

Incorporating Facial EMG Emotion Measures as Feedback in the Software Design Process

Joey Benedek

Microsoft Corporation
One Microsoft Way
Redmond, WA 98052
joeyb@microsoft.com

Richard L. Hazlett

Johns Hopkins University School of Medicine
2045 York Road, 3rd Floor
Baltimore, MD 21093
rlhazlet@jhmi.edu

ABSTRACT

Facial electromyography (EMG) has been used in marketing research to measure emotional reaction to brand and advertising. This paper reports on two studies that bring EMG to HCI. We demonstrate the usefulness of measuring involuntary emotional reaction at key product purchase evaluative stages: first-impression (aesthetic) and during use (interaction).

Incorporating Facial EMG Emotion Measures as Feedback in the Software Design Process

Joey Benedek

Microsoft Corporation
One Microsoft Way
Redmond, WA 98052
joeyb@microsoft.com

Richard L. Hazlett

Johns Hopkins University School of Medicine
2045 York Road, 3rd Floor
Baltimore, MD 21093
rlhazlet@jhmi.edu

ABSTRACT

This paper reports on the results of two studies that used facial EMG measures combined with verbal and performance measures to provide feedback in the software design process on the user's emotional state. The first study assessed 16 participant's emotional responses while they passively viewed mock ups of proposed new operating system features. The second study measured the emotional responses of 15 participants while they actively used one of two versions of a media player. This new multimodal assessment method was able to provide a sensitive measure of the desirability of the proposed software features, and a measure of emotional tension and mental effort expended in the interactive tasks. The results of the interactive study also demonstrated that total time on task was not a reliable measure of the tension and effort required for software that provides users with opportunities for enjoyment and entertainment.

Author Keywords

Emotion, user experience, software testing methods, desirability, physiologic measures

ACM Classification Keywords

H5.2. [Information interfaces and presentation]: User Interfaces – *Evaluation/Methodology, user-centered design.*

INTRODUCTION

There has been an increasing recognition of the influence of emotion in human computer interactions and consumer purchases of software [20,30,31]. Marketing research in general has identified that a positive emotional response to products increases desirability and likeability of the product, increases attachment and positive attitude toward the brand, and increases the likelihood of purchase [1,19].

Research on how the user's emotional state during software use affects the product's desirability and the user's emotional attachment to the product or brand is in its infancy.

On considering a user's experience, a positive emotional response to a software product can occur when first seen (an over the shoulder glance), or it can develop when using the product. This positive emotional response in the prospective user is likely produced in at least one of two ways: 1) the attractiveness of the visual appearance of the product's features, and 2) an appreciation of the new and improved functionality offered by the features of the product. After purchase, the software's desirability is affected by the pleasure the user gets out of using the product, and also by the minimizing of stress, tension and excessive mental effort that can lead to frustration during use. Though the importance of the user's emotion has been recognized, it is much less understood how to effectively incorporate feedback on the user's emotion into the multiple steps of an iterative design process, and even more basically how to best measure the user's emotional state. Our goal in this paper is to explore the background concepts and issues in emotion assessment in HCI, and to describe an approach to iterative emotion assessment of software design. To start this exploration we look at the literature on the nature of emotion and its measurement.

THE NATURE AND STRUCTURE OF EMOTION

Multi-variate research into a variety of emotional phenomenon such as affectively toned pictures, judgments, behavior, and word meanings have found that emotional phenomena can best be organized into two overall emotional/motivational dimensions [27,32,35,45]. These two dimensions or systems have been variously called Positive Activation and Negative Activation [45], Positivity and Negativity [6], and Energetic Arousal and Tense Arousal [40]. The first dimension encompasses the various discrete positive emotions such as joy, pleasure, love, interest, hope, etc., and the more attentional and motivational concepts of energy, flow, immersion, and engagement. The second dimension encompasses the discrete negative emotions of fear, sadness, anger, etc., and the more motivational concepts of stress and tension. Neurophysiologic studies support this dual organization

and have identified a separate neural system that evaluates and reacts to stimuli with positive significance for the individual, and a system for reacting to aversive stimuli [29,6]. The positive system is organized and controlled by areas in the left frontal lobe of the brain, and control of the evaluative process of aversive stimuli resides in the right frontal lobe [42].

