Computers in Human Behavior 28 (2012) 1072–1081

Contents lists available at SciVerse ScienceDirect

Computers in Human Behavior

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

An empirical evaluation of multi-media based learning of a procedural task

Doo Young Lee [1], Dong-Hee Shin *

Department of Interaction Science, Sungkyunkwan University, 110-745 Seoul, South Korea

ARTICLE INFO

Article history:
Available online 9 February 2012

Keywords:
Multi-media learning
Learning of procedural task
Visual representation
Static visualization
Animation
Spatial ability

ABSTRACT

The present study investigated the effects of multi-media modules and their combinations on the learning of procedural tasks. In the experiment, 72 participants were classified as having either low- or high spatial ability based on their spatial ability test. They were randomly assigned to one of the six experimental conditions in a 2×3 factorial design with verbal modality (on-screen text procedure vs. auditory procedure) and the format of visual representation (static visual representation vs. static visual representation with motion cues vs. animated visual representation). After they completed their learning session, the ability to perform the procedural task was directly measured in a realistic setting. The results revealed that: (1) in the condition of static visual representation, the high spatial ability group outperformed the low spatial ability group, (2) for the low spatial ability participants, the animated visual representation group outperformed the static visual representation group, however, the static visual representation with motion cues group did not outperform the static visual representation group, (3) the use of animated visual representation helped participants with low spatial ability more than those with high spatial ability, and (4) a modality effect was found for the measure of satisfaction when viewing the animated visual representation. Since the participants with low spatial ability benefited from the use of animation, the results might support an idea that people are better able to retrieve the procedural information by viewing animated representation. The findings also might reflect a preference for the auditory mode of presentation with greater familiarity with the type of visual representation.

1. Introduction

An increasing number of computer-based commercial products for e-learning including computer-based learning (CBL) and web-based learning (WBL) are available in the marketplace. Many educational institutions, organizations, and households now use several types of instructional modules intended to enhance learning activities. Such tools tend to use visual representation, along with verbal information. They include a wide variety of multi-media modules, such as text, audio, pictures, and animation. It seems logical that the visual representation of such tools plays an important role in helping people learn and apply concepts in real settings, and especially so for a step-by-step procedural task (Bhowmick, Khasawneh, Bowling, Gramopadhye, & Melloy, 2007). For example, one may be able to simply visit a product's website and learn a procedural instruction of an assembly task in a way that allows both text and visual information to be displayed together on a computer screen. A procedural task denotes performing a specific goal in a linear manner, followed by learning 'procedural motor knowledge' (Höffler & Leutner, 2007; Van Gog, Paas, Marcus, Ayres, & Sweller, 2009; Watson, Butterfield, Curran, & Craig, 2010; Wong et al., 2009) consisting of a series of executable actions (Bovair & Kieras, 1991; Eirìksdóttir & Catrambone, 2007; Ellis, Whitehill, & Irick, 1996; Konoske & Ellis, 1991). Developing understanding requires that people make a connection between structural configuration of the equipment and the functional sequence of the procedural task (Heiser & Tversky, 2006). It is important to present instructional information to people in the most effective and efficient manner to facilitate clear understanding during this process (Brunyé, Taylor, & Rapp, 2008; Rodriguez, 2002).

With the relatively simple computer technology available, the use of paper- or computer-based static visual representation received interest in learning of procedural tasks (Bieger, 1982; Booher, 1975; Patel, Eisemon, & Arocha, 1990; Stone & Glock, 1981). For example, Stone and Glock (1981) examined if the additional presence of illustrations corresponding to procedural text would enhance comprehension of tasks for loading cart assembly. The results indicated that participants who viewed the text with illustrations outperformed the stand-alone text group on the accuracy measure. They explained that illustration provided a redundant means of expression to text and that illustration might present spatial information more effectively than text. In a related set of studies, Bieger (1982) found that text-graphic instruction facilitated learning of a procedural assembly task. The superiority of presenting instructional information in both visual and verbal form to merely a verbal format can be explained by the theory of

* Corresponding author. Tel.: +82 2 740 1864; fax: +82 2 740 1856.
E-mail addresses: bugabuga@skku.edu (D.Y. Lee), dshin@skku.edu (D.-H. Shin).
[1] Tel.: +82 10 3649 5964; fax: +82 2 740 1856.

0747-5632/$ - see front matter
doi:10.1016/j.chb.2012.01.014

Dual-Coding (Paivio, 1990, 1991). The implication of this theory is that a person mentally forms referential connection, leading to cross-activation when information is coded through both verbal and non-verbal forms. Rieber (1990) stated that, "this redundant encoding increases the probability of retrieval, because if one memory trace is lost (whether visual or verbal), the other is still available" (p. 135). Thus, Dual-Coding theory justifies the cognitive benefits of presenting visual representation, along with verbal information.

