a balance between the intrinsic cognitive load imposed by the learning contents and extraneous cognitive load imposed by the time-compression.

4.0 Build Small Learning Objects

Build your audio-enhanced multimedia learning materials in short clips (several minutes) to avoid maximizing the extraneous cognitive load and to provide the learner a rest from the materials. This design consideration is also a recommendation from our literature base on learning object theory and practice (Boyle, 2003; Ritzhaupt, 2010) in which you build smaller learning objects that are loosely coupled and highly cohesive for all types of material (compressed or not). An unresolved issue is how long learners can observe a time-compressed multimedia message before their cognitive load is maximized. Both Ritzhaupt and colleagues (2008, 2011) as well as Pastore (2010, 2012) provided breaks in between sections of learning materials to reduce the cognitive load of the treatment. The research base does not currently have a threshold that tells us how long a learner can observe a time-compressed multimedia message.

5.0 Consider Minimal Verbal Redundancy

Consider adding some verbal redundancy via onscreen text emphasizing keywords and concepts (Ritzhaupt & Barron, 2008), but avoid 100% verbal redundancy (on screen redundancy not repetition of content) as this does not have a positive effect on learning (Pastore, 2012). The use of verbal redundancy is an age old question in multimedia learning research. Mayer and his colleagues suggest that we minimize the use of onscreen text when using narration because of cognitive load (Mayer, 2001; Mayer, 2003). This is further confirmed by Barron and Atkins (1994). However, the research by Ritzhaupt, Gomes, and Barron (2008) suggests that some verbal redundancy (limited redundant words both onscreen and in narration) might have a positive effect on learning outcomes. Further, research by Adesope (2010) suggest verbal redundancy leads to better learning outcomes. This is presently an unresolved issue in multimedia learning research.

6.0 Activate a Referential Process

Be sure there is a strong semantic relationship between the pictures used and narration used within your learning materials. The goal is to activate a referential process when the learner organizes the words and pictures in working memory (Mayer, 2001). Should an unrelated picture be used, it might diminish the capacity of the picture to serve as a secondary retrieval cue for the verbal information (Ritzhaupt, Barron, & Kealy, 2011). Levin (1981) suggests that images can serve as decorational, representational, organizational, and transformational. While decorational images serve no purpose and can actually hinder the learning process, representational, organizational and transformational images have been found to have effects on learning (Carney & Levin, 2002; Levin, 1981). The important detail for the instructional designer is to create and select appropriate pictures to use in time-compressed multimedia learning materials.

7.0 Provide Learners Speed Options

Learners like to have choices when learning or traversing multimedia contents. Multimedia products should provide learners with a choice of time-compression. This point is supported by Ritzhaupt, Gomes, & Barron (2008) and by Pastore (In review). When reading online news articles or textual web-based instruction, a learner has the capacity to scan or skim content. Learners viewing multimedia based content using video or audio are not always afforded this luxury. With the proliferation of video-based and audio-based multimedia content and the heightened popularity of these media online, the need to skim multimedia is of increasing importance (Omoigui, He, Gupta, Grudin & Sanocki, 1999). One technique used to empower learners with this ability is time-compression technology. Time-compression technology aims at reducing the amount of time that a learner listens to and/or watches multimedia content. This recommendation is to software developers as much as it is to instructional designers.

8.0 Allow Playback

Ritzhaupt and colleagues (2008, 2011) did not provide the learner the option to replay the multimedia messages during the experimental treatments. However, Pastore (2010, 2012) afforded the learners this feature in the design of the intervention. Further, he studied the review behavior as a variable in the statistical models. As one would anticipate, those that were in the accelerated treatments (speeds at which learning was significantly decreased) would more frequently choose to review the contents a second or third time (Pastore, 2010). We suggest that learners should always be provided the opportunity to review learning materials again for it to be an ecologically valid treatment.

Recommendations for Future Research

There remain many unanswered questions related to the use of digital time-compression in multimedia learning environments. First, we have been unable to locate any relevant articles that investigate the use of digital time-compression with instructional video or narrated screen recordings. With the explosion of YouTube and other relevant video sources like Lynda.com, we believe the next logical step is to rigorously examine time-compressed video learning materials, including narrated screen recordings. As a result, the authors have begun a research program to examine both time-compressed instructional video and narrated screen recordings using similar procedures to the ones presented in this review.