Research shows that positive emotion has effects that go well beyond just the positive feeling of the moment. While negative emotions have been found to narrow one's attention and action repertoires in preparation for fight or flight, positive emotions broaden one's thoughts and action repertoires, opening up the person to discover novel lines of thought and action. For example, joy creates the urge to play, and interest creates the urge to explore. Barbara Fredrickson has developed the 'Broaden and Build' theory of the function of positive emotion [21]. She and others have found that the positive emotions initiate an upward spiral in emotion that builds on itself, predicting future positive emotion and successful coping [22]. Increases in positive emotion broaden ones resources, building up reserves which can be drawn upon when needed. Studies have shown that increased positive emotion enhances creative, flexible and efficient thinking, leads to creative action, undoes negative emotion, enhances psychological resilience, and increases work productivity [36].

The negative emotion most likely to surface in using software is frustration. The pervasiveness of frustrating experiences for the user of computers has been well-documented and occurs on a frequent basis [9]. Lazar and colleagues [28] reviewed the literature on the concept of frustration and investigated it with a time diary approach. They found that computer users in the workplace reported wasting over 42% of their time due to frustrating experiences. Though frustration has been variously defined (see [28] for review), most definitions involve the blockage, delay or interference in obtaining a goal, and that the development of frustration is dependent upon both external task and internal individual factors. The major individual user factors that influence levels of frustration are the importance of the goal to the user and the expectation for success at goal attainment [4,16]. Task performance factors that influence frustration are the severity of the interference or goal obstacle, the length of the delay in goal attainment, and the amount of effort required to try and overcome the obstacle. Lazar and colleagues [28] found that when their users had a frustrating experience, most of the time they were able to complete their tasks, but at a cost in time, effort and mood. The user's frustration increases and the product's desirability decreases, therefore, as a user invests more effort than they expected into using a product.

EMOTION MEASUREMENT

Formal usability tests in a lab setting are an excellent method to evaluate whether users can complete tasks; however, the techniques employed have limited effectiveness for measuring the user's emotional experience and desirability of the product. One standard method used to evaluate these intangible aspects is a questionnaire with Likert scales. One problem with this method, particularly related to emotional assessment, is that the topics of the questions or anchors on the scales are assigned by the practitioner and often do not mean as much to a participant brought into the lab. Dawes [12] talks about 4 common challenges with creating strong scales which are not easily overcome: ambiguity of statements, ambiguity of scales, confounding and response bias. An inherent positive response bias related to implicit demands of the testing scenario has been noted [30], and this measurement bias is enhanced when testing is conducted in the lab of the organization whose product is being tested. Czerwinski, Horvitz and Cutrell [10] found that although less than 60% of tasks could be completed without experimenter intervention, an amazing 17 out of 19 satisfaction questionnaire items were rated as above average.

There are other approaches for collecting user experience data. An interview can result in useful data, but this approach can be time consuming and with some users it can be difficult to elicit their candid or more negative feedback. Think out loud techniques can interfere with the experience and are more a cognitive than an emotional measure. A fundamental problem with these verbal measures is that emotional experiences are not primarily language-based: cognitive effort is required to put emotional experience into words, and this effort can contaminate the measures. Also, post-use questions are retrospective and users may not remember how their responses unfolded over time while navigating or completing a series of tasks.

There have been attempts in HCI and product design fields to develop better self-report measures. Benedek and Miner [3] attempted a practical application of the card sort in place of rating scales called the Desirability Toolkit. Generic words which were not necessarily related to the product or the dimensions of measurement interest were placed on cards and users were able to select the words they would use to describe the product they used or how it made them feel. Users were not given ratings on predetermined scales but rather created their own scales with an opportunity to explain their answers. Desmet's [14] Product Emotion Measurement Instrument provides a method for measuring emotional reaction to product design. In place of typical self report techniques and scales participants select animated characters to represent their opinion of a product. The characters were validated as representative of a set of emotions. Both the

Desirability Toolkit and the Product Emotion Measurement Instrument avoid some of the problems with classic self report but both still require conscious decision making and evaluation and are susceptible to cognitive influences.