Given the multi-media technological advancements, the form of video-typed or animated visual representations has increased in recent years. It is believed that animated visual representation should be more able to demonstrate procedural content, compared to static visual representation (Anglin, Vaez, & Cunningham, 2004; Fischer, Lowe, & Schwan, 2008; Weiss, Knowlton, & Morrison, 2002; Wong et al., 2009). This follows, because animation can facilitate mental representation of a procedure clearly (Höffler & Leutner, 2007), present step-by-step sequential actions (Park & Hopkins, 1993) and show the specific behavior or dynamic movement of the equipment, as it changes over time (Arguel & Jamet, 2009; Boucheix & Schneider, 2009; Lin & Atkinson, 2011; Zacks & Tversky, 2003). Conversely, static visual representation is 'a snapshot of the specified behavior of the system at a particular time or static description of all possible behavior' (Dulac, Viguier, Leveson, & Storey, 2002, p. 73). Due to the nature of static visual representation, people may need to map external static representation into their internal dynamic representation, when they are required to understand a procedural task from an implicitly presented static visual representation (Lowe, 2004). By contrast, animated visual representation is more likely to explicitly provide procedural instruction, because it is able to directly demonstrate what a person needs to perform (Carroll & Wiebe, 2004). In particular, a recent meta-analysis conducted by Höffler and Leutner (2007) indicated that animation is more effective than static visual representation for learning procedural motor knowledge. However, these claimed advantages of using the animated visual representation over static visual representation in learning of a procedural task, have been investigated in only a small set of practical applications, such as a first aid task (Michas & Berry, 2000), injury prevention exercise (Van Hooijdonk & Krahmer, 2008), and handheld device assembly (Watson et al., 2010). Other studies suggest that adding a signaling or highlighting of motions (i.e., arrow) to static visual representation may provide the function of guiding attention to important aspects for learning a procedural task, such as the sequence of operations or actions (Heiser & Tversky, 2002, 2006; Tversky, Zacks, Lee, & Heiser, 2000) or temporal change (Lowe, 2004; Tversky, Morrison, & Betrancourt, 2002). Thus, both animated visual representation and the static images containing appropriate motion cues may be more effective than static visual representation in facilitating the understanding of a procedural task.

Another important issue that needs to be investigated in the context of multi-media based learning of a procedural task is the modality effect of audio-visual superiority to the text-visual. According to Baddeley's multi-component model of working memory (Baddeley, 1986, 1992), verbal information can be processed differently, depending on the nature of modality presented. For example, when written text information and corresponding visual information are both presented visually, then people's visual attention might be split in the visuospatial sketchpad in working memory. Therefore, visual attention is likely to become overloaded. In contrast, when the verbal information is presented aurally and the corresponding visual information is presented, people can process the internal representation of the verbal information within a phonological loop and the internal representation of the visualization in the visuospatial working memory separately, which reduces the load on visual attention (Guan, 2009). Thus, use of auditory text is possibly better than the use of on-screen text in explaining the visual representation (Mayer, 2005; Penney, 1989; Seufert, Schu, & Nken, 2009; Sweller, Van Merriënboer, & Paas, 1998). A number of studies have supported the superiority of dual-modality presentation (Brünken, Plass, & Leutner, 2004; Mayer & Moreno, 1998; Mousavi, Low, & Sweller, 1995; Tindall-Ford, Chandler, & Sweller, 1997). Past studies of the modality effect have mainly emphasized the learning of scientific concepts with declarative representation (i.e., formation of lightning, blood circulating system), but not on the learning of a procedural task.

2. Research objective and development of hypotheses

The objective of the current study is to examine the effects on the learning of a procedural task of variation in the format of visual representation and verbal modality. Three types of visual representation format include static visual representation, static visual representation with motion cue, and animated visual representation. Two types of verbal modality are on-screen text procedure (written procedure) and auditory procedure (audio). Besides studying the formats of visual representation and verbal modality, the role of spatial ability in learning from visual representation of procedural task was investigated. Spatial ability refers to the capability to mentally represent and manipulate objects in 2-D and/or 3-D space over time (Carroll, 1993; Stanney & Salvendy, 1995; Yang, Andre, & Greenbowe, 2003) and plays a significant role in constructing a mental model from visual representation (Höffler & Leutner, 2011; Münzer, Seufert, & Brünken, 2009). It is known that there are individual differences in spatial ability, which result in differences in the way different learners' process visual information. These differences may then influence the degree to which presentation of verbal information and the corresponding visual representation is effective for learning. In a series of studies (Hegarty, 1992; Hegarty & Just, 1993; Hegarty & Sims, 1994; Hegarty & Steinhoff, 1997), Hegarty and her colleagues investigated 'mental animation', which is constructing an internal representation or mental model of dynamic nature of the mechanical system from viewing the text and static visual representation. They found that 'mental animation' demands cognitive resource in working memory and that people with low spatial ability made more errors in 'mental animation' than those with high spatial ability. They explained that limited capacity of working memory might be the source of error in the people with low spatial ability for understanding from static visual representation. Thus, individual differences in spatial ability can account for differences in learning outcomes, such that people with high spatial ability outperform those with low spatial ability, in a case of viewing the static visual representation (Höffler, 2010). Conversely, several lines of studies in other settings imply that viewing the animated visual representation may lead to better performance than static visual representation for people with low spatial ability (Chanlin, 2000; Hays, 1996; Hegarty, 2005; Hegarty & Kriz, 2008; Höffler & Leutner, 2011; Höffler, Sumfleth, & Leutner, 2006; Lee, 2007). This view is consistent with Höffler's finding (2010). The author conducted a meta-analysis of 19 past studies with 59 pair-wise comparisons to investigate the association between individual differences in spatial ability (low vs. high) and type of visual representation (static vs. animated). The result indicated that the use of animated visual representation could significantly enhance learning for people with low spatial ability. Höffler (2010) stated that, "...learners with low spatial ability might be supported by dynamic visualizations because the visualization provides the learners with an external representation of a process or procedure that helps them to build an adequate mental model" (p. 249). In contrast, other studies argued that people with high spatial ability might benefit more from the animated visual representation than those with low spatial