

The Influence of Representational Formats and Learner Modality Preferences on Instructional Efficiency Using Interactive Video Tutorials

Chun-Ying Chen

General Education Center, National Taichung University of Education

No. 140, Minsheng Road, West District, Taichung City 40306

Email: cychen666@gmail.com

Received: July 24, 2020 Accepted: August 19, 2020 Published: August 21, 2020

doi:10.5296/jet.v7i2.17415

URL: <http://dx.doi.org/10.5296/jet.v7i2.17415>

Abstract

This study investigated how to create effective interactive video tutorials for learning computer-based tasks. The role of learner modality preferences was also considered. A 4×4 between-subjects factorial design was employed to examine the influence of instruction representational formats (noninteractive static, interactive static, interactive visual-only video with onscreen text, interactive video with audio narration) and learner modality preferences (visual, aural, read/write, multimodal) on instructional efficiency. Instructional efficiency was a combined effect of test performance and perceived cognitive load during learning. The results suggested that implementing interactivity into the video tutorials tended to increase transfer performance, and the role of modality preferences was related to learners' perceived cognitive load. The significant interaction effect on transfer efficiency indicated: (a) the auditory preference tended to exhibit better transfer efficiency with the narrated video, and (b) the read/write preference tended to exhibit better transfer efficiency with both the noninteractive static format and the captioned video. This study highlighted the importance of considering individual differences in modality preferences, particularly that of auditory and read/write learners.

Keywords: video tutorial, interactivity, instructional efficiency, learner modality preference, the VARK modalities

1. Introduction

As the increased interest in dynamic visuals, video tutorial becomes a popular multimedia instruction used for software training. This type of instruction involves a digital recording on a computer screen, in which an expert demonstrates the execution of interface procedures accompanied by narrated explanations. However, video may be ineffective due to information transience stemming from both video scenes and audio narrations. Given working memory's

limited capacity and duration, the transient information could impose a high cognitive load when learners need to hold and process information in working memory for extended periods. Compared to the written form of information (i.e., a series of static graphics and written commentaries) that is permanently available, the transient information effect (Leahy & Sweller, 2011) may lead to cognitive overload and that could hamper learning. The effect, therefore, assures the assumption that dynamic visualizations' general superiority over static graphics is not definitively supported (for reviews, see Höffler & Leutner, 2007; Tversky, Morrison, & Betrancourt, 2002).

Segmentation and learner control are techniques for overcoming the problems associated with the high cognitive load incurred when processing transient information (for a review, see Spanjers, van Gog, & van Merriënboer, 2010). These techniques are implemented to subdivide a lesson into manageable segments in which learners access the material at their own pace (Mayer, 2009). Interactivity is accordingly established by introducing content delivery control (e.g., pacing and sequencing control devices), which enables learners to adapt the instructional pace to their individual cognitive needs. However, findings on the cognitive benefits of introducing interactivity into multimedia instructions have been inconsistent (Lin & Hsieh, 2001; Höffler & Schwartz, 2011; Tabbers & Koeijer, 2010). Researchers, therefore, suggested that interactivity should be examined in the context of specific learning environments and learner variables (Höffler & Schwartz, 2011; Tabbers & Koeijer, 2010).

Besides interactivity, several moderating effects may arise from individual learner features, providing insight into specific design principles for effective learning. Do learner modality preferences for receiving instruction in different representational formats have any relation with instructional efficiency? Regarding individual learners, research on multimedia learning has focused mostly on cognitive styles in terms of verbalizer–visualizer differences (Chen & Sun, 2012; Höffler & Schwartz, 2011; Kollöffel, 2012; Massa & Mayer, 2006) or spatial ability (e.g., Kollöffel, 2012), but little is known about learner modality preferences (e.g., visual, aural, read/write, and multimodal).

2. Theoretical background

2.1 Segmentation, learner control, and interactivity

Both the cognitive theory of multimedia learning (CTML; Mayer, 2009) and cognitive load theory (CLT; Sweller, Ayres, & Kalyuga, 2011) underlie this study. Three assumptions underlying CTML are dual channels, limited capacity, and active processing: (a) humans possess independent, separate processing channels for pictorial and verbal information in working memory; (b) both of these channels are limited in duration and capacity; and (c) humans actively engage in cognitive processes (i.e., paying attention, organizing incoming information, and integrating incoming information with other knowledge) to take place in the pictorial and verbal channels (Mayer, 2009). The central tenet of CTML and CLT is that learners can engage in three kinds of cognitive processes within their available cognitive capacity. That is, both theories emphasize working memory constraints as determinants of instructional design effectiveness. Cognitive load can arise primarily from two sources: intrinsic and extraneous cognitive load. Intrinsic cognitive load is the complexity or difficulties

inherent in learning materials. Extraneous cognitive load is an unnecessary load caused by inadequately designed instruction. In this study, the extraneous cognitive load was determined according to how instruction representational formats facilitate or hinder the cognitive processes of learners. Another frequently used term, germane cognitive load, is derived from the effort that contributes to knowledge construction; it refers to working memory resources required to manage intrinsic cognitive load, thereby resulting in learning. Germane cognitive load is defined accordingly as intrinsic load because it is closely related to and dependent on intrinsic cognitive load (Sweller, 2010).