An alternative measurement approach is to measure a naturally occurring marker (behavioral or physiological) that accompanies changes in emotional state, which doesn't require cognitive effort or memory to produce. A number of physiologic measures have been tried as emotion markers. One notable study [34] to apply physiologic measures to HCI had individuals play a computer game that froze seemingly at random while measuring skin conductance, blood volume pressure and muscle tension. The authors found a correlation between the physiological signals and the frustrating events. Heart rate and skin conductance measures though are limited as emotional measures, as they are only able to indicate arousal, and not emotional valence. The study of the brain with imaging methods such as functional MRI (fMRI) is intriguing and potentially very informative, but currently it's not understood what brain activation patterns are desirable, and current theory is too rudimentary to provide direction [41]. At this point fMRI studies are also prohibitively expensive and time consuming.

FACIAL EXPRESSIONS AS NATURAL EMOTION BEHAVIORS

Emotion researchers have attempted to overcome the various stumbling blocks to the measurement of emotion for many years, and as the influential emotion theorist Silvan Tomkins wrote, the task of measuring the primary affects may not be that difficult, as they ... "seem to be innately related in a one-to-one fashion with an organ system which is extraordinarily visible" (Tomkins, [43], p. 204). Tomkins was of course referring to the face. Facial expressions are by far the most visible and distinctive of the emotion behaviors. Since Darwin [11], human facial displays have been thought to both reflect the individual's current emotional state and to be a means of communicating emotional information. Certain configurations of facial muscle movements have been identified with the expression of specific emotions across disparate cultures [17]. In order to measure changes in facial expressions that reflect emotional experience, Ekman and Friesen [18] developed the Facial Action Coding System (FACS), which codes observable facial muscle movements. There have been several attempts to measure viewers' emotional responses in the field of marketing and product development with the FACS system [13,23], but results were either not significant or there were too few content-related facial expressions to score. These results reflect the findings in emotion research at large: mild to moderate emotional stimuli are

often not accompanied by visually observable changes in facial expressions.

There is, however, a much more precise and sensitive method to measure changes in facial expressions than visual observation. Electromyography (EMG) measures minute changes in the electrical activity of muscles, which reflects minute muscle movements. Electromyographic techniques have been applied to certain facial muscles, and facial EMG has been shown to be capable of measuring facial muscle activity to weakly evocative emotional stimuli even when no changes in facial displays have been observed with the FACS system [8]. Even when subjects are instructed to inhibit their emotional expression facial EMG can still register the response [5]. Facial EMG studies have found that activity of the corrugator muscle, which lowers the eyebrow and is involved in producing frowns, varies inversely with the emotional valence of presented stimuli and reports of mood state; and activity of the zygomatic muscle, which controls smiling, is positively associated with positive emotional stimuli and positive mood state [27,8,15,25,27,37,38]. Corrugator muscle activity has been found to predict emotional state [7].

The corrugator muscle is also related to the more motivational components of the second dimension Negativity/Tension. In task performance research, activity of the corrugator muscle has been found to provide a sensitive index of the degree of exerted mental effort [44] and to increase with the perception of goal obstacles [33]. Since level of frustration is influenced by perceived interference by goal obstacles, corrugator EMG levels are also related to levels of frustration. In a study to validate the corrugator muscles as a measure of tension and frustration in HCI [26], corrugator muscle activity during web site usage was found to be greater for: (1) novices as compared to experienced users, (2) incorrect as compared to correct answered tasks, and (3) for the web site that was rated as more difficult.

