Computers & Education 80 (2015) 108–121

Contents lists available at ScienceDirect

Computers & Education

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

Effects of different video lecture types on sustained attention, emotion, cognitive load, and learning performance

Chih-Ming Chen*, Chung-Hsin Wu

Graduate Institute of Library, Information and Archival Studies, National Chengchi University, No. 64, Section 2, ZhiNan Road, Wenshan District, Taipei City 116, Taiwan, ROC

ARTICLE INFO

Article history:
Received 27 May 2014
Received in revised form
23 August 2014
Accepted 23 August 2014
Available online 3 September 2014

Keywords:
Media in education
Evaluation methodologies
Distance education and telelearning
Interactive learning environments

ABSTRACT

Although online courseware often includes multimedia materials, exactly how different video lecture types impact student performance has seldom been studied. Therefore, this study explores how three commonly used video lectures styles affect the sustained attention, emotion, cognitive load, and learning performance of verbalizers and visualizers in an autonomous online learning scenario by using a two-factor experimental design, brainwave detection, emotion-sensing equipment, cognitive load scale, and learning performance test sheet. Analysis results indicate that, while the three video lecture types enhance learning performance, learning performance with lecture capture and picture-in-picture types is superior to that associated with the voice-over type. Verbalizers and visualizers achieve the same learning performance with the three video types. Additionally, sustained attention induced by the voice-over type is markedly higher than that with the picture-in-picture type. Sustained attention of verbalizers is also significantly higher than that of visualizers when learning with the three video lectures. Moreover, the positive and negative emotions induced by the three video lectures do not appear to significantly differ from each other. Also, cognitive load related to the voice-over type is significantly higher than that with by the lecture capture and picture-in-picture types. Furthermore, the cognitive load for visualizers markedly exceeds that of verbalizers who are presented with the voice-over type. Results of this study significantly contribute to efforts to design of video lectures and also provide a valuable reference when selecting video lecture types for online learning.

1. Introduction

Although most university classes use traditional face-to-face instruction, many online courses are available in which video lectures are used in digital form. Created by simply uploading a video recording of a lecturer, a video lecture may be more complex, paired with slide presentations, interactive quizzes and demonstrations (Osborn, 2010). Online video lectures have become increasingly common in recent years, as evidenced by their use in many organizations, educational institutions, and open learning systems, such as Coursera, Khan Academy, and TED. Video lectures often provide students with additional time to fully understand classroom course materials by allowing them to review lectures repeatedly (Brecht & Ogilby, 2008). Additionally, online video lectures with audio and video instruction can enrich a learning experience, allowing students to see and listen as they would be in an actual classroom.

Common online learning media include lecture capture (or called the talking-head lecture) (Danielson, Preast, Bender, & Hassall, 2013; Ilioudi, Giannakos, & Chorianopoulos, 2013; Wiese & Newton, 2013), voice-over presentation (Griffin, Mitchell, & Thompson, 2009), picture-in-picture (Chorianopoulos & Giannakos, 2013), and Khan-style video lecture (Chorianopoulos & Giannakos, 2013), all of which present multimedia information in different styles. A video lecture must harness learning motivation, increase learning performance, satisfy individual learning needs with different learning styles, and select the most appropriate format to facilitate learning (Hornbæk, Engberg, & Gomme, 2002). Moreover, cognitive psychology commonly views attention as facilitating the selection of incoming perceptual information and limiting the number of external stimuli processed by the bounded cognitive system to avoid overloading (Driver, 2001). Importantly, a

* Corresponding author. Tel.: +886 2 29393091x88024; fax: +886 2 29384704.
E-mail address: chencm@nccu.edu.tw (C.-M. Chen).

http://dx.doi.org/10.1016/j.compedu.2014.08.015
0360-1315/

learning process without sustained attention lacks effective identification, learning, and memory (Broadbent, 1958). Restated, sustained attention to learning content is of priority concern for effective learning, explaining the need to determine whether different styles of video lectures affect sustained attention in online learning scenarios. Moreover, many studies have asserted that design of multimedia materials or video lectures should consider the affective state (i.e. a learner's emotional state) (Chen & Sun, 2012; Chen & Wang, 2011). However, exactly how video lecture types affect learning performance, learner emotions, and sustained attention has seldom been studied empirically, results of which would provide a valuable reference for video lecture design.

According to Sweller, van Merriënboer, and Paas (1998), limited working memory is a defining aspect of the human cognitive architecture and, accordingly, all instructional designs should be analyzed from a cognitive load perspective. Educational research has also confirmed that considering individual learning styles is more important than instructing all learners with one style (Dunn & Griggs, 2000). Individual differences in learning styles must be identified when learners process video lectures since they add to existing knowledge of processing preferences and predict personality variables accurately. Of the cognitive styles related to multimedia learning, the visualizer–verbalizer hypothesis is especially relevant to individual differences when using video lectures for online learning because they typically present information to learners using audio and video (containing slides, texts, and pictures simultaneously) (Mayer & Massa, 2003).

Although many educational organizations create and share video lectures, no conventional standard is available to create a video lecture. No guidelines are also available for the presentation style of video lectures (Ilioudi et al., 2013). Importantly, the merits and limitations of each video lecture type for online learning have not yet been thoroughly investigated. In sum, despite a growing number and variety of online educational video lectures, their effectiveness in terms of learning and usability is poorly understood (Chorianopoulos & Giannakos, 2013). Therefore, this study aims to explore how the three considered video lecture types including lecture capture, voice-over presentation, and picture-in-picture affect the sustained attention, emotion, cognitive load, and learning performance of verbalizers and visualizers in an autonomous online learning scenario by using a two-factor experimental design. Results of this study significantly contribute to efforts to select the most appropriate video lecture type for online learning that maximizes learning performance in an autonomous learning context.

2. Literature review

2.1. Video lecture design based on cognitive load, multimedia learning, and media richness theories

Learners generally process and remember images much more efficiently than what they read or hear (Shorter & Dean, 1994). Recent years have witnessed tremendous growth of available online educational video lectures, spanning K-12 tutorials to university lectures. Different video lectures (e.g., lecture capture, voice-over presentation, picture-in-picture, and Khan-style video lecture) present multimedia information differently (Chorianopoulos & Giannakos, 2013; Griffin et al., 2009; Ilioudi et al., 2013). Developing a video lecture can be a complex process that requires planning and an implementation procedure. Learning theories and their instructional implications are essential to learning the contents of a video lecture with the most appropriate delivery components (Chorianopoulos & Giannakos, 2013). As theoretical frameworks, cognitive load theory (CLT) (Chandler & Sweller, 1991) and the cognitive theory of multimedia learning (CTML) (Mayer, 2001) focus on characteristics related to multimedia materials and provide design guidelines for educationally effective multimedia materials. Notably, CLT suggests that multimedia materials should reduce cognitive load and optimize the use of working memory (Chandler & Sweller, 1991). Additionally, CTML (Mayer, 2001) asserts that multimedia narration and graphic images produce verbal and visual mental knowledge that is integrated with prior knowledge to construct new knowledge. CLT distinguishes between three cognitive loads: intrinsic, extraneous, and germane. Each load competes for limited resources of working memory when complex visual and verbal information is processed (Sweller, 1999). Intrinsic load is inherent to learning materials. An increasing material complexity implies an increasing intrinsic load. While referring to the mental effort imposed by instructional activities, their design, and presentation, extraneous load does not contribute directly to an understanding of material. Finally, germane load refers to the mental effort exerted by learners to process new information and then integrate it into existing knowledge structures. Although intrinsic cognitive load cannot be manipulated, extraneous and germane cognitive load can. Additionally, although instructional design can influence both extraneous and germane cognitive loads, extraneous load interferes with learning while germane load facilitates it. Moreover, from a theoretical perspective, Mayer's (2001) CTML suggests that information presented in the visual and auditory modalities simultaneously can enhance learning performance, especially in retention and information transfer, owing to reduced student cognitive loading and optimized use of working memory. Mayer (2001) asserts that multimedia learning in which animation and narration are combined generally more significantly improves student performance on retention tests than when information is presented as either text or narration. Despite the extensive use of CLT and CTML in multimedia design, the feasibility of applying CLT and CTML to lecture-style multimedia presentations warrants further study (Day, Foley, & Catrambone, 2006).

