Computers & Education 92-93 (2016) 64–75

Contents lists available at ScienceDirect

Computers & Education

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

# Embodied learning using a tangible user interface: The effects of haptic perception and selective pointing on a spatial learning task

Alexander Skulmowski [a, *], Simon Pradel [a], Tom Kühnert [b], Guido Brunnett [b], Günter Daniel Rey [a]

[a] E-Learning and New Media, TU Chemnitz, Chemnitz, Germany
[b] Computer Graphics and Visualization, TU Chemnitz, Chemnitz, Germany

A R T I C L E I N F O

*Article history:*
Received 5 May 2015
Received in revised form 7 October 2015
Accepted 9 October 2015
Available online 22 October 2015

*Keywords:*
Interactivity
Virtual and physical learning environment
Embodied Cognition
Cognitive load theory
Human–computer interface

A B S T R A C T

Tangible User Interfaces offer new ways of interaction with virtual objects, yet little research has been conducted on their learner-friendly design in the context of spatial learning. Although frameworks such as Embodied Cognition stress the importance of sensory perception and movement, studies have found that high interactivity can be overwhelming and may lead to a lower learning performance. In a 2 × 2 factorial design participants ($n = 96$) learned heart anatomy using a 3D model that was either controlled using a mouse or a tangible object, i.e. a motion tracked plastic model of the virtual heart. Secondly, we varied the interaction mode featuring either a selective pointing mode in which only the label that the user currently activated was displayed or permanent display of all labels. Retention performance, cognitive load scores, and motivation measures indicate that the tangible object leads to significantly higher learning outcomes. The effect of the label display mode is different for the two input devices: The performance with selective pointing in the mouse condition is better than the performance with permanent display in the mouse condition; in the TUI condition this is exactly the other way around. Based on these results, we propose extensions for Embodied Cognition and Cognitive Load Theory.



## 1. Introduction

Constructing spatial representations of three-dimensional objects is a complex cognitive task relevant in many areas of education including STEM fields and medicine. In recent years, researchers have begun to investigate the benefits of virtual learning environments as well as the cognitive foundations for their learner-friendly design (Nicholson, Chalk, Funnell, & Daniel, 2006; Pan, Cheok, Yang, Zhu, & Shi, 2006; Stull, Hegarty, & Mayer, 2009). Despite the introduction of theoretical frameworks aimed at explaining and predicting the effects of interactive and perceptually rich learning environments, studies in this field often lead to mixed results. We propose a more systematic approach for research on spatial cognition in virtual environments by contrasting predictions derived from Cognitive Load Theory (CLT; Sweller, 1999; Sweller, Merrienboer, &

* Corresponding author. E-Learning and New Media, Faculty of Humanities, Technische Universität Chemnitz, Straße der Nationen 12, 09111 Chemnitz, Germany.

*E-mail address:* alexander.skulmowski@phil.tu-chemnitz.de (A. Skulmowski).

http://dx.doi.org/10.1016/j.compedu.2015.10.011
0360-1315/

Paas, 1998) with current research on Embodied Cognition (EC; Barsalou, 1999; Shapiro, 2010). Based on this analysis, we conducted an experimental study focused on the influence of multimodal perceptual input and an interactive display mode using a tangible user interface (TUI). While the main focus of research on TUIs lies on devising more natural interaction modes and creating more engaging experiences (Hornecker & Buur, 2006; Shaer & Hornecker, 2010; Xie, Antle, & Motamedi, 2008), we are primarily concerned with the use of TUIs for generating spatial representations.

Spatial learning can be investigated using virtual learning environments that allow users to interact with three-dimensional objects or scenes. Although several studies have shown that spatial learning using interactive media can lead to improved learning performances (Barrett, Stull, Hsu, & Hegarty, 2015; Huk, 2006; Stull, Barrett, & Hegarty, 2013; Stull et al., 2009), some results indicate that high interactivity and perceptual richness can be detrimental to learning (Levinson, Weaver, Garside, McGinn, & Norman, 2007; Song et al., 2014). These results have been most commonly discussed within the framework of CLT and, in recent years, EC.