PROJECT GOALS AND DESIGN

The above literature suggests that maximizing the software user's positive emotion and minimizing their negative emotion leads to greater creativity, learning and productivity, and less frustration for the user. Marketing research referenced above suggests that a positive emotional experience for a user would lead to greater product desirability, greater emotional attachment of the consumer to the product and the brand, and increased purchase behavior. It would seem then that the process of creating a highly desirable software product could benefit from feedback on the positive and negative emotional state of the user. In considering the emotion research literature, a promising measure for feedback at critical junctures in the software design process could be facial EMG. It can provide real time emotion measurement

while participants are viewing new features for the first time, mimicking a pre-purchase ‘over the shoulder glance’, and also facial EMG can provide real time measurement during interactive usage. The EMG methodology is more intrusive than what is typical for usability testing, as participants do have to wear several sensors on their face while being tested, but facial EMG studies indicate that valid emotional responses are produced to the appropriate stimuli, and participants seem to stop attending to the sensors as they get involved in the study.

Therefore, we wanted to test the usefulness of facial EMG at multiple stages in the design process, combining it with verbal probes and performance measures to take advantage of multiple sources of feedback on the user’s emotional responses. We tested the usefulness of this multimodal assessment protocol of positive and negative emotional state using facial EMG and verbal measures at two different stages in a design process. During the first stage, the pre-code testing stage, participants passively viewed mock ups of new operating system features while facial EMG responses were continuously measured. This measurement probe is important because it captures the magnitude of the potential consumer’s initial positive emotional response to the feature before a costly investment in writing the code has occurred. This passive viewing design also mimics the ‘over the shoulder glance’ that a potential user might catch and that could be a motivating experience for purchasing the product.

The second measurement probe is after the code has been written and involves the comparison of the old version (A) and a new beta version (B) of the software. This assessment stage is important to give feedback on how well the new design performs, and a chance to catch any problems that could adversely affect a user’s experience. During this probe facial EMG measures were collected while participants actively worked on a series of common tasks with two versions of a media player. Since a post code version of the operating system features were not yet available, we used these two versions of the media player for the interactive comparison. A between-subject design was used to avoid learning effects from completing the same set of tasks twice. A drawback to using experimenter chosen tasks is that task importance is one of the user variables that influences level of frustration [28], and these tasks may not lead to as much frustration as self-chosen tasks. However, for our purposes, the variables of tension and effort that reflect the severity of the task interference and that lead to the experience of frustration are the important variables to study for feedback on design desirability.

STUDY 1: TESTING PRE-CODE MOCKUPS: MEASURING THE EMOTIONAL REACTION TO PASSIVE FIRST IMPRESSION

METHODS

Seventeen participants were tested in a usability lab set up with observer and participant rooms separated by a soundproof wall and one-way mirror. Participants worked on a Dell Dimension 8200 PC with an Intel Pentium 4, 2.53 GHz and Sony Multisync 17” monitor. The participants consisted of ten males and seven females whom were all current PC users with a minimum of one computer in the home with an internet connection. One participant did not have valid data for this phase of the experiment resulting in a total of 16 participants. There were six females and ten males. Participants were comfortably seated in front of a computer screen, while the experimenters viewed from behind a one-way mirror, and communicated with participant over an audio speaker system. The participant’s torso and computer screen were video taped for later analysis.

Participants passively viewed a demonstration of common PC tasks such as computer “boot up”, Start Menu use and file folder implementation. The tasks were demonstrated in a Macromedia Director presentation.

Facial EMG was continuously measured with Rochester miniature Silver/Silver Chloride surface electrodes placed over the right zygomaticus major and corrugator supercillii muscles. Standard skin preparation was performed and electrode placement followed recommended guidelines [39]. The raw EMG signals were amplified and filtered with two Psylab (Contact Precision Instruments) bioamplifiers and processing system. EMG detection bandpass was set at 30Hz-500Hz, and the analogue EMG signal was digitized at 1000 Hz.



Figure 1. Participant with facial EMG sensors

Self-report data were collected at the end of each presentation. Participants were asked to comment on what aspect of the system they particularly liked. Participants were also asked to list the features they liked without prompting. After unaided reports, notable EMG elevations during the demonstration were used as a guide for specific probing, and participants were asked to explain what they liked about the features they listed.