Sweller et al. (1998) proposed multimedia instructional design schemes, including the goal-free, worked example, completion problem, split-attention, modality, and redundancy effects, as well as the variability effect based on CLT. We believe that the redundancy, modality, and split-attention effects are closely related to video lecture design, ultimately affecting sustained attention, emotion, cognitive load, and learning performance. The redundancy effect occurs when information that can be fully understood in isolation, as either visual or auditory information, is presented to both channels as the same information (Sorden, 2005). Integrating redundant information in both working memories can increase cognitive load (Sorden, 2005). Conversely, Chandler and Sweller (1991) demonstrated that eliminating redundant information can reduce extraneous cognitive load. The modality effect refers to the ability to increase effective working memory capacity by using auditory and visual working memory together rather than using either one independently (Sorden, 2005). Mousavi, Low, and Sweller (1995) asserted that using dual-mode (visual-auditory) instructional schemes decreases cognitive load and processing information when using both the visual and auditory channels increase the limited capacity of working memory. Moreover, the split-attention effect occurs when learners must divide their attention into multiple information sources that must be integrated with each other before they can be understood. Restated, each information source is essential for understanding the learning material (Ayres & Sweller, 2005; Mayer & Moreno, 1998; Sorden, 2005). The split-attention effect increases the cognitive load because multiple information sources must be integrated with each other (Ayres & Sweller, 2005).

The media richness theory (MRT) of Daft and Lengel (1986) suggests that different media have different degrees of richness. In this paradigm, richness is based on the ability to reproduce the information that media transmit. Lee, Cheung, and Chen (2007) indicated that when a communication medium is rich, both uncertainty and equivocality associated with learning task decrease, explaining why effort required to use the medium decreases. According to Lim and Benbasat (2000), a medium that allows users to send and receive multiple cues is perceived by users as useful. Therefore, in addition to transmitting a sufficient amount of relevant information in order to reduce uncertainty, a rich medium should process rich information to reduce equivocality (Sun & Cheng, 2007). Based on MRT, Bassili (2008) attempted to explain why some students prefer watching lectures online rather than attending lectures. Therefore, whether the three video lecture types, each differing in media richness, affect learning performance warrants further study.

2.2. Effects of video lecture types on learning performance

Lecture capture, voice-over presentation, picture-in-picture, and Khan-style video lecture are commonly used video lecture types in online learning environments (Chorianopoulos & Giannakos, 2013; Griffin et al., 2009; Ilioudi et al., 2013). The lecture capture type simply records an instructor's presentation for online viewing. The video typically contains PowerPoint slides, instructor's voice, and occasionally a video recording of the instructor with text on a whiteboard. The voice-over type synchronizes audio recordings of a lecture to accompanying PowerPoint slides via specialized lecture recording software (e.g., Microsoft Producer or PowerCam). The picture-in-picture type displays an instructor's image and lecture slides and contains the instructor's voice, subtitles, or even flash animation. Thus, this type distinguishes between an instructor's video feed and slides, as well as requires elaborate post-production. Khan-style video lectures rely mainly on handwritten tutorials, as produced by using a digital pen and tablet, with an audio voice-over from a lecturer. However, to our knowledge, exactly how different video lecture types affect online learning performance has seldom been studied empirically. Of those few empirical studies, Ilioudi et al., 2013 explored whether lecture capture, Khan-style video lecture, and traditional paper book support mathematics self-study in secondary education. In addition to demonstrating that lecture capture is more effective than books for complex topics, their results also indicated that learning performance for lecture capture exceeds that with a Khan-style video lecture. Moreover, Griffin et al. (2009) evaluated the possible pedagogical benefits of different video lectures formats, including those that present PowerPoint slides and a lecturer's voice synchronously, as well as those that present PowerPoint slides and audio files asynchronously. According to their results, the synchronous mode is superior to the asynchronous mode in terms of learning performance. Homer, Plass, and Blake (2008) examined how two versions of a computer-based multimedia presentation, i.e. a video version (including a video of a lecture synchronized with slides) and no video (including only the slides and an audio recording of a lecture) affect learning performance. Their results suggested that having video as well as PowerPoint slides had a split-attention effect, subsequently increasing the cognitive load. Increased cognitive load typically reduces learning performance (Sweller, 1994). Moreover, Wiese and Newton (2013) summarized the advantages of lecture capture for students as increased satisfaction, enhanced understanding of content, clarification of difficult topics, improved accuracy of course notes, and increased accessibility for physically challenged and non-native English speaking students. For instructors, the number of requests for clarification decreased.

Several multimedia learning studies have investigated multiple-channel presentations and the resulting split attention phenomenon (Mayer & Moreno, 1998; Schmidt-Weigand, Kohnert, & Glowalla, 2010). In addition to slides or content on a board, most video lectures also transmit the instructor's image, leading to the divided attention problem. Restated, despite having only one locus of attention, learners must focus on two areas on a screen, i.e. the window showing the instructor, and the board or slides (Friedland, 2004). Chen and Wang (2011) investigated how multimedia materials presented in different styles affect learners' emotions and their performance. Their results suggested that video-based multimedia material generates the best learning performance and the most positive emotions among static text- and image-based multimedia materials, video-based multimedia materials containing moving images with audio, and animated interaction-based multimedia materials, which integrate text and animated images with interactive features. Additionally, Chen and Lin (2014) claimed that an appropriate text display type for mobile reading in different reading contexts should be deployed, allowing learners to read content on mobile devices with small screens. That study performed a mobile reading experiment with a two-factor experimental design to assess how the static, dynamic, and mixed-text types, which were respectively presented while learners sat, stood, and walked, affect reading comprehension, sustained attention, and the cognitive load of learners. According to their results, these three reading contexts with these three texts both have merits and limitations in terms of reading comprehension, sustained attention, and cognitive load. Consequently, the text display format for mobile reading on small screens should be tailored to the reading context in order to either improve reading comprehension and attention or reduce cognitive load.

Despite significant advances in online educational video lectures in recent years, their effectiveness in terms of learning and usability is still unclear (Chorianopoulos & Giannakos, 2013). Assessing how different video lecture formats affect sustained attention, emotion, cognitive load, and learning performance would significantly contribute to efforts to select an appropriate video lecture type for individualized online learning while maximizing their learning performance in an autonomous learning context.

2.3. Effects of learning styles on learning performance in multimedia leaning environments

In addition to that student ability (Butcher-Powell, 2004, pp. 60–72), student attitude (Kettanurak, Ramamurthy, & Haseman, 2001), instructional approach (Eysink et al., 2009), and learning motivation (Astleitner & Wiesner, 2004) affect learning performance in multimedia learning environments, various studies have demonstrated that student learning style is one of the factors affecting learning performance in multimedia learning environments (Chen & Sun, 2012; Cheng, Cheng, & Chen, 2012; Ocepek, Bosnić, Šerbec, & Rugelj, 2013). A learning style refers to students' characteristics that are manifested in their learning behavior, including how a student learns, should be taught, and interacts with a learning environment (Ocepek et al., 2013). While investigating perceptions of various learning environments, Carter (1985) found that students perceived environments based on their preferred learning style. Of the many learning styles addressed in the literature, the attribute–treatment interaction (ATI) hypothesis of Mayer and Massa (2003) asserts that some individuals process words most effectively (i.e. verbalizers), while others process pictorial representations most effectively (i.e. visualizers). The ATI hypothesis is especially relevant in the design of multimedia materials since multimedia materials often present information to learners using words and pictures simultaneously (Mayer & Massa, 2003). Further, Massa and Mayer (2006) tested the ATI hypothesis, but their results failed to support the ATI hypothesis. Moreover, Chen and Sun (2012) determined whether visual and verbal learning styles affect learners' emotions and performance with three multimedia materials, i.e. static text and image-based, video-based, and animated interactive. Their study demonstrated that video-based multimedia material generates the best learning performance and the most positive emotions for verbalizers. Moreover, dynamic multimedia materials containing video and animation are more appropriate for visualizers than static multimedia materials containing text and image. Namely, their findings do not also support the ATI hypothesis, which states that visualizers learn best with visual methods, whereas verbalizers learn best with verbal methods. Additionally, Kappe, Boekholt, Den Rooyen and Van der Flier's study (2009) argued that considering learning styles for college students to stimulate learning is debatable because their study found that although learning styles were matched to correspondingly suitable learning criteria, the Learning Style Questionnaire revealed no predictive validity.

