Computers in Human Behavior 26 (2010) 371–376

Contents lists available at ScienceDirect

Computers in Human Behavior

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

Neurophysiological correlates in interface design: An HCI perspective

Q.H. Mach [a,b,*], M.D. Hunter [a], R.S. Grewal [c,1]

[a] Behavioural Neuroscience Laboratory, Laurentian University, Sudbury, Ont., Canada P3E2C6
[b] School of Engineering, Laurentian University, Canada
[c] Dept. of Mathematics & Computer Science, Laurentian University, Canada

ARTICLE INFO

Article history:
Available online 23 December 2009

Keywords:
QEEG
HCI
Expert user
Novel interfaces

ABSTRACT

Objective: The current study examines the changes in functional connectivity that occurs when expert users adapt to an alternate mapping. *Background:* Research has indicated that interfaces that are similar will result in more errors and may contribute to confusion. *Methods:* Six volunteers were recruited to determine the neurophysiological changes that occur when users are exposed to an alternate mapping once an internal mental model is formed. *Results:* The results indicated a change in synchronization after alterations to the button mappings occurred. By altering the layout or order of the task, a difference in the activation pattern was observed. New areas became synchronized while synchronized activity that was present in the developed internal model became desynchronized. Altering the complexity of the task resulted in different patterns of activation recorded on the quantitative electroencephalogram (QEEG). *Conclusion:* Users often form a schema when learning a device and subsequent interactions are compared to the mental model formed during the initial learning phase. If the newer interface differs significantly a new schema is formed, resulting in a different pattern of synchronization recorded on the QEEG. *Application:* The use of this knowledge can assist in the development of new interfaces. If the intent is to create a similar interface design, the activation pattern should remain the same indicating that the old schema can be applied. An interface that displays a different cognitive pattern will indicate that a new schema was developed.

1. Introduction

One of the general themes in the field of human computer interaction is consistency of design across platforms (Ivory & Megraw, 2005; Shneiderman, 1997). It has been argued that having consistency in an interface will help facilitate learning, reducing the number of user errors (Holzinger, Stickel, Fassold, & Ebner, 2009; Rhee, Moon, & Choe, 2006). Consistency will also allow a user to transfer information from one application and seamlessly apply it in another situation or application (Maekawa, Itoh, Kawai, Kitamura, & Kishino, 2009). This will assist the user in transferring their knowledge without the need of relearning a new interface (Satzinger & Olfman, 1998). When users first learn a computer application they develop mental models of interface use. Sweller suggested that the distinguishing feature between novices and experts in problem-solving skills are the formation of schemas. The development of schemas allows the experts to determine the best method to solve the problem based on problem categorization. Novice users that had not developed a mental model did not possess sophisticated schemas and alternatively used methods that were often less effective (Sweller, 1988). Satzinger & Oflman studied mental model development when switching interfaces and found that the development of a global model for two visually similar tasks may have caused difficulty when switching between two tasks (Satzinger & Olfman, 1998). The lack of visual cues to distinguish the two tasks created one mental model as opposed to two separate mental models, the inability of the user to distinguish between the tasks led to increased errors. Despite the findings of this research, consistency across interfaces is still being introduced as the best method of human computer interaction (Bruce et al., 2006; Llanos & Munoz, 2007; Rodrigues, Junior, & Suarez, 2005). Applications that do not have consistency can be aided with the application of a script to accommodate for the inconsistencies (Rodrigues et al., 2005).

Personal universal controllers (PUC) with consistent user interfaces are currently being heralded as the newest method for users to transfer knowledge. The goal of the PUC is to present the user with an interface that they are accustomed to, regardless of the make or model of the application. It was found that models and devices from the same manufacturer did not necessarily have the same interface, making it difficult for end users to memorize all

* Corresponding author. Address: Behavioural Neuroscience Laboratory, Laurentian University, Sudbury, Ont., Canada P3E2C6. Tel.: +1 705 675 1151x2130; fax: +1 705 671 3844.

E-mail addresses: qh_mach@laurentian.ca (Q.H. Mach), md_hunter@laurentian.ca (M.D. Hunter), rsgrewal@cs.laurentian.ca (R.S. Grewal).

[1] Tel.: +1 705 675 1151x2351.

0747-5632/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.chb.2009.11.008

of the functions. The use of a PUC simplified the process, creating consistency across all devices (Nichols, Myers, & Rothrock, 2006). As was shown by Satzinger, having devices that are too similar may cause confusion as the user has difficulty distinguishing applications. The minor changes in the interface would not be noticeable, and users would create more errors by applying the wrong mental model. Research by Besnard & Cacitti indicated that too much change in an interface was better than too little change (Besnard & Cacitti, 2005). With a large amount of change, users would be forced to create a new mental model and would be able to distinguish between the interfaces with the end result being fewer errors. One of the problems associated with large amounts of visual change is the need to relearn the interface. Users may not be able to transfer their knowledge from device to device as easily.

The issue of devices with very similar interfaces has lead to a negative transfer of skills in the past. The aviation industry is a prime example of negative transfer with similar interfaces. Early in his career Chapanis noticed that a large number of accidents during landing were caused by the interface. The landing gear and flaps were side by side with no discernable difference. Pilots were often retracting the landing gear instead of manipulating the flaps (Vicente, 2003).

