

Evaluating Human Computer Interaction through Self-Rated Emotion

Danielle Lottridge¹

Advisor: Mark Chignell, chignell@mie.utoronto.ca

¹ Interactive Media Lab, Mechanical & Industrial Engineering, University of Toronto, 5 King's College Rd, Toronto ON M5S3G8, Canada
danielle.lottridge@utoronto.ca

Research Area. Affective self-report, Emotion, Evaluation methods.

Related Submissions. “Emotional Bandwidth: Information Theory Analysis of Affective Response Ratings Using a Continuous Slider” [12].

Research Topic. This thesis investigates the tools, psychometric characteristics, physiological relationships, and overall benefits and limits of continuous self-rated emotion as a measure of conscious affective experience.

1 Research Problem

Emotions are an important part of the user experience in human machine interaction [1]. User interface and product design requires satisfactory methods for evaluating and comparing experiences, and generally obtaining user feedback, including emotional feedback. In order to be widely used by designers, researchers, usability engineers and ergonomists, measures of users' emotional responses need to be standardized, and need to be relatively lightweight and easily usable.

Interviews and questionnaires are standard ways to assess emotion after system interaction. Some, such as the self-assessment manikin (SAM) [2], are used to rate websites. Some tools break reactions down into emotional categories for product feedback (e.g., [3]). However, summative emotions occurring after interaction has occurred are not suitable for evaluating changing interface aspects and interactions since they do not provide moment-to-moment ratings. Further, judgments may correspond to mode, average or end feelings [4]. In order to provide continuous feedback, some researchers have explored physiological measurements (e.g., [5]). These measures represent a fundamental component of emotion, and are often analyzed using self-report data, since physiological data are complex to interpret. Others [6] have explored qualitative methods to capture emotional responses in time. These investigations aim to deepen understanding of the human emotional experience, rather than provide a pragmatic way to evaluate emotional interactions.

Recent work has proposed sliders as a useful way to continuously measure self-reported emotion [7,8,9]. My thesis extends this work, to examine: what tools (sliders and variations) can we design to facilitate affective self-report? What are the

behavioural characteristics (psychometric properties) of how people use these tools? What are physiological correlates to the ratings? And what are relationships between self-reports to tasks and interaction performance? Throughout this thesis, we examine the benefits and limits of continuous affective ratings, which correspond to participants' impressions or perceptions of their emotional status.

1.1 Hypothesis/Claim

I believe that measures of conscious affective experience can give insight into mild, near neutral and generally "difficult to label" states, which may or may not correlate with physiological manifestations. I believe that different tools will facilitate different aspects of emotional self-reporting, and will hold unique psychometric properties. I posit that participants will display reliable individual differences in how they self-report their emotions. This work aims to produce rigorously validated emotion self-report tools with known psychometric properties for affective evaluation.

2 Methods

A series of experiments are proposed to progressively test different aspects of self-rated emotion. An initial step was to determine the discriminatory power of affective self-report. I collaborated in a study that used a single slider, and found a significant difference between two videoconference experiences [10]. We reanalyzed the data to uncover psychometric properties that differentiated four types of response patterns (submitted to Interact).

My first experiment, conducted in 2008, compared the three devices in a within-subjects study with 12 participants, to investigate the validity of the tools and the tradeoffs between the devices [11]. We presented 12 short videos. The ratings generally reflected the anticipated emotional content of the videos, with a certain amount of individual variation. Participants stated a preference for the touchscreen device. However, the 1 slider and 2 slider devices were found to be superior in terms of reliability measures. The results led to development of revised emotion rating tools and selection of more immersive videos to elicit stronger emotions in future studies.

The second experiment will be conducted in the spring 2009 with the goals to validate the iterated versions of the tools, discover the learning curves for emotional-self-reporting, individual differences and psychometric properties. The design is similar to the first experiment, except with longer videos and individual difference measures such as personality and emotional intelligence. The third experiment will be conducted in summer 2009. This third experiment will repeat the second experiment, but in Japan and with Japanese participants. The stimuli will be identical except that the instructions will be in Japanese and the videos will be narrated in Japanese. This experiment will give insight into the cultural differences that may exist in affective self-report.

The fourth and last experiment in this thesis will involve a health care-related task, such as browsing for chronic disease information online. The goal is to relate affective self-report to outcome measures such as browsing behaviour and retention of

information. Participants will report their affective state throughout the experience. This experiment is planned for fall 2009. It is hoped that feedback from the Interact Doctoral Consortium will facilitate insights into tradeoffs for the design of this study.