Additionally, Ocepek et al. (2013) attempted to minimize students' cognitive overload and stress by designing an adaptive learning system, capable of providing an accurate and reliable model that recommends different multimedia types for different individuals by associating combinations of learning styles with preferred multimedia materials. Homer et al. (2008) further found that an instructor's face affects learner's cognitive load differentially, depending on their cognitive preference for visual or verbal information. Different video lecture types generally display an instructor's image using different presentation styles. Lecture capture displays the entire body of a lecturer; voice-over presentation displays the lecturer's image in a separate window; and picture-in-picture displays the lecturer's image in a spotlight. Whether different video lecture types with various display styles of instructor's image affect learning performance is thus of worthwhile interest.

Although several studies (Chen & Sun, 2012; Kappe et al., 2009; Massa & Mayer, 2006) claim that support for the attribute–treatment interaction (ATI) hypothesis, indicating that verbalizers and visualizers should be taught using different multimedia instruction methods to improve their learning, is weak or non-existent. However, Mayer (2011) appealed that researchers and practitioners must search long and hard for the educational implications of learning styles research by taking an evidence-based approach because the implications of learning styles research for educational practice are still less clear. Particularly, online multimedia video lectures have become increasingly common in recent years, whether the three considered video lecture types with various combinations of multimedia elements are unfavorable to verbalizers in terms of learning performance, sustained attention, emotion, and cognitive load warrants further study. The findings of the study are helpful in confirming whether or not the ATI hypothesis exists in different types of multimedia materials and various aspects of learning performance.

3. Research methodology

3.1. Three video lecture formats

3.1.1. Lecture capture format

Despite the popularity of e-learning in recent years, most learning activities occur in traditional classrooms. Lecture capture records classroom lectures using a digital video camera, allowing students to watch the video online or via a computer or a mobile device. Lecture capture technology records a lecturer's voice and image simultaneously, as well as instructional aids, such as writing on a whiteboard and PowerPoint slides. Moreover, lecture capture is characterized by its ability to preserve interactivity in a classroom—students' questions and their reactions to new information. Lecture capture is used by most universities and online video-based learning platforms (e.g., Stanford, MIT Open Courseware, iTunes U). Fig. 1 illustrates an example of the lecture capture type.

3.1.2. Voice over presentation type

A voice-over presentation consists mainly of PowerPoint slides, supplemented with a voice-over explaining the information presented on the slides. A voice-over generally contains narration, PowerPoint slides, and table of contents (ToC) simultaneously, as separated in three windows. However, this method lacks the learning context and visual information of lecture capture, such as classroom activities in the

Fig. 1. An illustration of lecture capture type.

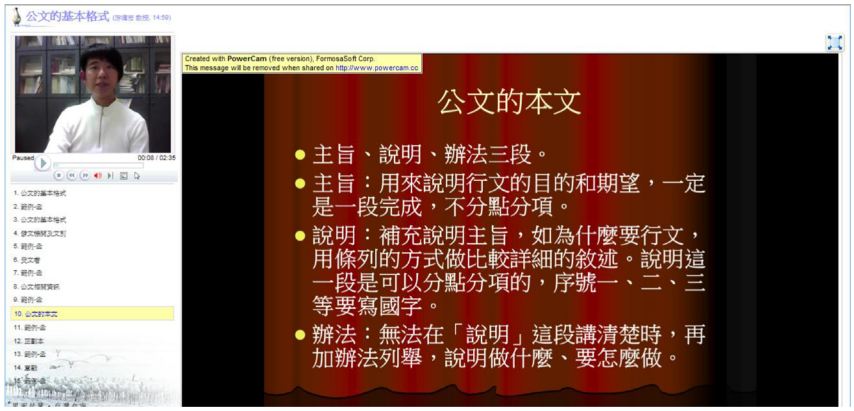

Fig. 2. An illustration of the voice-over type.

lecture capture format and the flash animation features in the picture-in-picture method. Fig. 2 illustrates an example of the voice-over type. The upper left pane contains streaming video displayed by a Windows Media Player, in which a lecturer's image is shown. The lower left pane is the ToC with a list of links for all slides used in the lecture; each link also contains a synchronized anchor point in the video. Thus, viewers can easily skip around the lecture simply by clicking on ToC entries. The right pane is the current PowerPoint slide.

3.1.3. *Picture-in-picture method*

As one of the most complex lecture video types, the picture-in-picture approach overlays an instructor's image and lecture slides. This approach presents an instructor's recorded image and voice, PowerPoint slides, subtitles, and other flash animation features. While this approach provides usable cuts between an instructor's video feed and the slides, elaborate post-production is required. Restated, the picture-in-picture type is a highly effective integration of a rich media presentation and informative video lesson. The picture-in-picture type is extensively adopted in massive open online courses (MOOCs), including those offered through Coursera. Fig. 3 illustrates an example of the picture-in-picture type.

Table 1 compares the three video lectures with different presentation methods, based on cost, production technology, conveyed learning context, multimedia elements, and media richness. Comparison results indicate that the picture-in-picture approach is the most expensive, uses the most production technology, and has the greatest media richness, followed by the lecture capture approach and then the voice-over method. The lecture capture method can convey the context of instruction in a physical classroom, allowing learners to experience the associated activities. That is, in addition to simultaneously presenting the lecturer's voice, image, and instruction aids (including writing on a whiteboard and PowerPoint slides), the lecture capture method preserves the interactivity between a teacher and students in a classroom. Additionally, the voice-over presentation method can be viewed as a speech-based lecture. The voice-over method simultaneously presents not only the elements of a lecturer's image, PowerPoint slides, and a table of contents for slides, but also these three elements in three separate windows. However, this method fails to create a learning context, such as classroom context captured by the video capture style, and visual information (*i.e*. non-verbal information), such as flash animation contained in the picture-in-picture type. In contrast, the picture-in-picture method contains the instructor's image and voice, PowerPoint slides, subtitles, as well as other flash animation features. Notably, this is the only one video lecture approach that uses subtitles of a lecturer's narration. According to Mayer (2001), computer-based multimedia learning environments support the concept that individuals learn more effectively when appropriate pictures (*i.e*. animation, video, and static graphics) are added to text or narration. Restated, as a highly effective combination of a rich media presentation and informative video lesson, the picture-in-picture method provides usable cuts between the instructor's video feed and slides. A video lecture attempts to facilitate learning experiences, thereby maximizing learning performance and satisfaction of learners while maintaining limited production costs.

Fig. 3. An illustration of the picture-in-picture type.

Table 1
Comparison of three considered video lectures with different information presentation types.

Comparison item	Lecture video type		
	Lecture capture	Voice over presentation	Picture-in-picture
Cost	Moderate	Low	High
Production technology	Moderate	Low	High
Conveyed learning context	Physical classroom instruction	Online speech-based instruction	Multimedia interaction instruction
Multimedia elements	1 Lecture slides 2 Lecturer's audio and video 3 Writing on the whiteboard 4 Interaction between lecturer and students	1 Lecture slides 2 Lecturer's audio and video 3 Table of contents of slides	1 Lecture slides 2 Lecturer's audio and video 3 Flash animation materials 4 Subtitle of slides
Media richness	Moderate	Low	High

3.2. Experimental design

To select appropriate learning materials, the study invited a university professor who had over 5 years of experience in online instruction in Taiwan's university to select three learning units with similar difficulty from his online course "Document Writing." Furthermore, he helped the study record the three learning units as three video lecture formats mentioned earlier. The pilot study recruited 48 undergraduate students from the Department of Chinese Literature at National Chengchi University, Taiwan to evaluate the difficulties of the three video lecture formats by the corresponding test sheets of the three learning units after students viewed the videos of three units presented by the same video lecture type. According to our statistical analysis results, the units in the three video lecture types have the same level of difficulty. Moreover, in addition to a brainwave detector, emWave stress detector, cognitive load scale, and learning performance test, this study designed an online learning experiment with a two-factor design to evaluate how verbalizers and visualizers presented with the three considered video lecture methods affect sustained attention, emotions, cognitive load, and learning performance. This study enrolled 37 undergraduate students who are different with the students in the pilot study from the same department to participate in the instructional experiment.