Several procedures were used to denoise the EMG signal and identify emotionally stimulating features. First, extreme values too high to be valid data points that were due to wire pulling, head movements etc. were filtered out of the data series. Second, the video of each participant was viewed to identify and eliminate any noise in the data series due to participant movement. The raw EMG values were then rectified and summated to 200 millisecond values, and the time series was graphed. For each participant mean corrugator and zygomatic values were calculated for the overall series and for each feature. A participant was considered to have a positive emotional response to a feature if they had a zygomatic response during the feature presentation that was one standard deviation above the series mean, and also that they had an absence of elevation in the corrugator level. We based this measure on the standard in psychological testing where values above one standard deviation are considered out of the normal range [2].

RESULTS AND DISCUSSION

Listed in Table 1 are the desirability rankings of the new OS features. EMG mean level was the zygomatic mean during feature presentation summated across all subjects whom had significant elevations. Percentage of participants was the percentage of the 16 participants whom had significant elevations to that feature.

Many of the features that had significant elevations in positive emotional response were later spontaneously identified by the participant as desirable. Out of the thirteen features shown to the participants they listed as desirable from 3 to 6 (median = 4) of these features without aid. Of the two highest ranked features identified by the EMG measures for each participant, 68% were subsequently verbally identified without prompting as desirable. Verbal probing at session end collected data on what aspects of the feature the participant responded to. Participants were often able to report whether they responded to the visual attractiveness of a feature, its functionality or both. After unaided rankings and responses from the participant were collected, EMG elevations that were noted during the session were probed; revealing participants’ reactions that otherwise would not have been reported. The median number of EMG assisted probes that yielded further verbal information was three per participant.

Rank	Feature	Mean EMG in microvolts	Percentage of Participants
1	Listview 1	17.16	.75
2	Option Menu 1	17.03	.62
3	Window Management 1	16.89	.62
4	Listview 2	14.32	.56
5	Desktop 1	13.83	.31
6	File Sharing	13.04	.25
7	Listview 3	12.68	.31
8	Desktop 2	12.55	.50
9	Desktop 3	11.88	.50
10	Desktop 4	10.43	.56
11	Option Menu 2	10.12	.18
12	Desktop 5	9.45	.37
13	Screensaver	8.09	.31

Table 1. Desirability Rank of OS Feature Based on Emotional Response

This passive viewing assessment was able to measure the magnitude of positive emotional response to the mock up of possible new features with real time, unobtrusive measurement. Features were able to be compared on desirability as an aid in decision making about the new operating system design. As a measure of concurrent validity, the users self-report was found to be related to the facial EMG measures. The multimodal assessment approach employed the physiologic measures to provide a sensitive and precise measure of positive emotional activation, and the verbal report provided qualitative information. Using the online EMG readout as a method to mark significant emotional reactions of the participant, and then following up with a verbal probe seemed a useful assessment combination, and lead to insights otherwise hidden from view.

STUDY 2: INTERACTIVE BETA TESTING: MEASURING THE EMOTIONAL REACTION DURING ACTIVE USE

METHODS

We tested 16 participants in a similar lab as in study 1. One participant had extreme difficulty in completing the tasks, partially due to a language problem, and as she was a clear outlier from the other participants her data were not used in the study. This left seven participants who

used version A of a media player, and eight who used version B. The version of the media player the participant used was alternated between participants. The participants consisted of six females and nine males whom were all current PC users with a minimum of one computer in the home with an internet connection. In addition, all participants had previous experience with the media player.

Participants were given nine common tasks to complete using the media player (see Table 2 for the list). After the procedure was explained to the participant the facial EMG sensors were attached. The participants were then left in the participant room alone and told to proceed with the task list. Interaction between the observer and the participant was kept to a minimum after the tasks were begun, with only occasional instructions from the observer when the participant was having difficulty understanding the task. A video of the participant's user interface and the user's torso was recorded during the session and copied to a DVD. The EMG data collection was similar to study 1 and the raw EMG values were denoised as in study 1.

