Ten Emotion Heuristics: Guidelines for assessing the user's affective dimension easily and cost-effectively

Eva de Lera
Universitat Oberta de Catalunya
Avinguda Tibidabo 39-43
08003 Barcelona
+34 3 253 23 00

edelera@uoc.edu

Muriel Garreta-Domingo
Universitat Oberta de Catalunya
Avinguda Tibidabo 39-43
08003 Barcelona
+34 3 253 23 00

murielgd@uoc.edu

ABSTRACT

Emotional appeal is a key dimension in user experience that often goes unmeasured in most user-centered design projects. This paper presents preliminary work for developing a set of guidelines for efficiently, easily and cost-effectively assessing the users' affective state by evaluating their expressive reactions during an interface evaluation process. The evaluation of this dimension complements the analysis of the objective and quantitative data gathered through usability tests and the subjective feedback provided through post-test questionnaires.

Categories and Subject Descriptors

H.5.2 [Information interfaces and presentation] User Interfaces

General Terms

Measurement, Documentation, Design, Experimentation, Theory.

Keywords

Affective computing, human-computer interaction, evaluation, usability testing, emotions, heuristics, user-centered design.

1. INTRODUCTION

Emotion is a key aspect in user experience since measuring it helps us understand the user's level of engagement and motivation. As Spillers [20] writes, "emotions govern the quality of interactions with a product in the user's environment and relate directly to appraisal of the user experience. Users generate emotion as a way to minimize errors, interpret functionality, or obtain relief from the complexity of a task."

Therefore, accounting for emotional cues during an interface evaluation process provides usability practitioners, researchers and interactive designers with valuable information. Nowadays, there are several software applications that automatically capture facial expressions and eye gaze which provide key information to the practitioner. However, this paper defines an observational system to help evaluate interactions in an easier, time and cost effective manner through observation of the users.

Our tool allows researchers and practitioners to take into account emotional measures such as gestures and oral expressions without the requirement of extra software and hardware. Measuring user's emotion is both difficult and costly [3, 11], therefore, most interface evaluation efforts focus on cognitive and subjective aspects, neglecting the affective dimension. To date, usability practitioners have mainly relied on performance user test data and on the subjective information from the post-test questionnaire to measure user satisfaction and emotions. Measuring errors, time and other objective measures provides key but partial information. Feedback surveys or questionnaires provide only partial and often unreliable data, especially considering that users tend to give a positive evaluation to avoid blaming the person who developed the application or simply to minimize the time spent on the evaluation. Moreover, analyzing a questionnaire is a subjective measure of the user's feelings and emotions; therefore it is not a dependable methodology to measure affect. As a result, most common evaluation methods (not considering facial recognition software and other advanced techniques as common methods) have some limitations, as objective data is mostly cognitive and a questionnaire's subjective data provides the evaluator's with the user's perception of his/her emotions and not actual state(s) during the test.

Human emotions and affect are essential to understanding users, as these can facilitate the development of persistence and deep interest in a subject or goal. The analysis of this affective dimension in empirical user-centered design (UCD) methods helps us ensure that our users will be engaged and motivated while using our systems. Therefore, analyzing and evaluating emotional cues will provide practitioners with a third dimension of analysis for collecting user data, supplementing typical and common evaluation methods and resulting in a more accurate understanding of the user's experience.

Nowadays, there are few techniques and methodologies for gathering affective data without asking the users what and how they feel. We can give computers affective perceptual abilities and measure physiological and behavioral signals such as body-worn accelerometers, rubber and fabric electrodes, for example [17]. We can also evaluate users' eye gaze and collect electro-physiologic signals, galvanic skin response (GSR), electrocardiography (EKG), electroencephalography (EEG) and electromyography (EMG) data, blood volume pulse, heart rate or respiration and, more recently, facial recognition software. As we have already seen, most of these methods have limitations as they can be intrusive for the user, costly and most require specific skills and additional evaluation time.