Historically, interfaces were developed with the intention of the users adapting to the system, newer approaches of interface design are human-centered. The idea is now to design systems that model a user's natural behaviour, creating a system that is more intuitive and easier to learn. This in turn would then lead to a reduction in performance errors. One of the suggested methods of error reduction is to minimize the cognitive load of the user. By designing a system that minimizes cognitive load, the user will have more mental resources available to perform other tasks (Oviatt, 2006).

The study of cognitive load has been used to develop instructional methods in learning, and has been defined as the multidimensional construct of the cognitive system during task performance (Baddeley, 2003; Paas & van Merriënboer, 1994; van Merrienboer, Kirschner, & Kester, 2003; Yaghoub Mousavi, Low, & Sweller, 1995). The intensity of the effort put forth has often been used as a measure of the cognitive load (Sawicka, 2008; Schmutz, Heinz, Metrailler, & Opwis, 2009; Seufert, Janen, & Brunken, 2007). One of the problems associated with studying cognitive load is the difficulty in assessing mental load, mental effort and performance. The same end result can be obtained by different people using different strategies with varied mental effort. A common technique of measuring cognitive load is the rating scale; the subject is required to introspectively assess their cognitive processes and report the mental effort expended (Gopher & Braune, 1984; Paas, Tuovinen, Tabbers, & Van Gerven, 2003). Research by Paas et al. has indicated that subjects could accurately and reliably rate their cognitive processing and scales were sensitive to small differences in cognitive load (Paas, van Merriënboer, & Adam, 1994).

An alternative to the rating scale is the dual-task analysis method. This method involves the performance of two tasks simultaneously and assessing the performance of one task with interference from the other. It is theorized that if two tasks utilizing the same resources are performed, the distribution of limited resources must be split among the two tasks (Brunken, Plass, & Leutner, 2003; Mayer & Moreno, 2003; Verwey & Veltman, 1996). The performance across the two tasks was measured in various configurations, with one approach analyzing performance in the primary task when a secondary task is added, or analyzing the primary task alone. The other method is to study the performance of the secondary task and to determine how the primary task alters performance. The rating scale and dual-task analysis rely on observations of the user or self-assessment of cognitive processing.

Physiological measurements such as heart activity, brain activity (Smith & Jonides, 1999) and eye activity have also been used to measure changes in cognitive functioning, though the most common model is the rating scale. Brookings and colleagues used subjective tests and all three physiological measures and found that the EEG was the only one that could accurately measure the differences among tasks (Brookings, Wilson, & Swain, 1996). The reliability and validity of using the EEG have been reproduced by other researchers, indicating that the EEG was sensitive enough to differentiate cognitive load with high precision (Gevins et al., 1998; Gundel & Wilson, 1992; Murata, 2005). Recent research by Berka et al. compared EEG measures with subjective and objective scores on vigilance and memory tasks. They found that the EEG was capable of monitoring dynamic fluctuations in cognitive states. An increase in EEG workload was observed with an increased difficulty in tasks and working memory load (Berka et al., 2007). As working memory load increased, more cortical networks were required to perform the task (Gevins et al., 1998). When the task complexity increased, it resulted in varied EEG modulation, most likely caused by the changes in cognitive requirements. The ability of the EEG to accurately analyze neurophysiological data in a small temporal window and the sensitivity to task differences makes this an attractive model for measuring cognitive workload in human–computer interfaces (HCI) (Buscema, Rossini, Babiloni, & Grossi, 2007; Minnery & Fine, 2009; Stevens, Galloway, & Berka, 2007).

2. Methodology

Six volunteers were recruited (3 males, 3 females) to determine how altering the button mappings in a touchscreen display would impact cognitive load. Past research by Besnard had indicated that small subtle changes would lead to increased errors (Besnard & Cacitti, 2005). The experiment was designed to elucidate how subjects process the changes in the interfaces. Subjects were informed that cognitive load would be measured over three days of interface use with the use of a quantitative electroencephalogram (QEEG). Monopolar recordings were obtained from the left and right frontal (F7, F8), parietal (P3, P4), temporal (T3, T4) and occipitals (O1, O2) using the international 10–20 standard. The interface consisted of 4 symbols centered in the bottom portion of the screen with a larger symbol displayed directly above the 4 symbols as seen in Fig. 1a and b.

The interface was displayed on a touchscreen display placed within arms length of the user. Each trial was composed of 12 sessions of 4 buttons displayed with a 5 s pause between sessions. The trial began with a 5 s pause, followed by a symbol flashed onto the screen for 3 ms. The subject would then be required to touch the corresponding symbol that appeared. Once the correct symbol was selected, the next symbol would appear for 3 ms. A sequence of four symbols would appear in this manner followed by a 5 s pause to disrupt the four symbol sequence. The sequence displayed was the same for all 12 sessions across all three days. Baseline EEG measures were taken each day to establish eyes closed conditions and for verification of the signal. During the first and second days, three trials were run each day to establish expert status. The third trial on the third day involved an alteration of button mappings and an extra fourth trial altered the button sequence and the order of the buttons displayed. EEG measurements were recorded for 6 frequency ranges (Table 1).

3. Results

The 6 frequency ranges for the QEEG data was combined to give overall cognitive workload for each trial and lobe. Synchronization was obtained by calculating correlations between the various brain areas. All correlations used to indicate synchronicity were greater