3 Proposed Solution

We designed three prototypes: a tangible one-slider model, representing valence; a two-slider model, representing valence and arousal; a two-dimensional touchscreen model, with an animated face that follows the participants' current location (Figure 1). In this case the expression on the face changed to match the position of the cursor within the two-dimensional space of valence and arousal; the curvature of the mouth changes with horizontal motion and the size (circular area) of the eyes change with vertical motion.



Fig. 1a. Two-slider Emotrace; 1b. Touchscreen Emotrace; 1c. Dynamic Facial Feedback: The mouth changes with horizontal motion; the eyes change size with vertical motion.

One outcome of this research is the discovery of psychometric properties to characterize use of sliders as scales for affective self-reports. Individuals can display varying affective capacities, which may change over time in response to the stimuli, environment and context. We used the Shannon-Weaver measure of informational entropy to quantify the rating usage bandwidth, which we refer to as Emotional Bandwidth [12].

As mentioned in the methods section, we aim to investigate the time taken to habituate to continuously self-report one's emotional state, and how individual and cultural differences affect ratings. Lastly, we plan to relate self-report to outcome measures such as behaviour, understanding and judgments of experience.

Contributions. This research is developing self-report tools for measuring emotional responses to user interfaces and interactions. The proposed thesis is original because it is the first to design and validate two dimensional self-report tools, discover their psychometric properties and test individual and cultural differences. The final experiment in this thesis will extend this research to a task scenario to investigate how ratings interact with behaviour and other outcomes. We aim for a rigorous, continuous method of measuring self-reported emotion, which is a valuable addition to traditional HCI measures such as response time, accuracy, usability, satisfaction, and workload.

References

1. Norman, D.: *Emotional Design: Why we Love (or Hate) Everyday Things*. Basic Books, New York (2004)
2. Russell, J.A.: A circumplex model of affect. *Journal of Personality and Social Psych*, 39, 1161–1178 (1980)
3. Desmet, P.: Measuring emotion: Development and application of an instrument to measure emotional responses to products. In M. A. Blythe, A. F. Monk, K. Overbeeke & P. C. Wright (eds.), *Funology: From Usability to Enjoyment*. pp. 111–123. Kluwer Academic Press, The Netherlands (2003)
4. Kahneman, D.: Objective happiness. In D. Kahneman, E. Diener, N. Schwarz (eds.), *Well-being: The foundations of hedonic psychology*. pp. 3--25. Russell Sage Foundation, New York (1999)
5. Mandryk, R.L., Inkpen, K.M.: Physiological Indicators for the Evaluation of Co-located Collaborative Play. In: *Proceedings of Computer Supported Collaborative Work CSCW'04*. Pp. 102--111. ACM Press, New York (2004)
6. Isbister, K., Höök, K., Sharp, M., & Laakso, J. (2006). The sensual evaluation instrument: developing an affective evaluation tool. In: *The Twenty-Fourth Annual SIGCHI Conference on Human Factors in Computing Systems CHI '06*. pp. 1163--1172. ACM Press, New York (2008)
7. Laurens, G. Desmet, P.: Designing a research tool: A case-study in how design and research can cross-fertilize each other, In: *DRS2008 Design Research Society Biennial Conference*, pp.362(1)--362(10). DRS digital archive (2008)
8. Baumgartner, H., Sujan, M., Padgett, D.: Patterns of Affective Reactions to Advertisements: The Integration of Moment-to-Moment Responses into Overall Judgments, *J of Market'g Research*, 34(2), 219-232 (1997)
9. Mauss, I.B., Levenson, R.W., McCarter, L., Wilhelm, F.H., Gross, J.J.: The tie that binds? Coherence among emotional experience, behavior, and autonomic physiology. *Emotion*, 5, 175--190 (2005)
10. Ranjan, A., Birnholtz, J., Balakrishnan, R.: Improving Meeting Capture by Applying Television Production Principles with Audio and Motion Detection. In: *The Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing Systems CHI '08*. pp. 227--236. ACM Press, New York (2008)
11. Lottridge, D., Chignell, M.: Emotrace: Tracing Emotions through Human-Product Interaction. In: *Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting HFES'09*. CD-ROM (2009)
12. Lottridge, D., Chignell, M.: Emotional Bandwidth: Information Theory Analysis of Affective Response Ratings Using a Continuous Slider. In: *Proceedings of INTERACT'09*. (2009)