Computers in Human Behavior 33 (2014) 242–255

Contents lists available at ScienceDirect

Computers in Human Behavior

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

Click versus drag: User-performed tasks and the enactment effect in an interactive multimedia environment

Ruth N. Schwartz *, Jan L. Plass

New York University, United States

ARTICLE INFO

Article history:
Available online 14 February 2014

Keywords:
Multimedia
Interactivity
Enactment effect
User-performed tasks
Embodiment

ABSTRACT

Does learner performance of specific tasks in an interactive multimedia environment affect learning outcomes? Within a multimedia environment, users may engage in a range of actions, or interactive tasks, from tapping a keyboard to executing large motor movements. To investigate the impact of particular performance tasks on learning, we first introduce an approach to classifying interactive multimedia tasks as enactive, iconic or symbolic. We then describe a study in which university students ($N = 112$) used a computer program that presented a series of action phrases in 4 different conditions, each condition requiring performance of a different task: *listen* (audio only), *look* (audio with static graphic), *click* (audio with animation triggered by a click), or *drag* (audio with click-and-drag graphic). Participants were tested on free recall and recognition of phrases immediately after treatment and again after 3 weeks. At immediate testing, recall was best for *drag* (iconic) items, followed by *click* (symbolic), *look*, and *listen* items, in that order, with significant differences between each pair of conditions. For immediate recognition, as well as for delayed free recall and delayed recognition, mean scores followed the same pattern, with some variations in significance. Results support our proposed classification of interactive behaviors, extend previous findings on the enactment effect into a computer environment, and suggest the importance of considering the design of interactive tasks in the development of multimedia learning materials.

1. Introduction

Does learner performance of specific tasks in an interactive multimedia environment affect learning outcomes? The potential of interactive environments to support learning has been a topic of discussion since the first appearance of computers for home use (Bransford, Brown, & Cocking, 1999; Hannafin & Peck, 1988; Plass, Schwartz, & Heidig, 2012), and interactivity has subsequently been explored from a number of perspectives. Some researchers have investigated interactivity as a construct that is *present* versus *absent* (e.g., Evans & Gibbons, 2007; Friedl et al., 2006) or *low* versus *high* (e.g., Haseman, Nuipolatoglu, & Ramamurthy, 2002; Jo & Kim, 2003), while others have examined specific features of interactive environments such as learner control of pacing and sequencing (Mayer & Chandler, 2001; Schwan & Riempp, 2004) as well as guidance and feedback provided by the environment (Moreno, 2004; Moreno & Mayer, 2005). We found that the majority of studies on interactivity that we reviewed focused either on the inclusion or exclusion of particular interactive features such as those above, or on users' cognitive responses to learning strategies, such as pre-training or worked examples, that were implemented through interactive features (Atkinson & Renkl, 2007; Moreno & Mayer, 2007). Surprisingly, although user-performed actions are definitional to interactive systems, we found few studies that examined the impact of particular actions carried out by learners within a multimedia environment. Research from other disciplines, however, suggests that the performance of specific tasks—what the learner must *do* in order to interact—may in fact play a role in learning (e.g. Clark, 1997; Cook, Mitchell, & Goldin-Meadow, 2008; Engelkamp, 1998; Glenberg, Gutierrez, Levin, Japuntich, & Kaschak, 2004). Furthermore, given the rapidly expanding range of options for learner control of multimedia environments, from touch-screens to full-body movements, consideration of the cognitive implications of different types of user-performed tasks is increasingly important. The goal of the present study was therefore to initiate exploration of whether different tasks performed by learners in a multimedia environment can affect learning outcomes.

In our theory section, below, we first discuss the role of learner-performed tasks in education. We then summarize theoretical approaches to multimedia learning and existing perspectives on learner interaction in multimedia environments, as well as

* Corresponding author. Address: CREATE – Consortium for Research and Evaluation of Advanced Technologies in Education, New York University, 196 Mercer Street, Suite 800, New York, NY 10012, United States. Tel.: +1 212 998 5511.
E-mail address: ruth.schwartz@nyu.edu (R.N. Schwartz).

http://dx.doi.org/10.1016/j.chb.2014.01.012
0747-5632/

relevant empirical research in multimedia and related areas. Finally, we propose an approach to framing research on user-performed tasks in interactive multimedia environments.

1.1. Educational implications of learner-performed tasks

Theorists and psychologists in the field of education have long recognized the benefits of learner-performed tasks and the effects of such tasks on cognition. Early educational leaders, such as Pestalozzi (1819/1898), Fröbel (1894/1904), Montessori (1914/1964), and Dewey (1916/1959), promoted the idea that performing specific tasks can affect a child's comprehension, recall, and conceptual development. More recently, a number of contemporary theorists have discussed the relationship between physical activities and learning. For example, Dual Coding Theory, which describes the role of verbal and nonverbal mental representations in processing information, specifically includes actions such as "drawing lines or pressing keys" as types of nonverbal input that may assist in encoding and retrieval (Clark & Paivio, 1991, p. 151). Recent work on embodied cognition has taken an even broader stance, suggesting that because mind and body are inextricably connected, embodied human experience determines the shape of all cognitive activity (e.g., Clark, 1997; Varela, Thompson, & Rosch, 1991).