Our aim was to find a non-invasive, cross-cultural, cost-efficient and easy to carry out method to help gain further

understanding about the affective state of a person during an interface evaluation with users. This observational technique does not replace the current and most common methods used during a UCD process, but complements the objective and subjective data gathered, therefore adding a third dimension to the evaluation process.

2. FACIAL EXPRESSIONS AND HUMAN-COMPUTER INTERACTION

Emotions are best treated as multifaceted phenomena consisting of behavioral reactions, expressive reactions, physiological reactions and subjective feelings [6]. However, to date, most instruments measure one component at a time or a group of specific components, such as facial recognition software, that does not yet gather other body gestures or vocal data. An exception is the AMUSE tool which helps practitioners to conduct interface evaluations by collecting and aggregating different sources of data including psychological and navigation data [2]. This approach is similar to ours in the sense that mixes different data from more than one source. In summary, our study aims at providing an observation instrument that helps account for affective events during common usability evaluations, providing more data than the one obtained through an intuitive and unstructured observation but without using a more complex, expensive, and costly technique.

Focusing on the expressive reactions component, facial expressions are central in the area of emotional research [1]. The first major scientific study of facial communication was published by Charles Darwin in 1872 [5], who concluded that many expressions and their meanings (e.g., for astonishment, shame, fear, horror, pride, hatred, wrath, love, joy, guilt, anxiety, shyness, and modesty) are universal. Other studies indicate that the facial expressions of happiness, sadness, anger, fear, surprise, disgust, and interest are universal across cultures [7]. Therefore, using facial expressions as a tool to evaluate the emotional dimension is a cross-cultural tool.

Several studies on emotions and human-computer interaction are based on the analysis of facial expressions. Nevertheless, most focus on the analysis of physiological data or facial recognition, omitting other non-verbal communication aspects. Hazlett [10] describes how facial EMG sensors were used to detect facial emotional responses while the subjects performed tasks on websites. Partala and Surakka [16] studied the effects of affective interventions by recording facial EMG responses from the muscle sites that control smiling and frowning. Branco et al. [1] approach was closer to ours in the sense that it complements the traditional methods of software usability evaluation by monitoring users´ spontaneous facial expressions as a method to identify the moment of occurrence of adverse events. However, they too used EMG sensors to do the monitoring. All of the above-mentioned methods are costly, require specific skills and are time consuming, a limitation for most usability and human computer interaction practitioners.

Our technique is based in observation and does not require extra implementation effort since most interface evaluations are conducted observing and recording the user as he or she interacts with the interface. In such a scenario, facial and body expressions are often observed and recorded, but generally not measured in a structured manner.

Another approach used to evaluate emotion is the instrument developed by Desmet [6]. PrEMO is a non-verbal self-report instrument that measures 14 emotions that are often elicited by product design. This tool requires respondents to report their emotions with the use of expressive cartoon animations. This method is closer to the self-assessment questionnaires at the end of the user test than to the analysis of emotions we had envisioned. The information gathered through this method is still subjective and does not provide and accurate understanding of the user's emotions.

3. THE 10 HEURISTICS

The ten emotion heuristics are based on theories that relate expressive reactions to distinct emotions. The heuristics are guidelines to help measure the affective state easily, cost-effectively and cross-culturally. One of the theories in which our work is based is the Facial Action Coding System (FACS) [8], currently a well-known standard to systematically categorize the physical expression of emotions. The Maximally Discriminative Facial Moving Coding System (MAX) [12] is another theory that links expression features to specific emotions.