CLT holds that our limited working memory capacity requires learning materials to be designed in a manner that leaves as much memory capacity as possible reserved for the actual learning contents (Sweller et al., 1998). Furthermore, all cognitive load resulting from content or interactions not directly contributing to learning should be avoided. While CLT is most commonly used in the fields of instructional and educational psychology, its basic architecture lends itself to a wide variety of research questions that involve the analysis of cognitive tasks. CLT proposes three types of cognitive load (Sweller et al., 1998): Intrinsic load is defined as the load resulting from the complexity of the learning material (or task). The more elements that are required to be kept in working memory simultaneously, the higher the intrinsic load. Therefore, learning materials that enable a segmentation of a learning task lead to better learning outcomes as the intrinsic load is lowered and more cognitive resources are available for learning. The second component, extraneous load, is a result of the design of the learning materials. The learner-friendly presentation of learning content lowers extrinsic load by facilitating the access to relevant information and thus aiding understanding. A series of fMRI investigations reviewed by Whelan (2007) suggest that extraneous load can be linked to processing constraints in brain areas that modulate attention across sensory modalities. More specifically, extraneous load activates modality-specific neural structures in the posterior parietal cortex and Wernicke's Area, which we interpret as an indication that resources for sensorimotor control, integration of spatial information and language processing can be affected through extraneous load. Lastly, germane load is described as the load used for the construction of mental representations from learning contents required for storing the information in long-term memory. While most CLT researchers use Baddeley's models of working memory (Baddeley, 1992, 2000) as their basis, there have been recent efforts to expand these models to include movements and related aspects of embodiment (Wong et al., 2009; Ayres, Marcus, Chan, & Qian, 2009).

One of the main conclusions that can be drawn from CLT research is that highly interactive and perceptually rich learning environments can be disadvantageous to learning due to the high extraneous load imposed on the user (e.g., Stull & Mayer, 2007). At the same time, it should be noted that CLT research often uses a dichotomy of "high" and "low" interactivity that can usually be translated into "active" (requires the use of more elaborate controls) or "passive" (requires little additional activity from the user) interaction designs. The well-established theoretical framework of CLT is empirically supported by a wide variety of studies in the fields of learning and human–computer-interaction (for an overview, see Plass, Moreno, & Brünken, 2010), however, the model is facing growing criticism for its poor ability to explain why some forms of additional interactivity, such as gesturing and pointing (De Koning & Tabbers, 2013; De Nooijer, van Gog, Paas, & Zwaan, 2013) can be beneficial for learning outcomes. While a "less is more" design approach derived from CLT often leads to better learning performance than highly interactive learning environments that overstrain learners, it is important to study how and why some forms of added perceptual and motor load have recently proven to be helpful.

Research on the role of multimodal perception and bodily movement in cognitive processes is predominantly conducted on the basis of the EC framework that stresses the importance of bottom-up processes and modality-based perceptual representations (e.g., Glenberg, 2010; Kirsh, 2013). Accounts of EC such as the Perceptual Symbol Systems model (Barsalou, 1999, 2008) aim to describe cognition as the re-enactment of perceptual states. Thus, thinking about an object cannot be adequately defined as the mental processing of abstract representations of that object, realized through non-perceptual mental contents such as semantic networks. Instead EC characterizes cognitive processes as the re-activation of neural states that were active during the perception and encoding of that object (Barsalou, Kyle Simmons, Barbey, & Wilson, 2003). This approach suggests that humans are able to store perceptually rich memories including information gathered through all sensory modalities, hence allowing tactile percepts and motor affordances to be memorized in conjunction with visual and auditory information in multimodal representations.

Although a rapidly growing body of research within cognitive psychology and neuroscience supports these claims, attempts to utilize these results in more applied fields of psychology have led to mixed results. A number of studies found positive effects of perceptually enhanced learning materials featuring haptic feedback (Minogue & Jones, 2006; Schönborn, Bivall, & Tibell, 2011; Wiebe, Minogue, Gail Jones, Cowley, & Krebs, 2009), and interactivity (for a review, see Plass, Homer, & Hayward, 2009). Consequently, researchers have asked for updates to the "less is more" stance of CLT to be compatible with findings from EC (Paas & Sweller, 2011). By contrast, a number of studies designed with the principles of EC in mind have revealed that highly interactive and multimodal user interfaces do not always lead to an improved learning performance or may even overburden the learner (e.g., Post, van Gog, Paas, & Zwaan, 2013; De Nooijer et al., 2013). A typical example of such studies was carried out by Song et al. (2014): Participants were assigned to one of four versions of a medical learning tool for stroke syndromes differing in their degree of interactivity. One group watched a non-interactive presentation of a tissue lump