Computers in Human Behavior 62 (2016) 9–18

Contents lists available at ScienceDirect

Computers in Human Behavior

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

Full length article

Multimedia recipe reading: Predicting learning outcomes and diagnosing cooking interest using eye-tracking measures

Ching-Yeh Wang [a], Meng-Jung Tsai [b, *], Chin-Chung Tsai [b]

[a] Graduate Institute of Applied Science and Technology, National Taiwan University of Science and Technology, No. 43, Sec. 4, Keelung Rd., Taipei 106, Taiwan

[b] Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, No. 43, Sec. 4, Keelung Rd., Taipei 106, Taiwan

ARTICLE INFO

Article history:
Received 26 December 2014
Received in revised form
24 November 2015
Accepted 20 March 2016

Keywords:
Multimedia learning
Eye movement
Retention
Cooking interest
Recipe reading

ABSTRACT

This study aimed to explore the relationships between students' visual behaviors and learning outcomes, and between visual behaviors and prior cooking interest in multimedia recipe learning. An eye-tracking experiment, including pretest, recall test, and retention test, was conducted with a sample of 29 volunteer hospitality majors in Taiwan. The multimedia recipe included a static page showing the ingredients in a text-and-picture representation and a dynamic page showing the knife skills in a text-and-video representation. Total reading time, total fixation duration, number of fixations and inter-scanning count were used to explore the students' visual attention distributions among the different representation elements and their visual strategies for learning the recipe. The results showed that all students paid more visual attention to the text than to the picture information for the static recipe, and paid more visual attention to the video than to the text on the dynamic page. In addition, the visual attention paid to the text on the dynamic page was negatively correlated with the retention of the episodic knowledge of knife skills. In contrast, the visual attention paid to the text on the static ingredient page was positively correlated with students' prior cooking interest. Finally, the inter-scanning count between text and video on the dynamic page was the best index to negatively predict students' learning retention. Total fixation duration on the text information on the static page was the best index to positively predict students' prior cooking interest. Future studies and applications are discussed.

1. Introduction

Reading skill is one of the basic literacy skills for all learners in all learning domains in the 21st century. In culinary education, students are required to read recipes before practicing or performing a cooking task. Recipe learning refers to the acquisition of both the semantic knowledge and the episodic knowledge (Tulving, 1972) before cooking via reading recipes. The semantic knowledge indicates the identification of the names and kinds of ingredient, while the episodic knowledge indicates the procedures and episodes of cooking such as the knife skills. These two kinds of knowledge are required for learners before they can successfully make a dish, and a recall or retention test such as a paper-and-pencil test is often used to assess students' memory of these two types of knowledge (Tulving, 1972). Meanwhile, digital multimedia technology has revolutionarily changed our daily lives and school teaching during the past decades. The learners usually learn cooking through multimedia channels such as television programs or Web-based information (Worsley, Wang, Ismail, & Ridley, 2014). In multimedia recipe learning, different kinds of knowledge of recipe learning may be best presented using different types of representation. For example, knife skills may best be displayed with a dynamic representation such as a video which can clearly show the details of the operational sequences. Based on the above, we are interested in exploring how students pay attention to reading multimedia recipes, how they acquire semantic knowledge and episodic knowledge, and how they process the static and dynamic information in the recipes.

Generally, personal factors and instructional media may contribute to students' visual attention distribution while learning. The personal factors such as learning interest, motivation, preference, prior knowledge and prior learning experience (Clinton & van den Broek, 2012; Uitto, 2014; Wild & Schiefele, 1994) may influence

* Corresponding author.
E-mail addresses: D10122302@mail.ntust.edu.tw (C.-Y. Wang), mjtsai99@mail.ntust.edu.tw, mjtsai99@gmail.com (M.-J. Tsai), cctsai@mail.ntust.edu.tw (C.-C. Tsai).

http://dx.doi.org/10.1016/j.chb.2016.03.064
0747-5632/

students' attention and performance. For example, students with high interest in cooking may pay more attention to reading recipes, while those with more prior cooking experience may adopt more efficient reading strategies to read recipe information. Meanwhile, instructional media design such as layout design and representation design may also affect students' attention and performance (Mayer, 2009). For example, students may be attracted to watch a video (dynamic display) to gather the episodic information of cooking and to view a snapshot of a scenario (static display) to identify the ingredients of a dish. In sum, different visual attention patterns may reveal the interplay of personal factors in recipe learning. Most high school culinary students are usually beginners with limited experience of cooking. Thus, this study was interested in exploring the role culinary students' cooking interest plays in their visual attention distribution in the process of recipe learning.