Another unresolved issue is how long a learner can observe a time-compressed multimedia message before maximizing their cognitive load. While instructional designers following good design practice build small digital learning objects, we do not have an optimal length for these learning objects under conditions of time-compression. Research might choose to tackle this unresolved issue by looking at the length of time learners can handle both simple and complex material in a multimedia environment.

A final consideration is the element of learner choice. This topic has been briefly examined in Ritzhaupt and Barron (2008) and Pastore (In review), however, we have yet to determine if our recommendations will hold as cognitive load is increased and various forms of media are introduced. As a result, research on the topics discussed previously in this section: instructional video and cognitive load should include a learner preference variable.

Dr. Raymond Pastore *is an Assistant Professor of Instructional Technology at the University of North Carolina Wilmington. He earned his Ph.D. in Instructional Systems with a minor in Educational Psychology from Penn State University. His research focuses on multiple representations (multimedia), gaming, mobile learning, and metacognitive strategies that support learning from a multimedia environment.*

Dr. Albert D. Ritzhaupt *is an associate professor of educational technology in the School of Teaching and Learning at the University of Florida. His primary research areas focus on the design and development of technology-enhanced learning environments, and technology integration in education. Direct correspondence regarding this article to him via email: aritzhaupt@gmail.com*

References

Adesope, O. O. (2010). *The cognitive effects of verbal redundancy and animated concept maps on learning* (Doctoral dissertation, Faculty of Education-Simon Fraser University).

Barabasz, A. F. (1968). A study of recall and retention of accelerated lecture presentation. *Journal of Communication, 18*(3), pp. 283-287.

Barron, A. E. (2004) Auditory Instruction. In D. H. Jonassen (ed., pp. 949-978), Handbook of research on educational communications and technology (Mahwah, NJ: Lawrence Erlbaum, 2004).

Barron, A. E., & Atkins, D. (1994). Audio instruction in multimedia education: Is textual redundancy important?. *Journal of Educational Multimedia and Hypermedia*, *3*(3), 295-306.

Benz, C. R. (1971). Effects of time compressed speech upon the comprehension of a visual oriented television lecture. Unpublished doctoral dissertation, Wayne State University.

Boyle, T. (2003). Design principles for authoring dynamic, reusable learning objects. *Australian Journal of Educational Technology*, *19*(1), 46-58.

Carney, R. N. & Levin J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review, 14*(1), 5-26.

Clark, J. M. (1983). Reconsidering research on learning from media. *Review of Educational Research, 53*(4), 445-459.

Clark, J. M. & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review, 3*(3), 149-170.

Dwyer, F. M., & Lamberski, R. (1983). The instructional effect of coding (color and black and white) in facilitating students' information acquisition and retrieval. *Educational Communication & Technology Journal, 31*, 9-21.

Fairbanks, G., Guttman, N., & Miron, M. S. (1957). Auditory comprehension of repeated high-speed messages. *Journal of Speech and Hearing Disorders, 22*, 23-32.

Foulke, E., A. & Sticht, T. G. (1967). The intelligibility and comprehension of accelerated speech. Proceedings of the Louisville Conference on Time Compressed Speech, Louisville, KY, pp. 21-28.

Goldhaber, G. M. (1970). Listener comprehension of compressed speech as a function of the academic grade level of the subjects. *Journal of Communication, 20*(1), 167-173.

Hede, A. (2002). An integrated model of multimedia effects on learning. *Journal of Educational Multimedia and Hypermedia, 11*(2), 177-191.

Honing, H. (2006). Evidence for Tempo-Specific Timing in Music Using a Web-Based Experimental Setup. *Journal of Experimental Psychology, 32(3), 780–786.*

Janse, E., Nooteboom, S., & Quene, H. (2001). Word-level intelligibility of time-compressed speech: Porsodic and segmental factors. The Netherlands: Utrecht Institute of Linguistics.

Jester, R. E., & Travers, R. M. W. (1967). The effect of various presentation patterns on the comprehension of speeded speech. *American Educational Research Journal*, 353-360.