3.3. Experimental procedures

The experiment consisted of three stages. In the first stage, in addition to receiving a 10-min orientation on the experimental procedures, the participants learned how to operate the three video lectures, along with a pretest administered to assess their prior knowledge of each unit in the three videos and identify their leaning styles by using the style of Processing (SOP) scale developed by Childers, Houston, and Heckler (1985). In the second stage, learning activities were performed. To prevent the order of the three video lectures affecting the learning performance, each participant viewed the three videos in a random order on a notebook computer while seated in an observation room. During the learning activities, an attempt was made to identify sustained attention and emotion based on human brainwave signals and heart rate variability (HRV) patterns, in which all participants simultaneously wore the MindSet headset developed by Neurosky and emWave stress detector developed by the Institute of HeartMath during learning activities (Fig. 4). Each participant viewed and attempted to learn from each of the three video lectures by autonomous learning, no instructor's support. Proponents of autonomous learning suggest that in addition to lecturing for no longer than 15–20 min, instructors should adopt a strategy that causes students to engage with the information just presented (Osborn, 2010). Therefore, the students viewed for each video lecture for about 15 min. In the third stage, after the participants finished viewing each video lecture, they immediately took a posttest that assessed their learning performance. The participants also filled out a cognitive load scale that identified the magnitude of cognitive load on a scale from each of the three video lectures. Finally, several research participants were interviewed to understand their learning satisfaction.

Fig. 4. The experimental scene of a learner who simultaneously wore a Mindset earphone and an emWave stress detector to assess sustained attention and emotion while performing an autonomous online learning by video lecture.

3.4. Research participants

All 37 participants provided written informed consent after the experimental details were explained. Of the 37 participants, 9 (24%) were male and 28 (76%) were female, with ages from 20 to 21 years old. Of the 37 participants, there are 20 and 17 students who were identified as verbalizer and visualizer, respectively. All participants had experience with e-learning on computers or electronic devices.

3.5. Research instruments

Sustained attention, emotion, cognitive styles of verbalizers and visualizers, cognitive load, and learning performance were evaluated using five research instruments.

3.5.1. Brainwave detection system

NeuroSky's MindSet headset records raw electroencephalography (EEG) data via a single contact sensor on a learner's forehead. Based on EEG data in real time, the headset can output two custom measures, i.e. attention and mediation. The attention values, which ranged from 0 to 100, denoted a learner's level of focus. Our previous study (Chen & Huang, 2014) confirmed that NeuroSky's MindSet headset has sufficient validity and reliability based on the correlation between Birdwatching scores, which is a visual attention-based cognitive training program developed by Lumosity (Hardy, Drescher, Sarkar, Kellett, & Scanlon, 2011), and attention meter values sensed by NeuroSky's MindSet headset. Restated, birdwatching scores were strongly and positively correlated with meter values (correlation coefficient is as high as .73). Moreover, Rebolledo-Mendez et al. (2009) also found a positive correlation between attention meter values measured by NeuroSky's MindSet headset and self-reported attention levels via a Second-Life assessment exercise. Analytical results demonstrated that the attention value measured by NeuroSky's MindSet headset had a satisfactory validity and reliability for identifying learner attention in a learning activity.

3.5.2. emWave system

As a stress detector developed by the Institute of HeartMath for emotional states, the emWave system determines heart rate variability (HRV) based on spectral density analysis of heart rate power by using an ear sensor (McCraty, Atkinson, Tiller, Rein, & Watkins, 1995). To identify emotional states, the emWave system provides an easy-to-use software with a heart rhythm monitor and an HRV-based emotion recognition algorithm. Several studies have indicated that HRV patterns directly respond to changes in emotional states (Latham, 2006; McCraty et al., 1995; Tiller, McCraty, & Atkinson, 1996). Latham (2006) indicated that two major theoretical frameworks, i.e. Polyvagal theory and Neurovisceral Integration theory, articulate elucidate the role of HRV in emotional responses. Several previous studies associated with multimedia learning and mobile learning assessed learners' emotional states using the emWave system (Chen & Lin, 2014; Chen & Sun, 2012; Chen & Wang, 2011). This study also attempted to identify how different video lectures affect individual learning emotions by using the emWave system.

3.5.3. Style of processing (SOP) scale for identifying a verbalizer and visualizer

This study attempted to identify verbalizers and visualizers by using the style of Processing (SOP) scale developed by Childers et al. (1985). The SOP scale has two subscales (i.e. the verbal and visual subscales) to create two categories: verbalizer and visualizer. Reliability of each subscale in the SOP scale in recognizing cognitive styles of visualizers and verbalizers is .81 and .86, respectively, by Cronbach's alpha. Global reliability of the SOP scale is .88. Overall, reliability of the SOP scale for identifying visual and verbal cognitive styles is sufficient. The participants with SOP scores that are higher and lower than the mean SOP score are identified as visualizers and verbalizers, respectively.

3.5.4. Cognitive load scale

Exactly how the three video lectures influence cognitive load was determined using the cognitive load scale developed by Sweller et al. (1998). With one subscale for mental load and one subscale for mental effort, this scale contains four items with responses on a seven-point Likert scale. Two items addressed mental load (i.e. intrinsic load) and two items addressed mental effort (i.e. extraneous and germane load); total score for each subscale was 14. Cronbach's α value for the cognitive load scale was .92; for mental effort and mental load, the Cronbach's α values were .86 and .85, respectively, demonstrating the high reliability of the scale.

3.5.5. Test sheet

Learning performance was evaluated by designing three tests based on the learning content in the three learning units. Each test consisted of ten items that assessed memory, comprehension, and application. Memory items identified whether learners memorized facts conveyed in a video lecture; comprehension items assessed whether learners understood facts conveyed in a video lecture, and organized or interpreted these facts correctly; in addition, application items assessed whether learners solved problems via their understanding. Each correct answer on the test received 1 point. Restated, total score for each test was 10. Based on the statistical analysis, the pilot study confirmed that the difficulties of the corresponding test sheets of the three learning units have the same level of difficulty. Also, estimation of item difficulty and discrimination by classic testing theory reveals that average difficulty of test items in each test sheet is moderate, and discrimination of each test item in the three tests is satisfactory. Analytical results suggest that the three tests are highly reliable in assessing learning performance.

4. Experimental results

4.1. Effects of students' sustained attention

This section aims to examine whether the three considered video lectures leaded to that sustained attention of students differed significantly as well as whether sustained attention of visualizers and verbalizers who were presented with the three considered video

Table 2
Descriptive statistics of sustained attention of verbalizers and verbalizers presented with three considered video lecture types.

Video lecture type	Cognitive style	Number of learners	Mean	Std.
Lecture capture	Verbalizers	20	48.36	8.26
	Visualizers	17	45.39	5.07
	Total	37	47.00	7.05
Voice over presentation	Verbalizers	20	53.29	6.68
	Visualizers	17	45.90	8.02
	Total	37	49.92	8.11
Picture-in-picture	Verbalizers	20	47.25	5.97
	Visualizers	17	44.39	3.33
	Total	37	45.93	5.08
Total	Verbalizers	20	49.63	7.40
	Visualizers	17	45.25	5.72
	Total	37	47.62	7.01

Table 3
A comparison of sustained attention of verbalizers and verbalizers presented with three considered video lecture types by two-way ANOVA with the Scheffe test.