Clark and Mayer (2008) recommended several principles for using multimedia instructions to optimize learning based on cognitive theories and research evidence. The principles relevant to this study are segmentation and learner control. Segmentation and learner control are techniques for managing transient information stemming from dynamic visuals and implementing interactivity into instruction. There are various segmentation methods, including labeling and pacing. Labeling involves assigning a “label” to every meaningful segment within a lesson to organize the target learning content (Ertelt, Renkl, & Spada, 2005). Pacing involves learner control over instruction that enables learners to control content sequencing and pacing by using options such as navigational menus and pacing devices. Interactivity is established by introducing (a) a navigational menu in which labeled segments constitute the structure of the instructions, and (b) pacing in interactive devices that enable playing, pausing, forwarding, and rewinding. Such controls enable learners to match the instructional pace to their individual cognitive needs, thereby engaging them in learning.

Research on the effectiveness of learner control in multimedia instructions has yielded mixed findings, indicating that interactivity is not always beneficial to learning (Lin & Hsieh, 2001; Höffler & Schwartz, 2011; Tabbers & Koeijer, 2010). Most studies have reported that using learner-controlled pacing and segmentation reduces cognitive load and increases transfer performance (e.g., Ertelt et al., 2005; Hasler, Kersten, & Sweller, 2007; Mayer & Chandler, 2001; Mayer, Dow, & Mayer, 2003; Tabbers & Koeijer, 2010). However, other studies have not observed these positive effects of learner control (e.g., Lowe, 2004; Schnotz, Böckheler, & Grzondziel, 1999). Learner control may not foster learning because novices do not know how to use interactive devices effectively (Lowe, 2004) or the interactive features increase cognitive load (Schnotz et al., 1999). By investigating differences in the effects of pacing among dynamic and nondynamic representations, Höffler and Schwartz (2011) observed that for animations, self-pacing was effective, whereas for static pictures, self-pacing was ineffective.

Therefore, learner control must be applied with caution in instructing novice learners. Studies have suggested reducing the extraneous cognitive load that learner control imposes on novices. The availability of too many interactive devices might lead to cognitive overload (Scheiter & Gerjets, 2007). Novices are suggested to have limited control over the appearance of the modeled performance, such as zoom in and zoom out on a specific part of the performance; or observing the model’s performance from different angles (Wouters, Tabbers, & Paas, 2007). Those control options over the appearance of the demonstrations other than those of controlling the pace of presentation may require extra cognitive resources from novice learners. Additionally, it was suggested that pacing with predefined segments may properly balance the

amount of information processed in working memory at one time, and reduce temporal split attention (Kalyuga, 2007; Wouters et al., 2007).

2.2 Learning modality preferences

Learning styles have at least four general dimensions in terms of cognitive, affective, physiological, and psychological aspects, as evidenced in the variety of instruments (for reviews, see Jonassen & Grabowski, 1993; Smith, 1982; Sternberg & Zhang, 2001). Advocates of learning style theory assert that instructional approaches should be adapted according to student learning styles (or preferences), although there is still little agreement about a precise definition of learning styles, and no firm conclusions were obtained from previous research (for a critique of learning styles, see Stahl, 1999).

Most modern style theories focus more on the cognitive aspect of learning style, such as those that take a visual/verbal approach to learning (visualizers vs. verbalizers) (Riding, 2001). Asserting that verbalizers learn better with verbal instructions, whereas visualizers learn better with visual instructions, that so-called aptitude-treatment interaction (ATI) effects have generated no firm conclusions (Massa & Mayer, 2006). Measuring three facets of the verbalizer-visualizer dimension in terms of cognitive ability (i.e., proficiency in creating, holding, and manipulating spatial representations), cognitive style (i.e., tendency to think with images or words), and learning preference (i.e., preference for receiving instruction involving images or words), Massa and Mayer obtained no support for the ATI hypothesis when providing online help in either a verbal format (presented as a text definition) or a pictorial format (presented as an illustration) to teach basic electronics. However, their results are consistent with the multimedia effect: people learn better from words and pictures than from words alone. Likewise, Kollöffel (2012) obtained no support for the ATI hypothesis after comparing a verbal version (a combination of text and arithmetic) with a visual version (a combination of tree diagrams and arithmetic) of simulation instruction on combinatorics and probability theory. Kollöffel reported that cognitive ability (particularly spatial visualization) and the extent to which a representational format affords cognitive processing might influence learning outcomes. Besides, Kollöffel asserted that a format designed to support cognitive processing more effectively facilitates learning than does a format consistent with the learners' preferred format. However, the pictorial materials that they used were only static visuals, such as illustrations or graphs, and the verbal materials used were only printed words. Dynamic visuals (i.e., animations or videos) and another form of verbal information (i.e., spoken words) were not used to test the ATI effect.

In contrast, some studies that utilized dynamic visuals into instructions found the ATI effect. Chen and Sun (2012) obtained partial support for the ATI hypothesis, reporting that dynamic multimedia materials containing video and animation are more appropriate for visualizers than are static instructions, whereas video-based multimedia materials are appropriate for verbalizers. Höffler and Schwartz (2011) found that (a) learners tending toward a visual cognitive style learned significantly better with animations than with static pictures; and (b) although not significantly different, learning outcomes of learners tending toward a verbal style were descriptively better with static pictures. In teaching the role of surfactants during the washing process, the study of Höffler and Schwartz included four versions of a computer-based

learning environment in which either system-paced narrated static pictures and animations or learner-paced narrated static pictures and animations were used. The instruction content was delivered either through a short animation (73 s) or a series of four static pictures representing the key moments of the process. The two learner-paced versions contained control functions, such as play, pause, rewind, and fast-forward buttons, to enable learners to pace their progress.

