Instrumentation for Automatic Monitoring of Affective State in Human-Computer Interaction

Jing Zhai [1] and Armando B. Barreto[1,2]

[1]Electrical and Computer Engineering Department
[2]Biomedical Engineering Department
Florida International University
{ jzhai002, barretoa }@fiu.edu

Abstract

In the modern world we interact with computers for many activities, such as working and learning. Computers are great tools for these tasks. However, some times our interaction with computers is degraded because they are inflexible and unaware of our own level of comfort in that interaction. The emerging field of 'Affective Computing' strives to give computers an awareness of the affective state of the user, and the ability to adjust to it. This research aims at sensing and recognizing typical negative emotional experiences, especially 'frustration', which may arise when people interact with computers. An integrated hardware – software setup has been developed to achieve real-time assessment of the affective status of a computer user. Three physiological signals: Blood Volume Pulse (BVP), Galvanic Skin Response (GSR) and Pupil Diameter (PD) were collected and synchronized from each computer user participating in our experiments. Several signal processing approaches were then applied to recognize the 'frustration' state of computer users. Preliminary results indicate that there exists a strong correlation between changes in these physiological signals and the shift in the emotional states of a computer user when a frustrating stimulus is applied to the interaction environment.

Introduction

Undoubtedly, computers have taken a central role in our lives, influencing many aspects of human activity, such that our daily lives cannot be imagined without computers. But at present, the mechanisms for human-computer interaction do not yet include the complexities of naturalistic social interaction, as it occurs in human-human interaction. The pursuit of affective computing is the exploration of techniques that may lead, in the future, to the development of computer systems enriched with affective understanding. Affective computing, a relatively new area of computing research, has been described as "computing which relates to, arises from, or deliberately influences emotions"(Picard 1997). It involves not only emotion detection, but extends to the implementation of emotions, and attempts to give the computer the ability to recognize and express emotions. Therefore it can develop the computers' ability to respond intelligently to human emotion and help the computer user relax from negative emotional states. This study attempts to explore a way to visualize and evaluate the emotional state identified with 'frustration' of the users, through several physiological signals that can be measured non-invasively and non-obtrusively, and its preliminary results may come in support of the concept of affective computing.

Problem Formulation

Researchers have investigated many aspects of the human body responses to emotions, in the past. Attempts have also been made to apply the knowledge gained in those studies to enrich the communication of humans to computers, for example, by the identification of facial expressions, in isolation, or in combination with speech understanding techniques. An underlying rationale for these attempts can be found in the position postulated by Reeves and Nass at Stanford: Human-computer interaction should be inherently natural and social, following the basics of human-human interaction(Reeves and Nass 1996). In human-human interaction, many of non-verbal communication cues are generated, to some extent, under conscious control of the subject. Therefore, they may be purposely manipulated ("faked") by the subject, and not reveal the true emotional state of the subject, which is the starting point for the implementation of affective computing. Physiological signals, on the other hand, are not easily controlled by the subject, and can, therefore, be seen as a truly reliable indicator of the emotional state of the subject. A vast amount of previous research focused on the assessment of mental workload through the measurement of physiological signals has also provided knowledge on the relationship between emotional responses (such as frustration) and these signals.

Many publications in the psychophysiological literature e.g., (Grings and Dawson 1978, Hugdahl 1995) indicate that there are many bodily reactions that accompany changes in the emotional state of a subject. Most previous reports from the literature on Affective Computing include measurements of Blood Volume Pulse (BVP), Galvanic

Skin Response (GSR) and Heart Rate Variability (HRV). Variation of the Pupil Diameter (PD) has also been investigated during and after auditory emotional stimulation(Partala and Surakka 2003). But, the joint analysis of pupil diameter changes with other physiological signals to indicate the emotional state of a subject has not been fully investigated. Our study explores one method that could coalesce three physiological signals for the assessment of a possible state of temporary frustration in a computer user.

From human physiology studies, it is known that the Sympathetic Division of the Autonomic Nervous System (ANS) significantly influences these physiological variables. The heart rate, skin resistance, blood pressure and pupil diameter are all affected by branches of the Sympathetic Division of the ANS. In this study, we monitor three physiological variables (GSR, BVP, PD) simultaneously to analyze potential concurrent changes that may be due to sympathetic activation associated with a multifaceted emotional state — 'frustration'.