Item	Sum of squares	Degree of freedom	Sum of mean squares	*F*	Sig.	Result of Scheffe test
Video lecture styles	283.73	2	141.86	3.35*	.039	Voice over presentation > Picture-in-picture
Cognitive style	529.22	1	529.22	12.50**	.001	Verbalizer > Visualizer
Video lecture styles × Cognitive style	119.08	2	59.54	1.40	.249	–
Error	4443.16	105	42.31			–
Total	257,150.05	111				–

* indicates $p < .05$; ** indicates $p < .01$.

lectures differed significantly. Table 2 summarizes statistical results for the sustained attention of verbalizers and verbalizers when students were viewing the three video lectures. Analysis by two-way analysis of variance (ANOVA) reveals that cognitive styles and video lecture types do not have a significant interaction effect ($F = 1.40$, $p = .249 > .05$) (Table 3). Main effect analysis reveals not only that video lecture type significantly affected sustained attention ($F = 3.35$, $p = .039 < .05$), but also that cognitive style significantly affected sustained attention ($F = 12.50$, $p = .001 < .05$). Scheffe's multiple comparison shows that sustained attention with the voice-over method was significantly higher that with the picture-in-picture method. Moreover, sustained attention of verbalizers was significantly higher than that of visualizers when students were viewing the three video lectures. Further, compared with the lecture capture and picture-in-picture methods, voice-over presentation generates not only the highest mean sustained attention score (mean = 49.92), but also the highest standard deviation of sustained attention (Std. = 8.11). Thus, the split-attention effect likely occurs to some degree with the voice-over lecture method.

4.2. Effects of students' emotion

Based on the emWave stress detector, this section aims to examine whether the three considered video lectures leaded to that emotions of students differed significantly as well as whether emotions of visualizers and verbalizers who were presented with the three considered video lectures differed significantly.

4.2.1. Negative emotions

Table 4 summarizes statistical results for the negative emotions of verbalizers and verbalizers. Two-way ANOVA results indicate that cognitive styles and video lecture methods do not have a significant interaction effect ($F = .09$, $p = .905 > .05$) (Table 5). According to analytical results of the main effect, the video lecture methods negligibly affected negative emotion ($F = .92$, $p = .399 > .05$); in addition, cognitive styles also negligibly affected negative emotion ($F = .65$, $p = .419 > .05$).

4.2.2. Positive emotions

Table 6 summarizes the statistical results for positive emotions of verbalizers and verbalizers. Two-way ANOVA results show that cognitive styles and the video lecture methods do not have a significant interaction effect ($F = .07$, $p = .931 > .05$) (Table 7). According to

Table 4
Descriptive statistics of negative emotion of verbalizers and verbalizers presented with three considered video lecture types.

Video lecture type	Cognitive style	Number of learners	Mean	Std.
Lecture capture	Verbalizers	20	72.67	23.73
	Visualizers	17	67.23	23.73
	Total	37	70.17	21.47
Voice over presentation	Verbalizers	20	75.52	22.3
	Visualizers	17	74.40	17.08
	Total	37	75.00	19.81
Picture-in-picture	Verbalizers	20	70.27	23.20
	Visualizers	17	67.1	17.07
	Total	37	68.85	20.40
Total	Verbalizers	20	72.82	22.79
	Visualizers	17	69.61	17.63
	Total	37	71.34	20.56

Table 5
A comparison of negative emotion of verbalizers and verbalizers presented with three considered video lecture types by two-way ANOVA with the Scheffe test.

Item	Sum of squares	Degree of freedom	Sum of mean squares	*F*	Sig.	Result of Scheffe test
Video lecture type	801.03	2	400.51	.92	.399	–
Cognitive style	284.14	1	284.14	.65	.419	–
Video lecture type × Cognitive style	85.90	2	42.95	.09	.905	–
Error	45,353.93	105	431.94			–
Total	611,529.88	111				–

Table 6
Descriptive statistics of positive emotion of verbalizers and verbalizers presented with three considered video lecture types.

Video lecture type	Cognitive style	Number of learners	Mean	Std.
Lecture capture	Verbalizers	20	18.24	22.37
	Visualizers	17	22.53	15.38
	Total	37	20.21	19.3
Voice over presentation	Verbalizers	20	15.26	18.45
	Visualizers	17	16.58	14.77
	Total	37	15.87	16.64
Picture-in-picture	Verbalizers	20	19.87	20.95
	Visualizers	17	21.55	14.95
	Total	37	20.64	18.21
Total	Verbalizers	20	17.79	20.39
	Visualizers	17	20.22	14.96
	Total	37	18.91	18.06

analytical results of the main effect, the video lecture methods negligibly affected positive emotion ($F = .78$, $p = .458 > .05$); in addition, cognitive styles negligibly affected positive emotion ($F = .48$, $p = .488 > .05$).

4.3. Effects of students' cognitive load

This section aims to examine whether the three considered video lectures leaded to that cognitive load of students differed significantly as well as whether cognitive load of visualizers and verbalizers who were presented with the three considered video lectures differed significantly. Table 8 summarizes the statistics for the cognitive load. Two-way ANOVA indicates that video lecture methods and cognitive styles for cognitive load have a significant interaction effect ($F = 3.98$, $p = .021 < .05$). According to analysis results of the simple main effect, the three video lecture approaches did not significantly affect the cognitive load of verbalizers ($F = 2.99$, $p = .058 > .05$); those approaches significantly affected cognitive load when visualizers were presented with three considered video lectures ($F = 9.86$, $p = .000 < .05$). Scheffe's test results indicate that the cognitive load of visualizers viewing the voice-over video lecture was significantly higher than that when students were viewing the lecture capture video and picture-in-picture video (Table 9). Moreover, according to analysis results of the simple main effect, cognitive load was not significantly affected when verbalizers and visualizers viewed the video capture lecture ($F = .98$, $p = .327 > .05$); cognitive load was significantly affected when verbalizers and visualizers viewed the voice-over type ($F = 7.62$, $p = .009 < .05$); and cognitive load was not significantly affected when verbalizers and visualizers viewed the picture-in-picture video ($F = 1.56$, $p = .219 > .05$). Scheffe's test results demonstrate that the cognitive load of visualizers viewing the voice-over presentation was significantly higher than that of verbalizers presented viewing the voice-over method (Table 10).

4.4. Effects of students' learning performance

By using the paired-sample *t* test, this section determines whether the three considered video lectures generated good learning performance based on pretest and posttest scores. Also, this section examines whether the three considered video lectures leaded to that learning performance of students differed significantly as well as whether learning performance of visualizers and verbalizers who were presented with the three considered video lectures differed significantly. Table 11 shows the paired-sample *t* test results of learning performance of all students presented with three considered video lecture types. Analytical results indicate that the three lecture types significantly improved learning performance. That is, the three considered video lectures all generated good learning performance. Table 12 lists the statistical results for the learning performance of verbalizers and visualizers. Two-way ANOVA results indicate that cognitive styles and video lecture types do not have a significant interaction effect ($F = .57$, $p = .565 > .05$) (Table 13). Analysis results of the main effect

Table 7
A comparison of positive emotion of verbalizers and verbalizers presented with three considered video lecture types by two-way ANOVA with the Scheffe test.

Item	Sum of squares	Degree of freedom	Sum of mean squares	*F*	Sig.	Result of Scheffe test
Video lecture type	526.61	2	263.30	.78	.458	–
Cognitive style	162.35	1	162.35	.48	.488	–
Video lecture type × Cognitive style	48.18	2	24.09	.07	.931	–
Error	35,183.82	105	335.08			–
Total	75,604.46	111				–

Table 8
Descriptive statistics of cognitive load of verbalizers and verbalizers presented with three considered video lecture types.

Video lecture type	Cognitive style	Number of learners	Mean	Std.
Lecture capture	Verbalizers	20	13.70	4.96
	Visualizers	17	12.23	3.80
	Total	37	13.02	4.46
Voice over presentation	Verbalizers	20	13.55	5.04
	Visualizers	17	18.00	4.69
	Total	37	15.59	5.31
Picture-in-picture	Verbalizers	20	10.55	3.63
	Visualizers	17	12.23	4.56
	Total	37	11.32	4.11
Total	Verbalizers	20	12.60	4.74
	Visualizers	17	14.15	5.08
	Total	37	13.31	4.94

Table 9
The one-way ANOVA results of cognitive load of visualizers presented with three considered video lecture types.