Using these theories and other research as a starting point, we selected a subset of features that allowed us to partially assess the emotional reaction of the users as they interacted with an application. These set of features were identified from previous user evaluations and helped us build a list of the most common expressions taking place during user evaluations. We correlated the emotional cues identified with an emotional state and ensuring that these could be easily identified and measured during a user evaluation. A total of 10 emotional cues were selected. Better than any body parts, our faces reveal emotions, opinions, and moods. However, we use all of our body to communicate nonverbally. Thus, our list of heuristics includes some features that are not directly related to facial expressions. Our study focused on recording one measure for the 10 emotional cues, as more work and experimentation would be required to provide a specific measure for each cue. The goal of this study was to provide a positive, neutral or negative value to the overall user experience.

Again, it is important to note that this emotional data needs to be analyzed in conjunction with the user evaluation's other objective measures (time, errors, etc.) as well as the subjective measures gathered through the feedback questionnaire. In conclusion, the ten heuristics focus on taking into account the user's instantaneous emotional reactions, while the performance and navigation data provides objective data and, the *a posteriori* self-assessment provides the user's perception of his/her emotional state.

The 10 heuristics help measure the affective dimension when the product designers are looking for a neutral and relaxed interaction with the application. Therefore, these do not apply when evaluating a game or music website, for example. The smile heuristic included in this method represents the goal of the evaluation: to see a user with relaxed facials, therefore, without experiencing negative reactions or frustration.

1. **Frowning**. Frowning can be a sign of a necessity to concentrate, displeasure or of perceived lack of clarity. Darwin [5] wrote about how frowning is one of the signs of deep and "perplexed reflection". In their study, Partala and Surakka [16] found that the frowning activity attenuated significantly after the positive interventions than the no intervention condition.

2. **Brow Raising.** Brow raising should also be considered a negative expressive reaction. To lift the arch of short hairs above the eye is a sign of uncertainty, disbelief, surprise and exasperation [9].

3. **Gazing Away.** The gazing away from the screen may be perceived as a sign of deception. For example, looking down tends to convey a defeated attitude but can also reflect guilt, shame or submissiveness [9].

4. **Smiling.** A smile, or elevation of the cheeks, is a sign of satisfaction. The user may have encountered an element of joy during the evaluation process. Partala and Surakka [16] found that smiling activity was significantly higher during the positive condition.

5. **Compressing the Lip**. Seeing the user compress his or her lips should be perceived as a sign of frustration and confusion. Lip and jaw tension clearly reflects anxious feelings, nervousness, and emotional concerns. [9]

6. **Moving the Mouth.** If the user is seen mouth gesturing or speaking to himself / herself, this is associated with a sign of being lost and of uncertainty.

7. **Expressing Vocally.** Vocal expressions such as sighs, gasps, coughs, as well as the volume of the expression, the tone or quality of the expression may be signs of frustration or deception.

8. **Hand Touching the Face.** Elevating the hand that is placed on the mouse to his / her face is a sign of confusion and uncertainty, generally a sign of the user being lost or tired.

9. **Drawing Back on the Chair.** The user may be experiencing negative or refusing emotions. By drawing back the chair, he / she may be showing a desire to get away from the present situation.

10. **Forward Leaning the Trunk.** Leaning forward and showing a sunken chest may be a sign of depression and frustration with the task at hand. Like with the previous heuristic, the user might be encountering difficulties but instead of showing refusal, leaning forward is a sign of attentiveness, of "getting closer".

4. OUR PILOT STUDY

Our pilot study aimed at demonstrating the validity of the 10 heuristics as a one measure observational system to help evaluate, besides the traditional data gathered in user testing, a set of expressions or emotional cues that a user may demonstrate while interacting with a system. The 10 heuristics are easy to identify during a typical user evaluation and can be quickly analyzed in conjunction with the other gathered data. Each of these cues was assigned a positive, neutral and negative value and they are primarily aimed at evaluating the negative or frustrated emotional state [19]. For example, as mentioned, frowning is related to obstacles while the movement of the cheeks with pleasantness [18]. However, since the study did not experiment with each heuristic individually, the result of our evaluation is either that the user had a positive user experience, a neutral or negative one. In the study, identifying five negative heuristics provided a negative experience value that would later be evaluated in conjunction with the other data gathered.