From the corrugator EMG values a measure of participant elevated tension and undesirability of product was developed. The elevated tension measure was a count of every second that the participant had a corrugator EMG value greater than one standard deviation above the participant's overall corrugator mean. The number of seconds above one standard deviation is related to the amount of positive skewness in the EMG data set [24], which can vary depending on how much tension the participant is experiencing. The greater the tension or responsiveness of the participant the greater the positive skewness in the EMG data. The use of the participant's standard deviation as a threshold controls for differences in both mean and variance of the EMG data between participants, and counts how much time the participant had substantial elevations in effort and tension: factors that influence the development of frustration.

RESULTS AND DISCUSSION

As can be seen from Table 2 the two versions were fairly similar on the first four tasks in seconds of elevated tension, and that on the last five tasks version B had substantially less time of elevated tension. In looking at the results from these tasks, it would seem that version B achieved a greater desirability than version A. However, a number of the tasks had prolonged levels of tension for both versions, which may be long enough to cause the development of frustration in natural usage. In particular, task 7, Editing Metadata, was the most frustrating task, and sustained elevated levels of tension for over a minute in version A. Episodes of elevated tension identified through the EMG measure were then explored in the video record of the participant's UI, and particular

features of the software were often identified that helped or hindered in a user's experience. There were not any media player features that evoked notable differences in zygomatic activity between versions, perhaps because the differences in the user interface between the two versions were fairly subtle.

One notable pattern illuminated by Table 2 is that total time on task and seconds of elevated tension were often not congruent, suggesting that total time on task is not a reliable measure of mental effort and emotional tension. We explored this relationship and found that elevated tension and total time on task were only modestly correlated ($r = .55, p = .05$). Table 2 shows that elevated tension and total time on task only agreed for six out of the nine tasks on which of the two versions was most stressful; indicating that for a third of the tasks total time on task incorrectly identified the software version that was most stressful and that required the most effort. Observation showed that when using the media player there are often moments when the user is "on task" but yet they are enjoying the music or just browsing, and speed is not important to the user.

Task		Elevated Tension		Total Task Time	
		Version A	Version B	Version A	Version B
1	Play CD	8.5	10.0	62	113
2	Play Album	10.0	12.6	169	118
3	Burn CD	40.4	35.2	368	360
4	Make New Playlist	22.8	26.7	244	283
5	Play all Classical Music	25.1	14.4	81	111
6	Sync with device	48.1	32.7	369	207
7	Edit Metadata	72.4	39.8	266	259
8	Play Playlist	23.0	5.4	82	75
9	Rip CD	53.0	34.2	270	332

Table 2. Mean Number of Seconds of Elevated Tension and Total Task Time in Seconds

This study compared two versions of a media player with an interactive task methodology, and was able to discriminate between the two versions in desirability based on amount of elevated tension of the user. The methodology was able to identify specific UI features that affected levels of tension and effort and contributed to version B being a more desirable product. We were also able to identify features in the new version, B, that could be improved upon for a more positive and successful experience for the user.

CONCLUSION

In these two studies we described and tested a novel approach to evaluating software design that could contribute to the development process at multiple stages. Facial EMG measures were able to provide a real time measure of the magnitude of the user's positive emotional responses during a passive viewing of mock ups of proposed new features, and the measure was able to discriminate between those features. The interactive beta testing stage gave feedback on levels of tension and effort, variables that contribute to the development of user frustration and the lowering of product desirability. This interactive testing allowed comparisons between old and new versions of the software, and gave feedback on how the new version had improved a user's experience, and how it could be changed to further improve a user's experience. The second study also illuminated the potential limitations of total time on task as a measure of mental effort and emotional tension. We speculate that the more that a software product provides the user with opportunities for enjoyment, exploring and entertainment while completing tasks, the more limited time on task becomes as a measure of emotional tension and effort. Future research with software products more designed for work tasks than entertainment will be needed to investigate whether total time on task is also a limited measure of emotional tension and mental effort for tasks where speed is a priority for the user.