	Sum of squares	Degree of freedom	Sum of mean squares	*F*	Sig.	Result of Scheffe test
Cognitive load	376.62	2	188.31	9.86***	.000	Voice over presentation > Lecture capture; Voice over presentation > Picture-in-picture

*** indicates $p < .001$.

indicate that video lecture types significantly affected learning performance ($F = 35.77$, $p = .000 < .05$). Moreover, according to Scheffe's multiple comparison test, learning performance with the lecture capture and picture-in-picture methods was significantly better than that with the voice-over presentation method. Also, cognitive styles negligibly learning performance ($F = 3.52$, $p = .063 > .05$). That is, verbalizers and visualizers did not significantly differ in learning performance.

Table 14 summarizes how different cognitive styles of learners affect sustained attention, emotions, cognitive load, and learning performance when students were viewing the three video lectures.

5. Discussion

This study examines the effects of verbalizers and visualizers presented with three considered different video lectures on sustained attention, emotion, cognitive load, and learning performance in an autonomous online learning scenario. Analytical results confirm that although the three video lecture methods significantly promote learning performance, learning performance with the lecture capture and picture-in-picture types is superior to that of the voice-over type. This observation is consistent with the findings in several studies (Griffin et al., 2009; Ilioudi et al., 2013), confirming that the lecture capture approach is more effective than the Khan-style video lecture and asynchronously presenting PowerPoint slides and audio files. We can infer that adopting an inappropriate presentation layout for a computer screen and with little visual information integrated with learning material, the voice-over presentation method, results in the poorest learning performance among the three considered video lecture styles. Restated, the lecture capture and picture-in-picture types have a higher degree of media richness, integrate verbal and non-verbal (*i.e*. visual) elements more appropriately, and have better visual layout presenting multiple multimedia elements than those of the voice-over presentation type, ultimately improving learning performance.

Furthermore, the ATI hypothesis asserts that visualizers perform best on learning tests when they receive visual rather than verbal instruction; whereas verbalizers perform best on learning tests when they receive verbal rather than visual instruction (Massa & Mayer, 2006). Smith and Woody (2000) claimed that multimedia material benefits students who prefer visual representations. However, according to our results, the three considered video lecture types result in the same learning performance for verbalizers and visualizers. Restated, our results do not support the ATI hypothesis in that well-designed video lectures provide equivalent benefits in promoting the learning performance of verbalizers and visualizers, despite the fact that most video lecture types contain rich and multiple multimedia elements. These comparison results are consistent with those in several studies (Chen & Sun, 2012; Karakaya, Ainscough, & Chopoorian, 2001; Kollöffel, 2012), which examined how cognitive styles affect multimedia learning performance. Chen and Sun (2012) confirmed that video-based multimedia material induces the best learning performance and most positive emotions for verbalizers among the static text- and image-based, video-based, and animated interactive multimedia materials. Additionally, video-based and animated interactive multimedia materials are more appropriate for visualizers than static text and image-based multimedia materials. Their study does not support the ATI hypothesis. Moreover, Kollöffel (2012) examined the relationships between cognitive style (*i.e.* visualizers and verbalizers), cognitive abilities (spatial and verbal abilities), and learning performance when learning from multimedia materials. Analytical results indicate that the visualizer and verbalizer cognitive styles are unrelated to learning outcomes. Their study concluded that cognitive ability

Table 10
The one-way ANOVA results of cognitive load of verbalizers and visualizers presented with the voice over presentation type.

	Sum of squares	Degree of freedom	Sum of mean squares	*F*	Sig.	Result of Scheffe test
Cognitive load	181.96	1	181.96	7.62**	.009	Visualizer > Verbalizer

** indicates $p < .01$.

Table 11
Paired-sample *t* test results of learning performance of all students presented with three considered video lecture types.

Video lecture type	Learning performance	Number of learners	Mean	Std.	*t*	Sig.
Lecture capture	Pretest	37	2.35	1.33	21.69***	.000
	Posttest	37	8.62	1.13		
Voice over presentation	Pretest	37	2.51	1.40	11.53***	.000
	Posttest	37	6.40	1.60		
Picture-in-picture	Pretest	37	2.94	1.02	22.39***	.000
	Posttest	37	8.75	1.32		

*** indicates $p < .001$.

Table 12
Descriptive statistics of learning performance of verbalizers and visualizers presented with three considered video lecture types.

Video lecture type	Cognitive style	Number of learners	Mean	Std.
Lecture capture	Verbalizers	20	8.80	1.05
	Visualizers	17	8.41	1.17
	Total	37	8.62	1.11
Voice over presentation	Verbalizers	20	6.80	1.60
	Visualizers	17	5.94	1.51
	Total	37	6.40	1.60
Picture-in-picture	Verbalizers	20	8.85	1.49
	Visualizers	17	8.64	1.11
	Total	37	8.75	1.32
Total	Verbalizers	20	8.15	1.68
	Visualizers	17	7.66	1.76
	Total	37	7.92	1.73

(especially spatial visualization) and the extent to which a format allows cognitive processing influence learning outcomes, rather than a match between a preferred format and the format administered. Additionally, Karakaya et al. (2001) examined how student learning styles (i.e. convergers, divergers, assimilators, and accommodators) impact student learning performance in a multimedia lecture presentation setting, According to their results, students with different learning styles did not significantly differ in learning performance.

Additionally, according to our results, the voice-over presentation type generates the highest sustained attention and highest cognitive load of the three considered video lecture types. Our results further demonstrated that the effect of high sustained attention is likely derived from learners' high mental load or stress while learning via the voice-over presentation type, thus increasing their cognitive load. According to the cognitive load theory, the voice-over presentation type might inhibit learning when learners must split or distribute their attention to integrate three windows, respectively containing the lecturer's image, PowerPoint slides, and a table of content for slides. This integration process might exceed limited working memory capacity (Kalyuga, Chandler, & Sweller, 1999). This theory is supported by interview results of this study. Several interviewees mentioned that their visual attention was split by three windows while learning via the voice-over presentation type. When learning with the voice-over presentation method, those interviewees stated that they had to switch frequently between the lecturer's image and slides to compare the narration to what is written on slides, thereby increasing their mental load. Several interviewees also expressed that the three windows meant that the text in PowerPoint slides was small, thus increasing visual burden. Moreover, the multimedia richness theory indicates that when a communication medium is rich, little uncertainty and equivocality are associated with the learning task; little effort is thus required to use it (Lee et al., 2007; Sun & Cheng, 2007). The picture-in picture video is the richest, followed by the lecture capture type, and the voice-over video. This perspective indicates why cognitive load with the lecture capture and picture-in picture videos is lower than that with the voice-over presentation type. Additionally, this study found that sustained attention of verbalizers was significantly higher than that of visualizers when students were viewing the three video lectures, despite no difference in learning performance. The possible reason is that the three considered video lecture types tend to be multimedia materials. This may lead to that verbalizers who process words more effective than pictorial representations need to make more effort in keeping sustained attention for effective learning. In contrast, this study also found that the cognitive load for visualizers markedly exceeds that of verbalizers who are presented with the voice-over presentation type. Similarly, the possible reason is that the voice-over type contains less multimedia elements, compared to lecture capture and picture-in picture types. This may lead to that visualizers who process pictorial representations more effective than words need to make more mental effort in effective learning, thus generating higher cognitive load than verbalizers.

Our results further indicated that in addition to generating the highest mean sustained attention, the voice-over presentation type has the highest standard deviation for sustained attention. The voice-over presentation type likely generates the split-attention effect to some

Table 13
A comparison of learning performance of verbalizers and visualizers presented with three considered video lecture types by two-way ANOVA with the Scheffe test.

Item	Sum of squares	Degree of freedom	Sum of mean squares	*F*	Sig.	Result of Scheffe test
Video lecture type	130.75	2	65.37	35.77***	.000	Lecture capture > Voice over presentation; Picture-in-picture > Voice over presentation
Cognitive style	6.44	1	6.44	3.52	.063	–
Video lecture type × Cognitive style	2.10	2	1.05	.57	.565	–
Error	191.89	105	1.82			–
Total	7306.00	111				–

*** indicates $p < .001$.

Table 14
The summary of the effects of verbalizer and visualizer presented with three considered video lecture types on sustained attention, emotion, cognitive load, and learning performance.