Focusing on the physiological aspect of learning style, the Visual, Aural, Read/write, Kinesthetic (VARK) model (Fleming, 2001) that deals with perceptual modes to identify individual's instructional preference suits the purpose of this study. Fleming suggested that individuals receive information through sensory modalities and have sensory modality preferences. Visual learners prefer symbolic representations of information, such as graphs, charts, maps, logos, and diagrams; aural learners prefer aural forms of information, such as lectures, tutorials, and discussion with other students and instructors; read/write learners prefer information printed as words; and kinesthetic learners are multimodal and have perceptual preferences related to the use of experience and practice (Fleming, 2001; Fleming & Mills, 1992). Notably, Fleming defined the visual preference as not including "real" pictures such as photographs, movies, videos or PowerPoint because he argued that these presentations also involve aural, read/write, and kinesthetic perceptual modes.

3. Experiment design and hypotheses

A 4 x 4 between-subjects factorial design was employed to explore the relation between representational formats and learner modality preferences (visual, aural, read/write, multimodal) on instructional efficiency. Instructional efficiency was a combined effect of test performance and perceived cognitive load.

Four instructional formats were developed and compared: (a) noninteractive static instruction + onscreen text (noninteractive static); (b) interactive static instruction + onscreen text (interactive static); (c) interactive video + onscreen text (interactive demo-text); and (d) interactive video with audio narrations (interactive demo-spoken). Onscreen text refers to verbal explanations in the form of printed words whereas audio narrations in the form of spoken words. Two formats were static instructions (manuals containing static graphics, namely screenshots obtained from the demonstrations of the execution of interface procedures and accompanied by explanatory text), whereas the second two were videos. The noninteractive static, interactive static, and interactive demo-text formats were visual-only instructions, whereas the interactive demo-spoken format was a dual-modality instruction. The noninteractive static format designed without interactivity was used as a control condition that imitated a traditional text manual with a linear content sequence.

Interactivity was implemented by applying the segmentation and learner control techniques that were assumed to increase video effectiveness by alleviating the transient information effect and mimicry learning. Therefore, the main effect of representational formats on instructional efficiency in favor of video formats as compared to static formats was expected.

Research has indicated that self-pacing is effective for narrated animations, whereas it is ineffective for narrated static pictures (Höffler & Schwartz, 2011). Höffler and Schwartz reported that the unexpected combination of static graphics with audio narrations and control

devices associated with animations engendered extraneous cognitive load. To avoid this deficiency, the interactive static format in this study was implemented without audio narrations and content pacing devices (i.e., play, pause, forward, and rewind), but employed only the segmenting principle for content navigation (i.e., a navigational menu). Accordingly, the interactive static format was expected to be superior to the noninteractive static format.

Based on the inconsistent results obtained in multimedia studies regarding presentation modality (Leahy & Sweller, 2011; Tabbers, Martens, & van Merriënboer, 2004; Wong, Leahy, Marcus, & Sweller, 2012), an interactive dual-modality format containing audio narration (the demo-spoken condition) and another interactive visual-only format containing onscreen text (the demo-text condition) were created for comparison. To avoid the redundancy effect (Clark & Mayer, 2008), a format in which auditory and visual modalities are used simultaneously for identical explanatory words was considered an ineffective technique and was not created. Because each video scene contained visual complexities created by manipulating the graphical user interface elements of the software application, a split attention effect (Sweller et al., 2011) was more likely to occur in the demo-text condition because participants' visual attention is split between viewing the video scenes and reading the concurrent onscreen text. Therefore, participants in the demo-spoken condition were expected to perceive lower cognitive load than those in the demo-text condition and, thus, exhibit better instructional efficiency. Similarly, the interactive static condition that relied only on the visual channel could lead to a higher perceived cognitive load during learning. Accordingly, participants provided with audio narration were expected to exhibit better instructional efficiency than participants provided with onscreen text in both interactive static and video conditions.

Based on the assumption that matching teaching to learning is effective, it was hypothesized that a match between representational formats and learner modality preferences would enhance learning. Therefore, if an interaction between representational formats and participant modality preferences would occur, there might be two expectations. First, if video formats are more effective for auditory preference, then a modality effect might occur because of the preference of auditory learners for audio narration. Second, if static formats are more effective for read/write learners, a modality effect might not occur because of the preference of read/write learners for processing information in text form. For multimodal learners, whose preference for a certain modality is not strong; and for visual participants, who prefer symbolic presentation of information that was different from either the static or dynamic demonstrations examined in this study, the interaction effect might not occur to them.

The aforementioned discussion led to the following hypotheses:

Hypothesis 1. Participants presented with video conditions would exhibit better instructional efficiency than participants presented with static conditions.

Hypothesis 2. Participants presented with the interactive static condition would exhibit better instructional efficiency than participants presented with the noninteractive static condition.

Hypothesis 3. Participants presented with audio narration would exhibit better instructional efficiency than participants presented with onscreen text in both interactive static and video conditions.

Hypothesis 4. For auditory participants, the superiority of video formats over static formats was expected to be more pronounced for audio narration than for onscreen text.

Hypothesis 5. For read/write participants, the superiority of static formats over video formats was expected to be more pronounced for onscreen text than for audio narration.

4. Method

4.1 Participants

The participants were 358 university students (19–25 years of age; 207 females and 151 males) in Taiwan, where the official language is traditional Chinese. The participants represented various academic disciplines excluding computers and graphic design majors. None of the participants had previous experience with the Adobe Illustrator software. All participants were accustomed to working with computers, and the experiment was conducted in the computer lab where they normally attended computer-based classes.