In both the pre-code and beta testing stages the multimodal assessment approach provided precise real time quantitative biologic data combined with both quantitative performance data and qualitative verbal feedback. These multiple data sources were used in combination to uncover and probe the participants' emotional reactions that at times would otherwise have been hidden from view, or ran counter to what would have been indicated from traditional measures, such as task time data. This method also appears promising for measuring emotion in both the first-impression and active-use of software, which would be helpful feedback in understanding these different facets to the desirability of a software product.

REFERENCES

1. Allen, C.T., Machleit, K.A., & Kleine, S. A comparison of attitudes and emotions as predictors of behavior at diverse levels of behavioral experience. *Journal of Consumer Research*, 18, (1992), 493-504.
2. Anastasi, A. *Psychological testing*, 6th edition. Macmillan, New York, NY, 1988.
3. Benedek, J., Miner, T. Measuring Desirability: New methods for evaluating desirability in a usability lab setting. *In Proceedings of Usability Professionals Association*, (2002), Orlando, July 8-12.
4. Berkowitz, L. Whatever happened to the frustration aggression hypothesis? *American Behavioral Scientist*, 215 (1978), 691-708.
5. Cacioppo, J.T., Bush L.K., & Tassinary, L.G. Microexpressive facial actions as a function of affective stimuli: Replication and extension. *Psychological Science*, 18, (1992), 515-526.
6. Cacioppo, J.T., Gardner, W.L., & Berntson, G. G. The affect system has parallel and integrative processing components: Form follows function. *Journal of Personality and Social Psychology*, 76, (1999), 839-855.
7. Cacioppo, J. T., Martzke, J. S., Petty, R. E., & Tassinary, L. G. Specific forms of facial EMG response index emotions during an interview: From Darwin to the continuous flow hypothesis of affect-laden information processing. *Journal of Personality and Social Psychology*, 54, (1988), 592-604.
8. Cacioppo, J.T., Petty, R.E., Losch, M.E., & Kim, H.S. Electromyographic activity over facial muscle regions can differentiate the valence and intensity of affective reactions. *Journal of Personality and Social Psychology*, 50, (1986), 260-268.
9. Ceaparu, I., Lazar, J., Bessiere, K., Robinson, J. & Shneiderman, B. *Determining Causes and Severity of User Frustration*. <http://www.cs.umd.edu/hcil/newcomputing/>
10. Czerwinski, M., Horvitz, E. & Cutrell, E. Subjective Duration Assessment: An Implicit Probe for Software Usability. *In Proceedings of IHM-HCI 2001 Conference, Volume 2, (September, 2001), p. 167-170.*
11. Darwin, C. *The expression of emotions in man and animals*. Murray, London, 1872
12. Dawes, R. M. *Fundamentals of Attitude Measurement*. John Wiley and Sons, New York, NY, 1972.
13. Derbaix, C.M. The impact of affective reactions on attitudes toward the advertisement and the brand: A step toward ecological validity. *Journal of Marketing Research*, 32, (1995), 470-479.
14. Desmet, P.M.A. *Designing Emotions*. Unpublished doctoral thesis. ISBN 90-9015877-4, 2002.