Sustained attention	Video lecture type		Voice over presentation > Picture-in-picture
	Cognitive style		Verbalizer > Visualizer
Positive emotion	Video lecture type		–
	Cognitive style		–
Negative emotion	Video lecture type		–
	Cognitive style		–
Cognitive load	Video lecture type	Lecture capture	–
		Voice over presentation	Visualizer > Verbalizer
		Picture-in-picture	–
	Cognitive style	Verbalizer	–
		Visualizer	Voice over presentation > Lecture capture; Voice over presentation > Picture-in-picture
Learning performance	Video lecture type		Lecture capture > Voice over presentation; Picture-in-picture > Voice over presentation
	Cognitive style		–

degree since learners must switch their attention between the three windows. These analytical results are consistent with those in several studies (Kalyuga et al., 1999; Mayer & Moreno, 1998; Mousavi et al., 1995), indicating that presentations typically involve an audio and video recording of a lecturer, along with a visual presentation of slides. This circumstance is not ideal for learning since it divides visual attention, thus creating a split-attention effect. Namely, the video of a lecturer in the voice-over presentation type likely overloads a learner's visual channel, subsequently invoking the split-attention effect. Conversely, several studies (Church, Ayman-Nolley, & Mahootian, 2004; Gunawardena, 1995; Valenzeno, Alibali, & Klatzky, 2003) claimed that a video of a lecturer may improve learning in other ways. Gunawardena (1995) asserted that the video of a lecturer may give learners a sense of interacting with an actual person while watching a video lecture. In other words, the video of a lecturer may foster a sense of social presence. The social-cue hypothesis asserts that the outcome of most communications is most effective if they contain natural social cues (Rutter, 1984). In video lecture presentations, a pedagogical agent (e.g., an animated cartoon character or a video image of an actual instructor) occasionally provides social cues. Church et al. (2004) also demonstrated that gestures and other forms of nonverbal communication (i.e. visible in the video of a lecturer) enhance learning (Church et al., 2004; Valenzeno et al., 2003). Moreno, Mayer, Spires, and Lester (2001) also indicated that the presence of an on-screen agent does not produce the split-attention effect. However, analytical results of this study for the voce-over presentation type do not support the social-cue hypothesis. We infer that displaying a video of a lecturer in a separate window is the main factor decreasing the effects of natural social cues from an actual instructor.

Our results further demonstrated that the three video lecture types generate the same positive and negative emotions in verbalizers and visualizers. These analytical results contradict those acquired Chen and Wang (2011), demonstrating that video-based multimedia material generates the more positive emotions than static text- and image-based, video-based, and animated interaction-based multimedia materials. We infer that the main factor is that multimedia elements in the three video lecture types resemble each other, thus generating similar emotional responses.

Despite its contributions, this study has certain limitations. First, while evaluating only three of the most common video lecture methods, this study did not consider other video lecture types, such as the Khan-style video lecture (Chorianopoulos & Giannakos, 2013), which combines an interactive board with voice over. Moreover, this study was conducted in a highly controlled environment, possibly inducing demand characteristics that altered participants' performance. Specifically, participants were told they would be tested on the material in the video, possibly increasing their apprehensiveness and desire to be a 'good participant.' Moreover, the video lecture was shown in this study without the opportunity to pause, rewind, or take notes, possibly increasing the cognitive load to a level exceeding that without such constraints. Also, pedagogical strategies generally used in video lectures include receptive viewing, problem solving, and video podcasts (Kay, 2012), in which the pedagogical strategy used in the three video lectures was receptive viewing. This study did not consider other pedagogical strategies.

6. Conclusions and future work

This two-factor based study investigates whether the three considered video lecture styles, which present information differently, verbalizers and visualizers significantly differ in sustained attention, emotions, cognitive load, and learning performance. Analytical results confirm that although the three considered video lectures markedly promote learning performance, learning performance with the lecture capture and picture-in-picture types is significantly better than that with the voice-over type. Importantly, verbalizers and visualizers have the same learning performance with the three video types. According to our results, the voice-over presentation type generates the highest sustained attention, while the sustained attention of verbalizers is significantly higher than that of visualizers with each of the three considered video lectures. However, the voice-over presentation type generated the highest cognitive load among the three considered video lecture styles. Moreover, cognitive load of visualizers was significantly higher than that of verbalizers when students were viewing the voice-over type. Furthermore, the positive and negative emotions of learners induced by the three considered video lectures do not significantly differ while learning from the three video lectures. We conclude that compared to the voice-over video, with its relatively low production cost, both the lecture capture and picture-in-picture videos may be worthwhile for online learning from the perspectives of improved learning performance and reduced cognitive load, despite the fact that both videos are expensive and production intensive.

We recommend future studies in the following areas. First, with the rapid development of mobile learning in recent years, the extent to which different video lecture methods affect sustained attention, emotion, cognitive load, and learning performance for mobile devices with small screens warrant further study. Second, this study enrolled 37 undergraduate students from the Department of Chinese Literature at National Chengchi University, Taiwan. Future research should confirm whether research subjects in different academic levels have different learning experiences. Moreover, Mayer (2001) found that the redundancy effect may be inapplicable to lecture-style presentations, whereas notes or outlines (like those provided with PowerPoint text) could aid learner processing. Future research should investigate whether the redundancy or modality effect occurs with these three video types since they all contain multiple multimedia elements. Finally, future study should confirm, using eye-tracking technology (van Gog, & Scheiter, 2010), whether the split-attention effect occurs with the voice-over video method, which displays multimedia instructional elements in three windows. Applying eye movement measures to explore visual attention on the determined area of interest (AOI) in a video lecture is a highly effective means of clarifying how video lectures affect the sustained attention, cognitive load, and overall performance of learners.

Acknowledgements

The authors would like to thank the Ministry of Science and Technology of the Republic of China, Taiwan for financially supporting this research under Contract No. NSC 100-2511-S-004-001-MY3.

