Web site quality evaluation combining eyetracking and physiologicial measures to self-reported emotions: an exploratory research

S. Prom Tep[1], A. Dufresne[2], .J. Saulnier[3], and M. Archambault[3]
[1]*RBC Financial Group Chair of E-commerce, HEC Montréal, Montreal (QC), Canada, Sandrine.prom-tep@hec.ca*
[2]*Department of Communications, Université de Montréal, Montreal (QC) Canada, Aude.dufresne@umontreal.ca*
[3]*User Experience Group, Web Bell Solutions, Montreal (QC), Canada, Julie.Saulnier@bell.ca, Marion.archambault@bell.ca*

Context

Studying Human Computer interaction, standard usability measures have focused on effectiveness, efficiency and satisfaction and do not incorporate emotions [1]. Recently, the importance of the affective quality in the design and in the evaluation of interactive systems has been stressed [2]. With e-commerce and the multiplication of retail Web sites, taking emotions into considerations become crucial as it has been clearly established that consumers are not always rational in their choices, nor utilitarian in their motivations [3]. If there is much more than cognitive information processing going on while consumers are online [4], measuring behavior for interactive systems evaluation has almost exclusively limited its focus to cognitive activities. Models combining emotions with human information processing variables are scarce and when available, they rely on reported data [5]. Research has shown though, that emotions are experienced both at the conscious and preconscious levels [6]. We will present an exploratory project, where we attempted to tackle this complex phenomenon. We triangulated non-consciously informed data derived from eyetracking and physiological measures, with post activity self-report of emotions on a scale to measure emotional reaction of users who are exploring a commercial Web site [7].

Objectives

Four working assumptions were made:

- The user profile (aesthetic and interest in the content) has an impact on how they will react to the imagery.
- There is a correlation between the user reported feelings and the variations in their physiological measurement during the interaction with the interface.
- Zones of interest in the interface can be detected using physiological measurement.
- The user's stress level in relation to the sensor diminishes throughout the session (habituation).

Methodology

Holbrook and Batra [7], present a set of 29 a priori multi-item emotional indices which categorizes emotions according to their type. While we did not seek to analyze all of these emotions through this experience, we selected some that are considered relevant in the field of human-machine interaction: joy (feeling pleased), surprise (feeling amazed), activation (feeling aroused, excited), hypoactivation (feeling bored), surgency (feeling playful), conflict (feeling tense), desire (feeling full of craving), involvement (feeling informed), déjà vu (feeling unimpressed), interest (feeling curious). We developed a questionnaire with these emotions and asked users to rate each emotion the interface provoked on a scale from 1 to 7 (e.g.: Did this interface make you feel bored? 1 meaning not at all bored and 7 meaning very bored). We also asked if they had felt any other type of emotion other than the ones proposed.

In order to explore the correlation between the self-reported emotional and behavioral reactions, we added physiological response to the usability testing of a commercial Web site, adapting the protocol to include physiological and eye-tracking measures of the user reactions. Among the range of sensors, we have chosen the least intrusive and the least dependent on the environment. So we used measures of respiration, skin conductance, temperature and pulse (BVP). Three scenarios were tested with six participants. In each session the participants were asked to verbalize their actions and reactions while completing the tasks. Each session was done in the same control environment and lasted ninety minutes. We wanted to measure their first impression as they saw the interface, so the experiment was organized with the following tasks:

- The users were greeted, received explanations and signed the agreement form.
- The sensors were installed and calibrated.
- Then for three sections of the site, the users were presented a page for 30 seconds, while physiological measures were taken; they were asked to fill the emotional self-report questionnaire, before moving to tasks related to that section of the site.
- We ended with a debriefing to get their impression on the experiment.

Testing environment

User testing was conducted in the Bell Solutions Web Laboratory. We used the MS Internet Explorer® Web browser (version 6.0) on a TOBII T120 unit (eye tracking system). We also used a 17" flat panel control monitor to observe the gaze movement. The processor was Intel® Core™2 CPU 6600@2.40Ghz 2GB of RAM, and the operating system was Windows® XP Pro2002 SP2. The software Tobii Studio 1.0 from Tobii was used to conduct the test and Biograph Infiniti from Thought Technology was used to capture the physiological measures. The testing room set-up conformed to the classic usability testing lab design with separate testing and observation rooms, with the control monitor connected to a computer through a VGA2USB adapter. We recorded comments from the user and the output of the control monitor, with an overlay of the gaze movement.

Analysis

Different analyses were conducted in relation to the hypotheses. First we compared self assessment measures with the different profiles of the subjects in terms of interest for a content and orientation toward aesthetics. Then we analyzed the evolution of the stress based on each physiological measure (i.e. increase in SC and pulse, temperature decrease, amplitude and frequency). We wanted to see whether stress was decreasing with habituation. We tested if we could find zones of interest where some of the physiological measures, especially those related to arousal, would be more affected. We also explored the evolution of interest as measured by the

eyetracking system: pupil dilatation, fixation time, and correlated it with physiological measures of stress, globally and also in relation to specific zones of interest. Last but not least, we compared the measures trying to relate them to the self-assessment measures, correlating them with the global assessment, but also to each specific emotion, since they may differ in effect.

This is an exploratory research with limitations in the number of subjects and the length of the experiment, but the main objective was to develop methods and tools to combine the different view points from which to assess the quality of the user experience. One of our goals also seeked to explore the hypotheses and how they could be operationalized using the different measures we had in the context of the evaluation of a commercial Web site.

References

1. Norman, D. (2004). Emotional Design: Why we love (or hate) everyday things. New York. Basic Books.
2. Zhang, P., & Li, N. (2005). The importance of Affective Quality. *Communications of the ACM,* **48(9)**, 105-108.
3. Childers, T. L., Carr, C.L., Peckc, J., and Carson, S. (2001). Hedonic and utilitarian motivations for online retail shopping behavior. *Journal of Retailing,* **77(4)**, 511-535.
4. Ethier, J., Hadaya, P., Talbot, J., Cadieux, J**.** (2006). B2C web site quality and emotions during online episodes: An empirical study, *Information & Management,* **43**, 627–639.
5. Chamberlain, L. and Broderick, A.J. (2007). Qualitative Market Research: An International Journal **10(2)**, 199-216.
6. Hirschman, E. C., & Holbrook, M. B. (1982). Hedonic Consumption: Emerging Concepts, Methods and Propositions. *Journal of Marketing,* **46**, 92-101.
7. Holbrook, M.B., Batra, R. (1987). Assessing the Role of Emotions as Mediators of Consumer Responses to Advertising, *Journal of Consumer Research,* **14(3),** 404-419.