In order to begin evaluating our methodology, we conducted a test with 8 participants. Four participants were assigned to carry out some tasks at an intentionally frustrating online supermarket, and 4 other participants were asked to carry out the same exact tasks at a much less frustrating online supermarket. This helped us identify whether the emotional cues gathered were in fact related to the difficulty of the task (frustration event) or another variable. Our study gathered the objective data (time, errors and number of clicks) and the

values for the emotional cues we noted in the 10 emotional heuristics guideline (positive, negative and neutral). At the end of the test, users were also asked to fill in a feedback questionnaire that included questions about the difficulty of the accomplished tasks and their overall satisfaction.

Our pilot study had a total of eight participants aged from 28 to 47. Half were men while the other half were women. All participants used computers on a daily basis and had not previously conducted their home shopping online. Half of the participants carried out the tasks in one supermarket while the other half in another online supermarket. While this is an initial study and further research and testing needs to be done to fully validate this technique, the results showed that the emotional cues identified always accompanied moments of errors or difficulty and, most importantly, that sometimes the emotional cues would come unrelated to a specific negative event providing us with new information about the user experience that we would not have collected if we were just gathering the other common data. When this happened, we would evaluate the event as negative, and when five of these events would happen during the user evaluation we would give the evaluation a negative measure for user experience.

For our usability laboratory, we used Morae software [15]. Using this software, we captured a video image mixing the PC screen and the participants' faces. Additionally, Morae saved all clicks and keyboard actions in a file. Capturing the user's video and audio allowed us to review specific moments as needed, and to involve other observers that could not be present at the time of the evaluation.

5. RESULTS AND FUTURE WORK

The interface evaluation analysis considered the three dimensions. The cognitive dimension was analyzed through the number of clicks, the time needed to accomplish the task, and the number of errors. The user's emotional perception was gathered through the feedback questionnaire and the affective dimension was observed through the 10 emotional cues. Users' expressions and comments were also noted to support the evaluation results. The tests were observed by a multidisciplinary team; one user experience director, one psychologist and one graphic designer.

The observers' analysis of both the interactions and observations recorded concluded that the emotional cues, together with the objective data, provided a more accurate understanding of the user experience and level of satisfaction than the questionnaires completed at the end of the test. Participants that encountered errors and took longer time showed signs of frustration (emotional cues) during the evaluation but did not mention them in the users' feedback questionnaire. Often participants want to please the practitioner, avoiding criticism, and they do not want to provide an overall negative evaluation. Besides, they tend to think that the cause was their lack of ability, instead of it being a design problem.

The analysis of these signals or emotional cues also showed consistency of emotional state throughout all participants as they all expressed frustration in similar ways. However, these cues were harder to identify when the participants did not display many facial or bodily expressions. Some participants were more expressive than others but the emotion heuristics allowed us to provide a positive, neutral or negative value for their overall experience. Identifying several emotional cues during the user evaluation helped us understand if the participant had an overall emotionally positive or negative experience.

In summary, and considering this as preliminary work that requires further evaluation, measuring the affective dimension with our observational system in conjunction with the other data provided a better understanding of the user's experience. At the same time, it is important to note that not all emotions can be identified through facial or bodily movements, so we may not be able to evaluate all moments of frustration, anxiety, or satisfaction, but the overall experience. Further research should be conducted in evaluating emotions when these occur without expression [4].

Our pilot study aims at providing a preliminary guideline to help conduct a structured observation to evaluate the emotional dimension during a user evaluation. The 10 emotion heuristics provide researchers and practitioners with a set of guidelines that can help them to begin incorporating the affective dimension in their user evaluations.

6. ACKNOWLEDGMENTS

7. REFERENCES

[1] Branco, P., Firth, P., Encarnao, L. M. & Bonato, P. Faces of emotion in human-computer interaction. In *Ext. Abstracts CHI 2005* (Portland, OR, April 2-7,2005), ACM Press, 1236-1239.