4.2 Learning Materials

The web-based instructions for this experiment that taught the participants how to draw with Illustrator's pen tool (a Bezier tool) were identical to the one used in the study of Chen and Yang (2020). The instructions began with a starting page, which was a short textual introduction to the content. The content comprised five parts: (a) an outline of the instructions (the starting page), (b) pen tool basics, (c) layers and placing images, (d) beginning to trace, and (e) editing paths. The instructions were divided accordingly into five segments, and each segment was labeled. The labeled segments were presented as a navigational menu on the left side of the screen, whereas the instructional content was displayed on the right side. The link of the selected working segment was highlighted in a different color from other links on the navigational menu.

Interactivity was implemented by using the following control options for content sequencing and pacing: (a) the navigational menu, (b) stop and play buttons and a slide bar for fast-forwarding and rewinding, (c) a volume button for adjusting the narration volume, and (d) a full-screen button to display the content. The noninteractive static format, as a control condition, was not segmented and used a scrolling-page design. The interactive static format employed the same scrolling-page design and contained the navigational menu. Interactive demo-text and demo-spoken formats were similar in information segmentation to the interactive static format and were designed with the same navigational menu. Also, the interactive video formats contained control options for content pacing and sequencing, but the volume button was disabled in the demo-text condition. After selecting a segment of the interactive video by clicking a link on the navigational menu, the participants were required to click the play button to start the demonstration. The demonstration automatically stopped after each segment and the participants were required to click any link on the navigational menu to continue the instructions. All formats were accompanied by identical commentaries, but the demo-text version presented the commentaries as written captions at the bottom of the screen to the matched demonstrations simultaneously. For video formats, the total length of all segments was 13:08 min (788 s), and each segment was displayed for 111–284 s, representing an average

of 3:17 min (197 s) per segment.

4.3 Measures

4.3.1 VARK questionnaire

The complete VARK questionnaire, version 7.1 (the version used in this study) can be accessed from the VARK website (www.vark-learn.com). The VARK contains 16 questions with four response options. Each option is associated with a different modality preference. An example question is shown below:

You have a problem with your heart. You would prefer that the doctor:

- described what was wrong.
- showed you a diagram of what was wrong.
- gave you something to read to explain what was wrong.
- used a plastic model to show what was wrong.

The VARK's popularity comes from its face validity, simplicity, and ease of use (Leite, Svinicki, & Shi, 2010). The research by Leite et al. indicated that the reliability estimates for the VARK subscale scores were 0.85, 0.82, 0.84, and 0.77 for the visual, aural, read/write, and kinesthetic subscales, respectively, which are considered adequate.

4.3.2 Recall test and transfer task

A recall test and a transfer task that were used to measure learning outcomes were identical to those used in the study of Chen and Yang (2020). The recall test comprised five true-or-false questions and five multiple-choice questions. One point was awarded for each correct answer, and the maximum score was 10. The transfer task required participants to trace an image with a simple contour by using the Adobe Illustrator pen tool. No instruction was provided on how to perform this task. The task required 14 major components to be traced, and one point was awarded for each correctly completed component; the maximum score was 14. No access to the learning material was allowed during the test phases.

4.3.3 Perceived cognitive load

Cognitive load ratings were assessed using one item based on a scale developed by Paas (1992). The item was on a 9-point-scale and that measured the perceived invested mental effort of the participants, who indicated the amount of effort they exerted to follow the instructions (1 = *lowest* and 9 = *highest*).

4.3.4 Instructional efficiency

To determine the instructional efficiency of the four instruction formats, efficiency scores were calculated as the joint function of test performance (the recall test or the transfer task) and perceived cognitive load to attain the test performance by adopting Paas and van Merriënboer's

(1993) technique. This approach is based on the z-transformation of perceived cognitive load values and raw performance scores. Instructional efficiency (E) was determined using the formula $E = (Z_{\text{Performance}} - Z_{\text{cognitive load}}) / \sqrt{2}$. Better efficiency indicated that high performance is obtained with relatively low perceived cognitive load.

4.4 Procedures

The participants were initially classified into the four modality preference styles using the VARK questionnaire; the kinesthetic learners were regarded as multimodal learners according to Fleming and Mills (1992). The data were collected in a group setting in which each participant worked independently on a computer. The experiment was conducted for a total of 10 sessions. In each session, approximately 40 participants with various modality preferences were assigned randomly to one of the four learning environment formats and were tested simultaneously.

The experiment comprised an orientation and a learning phase followed by two test phases. Self-reported cognitive load was administered to the participants immediately after the learning phase. The experiment began with an orientation in which the authors described the procedures and demonstrated how to use each version of the learning material for each group individually (each group did not see other versions except the format of their own). Time-on-task was restricted in both the learning phase and the test phases. The total length of the video was 13:08 min. Thus, it was estimated that viewing all of the materials required approximately 15 min. For the participants to have sufficient time to view all of the learning materials, they were allotted 30 min to complete the learning phase. The participants were allotted 10 min to complete the recall test and 40 min to complete the transfer task.

5. Results

The data were analyzed using a 4×4 between-subjects ANOVA design. The dependent variables were test performance regarding recall and transfer, perceived cognitive load during learning, and instructional efficiency. ANOVA was conducted using the general linear model function in SPSS software. Analyses were separately conducted for dependent measures. Fisher's least significant difference (LSD) test was conducted as a post hoc test when a main effect was significant. Simple main effects were further conducted when an interaction effect was significant. The alpha was set at 0.05 for all statistical tests.

Learning outcomes measured in this study included the recall test and the transfer task. For recall performance, neither the main effects of both representational formats and modality preferences nor the interaction effect between the two factors was observed ($p > .05$). Accordingly, only the results regarding transfer performance, perceived cognitive load, and transfer efficiency were reported. Table 1 shows the descriptive statistics for dependent variables with respect to representational formats and modality preferences.