Findings from several lines of research provide empirical support for the contention that user-performed tasks may have an impact on cognition. Investigations of the role of gesture in learning have suggested that using body movements in the representation of ideas can help students learn and retain conceptual information (Cook et al., 2008), while other studies have shown that having children manipulate relevant toys while reading specific passages enhanced comprehension (Glenberg, Brown, & Levin, 2007; Glenberg et al., 2004). Another series of investigations on the cognitive effects of different learning tasks, applied to the memorization of lists of action phrases, revealed an *enactment effect*: Participants who performed the actions specified in those phrases (e.g., "pick up the pen") scored better on tests of retention than those who simply listened to a reading of the phrases (Engelkamp & Dehn, 2000; Engelkamp & Jahn, 2003; Engelkamp & Zimmer, 1994, 1997). The Multimodal Theory of Memory offers an explanation for this effect, positing that encoding involves sensory, conceptual, and motor phases (Engelkamp, 1998). Hearing an instruction such as "pick up the pen" activates verbal information at the sensory-motor level as well as conceptual information at the conceptual level. Performing the referenced action results in motor encoding, which in turn facilitates the subsequent retrieval of information (Engelkamp, 1998; Engelkamp & Zimmer, 1994). The Multimodal Theory of Memory differs from models such as Dual Coding (Clark & Paivio, 1991) in that it includes an active role for sensory output, rather than input alone, as an element of encoding: It is not simply a question of what the user *perceives*, but what the user *does*, with particular emphasis on motor responses.

A number of studies have investigated nuances of the enactment effect. For example, investigations showed that actions performed by research participants themselves (*self-performed tasks*) had a stronger positive effect on recall than *experimenter-performed tasks*, which participants simply observed (Engelkamp & Krumnacker, 1980, as cited in Engelkamp & Zimmer, 1994, 1997). An enactment effect was found not only when participants used real objects to perform actions, but for the use of imaginary objects as well, clarifying that the effect is motor-based rather than sensory-based (Engelkamp & Krumnacker, 1980). The importance of the motor component is further reinforced by studies showing that imagining the performance of actions is not as effective in enhancing memory as actual performance (e.g. Goff & Roediger, 1998).

Other research on the enactment effect has examined differences between types of performance, demonstrating that individuals who initiated their own actions did better on recall tasks than those who mimicked the actions of others (Zimmer & Engelkamp, 1996). One explanation for this finding is that initiating and carrying out a motor action requires planning, which involves the conceptual level. Mimicked actions, however, originate and are executed from within the sensory-motor level, without rising to the conceptual level. The implication is that motor performance, absent conceptual activation, is not sufficient to evoke an enactment effect (Engelkamp & Zimmer, 1994). Yet another line of investigation found that the advantage in recall for subject-performed tasks over tasks requiring no performance was resilient over a 1-week time period (Nilsson, Cohen, & Nyberg, 1989).

Together, these theoretical approaches and empirical findings provide strong support for the contention that, because body and mind are so closely connected, the performance of specific tasks ranging from concrete manipulation of objects (Glenberg et al., 2004, 2007) to more abstract gestures (Cook et al., 2008) can affect cognition, and that sensory and motor output, rather than cognitive input alone, may play a role in encoding, retrieval, and conceptual development. Extensive research on the enactment effect (Engelkamp & Dehn, 2000; Engelkamp & Jahn, 2003; Engelkamp & Zimmer, 1994, 1997) suggests further that relatively subtle differences in performance tasks can affect learner outcomes by influencing processes of encoding and retrieval. However, previous research into the role of specific performance tasks has been conducted largely in non-mediated settings. We asked whether any effects might be observed in the context of computer-based interactive environments, which afford user-performed tasks of many different types.

1.2. User-performed tasks in interactive multimedia environments

In an interactive multimedia environment, learners may be able to choose a topic to investigate, control the pace of a presentation, or answer questions and receive feedback about the content presented. They may perform simulated experiments in an on-screen chemistry lab by controlling parameters such as temperature and pressure, or investigate physics by manipulating virtual materials (Adams et al., 2008; Plass et al., 2009). Students may collaborate to solve puzzles by sliding shapes across a screen (Scott, Mandryk, & Inkpen, 2003) or explore a virtual landscape, conversing with computer-generated characters in the process of learning (Moreno & Mayer, 2007). Users accomplish these tasks through a variety of actions supported by a variety of interfaces. For example, they may enter text on a keyboard, click a mouse, drag and drop objects on a touch screen, wave a controller, or jump up and down in front of a Kinect sensor. In each case, it is user-performed tasks, in conjunction with system responses, which drive the action. Learners have the opportunity not just to *act*, but to *interact*, and it is often assumed that these interactions provide rich opportunities for learning. Though this assumption frequently goes unchallenged and under-investigated, both theory and research in multimedia have examined certain aspects of how interactivity may contribute to the educational efficacy of a learning environment (Domagk, Schwartz, & Plass, 2010; Plass et al., 2012).

Predominant theoretical approaches to learning in multimedia environments include the Cognitive Theory of Multimedia Learning (CTML; Mayer, 2001) and the Integrated Theory of Text and Picture Comprehension (Schnotz, 2005), both of which focus on the cognitive processes involved in learning from words and pictures. CTML, for example, is based on three key assumptions: that auditory and visual information are perceived and processed in separate channels (Clark & Paivio, 1991); that each of these channels has a limited capacity and can process only a finite amount of