To properly analyze several signals on the same time scale, synchronization must be carefully achieved. Since synchronizing multiple signals is much harder than monitoring signals separately, why do we use three signals rather than one? Cacioppo and Tassinary(Cacioppo and Tassinary 1990) explore the nature of psychophysiology-emotion relationships, considering several categories of connections: one-to-one (i.e., one physiological signal maps to one particular emotion), many-to-many, one-to-many , and many-to-one. Because each practical measurement has a unique sensitivity for certain aspect of the body's response system, combining multiple physiological measures in the analysis makes it stronger and more reliable (Wilson and Eggemeier 1991). Accordingly, in this study, we attempt to obtain the maximum knowledge about the emotional status of the subject by synchronizing three physiological signals and analyzing them as a whole, as different external manifestations of the same internal emotional phenomenon.

Physiologic Signals and Sensors

It is well known that when a subject experiences stress and nervous tension, the palms of his/her hands become moist. This is because sweat glands in these areas of the body are activated and fill with sweat (a hydrate solution of water and salt). Skin with higher water content will conduct an electric current more easily than dry skin, because a hydrated sweat gland provides less resistance for the passage of an electric current. Increased activity in the sympathetic nervous system will cause increased hydration in the sweat duct and on the surface of the skin. The resulting drop in skin resistance (increase in conductance) is recorded as a change in electrodermal activity. So, in everyday language, electrodermal responses can indicate 'emotional sweating'(Hansen, Johnsen and Thayer 2003). The GSR is measured by passing a small current through a pair of electrodes placed on the surface of the skin and measuring the conductivity level. Although the recording of GSR as a function of emotional or cognitive activation is a rather simple and 'low-tech' technique, it is still considered one the most sensitive physiological indicators of psychological phenomena, both in the laboratory and in the clinic. GSR is also one of the signals used in the polygraph or 'lie detector' test.

A GSR2 module (white device in Figure 1), by Thought Technology LTD (West Chazy, New York) was used in this project to measure GSR. The two elongated electrodes and molded plastic case are ideal to have the hand of the subject rest on it and achieve contact over large portions of the volar surfaces of the medial and distant phalanges in two of the subject's fingers, as recommended by Dawson, et al(Dawson, Schell and Filion 2000). Also, a 9V battery powers the device to guarantee that only safe levels of current will be sourced from one electrode and sunk into the other for skin conductivity measurements.

The resistance found in between the two electrodes is treated as the resistive value to determine the oscillation frequency of a square-wave oscillator inside this device. Originally, the square wave oscillation is fed to a speaker, creating a sound whose fundamental frequency is proportional to the skin conductance measured (for the purpose of biofeedback training). Fortunately, an earphone output is available, which we have used to feed the square wave signal into an "frequency-to-voltage converter" integrated circuit (LM2917N), which then yields a voltage proportional to instantaneous skin conductance. This modified device was calibrated by connecting several resistors of known resistance to it and measuring the output voltage of the frequency-to-voltage converter in each case.

Figure 1: GSR and BVP sensors

The Blood Volume Pulse measurements in this project were obtained using the technique called photoplethysmography (PPG), to measure the blood volume in skin capillary beds, in the finger. PPG is a non-invasive monitoring technique which relies on the light absorption characteristics of blood, so it does not require costly equipment or specialized personnel. Traditionally, the Blood Volume Pulse (BVP) has been used to determine the heart rate. If measured precisely enough, it can be used to extract heart-rate variability, which is another indicator of user affective state to be considered for human-computer interfaces (Dishman, et al. 2000, Picard and Klein 2002). Some researchers have questioned the use of BVP, as compared with ECG (electrocardigram) to measure the heart rate variability (HRV) during experimental situations(Giardino, Lehrer and Edelberg 2002), but our objective in this study is to extract the relationship between HRV and a state of frustration in the context of human-computer interaction rather than for clinical purposes. While photoplethysmographic BVP can be obtained from a normal computer user non-obtrusively, placement of ECG sensors in a normal computer user is not a realistic operational assumption. Based on this consideration, we accepted the high resolution BVP signal as adequate to derive HRV information. In our experiment, the sampling rate used to record the BVP signal is 360Hz.

The pupil of the human eye can constrict and dilate such that its diameter can range from 1.5 to more than 9mm. It is well known that pupil dilations and constrictions are governed by the Autonomic Nervous System (ANS) in humans. Several researchers have established that pupil diameter increases due to many factors. Anticipation of solving difficult problems, or even thinking of performing muscular exertion will cause slight increases in pupil size. Hess(Hess 1975) referred that other kinds of anticipation may also produce considerable pupil dilation. Previous studies also have suggested that pupil size variation is also related to cognitive information processing. This, in turn, relates to emotional states (such as frustration) since the cognitive factors play an important role in emotions(Grings and Dawson 1978). Partala and Surakka have found that using auditory emotional stimulation, the pupil size variation can be seen as an indication of affective processing(Partala and Surakka 2003). All these previous results found in the literature prompted us to attempt to use the pupil size variation to detect affective changes during human-computer interactions.