[2] Chateau, N. and Mersiol, M. *AMUSE: A tool for evaluating affective interfaces*. Unpublished paper. Retrieved on 15 February, 2007 from http://www.sics.se/~kia/evaluating_affective_interfaces/Chateau.pdf

[3] Chin, J.P., Diehl, V.A., & Norman, K. Development of an instrument measuring user satisfaction of the human-computer interface, in *Proceedings of CHI 1988* (Washington DC, May 1988), ACM Press, 213-218.

[4] Dalgleish, T. and Power, M. (Eds.). *Handbook of Cognition and Emotion*. John Wiley & Sons, Ltd., Sussex, U.K., 1999.

[5] Darwin, C. *The expression of the emotions in man and animals*. Oxford University Press, New York, NY, 1872/1998.

[6] Desmet, P. M. A. Measuring emotions: Development of an instrument to measure emotional responses to products. In Blythe, M.A., Overbeeke, K., Monk, A.F. and Wright, P.C. (Eds.), *Funology: from usability to enjoyment*. Kluwer Academic Publishers. Dordrecht, Boston, London, 2003.

[7] Ekman, P. and Friesen, W. V. Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, 17(2) (1971) 124-129.

[8] Ekman, P. and Friesen, W. V. *Facial Action Coding System: A technique for the measurement of facial movement*. Consulting Psychologists Press, Palo Alto, CA, 1978.

[9] Givens, D. B. The nonverbal dictionary of gestures, signs and body language cues http://members.aol.com/nonverbal2/diction1.htm#The%20 NONVERBAL%20DICTIONARY.

[10] Hazlett, R. Measurement of User Frustration: A Biologic Approach. *Ext. Abstracts CHI 2003* (Florida, FL, April 5-10, 2003), ACM. Press, 734-735.

[11] Ives, B., Olson, M. H., and Baroudi, J. J. The measurement of user information satisfaction. *Communications of the ACM*, 26 (1983) 785-793.

[12] Izard, C. E. *The Maximally Discriminative Facial Movement Coding System (MAX)*. Newark: Instructional Recourses Centre, University of Delaware, Newark, DL, 1979.

[13] Mahlke, S. and Minge, M. *Emotions and EMG measures of facial muscles in interactive contexts*. Unpublished paper. Retrieved on 24 February, 2007 from www.bartneck.de/workshop/chi2006/papers/mahlke_hcif06.pdf

[14] Mandryk, R.L., Atkins, M.S. and Inkpen, K.M.. A continuous and objective evaluation of emotional experience with interactive play environments. In *Proceedings CHI 2006* (Montréal, Québec, Canada, April 22-27, 2006), ACM Press, 1027-1036.

[15] Morae Techsmith: Usability Testing for Software and Websites, http://www.techsmith.com/morae.asp

[16] PartalaT. and Surakka, V. The effects of affective interventions in human-computer interaction. *Interacting with Computers*, 16 (2004) 295-309.

[17] Picard, R. W., and Daily, S.B. *Evaluating affective interactions: Alternatives to asking what users feel*. Presented at CHI 2005 Workshop 'Evaluating Affective Interfaces' (Portland, OR, April 2-7,2005).

[18] Pope, L. K. and Smith, C. A. On the distinct meanings of smiles and frowns. *Cognition and Emotion*, 8 (1994) 65-72.

[19] Scherer, K., Wallbot, H.G. and Summerfield, A. *Experiencing Emotion. A cross-cultural study*. Cambridge University Press, Cambridge, MA, 1986.

[20] Spillers, F.: *Emotion as a Cognitive Artifact and the Design Implications for Products That are Perceived As Pleasurable*. Retrieved on 18 February, 2007 from http://www.experiencedynamics.com/pdfs/published_works/Spillers-EmotionDesign-Proceedings.pdf