References

Ayres, P., & Sweller, J. (2005). The split-attention principle in multimedia learning. In R. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 135–146). New York: Cambridge University Press.
Astleitner, H., & Wiesner, C. (2004). An integrated model of multimedia learning and motivation. *Journal of Educational Multimedia and Hypermedia, 13*(1), 3–21.
Bassili, J. N. (2008). Media richness and social norms in the choice to attend lectures or to watch them online. *Journal of Educational Multimedia and Hypermedia, 17*(4), 453–475.
Brecht, H. D., & Ogilby, S. M. (2008). Enabling a comprehensive teaching strategy: Video lectures. *Journal of Information Technology Education, 7*, 71–86.
Broadbent, D. E. (1958). *Perception and communication*. London: Pergamon press.
Butcher-Powell, L. M. (2004). *Teaching, learning and multimedia. interactive multimedia in education and training*.
Carter, S. (1985). Is this the stuff that marketers are made of? *European Journal of Marketing, 19*(6), 53–64.
Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*, 293–332.
Chen, C. M., & Lin, Y. J. (2014). Effects of different text display types on reading comprehension, sustained attention and cognitive load in mobile reading contexts. *Interactive Learning Environments*. http://dx.doi.org/10.1080/10494820.2014.891526.
Chen, C. M., & Sun, Y. C. (2012). Assessing the effects of different multimedia materials on emotions and learning performance for visual and verbal style learners. *Computers & Education, 59*(4), 1273–1285.
Chen, C. M., & Wang, H. P. (2011). Using emotion recognition technology to assess the effects of different multimedia materials on learning emotion and performance. *Library & Information Science Research, 33*(3), 244–255.
Childers, T. L., Houston, M. J., & Heckler, S. E. (1985). Measurement of individual differences in visual versus verbal information processing. *Journal of Consumer Research, 12*, 125–134.
Chen, C. M., & Huang, S. H. (2014). Web-based reading annotation system with an attention-based self-regulated learning mechanism for promoting reading performance. *British Journal of Educational Technology, 45*(5), 959–980.
Cheng, Y. H., Cheng, J. T., & Chen, D. J. (2012). The effect of multimedia computer assisted instruction and learning style on learning achievement. *WSEAS Transactions on Information Science and Applications, 9*(1), 24–35.
Chorianopoulos, K., & Giannakos, M. N. (2013). Usability design for video lectures. In *The 11th European Conference on interactive TV and video, June 24–26, Como, Italy* (pp. 163–164).
Church, R. B., Ayman-Nolley, S., & Mahootian, S. (2004). The role of gesture in bilingual education: does gesture enhance learning? *International Journal of Bilingual Education and Bilingualism, 7*, 303–319.
Daft, R. L., & Lengel, R. H. (1986). Organizational information requirements, media richness and structural design. *Management Science, 32*(5), 554–571.
Danielson, J., Preast, V., Bender, H., & Hassall, L. (2013). Is the effectiveness of lecture capture related to teaching approach or content type? *Computers & Education, 72*, 121–131.
Day, J., Foley, J., & Catrambone, R. (2006). *Investigating multimedia learning with web lectures*. GVU Tech. Report GIT-GVU-06–25. GA: Georgia Institute of Technology.
Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology, 92*, 53–78.
Dunn, R. S., & Griggs, S. A. (2000). *Practical approaches to using learning styles in higher education* (2nd ed.). Westport: Bergin & Garvey.
Eysink, T. H., de Jong, T., Berthold, K., Kolloffel, B., Opfermann, M., & Wouters, P. (2009). Learner performance in multimedia learning arrangements: an analysis across instructional approaches. *American Educational Research Journal, 46*(4), 1107–1149.
Friedland, G. (2004). Solving the divided attention problem in lecture recordings. In C. Knauer (Ed.), *2. Doktorandenworkshop des Instituts für Informatik*. Department of Computer Science, Freie Universität Berlin. Technical Report B-04–15, 24-28.
Griffin, D. K., Mitchell, D., & Thompson, S. J. (2009). Podcasting by synchronising PowerPoint and voice: what are the pedagogical benefits? *Computers & Education, 53*(2), 532–539.
Gunawardena, C. N. (1995). Social presence theory and implications for interaction and collaborative learning in computer conferences. *International Journal of Educational Telecommunications, 1*(2/3), 147–166.
Hardy, J. L., Drescher, D., Sarkar, K., Kellett, G., & Scanlon, M. (2011). Enhancing visual attention and working memory with a web-based cognitive training program. *Mensa Research Journal, 42*(2), 13–20.
Homer, B. D., Plass, J. L., & Blake, L. (2008). The effects of video on cognitive load and social presence in multimedia-learning. *Computers in Human Behavior, 24*(3), 786–797.
Hornbæk, K., Engberg, D. T., & Gomme, J. (2002). Video Lectures: HCI and e-learning challenges. In *Workshop on Human-computer interaction and E-learning*.
Ilioudi, C., Giannakos, M. N., & Chorianopoulos, K. (2013). Investigating differences among the commonly used video lecture styles. In *WAVe 2013 the Workshop on Analytics on video-based learning* (pp. 21–26).
Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia learning. *Applied Cognitive Psychology, 13*, 351–371.
Karakaya, F., Ainscough, T. L., & Chopoorian, J. (2001). The effects of class size and learning style on student performance in a multimedia-based marketing course. *Journal of Marketing Education, 23*, 84–90.
Kay, R. H. (2012). Exploring the use of video podcasts in education: a comprehensive review of the literature. *Computers in Human Behavior, 28*(3), 820–831.
Kollöffel, B. (2012). Exploring the relation between visualizer–verbalizer cognitive styles and performance with visual or verbal learning material. *Computers & Education, 58*(2), 697–706.
Kettanurak, V. N., Ramamurthy, K., & Haseman, W. D. (2001). User attitude as a mediator of learning performance improvement in an interactive multimedia environment: an empirical investigation of the degree of interactivity and learning styles. *International Journal of Human-Computer Studies, 54*(4), 541–583.
Kappe, F. R., Boekholt, L., Den Rooyen, C., & Van der Flier, H. (2009). A predictive validity study of the Learning Style Questionnaire (LSQ) using multiple, specific learning criteria. *Learning and Individual differences, 19*(4), 464–467.
Lee, M. K. O., Cheung, C. M. K., & Chen, Z. (2007). Understanding user acceptance of multimedia messaging services: an empirical study. *Journal of the American Society for Information Science and Technology, 58*(13), 2066–2077.
Lim, K. H., & Benbasat, I. (2000). The effect of multimedia on perceived equivocality and perceived usefulness of information systems. *MIS Quarterly, 24*(3), 449–471.
Latham, C. (2006). Heart rate variability as an index of regulated emotional responding. *Review of General Psychology, 10*, 229–240.
Massa, L. J., & Mayer, R. E. (2006). Testing the ATI hypothesis: should multimedia instruction accommodate verbalizer-visualizer cognitive style? *Learning and Individual Differences, 16*, 321–335.
Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.
Mayer, R. E., & Massa, L. J. (2003). Three facets of visual and verbal learners: cognitive ability, cognitive style, and learning preference. *Journal of Educational Psychology, 95*(4), 833–846.

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: evidence for dual processing systems in working memory. *Journal of educational psychology, 90*(2), 312.

Mayer, R. E. (2011). Does styles research have useful implications for educational practice? *Learning and Individual Differences, 21*(3), 319–320.

Moreno, R., Mayer, R., Spires, H., & Lester, J. (2001). The case for social agency in computer-based teaching: do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction, 19*, 177–213.

Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology, 87*, 319–334.

McCraty, R., Atkinson, M., Tiller, W. A., Rein, G., & Watkins, A. D. (1995). The effects of emotions on short-term power spectrum analysis of heart rate variability. *American Journal of Cardiology, 76*(14), 1089–1093.

Ocepek, U., Bosnić, Z., Šerbec, I. N., & Rugelj, J. (2013). Exploring the relation between learning style models and preferred multimedia types. *Computers & Education, 69*, 343–355.

Osborn, D. S. (2010). *Using video lectures to teach a graduate career development course*. Retrieved from http://counselingoutfitters.com/vistas/vistas10/Article_35.pdf.

Rebolledo-Mendez, G., Dunwell, I., Martínez-Mirón, E., Vargas-Cerdán, M., de Freitas, S., Liarokapis, F., et al. (2009). Assessing NeuroSky's usability to detect attention levels in an assessment exercise. *HCI, new Trends* (pp. 149–158).

Rutter, D. R. (1984). *Looking and seeing: the role of visual communication in social interaction*. Chichester: John Wiley & Sons.

Schmidt-Weigand, F., Kohnert, A., & Glowalla, U. (2010). A closer look at split visual attention in system-and self-paced instruction in multimedia learning. *Learning and Instruction, 20*(2), 100–110.

Shorter, J., & Dean, R. (1994). Computing in collegiate schools of business: are mainframes & stand-alone microcomputers still good enough? *Journal of Systems Management, 45*(7), 36–41.

Smith, S. M., & Woody, P. C. (2000). Interactive effect of multimedia instruction and learning styles. *Teaching of Psychology, 27*(3), 220–224.

Sun, P. C., & Cheng, H. K. (2007). The design of instructional multimedia in e-learning: a Media richness theory-based approach. *Computers and Education, 49*(3), 662–676.

Sweller, J. (1999). *Instructional design in technical areas*. Camberwell, Victoria, Australia: Australian Council for Educational Research.

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

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*, 295–312.

Sorden, S. D. (2005). A cognitive approach to instructional design for multimedia learning. *Informing Science Journal, 8*, 263–279.

Tiller, W. A., McCraty, R., & Atkinson, M. (1996). Cardiac coherence: a new, noninvasive measure of autonomic nervous system order. *Alternative Therapies in Health and Medicine, 2*(1), 52–65.

Valenzeno, L., Alibali, M. W., & Klatzky, R. (2003). Teacher's gesture facilitate students' learning: a lesson in symmetry. *Contemporary Educational Psychology, 28*, 187–204.

van Gog, T., & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction, 20*(2), 95–99.

Wiese, C., & Newton, G. (2013). Use of lecture capture in undergraduate biological science education. *The Canadian Journal for the Scholarship of Teaching and Learning, 4*(2). Article 4.