Table 1. Descriptive statistics for dependent variables with respect to representational formats and modality preferences

Representations	Modalities	Transfer performance (0-14)	Perceived cognitive load (1-9)	Transfer efficiency
G1	V ($n = 16$)	10.00 (2.58)	6.38 (1.59)	-0.37 (0.91)
	A ($n = 20$)	9.55 (2.35)	6.90 (1.25)	-0.71 (0.83)
	R ($n = 17$)	11.53 (1.33)	6.00 (1.73)	0.27 (0.87)
	M ($n = 34$)	11.15 (2.11)	6.76 (1.48)	-0.14 (0.96)
	Total ($n = 87$)	10.64 (2.24)	6.57 (1.52)	-0.24 (0.95)
G2	V ($n = 16$)	11.56 (2.28)	6.00 (1.51)	0.26 (0.82)
	A ($n = 24$)	10.46 (1.87)	6.92 (1.18)	-0.46 (0.79)
	R ($n = 18$)	10.44 (2.43)	5.83 (1.30)	0.05 (1.01)
	M ($n = 34$)	11.41 (2.38)	5.79 (1.65)	0.31 (1.05)
	Total ($n = 92$)	11.00 (2.27)	6.13 (1.50)	0.05 (0.98)
G3	V ($n = 17$)	11.35 (2.47)	6.12 (1.73)	0.14 (0.94)
	A ($n = 19$)	11.42 (2.71)	7.32 (1.34)	-0.32 (1.04)
	R ($n = 16$)	12.06 (2.18)	6.25 (2.05)	0.36 (1.26)
	M ($n = 41$)	11.61 (2.49)	6.39 (1.76)	0.14 (0.85)
	Total ($n = 93$)	11.60 (2.46)	6.51 (1.76)	0.08 (0.99)
G4	V ($n = 17$)	10.71 (2.42)	5.88 (1.36)	0.02 (0.89)
	A ($n = 22$)	11.23 (1.88)	5.59 (1.50)	0.36 (0.77)
	R ($n = 21$)	12.10 (2.05)	7.00 (1.05)	0.02 (0.71)
	M ($n = 26$)	11.23 (2.16)	6.46 (1.86)	-0.02 (0.86)
	Total ($n = 86$)	11.34 (2.13)	6.26 (1.57)	0.10 (0.81)
Total	V ($n = 66$)	10.91 (2.46)	6.09 (1.53)	0.01 (0.90)
	A ($n = 85$)	10.66 (2.28)	6.66 (1.45)	-0.28 (0.93)
	R ($n = 72$)	11.54 (2.11)	6.31 (1.58)	0.16 (0.95)
	M ($n = 135$)	11.37 (2.29)	6.35 (1.70)	0.08 (0.94)
	Total ($n = 358$)	11.15 (2.30)	6.37 (1.59)	-0.00 (0.94)

Note. G1 = noninteractive static; G2 = interactive static; G3 = interactive demo-text; G4 = interactive demo-spoken; V = visual; A = auditory; R = read/write; M = multimodal.

5.1 Transfer performance

The main effect of representational formats on transfer performance was significant ($F(3, 342) = 3.267, p = .022$, partial $\eta^2 = .03$), but the main effect of modality preferences on transfer performance was not significant ($F(3, 342) = 2.587, p = .053$, partial $\eta^2 = .02$). There was no significant interaction between these two factors ($F(9, 342) = 1.300, p = .235$, partial $\eta^2 = .03$). LSD showed that the interactive demo-text and demo-spoken conditions scored significantly higher than the noninteractive static condition ($p = .005$ and $.044$, respectively); significant differences from others were not observed ($p > .05$).

5.2 Perceived cognitive load during learning

The main effects of both representational formats and modality preferences were not significant ($p > .05$), but the interaction effect between the two factors was observed ($F(9, 342) = 2.684, p = .005$, partial $\eta^2 = .07$). Further analysis of simple main effects revealed that (a) for the

interactive static condition, the auditory preference perceived significantly higher cognitive load than the read/write and multimodal preferences ($p = .026$ and $.007$, respectively); (b) for the demo-text condition, the auditory preference perceived significantly higher cognitive load than the visual, read/write, and multimodal preferences ($p = .021$, $.044$ and $.032$, respectively); and (c) for the demo-spoken condition, the read/write preference perceived significantly higher cognitive load than the visual and auditory preferences ($p = .028$ and $.003$, respectively). Figure 1 showed that (a) the auditory preference tended to perceive less cognitive load in the demo-spoken condition, but perceive higher cognitive load in the demo-text condition; and (b) the read/write preference tended to perceive less cognitive load in both static conditions, but perceive higher cognitive load in the demo-spoken condition.