2. Literature review

2.1. Multimedia representations for learning

Students construct knowledge from information encrypted in various multimedia representations. The cognitive process of multimedia learning has been discussed extensively with regard to the cognitive theory of multimedia learning (Mayer, 2009) which involves three main assumptions: the dual channels theory (Paivio, 1986), the limited capacity theory (Baddeley, 1992) and the active processing assumption (Wittrock, 1989). In this multimedia learning model, the verbal information (e.g., text or narration) and pictorial information (static graphics/pictures or dynamic animations/videos) are processed by two sensory channels: the auditory channel and the visual channel. Narrations are processed by the auditory channel, whereas text (both printed and on-screen text) is processed by the visual channel. Therefore, the redundancy principle proposed in the multimedia learning theory indicates that people learn better from graphics and narration than from graphics, narration and printed-text. The redundancy causes extraneous processing because the visual channel can become overloaded when people visually scan between the printed text and the graphics, and they need to put more effort into comparing the spoken and printed texts (Mayer, 2009). In this case, the redundant information is the printed text when it is identical to the narration (spoken text). Recently, based on a broader view of the cognitive load theory (Chandler & Sweller, 1992; Sweller, Ayres, & Kalyuga, 2011), the redundancy effect for similar representations has been found in a mobile learning environment using three static visual representations including text, pictures and real objects (Liu, Lin, Tsai, & Paas, 2012).

Static text-and-picture representation is one of the main formats presented in various multimedia instructional materials, including recipes for learning cooking. How students process textual and graphic information in conceptual learning domains (Ho, Tsai, Wang, & Tsai, 2014; Liu & Chuang, 2011; Mason, Pluchino, Tornatora, & Ariasi, 2013) may provide some clues as to how they process text-picture-based multimedia recipes for semantic knowledge learning. For example, Schmidt-Weigand, Kohnert, and Glowalla (2010) explored university students' attention to text and picture information regarding the formation of lightning. Their results showed that students could switch between the textual and pictorial information while reading the materials, indicating that they may have tried to connect the two types of information for better comprehension. Ho et al. (2014) explored how students with different prior knowledge dealt with a web-based text-and-graphic scientific report; their results demonstrated that students with higher prior knowledge switched more frequently between the text and the graphic. The inter-scanning behavior between text and data diagrams is regarded as evidence of scientific inquiry skills. Mason, Tornatora, and Pluchino (2013) highlighted the significance of learners' revisited fixation duration spent on comprehending text-and-graphic information illustrated in traditional scientific textbooks. According to the above, students' visual behaviors are not only affected by prior knowledge, but are also highly correlated with their comprehension learning outcomes.

In addition, dynamic multimedia representation such as video, animation, simulation and interactive games is another format which can visualize abstract scientific concepts or demonstrate episodic knowledge of hands-on activities. Narration or captions are usually provided along with dynamic representations to explain the details or provide some important guidelines to improve students' understanding. Schmidt-Weigand et al. (2010) indicated that students spent more time reading the animation with spoken text than those with on-screen text for learning the formation of lightning. In contrast, culinary learning requires the knowledge of hands-on procedural skills. Dynamic multimedia representations such as videos with narration could benefit learners' construction of episodic knowledge. Do people process the dynamic representations in multimedia recipes in the same ways as in other learning domains? This study explored how students read the text and video with narration while learning the episodic knowledge of a recipe in order to answer this question.

2.2. Visual behaviors in multimedia learning

According to the eye-mind assumption (Just & Carpenter, 1980), human's gaze distribution is closely connected to the attention process. That is, based on learners' eye fixation locations, researchers can analyze the process of how they read to learn (Rayner, 1998; Rayner, Chace, Slattery, & Ashby, 2006). The time-, spatial- and frequency-based eye movement measures of fixation, saccades or even scan paths can imply the attention controls and implicit cognitive strategies utilized for learning (Lai et al., 2013). Fixation is the relatively static state of eye movement, while a saccade is the rapid movement between each two consecutive fixations. A scan path consists of a sequence of fixations and saccades while reading or viewing. It is believed that external visual stimuli are processed by the brain during fixations, while saccades are responsible for moving the eye fovea focus onto the target information. Generally, fixation location reveals the information being processed, fixation duration reveals the mental efforts made regarding the target, and the scan path reveals the metacognitive control strategies for reading and learning. However, the use and interpretation of eye-tracking measures depends on the research questions and the learning tasks defined in the studies (Lai et al., 2013; Rayner et al., 2006).

In the cognitive domain, several studies have used eye-tracking technology to examine students' visual behaviors related to multimedia reading in different learning domains, such as science (Canham & Hegarty, 2010; Chen & Yang, 2014; Tsai, Hou, Lai, Liu, & Yang, 2012), medical science (de Koning, Tabbers, Rikers, & Paas, 2010), biology (Hung, 2014; Jarodzka, Scheiter, Gerjets, & van Gog, 2010; She & Chen, 2009; Yang, Chang, Chien, Chien, & Tseng, 2013), mathematics (Susac, Bubic, Kaponja, & Planinic, 2014), and engineering (Meyer, Rasch, & Schnotz, 2010). All of the above studies are involved with conceptual learning of domain knowledge. In culinary education, although episodic knowledge is focused on more than semantic knowledge, cognitive control of attention is also required to learn it. Although some researchers (e.g., Yang, 2012) have utilized eye-tracking technology to examine consumers' visual scan paths when reading restaurant menus, little research has explored students' visual behavior in relation to the gains of episodic knowledge such as recipe reading. Recently, eye-