Junor, L. (1992). Teaching by tape: Some benefits, problems, and solutions. *Distance Education*, *13*(1), 93-107.

Kullhavey, R. W., Lee, B. J., & Caterino, L. C. (1985). Conjoint retention of maps and related discourse. *Contemporary Educational Psychology, 10*, 28-37.

Levin, J. R. (1981). On functions of pictures in prose. In Pirozzolo, F. J., and Wittrock, M. C. (eds.), *Neuropsychological and Cognitive Processes in Reading*, Academic Press, New York, pp. 203–228.

Mayer, R. E. (2001). Multimedia Learning. New York: Cambridge University Press.

Mayer, R. E. (2002). Multimedia learning. *Psychology of Learning and Motivation*, *41*, 85-139.

Mayer, R. E. (2003). Elements of a science of e-learning. *Journal of Educational Computing Research, 29*(3), 297-313.

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

McDaneial, M. A., Pressley, M. (1987). *Imagery and Related Mnemonic Processes: Theories, Individual Differences, and Applications*. New York City, New York: Springer-Verlag.

Moore, D. L., Hausknecht, D., & Thamodaran, K. (1986). Time compression, response opportunity, and persuasion. *Journal of Consumer Research*, 85-99.

Nichols, R. G., & Stevens, L. A. (1957). Are you listening? New York: McGraw-Hill.

Omoigui, N., He, L., Gupta, A., Grudin, J., & Sanocki, E. (1999). *Time-compression: systems concerns, usage, and benefits*. Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit, Pittsburgh, Pennsylvania, pp. 136 – 143.

Paivio, A. (1986). Mental representations. New York: Oxford University Press.

Paivio, A. (1990). Mental representations: A dual coding approach. New York: Oxford University Press.

Pastore, R. S. (2010). The effects of diagrams and time-compressed instruction on learning and learners' perceptions of cognitive load. *Educational Technology Research and Development*, *58*(5), 485-505.

Pastore, R. (2012). The effects of time-compressed instruction and redundancy on learning and learners' perceptions of cognitive load. *Computers & Education*, *58*(1), 641-651.

Pastore, R. (In Review) Learners' Perceptions of Time Compressed Instruction.

Richaume, A., Steenkeste, F., Lecocq, P., & Moschetto, Y. (1988, November). Intelligibility and comprehension of French normal, accelerated and compressed speech. In *Engineering in Medicine and Biology Society, 1988. Proceedings of the Annual International Conference of the IEEE* (pp. 1531-1532). IEEE.

Reid, R. H. (1968). Grammatical complexity and comprehension of compressed speech. *Journal of Communication*, *18*(3), 236-242.

Ritzhaupt, A. D. (2010). Learning Object Systems and Strategy: A description and discussion. *Interdisciplinary Journal of E-Learning and Learning Objects*, *6*(1), 217-238.

Ritzhaupt, A. D., Gomes, N. D., & Barron, A. E. (2008). The effects of time-compressed audio and verbal redundancy on learner performance and satisfaction. *Computers in Human Behavior*, *24*(5), 2434-2445.

Ritzhaupt, A. D., & Barron, A. (2008). Effects of time-compressed narration and representational adjunct images on cued-recall, content recognition, and learner satisfaction. *Journal of Educational Computing Research*, *39*(2), 161-184.

Ritzhaupt, A. D., Barron, A. E., & Kealy, W. A. (2011). Conjoint processing of time-compressed narration in multimedia instruction: The effects on recall, but not recognition. *Journal of Educational Computing Research*, *44*(2), 203-217.

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

Schnotz, W., Bannert, M. (1999). Supports and interference effects in learning from multiple representations. In Bangera, S. (Ed.), European Conference on Cognitive Science, 27-30 Oct. 1999, Instituto di Psicologia, Nazionale delle Ricerche, Rome, Italy, pp.447-452.

Standing, L. G., Conezio, J., Haber, N. (1970). Perception and memory for pictures: Single trial learning for 2500 visual stimuli. *Psychonomic Science, 19*, 73-74.

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

Taylor, S. E. (1964). Listening: What research says to the teacher. Washington, DC: National Education Association of the United States.