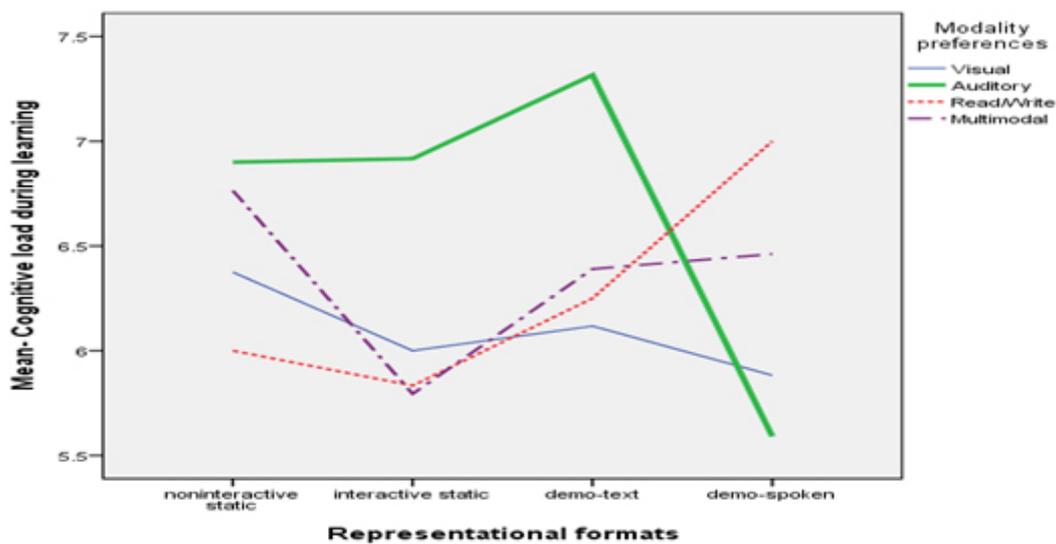


Figure 1. Effects of representations and modalities on perceived cognitive load

5.3 Transfer efficiency

The main effect of representational formats on transfer efficiency was not significant ($F(3, 342) = 2.373$, $p = .070$, partial $\eta^2 = .02$), but the main effect of modality preferences on transfer efficiency was significant ($F(3, 342) = 3.832$, $p = .010$, partial $\eta^2 = .03$). Significant interaction between these two factors was also observed ($F(9, 342) = 2.089$, $p = .030$, partial $\eta^2 = .05$).

Further analysis of simple main effects revealed that (a) for the noninteractive static condition, the read/write preference obtained significantly better transfer efficiency than the visual and auditory preferences ($p = .046$ and $.001$, respectively); and the multimodal preference obtained significantly better transfer efficiency than the auditory preference ($p = .029$); (b) for the interactive static condition, the visual and multimodal preferences obtained significantly better transfer efficiency than the auditory preference ($p = .015$ and $.002$, respectively); (c) for the demo-text condition, the read/write preference obtained significantly better transfer efficiency than the auditory preference ($p = .031$); and (d) for the demo-spoken condition, no significant differences were observed ($p > .05$). Figure 2 showed that (a) although not statistically significant, the auditory preference was more likely to exhibit better transfer efficiency with the demo-spoken condition; and (b) the read/write preference was more likely to exhibit better

transfer efficiency for both the noninteractive static and demo-text conditions as compared to other modality preferences.

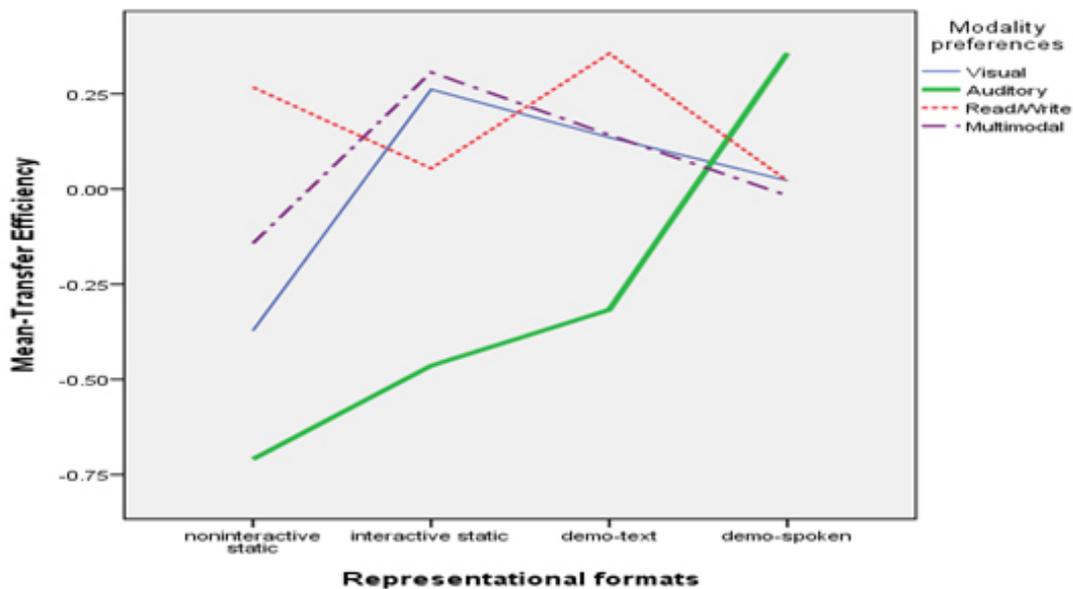


Figure 2. Effects of representations and modalities on transfer efficiency

Nevertheless, one-way ANOVA was further conducted to test Hypotheses 4 and 5. The main effect of representational formats on transfer efficiency for auditory participants was significant ($F(3, 81) = 6.083, p = .001$). For auditory participants, LSD showed that the demo-spoken condition tended to have better transfer efficiency than the noninteractive static, interactive static, and demo-text conditions ($p = .000, .002, \text{ and } .014$, respectively); significance differences from others were not observed ($p > .05$). Unexpectedly, for read/write participants, the main effect of representational formats on transfer efficiency was not significant ($F(3, 68) = 0.506, p = .679$).