15. Dimberg, U. Facial electromyography and emotional reactions. *Psychophysiology*, 27, (1990), 481-494.
16. Dollard, J., Doob, L. W., Miller, N. E., Mowrer, O. H., & Sears, R. R. *Frustration and Aggression*. New Haven: Yale University Press 1939.
17. Ekman, P., & Friesen, W.V. *Unmasking the face*. Prentice-Hall, Englewood Cliffs, NJ, 1975.
18. Ekman, P., & Friesen, W.V. *Facial coding action system (FACS): A technique for the measurement of facial actions*. Consulting Psychologists Press, Palo Alto, Ca, 1978.
19. Erevelles, S. The role of affect in marketing. *Journal of Business Research*, 42, (1998), 199-215.
20. Forlizzi, J. , Panel Chair. *Emotion and the Design of New Technology*. CHI 2003, Fort Lauderdale, FL, 2003.
21. Frederickson, B.L. What good are positive emotions? *Review of General Psychology*, 2, (1998), 300-319.
22. Frederickson, B.L. & Joiner, T. Positive emotions trigger upward spirals toward emotional well-being. *Psychological Science*, 13, (2002), 172-175.
23. Graham, J. L. A new system for measuring nonverbal responses to marketing appeals. *AMA Educator's Conference Proceedings*, 46, (1980), 340-343.
24. Hazlett, R.L., McLeod, D.M., & Hoehn-Saric, R. Muscle tension in generalized anxiety disorder: Elevated muscle tonus or agitated movement? *Psychophysiology*, 31, (1994), 189-195.
25. Hazlett, R.L., & Hazlett, S.Y. Emotional response to television commercials: Facial EMG vs. self-report. *Journal of Advertising Research*, 39, (1999), 7-23.
26. Hazlett, R.L. Measurement of user frustration: A biologic approach. *Proceedings of CHI 2003 Conference on Human Factors in Computing Systems v2* (2003), ACM Press, 734-735.
27. Lang, P.J., Greenwald, M.K., Bradley, M.M, & Hamm, A.O. Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, 30, (1993), 261-273.
28. Lazar, J., Jones, A., Ceaparu, I., Bessiere, K., & Shneiderman, B. *User Frustration with technology in the workplace*. <http://www.cs.umd.edu/hcil/newcomputing/>
29. LeDoux, J. E. Cognitive-emotional interactions in the brain. *Cognition and Emotion*, 3, (1989), 267-289.
30. Nielsen, J., & Levy, J. Measuring usability: Preference vs. performance. *Communications of the ACM*, 37 4, (1994), 66-75.
31. Norman, D.A. *Emotional Design: Why We Love (Or Hate) Everyday Things*. Basic Books, New York, NY, 2004.
32. Osgood, C., Suci, G., & Tannenbaum, P. *The measurement of meaning*. University of Illinois, Urbana, IL, 1957.
33. Pope, L.K. & Smith, C.A. On the distinct meanings of smiles and frowns. *Cognition and Emotion*, 8, (1994), 65-72.
34. Riseberg, J., Klein, J., Fernandez, R., & Picard, R. W., et. al. Frustrating the User On Purpose: Using Biosignals in a Pilot Study to Detect the User's Emotional State, in *Proceedings of CHI '98 Conference on Human Factors in Computing Systems v2* (1998), ACM Press, 227-228.
35. Russell, J. A circumplex model of affect. *Journal of Personality and Social Psychology*, 39, (1980), 1161-1178.
36. Seligman, M. *Authentic Happiness*. Free Press, New York, NY, 2002.
37. Schwartz, G.E., Ahern, G.L., & Brown, S.L. Lateralized facial muscle response to positive versus negative emotional stimuli. *Psychophysiology*, 16, (1979), 561-570.
38. Sirota, A.D., & Schwartz, G.E. Facial muscle patterning and lateralization during elation and depression imagery. *Journal of Abnormal Psychology*, 91, (1982), 25-34.
39. Tassinary, L.G., Cacioppo, J.T., & Geen, T.R. A psychometric study of surface electrode placements for facial electromyographic recording: I. The brow and cheek muscle regions. *Psychophysiology* 26, (1989), 1-16.
40. Thayer, R.E. *The biopsychology of mood and arousal*. Oxford University Press, New York, NY, 1989.
41. Tierney, J. Politics on the Brain? Resorting to MRIs for Partisan Signals. *The New York Times*, (April 20, 2004), A1.
42. Tomarken, A.J., Davidson, R.J., Wheeler, R.E., & Doss, R.C. *Journal of Personality and Social Psychology*, 62, (1992), 676-687.
43. Tomkins, S.S. *Affect, imagery, consciousness: The positive affects. (Vol 1)*. Springer, New York, NY, 1962.
44. Waterink, W., & Van Boxtel, A. Facial and jaw-elevator EMG activity in relation to changes in performance level during a sustained information task. *Biological Psychology*, 37, (1994), 183-198.
45. Watson, D., Wiese, D., Vaidya, J., & Tellegen, A. The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. *Journal of Personality and Social Psychology*, 76, (1999), 820-838.