Computers in Human Behavior 45 (2015) 222–227

Contents lists available at ScienceDirect

Computers in Human Behavior

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

Effects of captions and time-compressed video on learner performance and satisfaction

Albert D. Ritzhaupt [a,*], Raymond Pastore [b], Robert Davis [a]

[a] University of Florida, United States
[b] University of North Carolina Wilmington, United States

ARTICLE INFO

Article history:
Available online 2 January 2015

Keywords:
Video
Time-compression
Learning
Satisfaction
Captions
Verbal redundancy

ABSTRACT

Digital video is becoming increasingly popular in higher education with faculty digitally recording and broadcasting lectures for students to learn-on-demand, such as iTunes University or YouTube. Students have discovered accelerated playback features in popular computer software and use it to reduce the amount of time watching video-enhanced instruction. In the current study, 147 undergraduate students were randomly assigned to one of six video treatments based on a 3 Video Speed (1.0 = Normal vs. 1.25 = Fast vs. 1.50 = Very Fast) × 2 Captions (Captions Present vs. Captions Absent) × 2 Trial (Trial 1 vs. Trial 2) design. Results show no significant difference on learner performance across treatments based on Video Speed. Captions were found to have a significant negative effect on learner performance. A significant difference was found on learner satisfaction in favor of a normal Video Speed. The findings suggest that learners might be able to accelerate Video Speeds up to 1.5 times the normal speed, but are generally less satisfied with the learning experience.

1. Introduction

Video has been well established as a tool that can lead to high-level learning and can increase student motivation (assuming its well-designed) (Chen & Sun, 2012). Recently, with the availability of video recording devices such as cell phones, tablets, and augmented reality devices, such as Google Glass, streaming video has become extremely popular in the classroom, workplace, and social networking communities. It is easy to find, upload/download, and create, using websites and software like Youtube.com and iTunes University. These technologies make it easy for anyone to create and share videos. As a result, the use and popularity of video by instructors in the traditional and online classrooms has never been greater. There has never been a time when there was greater access to video lectures, most of which are free and available to anyone. However, these free videos create the need to sort through media to find good quality material. Thus, instructors and designers aim to find and/or create videos that are well-designed to support learning that they can use in their classes and can be shared online. Good video media design stems from theories related to the multimedia principle.

The multimedia principle suggests that presenting multiple representations, verbal and nonverbal, that explain for one another, is better for learning than just one (Clark & Mayer, 2011). As a result, presenting learners text and images, narration and images, or narration and animation/video is better for learning than just text alone. This has been demonstrated in a number of experimental research studies (Eilam & Poyas, 2008; Scheiter, Schüler, Gerjets, Huk, & Hesse, 2014; Vogel-Walcutt, Abich, & Schatz, 2012). Once it was established that multimedia is better for learning than text alone, other principles that help refine this recommendation were discovered to help maximize and enhance the learning experience. For instance, are text and images better for learning than narration and images? The modality principle, states that audio narration and images is better for learning than text and images (Sweller, 2005). This is due to the fact that learners can focus on the images and listen without having to split their attention between representations. However, there is a disadvantage to this dual modal principle. It takes significantly longer to listen to narration or to watch a video then it does to read a text document (with or without images). In fact, the average human speaks at 150 words per minute (wpm) but reads at 287 (wpm) (Benz, 1971). Given that time is a valuable commodity (and can involve dollars earned and lost), anything that can be done to

* Corresponding author at: The University of Florida, School of Teaching and Learning, College of Education, University of Florida, 2423 Norman Hall, PO Box 117048, Gainesville, FL 32611, United States. Tel.: +1 (352) 273 4180 (O); fax: +1 (352) 392 9193.

E-mail address: aritzhaupt@gmail.com (A.D. Ritzhaupt).

http://dx.doi.org/10.1016/j.chb.2014.12.020
0747-5632/

reduce time is highly desired by both learners and instructional designers. As a result, researchers have begun exploring the time-compression of media.

Time-compression refers to media that has been increased in speed yet has preserved quality (Orr & Friedman, 1967). While this topic has been explored for decades the use of multimedia and time compression has only recently begun to be explored (e.g., Pastore, 2010, 2012; Ritzhaupt & Barron, 2008; Ritzhaupt, Gomes, & Barron, 2008). Current research on multimedia and time-compression has shown that at 25% learning of high level material is not significantly different than those going through the multimedia content at a regular pace (Pastore, 2012; Ritzhaupt & Barron, 2008; Ritzhaupt et al., 2008). Thus, an hour of instruction can be delivered in 45 min and still meet the same learning objectives and levels of learner achievement. This small time savings can be significant if compounded over many employees by adding dollars and time saved. However, current time-compression research has only just begun to explore the many of the facets of multimedia learning. For instance, how does time-compression affect animation or video? Currently there are very few recommendations for these delivery mechanisms. In light of that, video is very popular. It is being used for instruction in both online and face-to-face learning environments. It is also being used in corporate and government training. As a result, the following study seeks to examine time-compressed video and captions (verbal redundancy) on learning.