Accordingly, for transfer efficiency, only Hypothesis 4 was supported whereas other hypotheses were not supported. The main effect of representational formats on transfer efficiency was not significant, so Hypothesis 1, 2, and 3 were not supported. However, the main effect of this factor on transfer performance was significant and the results indicated that participants presented with video conditions outperformed participants presented with the noninteractive static condition. For Hypothesis 2, although not statistically significant, the interactive static format descriptively exhibited better transfer efficiency than the noninteractive static format. Additionally, the former obtained positive transfer efficiency, whereas the latter obtained a negative one. As to Hypothesis 3, video with audio narrations did not exhibit better transfer efficiency than both the interactive static and video with captions. Further, a significant interaction effect of the two factors on transfer efficiency was observed and the results supported Hypothesis 4. Hypothesis 5 was not supported because no significant differences were observed among the four representational formats for read/write participants.

6. Discussion

This study aimed to explore the relation between representational formats and learner modality preferences on the effectiveness of interactivity when learning with video tutorials for software training. The main effect of representational formats for transfer performance in favor of interactive videos as compared to the noninteractive static format was found but was not observed for transfer efficiency. A significant interaction on transfer efficiency between the two factors was found, indicating that the role of modality preferences was related to learners' perceived cognitive load during learning.

For transfer performance, a significant main effect was found, but no ATI effect was obtained. In this case, the findings support Kollöffel's (2012) assertion that the extent to which a representational format affords to cognitive processing tended to determine learning outcomes. The main effect was in favor of both the interactive narrated and captioned videos as compared to the noninteractive static format, but not to the interactive static format. The video conditions did not outperform the interactive static condition probably due to the relatively long length of each video segment (averagely 3:17 min per segment in this study). The segment length interaction effect proposed by Wong et al. (2012) may help to explain the results. That is, for transient information including dynamic visuals and auditory input presented in long segments, (a) dynamic visuals could lose their superiority over static visuals, and (b) the modality effect under which audio-visual presentation is superior to visual-only presentation could disappear or reverse, due to working memory overload associated with large amounts of transient information (Wong et al., 2012). The cognitive demands coming from the transient video scenes or audio narrations were mitigated in the permanent written form of the interactive static format when the length of video segments was relatively long. Nevertheless, it may not be practical to divide related instruction into smaller segments that may hinder learners from seeing the interactions between segments. In this case, other cognitive support (e.g., design of memory aid) besides segmentation and learner control may be needed to defeat the effect of information transience (Chen, 2016). Further research is necessary to assure this issue.

The significant interaction effect illuminates the relation between representational formats and learner modality preferences on transfer efficiency when learning with interactive videos. Some observations regarding the perceived cognitive load should be discussed beforehand. First, it was observed that the auditory preference tended to perceive lower cognitive load in the demo-spoken condition whereas the read/write preference was the opposite. Additionally, the auditory preference tended to perceive a higher cognitive load in the demo-text condition. The observations imply that (a) aural forms of information was beneficial to the auditory preference, but tended to demand the read/write preference more cognitive capacity; and (b) text information such as captions tended to demand the auditory preference more cognitive capacity. Second, it was observed that the read/write preference tended to perceive less cognitive load in both static conditions. It implies that the static format, no matter interactive or not interactive, is a particularly suitable representational format for the read/write preference. As for the visual and multimodal preferences, perceived cognitive load among the four conditions did not differ significantly, so the main effect of representational formats was more applicable to these types of modality preference.

Accordingly, the interaction effect on transfer efficiency was particularly pronounced for auditory and read/write preferences. With the demo-spoken condition, the auditory participants tended to score high in transfer performance and perceive less cognitive load, thereby resulting in superior transfer efficiency. It implies that the narrated video tended to be effective for the auditory preference. For read/write participants, the noninteractive static format seems to be effective because they obtained good transfer efficiency by scoring high in transfer performance and perceiving lower cognitive load with this format. Read/write participants also obtained good transfer efficiency with other formats, particularly with the two video formats, but by scoring high in transfer performance as well as perceiving high cognitive load. They might compensate for an increase in mental load by investing more mental effort to maintain performance at a constant level (Sweller, van Merriënboer, & Paas, 1998). For visual and multimodal preferences, the results indicated that the three interactive formats were effective to them as compared to the noninteractive format, with which they obtained negative efficiency.

Concluding, the interaction effect indicated that (a) the auditory preference tended to exhibit better transfer efficiency with the narrated video; and (b) the read/write preference tended to exhibit better transfer efficiency with both the noninteractive static format and the captioned video. Unexpectedly, the read/write participants in the interactive static condition did not perform as well as their counterparts in the noninteractive condition. It was suspected that they might tend to be familiar with the traditional linear structure of textbooks when using static instructions. Further research is needed to resolve this issue.

7. Conclusion and limitations

Regarding the interaction effect on transfer efficiency, the results provide support elucidating the importance of considering learner modality preferences in developing multimodal instructions. The combined effect of representational formats and modality preferences tended to be stronger among auditory and read/write participants than among visual and multimodal participants. In this study, spoken words tended to be more beneficial to auditory learners, whereas written text tended to be more beneficial to read/write learners. When employing effective design principles in multimedia learning to reduce extraneous cognitive load, a representational format designed to match learners' preferred format should not be ignored, particularly for auditory and read/write learners. To substantiate the findings that the interaction effect was more pronounced among auditory and read/write participants, future studies should examine other learning tasks and various types of the learning environment.

There were limitations to this study. First, to obtain reliable assessments of cognitive load and learning outcomes, participants' level of familiarity with that task setting and environment features should be further controlled in this study. Second, this study used a self-reported one-item mental effort rating scale to measure the overall cognitive load experienced by learners. Measurement of different types of cognitive load separately can advance a more fine-grained examination on the sources of cognitive load. Third, the VARK questionnaire was suggested to be used with caution (Leite et al., 2010). That is, the prevalence of examinees classified as having multiple learning preferences is high because it allows multiple answers to each question.

References

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*, 1273-1285.

Chen, C-Y. (2016). Cognitive support for learning computer-based tasks using animated demonstration. *Interactive Learning Environments, 24*(4), 859-874.

Chen, C-Y., & Yang, Y-H. (2020). Investigation of the effectiveness of common representational formats in online learner-paced software training materials. *Innovations in Education and Teaching International, 57*(1), 97-108.

Clark, R.C., & Mayer, R.E. (2008). *E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning* (2nd ed.). San Francisco, CA: Pfeiffer.

Ertelt, A., Renkl, A., & Spada, H. (2005). Learning a new computer application using onscreen videos. In A. Méndez Vilas, B.G. Pereira, J.M. Gonzáles, & J.A.M. Gonzáles (Eds.), *Recent research developments in learning technologies Vol.2* (pp.757-762). Badajoz : Formatex.

Fleming, N.D. (2001). *Teaching and learning styles: VARK strategies*. Christchurch, New Zealand: N.D. Fleming.

Fleming, N.D., & Mills, C. (1992). Not another inventory, rather a catalyst for reflection. *To Improve the Academy, 11*, 137-155.

Hasler, B.S., Kersten, B., & Sweller, J. (2007). Learner control, cognitive load and instructional animation. *Applied Cognitive Psychology, 21*, 713-729.

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

Höffler, T.N., & Schwartz, R.N. (2011). Effects of pacing and cognitive style across dynamic and non-dynamic representations. *Computers & Education, 57*, 1716-1726.

Jonassen, D.H., & Grabowski, B.L. (1993). *Handbook of individual differences, learning, and instruction*. Hillsdale, NJ: Lawrence Erlbaum.

Kalyuga, S. (2007). Enhancing instructional efficiency of interactive e-learning environments: A cognitive load perspective. *Educational Psychology Review, 19*, 387-399.

Kollöffel, B. (2012). Exploring the relation between visualizer-verbalizer cognitive styles and performance with visual or verbal learning material. *Computers & Education, 58*, 697-706.

Leahy W., & Sweller, J. (2011). Cognitive load theory, modality of presentation and the transient information effect. *Applied Cognitive Psychology, 25*, 943-951.

Leite, W.L., Svinicki, M., & Shi, Y. (2010). Attempted validation of the scores of the VARK: Learning styles inventory with multitrait-multimethod confirmatory factor analysis models. *Educational and Psychological Measurement, 70*, 323-339.

- Lin, B., & Hsieh, C. (2001). Web-based teaching and learner control: A research review. *Computers & Education, 37*, 377-386.
- Lowe, R.K. (2004). Interrogation of a dynamic visualization during learning. *Learning and Instruction, 14*, 257-274.
- 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. (2009). *Multimedia learning* (2nd Ed.). New York, NY: Cambridge University Press.
- Mayer, R.E., & Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology, 93*, 390-397.
- 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., Dow, G.T., & Mayer, S. (2003). Multimedia learning in an interactive self-explaining environment: What works in the design of agent-based microworlds? *Journal of Educational Psychology, 95*, 806-813.
- Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive load approach. *Journal of Educational Psychology, 84*, 429-434.
- Paas, F., & van Merriënboer, J.J.G. (1993). The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Human Factors, 35*(4), 737-743.
- Riding, R. (2001). Nature and effects of cognitive style. In R. Sternberg & L.-F. Zhang (Eds.), *Perspectives on thinking, learning, and cognitive styles* (pp. 47-72). Mahwah, NJ: Lawrence Erlbaum.
- Scheiter, K., & Gerjets, P. (2007). Learner control in hypermedia environments. *Educational Psychology Review, 19*, 285-307.
- Schnotz, W., Böckheler, J., & Grzondziel, H. (1999). Individual and co-operative learning with interactive animated pictures. *European Journal of Psychology of Education, 14*, 245-265.
- Smith, R.M. (1982). *Learning how to learn*. Chicago: Follett.
- Spanjers, I.A.E., van Gog, T., & van Merriënboer, J.J.G. (2010). A theoretical analysis of how segmentation of dynamic visualizations optimizes students' learning. *Educational Psychology Review, 22*, 411-423.
- Stahl, S.A. (1999). Different strokes for different folks? A critique of learning styles. *American Educator, 23*(3), 27-31.

Sternberg, R., & Zhang, L.-F. (Eds.). (2001). *Perspectives on thinking, learning, and cognitive styles*. Mahwah, NJ: Lawrence Erlbaum.

Sweller, J. (2010). Element interactivity and intrinsic, extraneous and germane cognitive load. *Educational Psychology Review*, 22, 123-128.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York: Springer.

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

Tabbers, H., Martens, R., & van Merriënboer, J.J.G. (2004). Multimedia instructions and cognitive load theory: Effects of modality and cueing. *British Journal of Educational Psychology*, 74, 71-81.

Tabbers, H.K., & Koeijer, B. (2010). Learner control in animated multimedia instructions. *Instructional Science*, 38, 441-453.

Tversky, B., Morrison, J.B., & Betrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human-Computer Studies*, 57, 247-262.

Wong, A., Leahy, W., Marcus, N., & Sweller, J. (2012). Cognitive load theory, the transient information effect and e-learning. *Learning and Instruction*, 22, 449-457.

Wouters, P., Tabbers, H.K., & Paas, F. (2007). Interactivity in video-based models. *Educational Psychology Review*, 19, 327-342.

Copyright Disclaimer

Copyright reserved by the author(s).

This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).