2. Conceptual framework

2.1. Multimedia learning

Multimedia refers to a combination of verbal (words, narration, or text) and non-verbal (images, icons, graphs, or charts) representations, which explain for one another, and are used for learning and/or communication (Clark & Mayer, 2011; Mayer, 2005). A myriad of research studies have demonstrated that multimedia instruction can aid student learning, is an effective means of content delivery, and is preferred by learners (Eilam & Poyas, 2008; Mayer, 2001; Pastore, 2012). From this concept, Mayer and colleagues have developed the cognitive theory of multimedia learning (CTML) (Mayer, 2005). The CTML hypothesizes that one can process multiple external representations (MERs) simultaneously in working memory (Mayer, 2005). As a result, representations are used, created, and then stored in long-term memory (unlimited storage capacity) for later use (Seufert, 2003). The CTML is based on three assumptions (1) working memory is made up of a dual modality (dual coding) input channel system, (2) there is a limited capacity in working memory and (3) that learners engage in active processing. As a result of CTML, countless research studies have sought to examine if learning from verbal and nonverbal representations in a multimedia environment is better for learning than just text or images alone. For instance, Eilam and Poyas (2008) examined learning (achievement) in a multimedia environment (text and picture) vs. a single representation (text only). Participants in this study were given three homework assignments on cell phone use and assessed via recall and transfer. On each measure, the multimedia group performed significantly better. This finding helps confirm the multimedia principle. Additionally, this principle applies to static images and text, narration and images, and animation/video (Mayer & Moreno, 2003). Thus, there have been a plethora of research studies which examined this principle and how it is affected in various conditions and environments. As a result, a number multimedia principles have been established, which include but are not limited to the modality, redundancy, and split-attention, cueing, and coherence principles. For purposes of this paper, only the redundancy principle will be explored as it is the principle examined in this study.

2.2. Verbal redundancy

The redundancy principle occurs when learners are presented with redundant representations at the same time, for instance, duplicate text and narration. In the context of video-enhanced instruction, this might be manifested by the use of captions in a video with narration. However, verbal redundancy has been found to increase the burden on working memory and inhibit learning under certain circumstances (Pastore, 2012). According to Sweller (2005) there are two types of redundancy that can occur (1) two types of representations that contain duplicate information (this is the type used in this study) and (2) two types of media where one contains elaborate text and the other is summarized. This phenomenon has been examined in a number of experimental studies (Mayer, Heiser, & Lonn 2001; Pociask & Morrison 2008). For instance, Mayer et al. (2001) conducted several experiments on the redundancy principle and its effects. In their first study, 78 university students were presented with animations and audio and then either on-screen text that summarized the narration or extraneous (extra) details. Results indicated that students presented duplicate representations performed significantly worse than those that did not on recall and transfer tests. The second study consisted of 109 university students who were placed into treatments of animation with narration and then either no text, summary text, or full text. Participants in the no text group score significantly higher on recall and transfer tests than those in the other conditions, whose scores did not differ. As a result, redundancy appears to inhibit learning. Thus, it becomes apparent that the multimedia and redundancy principles are well established, however, how do these principles interact with learning when presented in a time-compressed video environment?

2.3. Time compression and multimedia

Time-compression refers to media that has been increased in speed yet has retained quality (Orr & Friedman, 1967). Compression allows the speed of video, multimedia, or narration to be increased without producing an audio chipmunk like effect (high pitched) but rather sounding normal with less pauses in between words. Thus the audio, video, and multimedia should sound very similar to their regular paced counterparts. In fact, Honing (2006) revealed that users could not distinguish between normal and compressed audio (at low levels). Reed (2003) came to a similar conclusion in a study which found that users could not distinguish between normal paced narration and narration compressed at 20%. Thus, users might not even know if media was compressed at low levels of compression.

Current time-compression literature indicates that multimedia can be compressed up to 25% without sacrificing factual and problem solving knowledge (Pastore, 2010; Ritzhaupt & Barron, 2008; Ritzhaupt et al., 2008). For instance, Pastore (2010) found that learners could retain high level knowledge without increasing working memory burden when media was compressed 25% in a multimedia environment. Pastore (2012) analyzed 154 university students who were presented with multimedia instruction compressed at 0% (normal paced), 25%, or 50% compression with redundant text and narration or narration only. The study found that those presented 0% and 25% compression speeds did not differ on tests measuring factual and problem solving knowledge. The 50% compression condition performed significantly worse, however. Similar results were uncovered in Ritzhaupt et al. (2008) who found that learning was not affected at 1.4 or 1.8 times normal paced instruction. Thus, it appears that learners from mul-