There are several techniques available to quantify pupil size variations(Grings and Dawson 1978). Currently, automatic instruments, such as eye-tracking systems, can be used to record the eye information including pupil diameter and point of gaze. In our study, the subject's left eye was monitored with an Applied Science Laboratories series 5000 eye tracking system running on a PC computer to extract the values of pupil diameter. The sampling rate of the system was 60 samples/second.

Experimental Setup

In order to measure the changes in GSR, BVP and PD that take place when frustration sets in, a hardware / software system was developed to: a) Provide an appropriate stimulus, capable of eliciting frustration in the subjects participating in the experiment; b) Provide synchronization signals for the rest of the instrumental setup, so that the segments of the several signals recorded under the frustration state can be identified and analyzed as a whole; and c) Record the signals with all the necessary time markers.

Fifteen healthy subjects participated in this study, recruited through flyers distributed in the FIU Engineering Center, including two females and thirteen males aged from 18 to 29 years old.

Stimulus Software and Context

In this research, the stimulus for the subject should elicit a state of frustration at known times, so that the changes in the physiological variables monitored can be assessed. Lawson(Lawson 1965), after Rosenzweig, defined frustration as "the occurrence of an obstacle that prevented the satisfaction of a need." The 'need', in this case, can be interpreted to mean either a need or a goal. So our objective for this work was to give the subject a goal, and prevent him/her from achieving that goal. To realize that, a protocol similar to the one used by Scheirer, et al.(Scheirer, et al. 2002)was followed. A computer game was designed with a series of 30 similar visual puzzles, each on a separate screen, presented one after the other. In each puzzle, the subject's task was to use the mouse to click on the box at the bottom of the screen that corresponded to the type of symbol (circle, square, triangle or star) that was repeated the most times in the screen shown (e.g., Figure 2). This mouse-click also advanced the current screen to the next one.

In an attempt to enhance the potential emotional responses of the experimental subject, the game was made into a competition. The individual who achieved the best overall score for accuracy and speed at the end of the data collection period would receive a 5-CD-R/RW playable/Dual Cassette Stereo System, which was given at the conclusion of this study.

To elicit user frustration in a controlled fashion during the playing of the game, the software was modified intentionally to make the screen cursor unresponsive to the movement of the mouse (frozen cursor), in three different visual puzzles. The scheduling of these delays, with respect to the 30 frames, or individual visual puzzles, such as the one shown in Figure 2, is diagrammed in Figure 3. These delays occurred after the subjects selected one box on the screen, and hindered their progress to the next puzzle. The stimulus sequence described, implemented with Macromedia Flash ®, was also programmed to output bursts of a sinusoidal tone through the sound system of the laptop used for stimulation, at selected timing landmarks through the protocol. The burst would be played out of the left channel only (binary $1 = 01_b$) at the beginning of the game; through the right channel only (binary $2 = 10_b$)

when each of the three cursor-freezing delays took place; and through both channels (binary 3 = 11_b) when the user clicked on a stop button to end the game. These timing bursts were used to introduce synchronization marks in the recordings of GSR, BVP and PD.

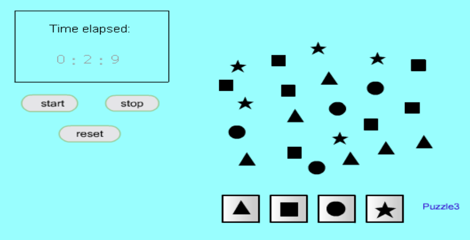

Figure 2: Stimulus Program Screenshot

The left and right channel signals from the stimulus laptop were connected to retriggerable monostable multivibrators to convert each one of those bursts (whenever they occurred) to TTL-level pulses of the same duration. Furthermore, both pulse signals were connected to a binary decoder, so that for any of the three active combinations described above, a negative digital pulse would appear in the corresponding output line of the decoder. These lines were used to activate analog switches connected in parallel with 3 individual keys ("1", "2" and "3") in the numeric pad of the keyboard controlling the eye gaze tracking system. This was done to emulate those keystrokes, which are accepted by the eye tracking system program to add timing markers in the eye gaze and pupil diameter file collected by the system. With the arrangement described above these markers would be automatically introduced when 1) the user starts the session, 2) each of the three intentional cursor delays takes place, and 3) the user concludes the session.

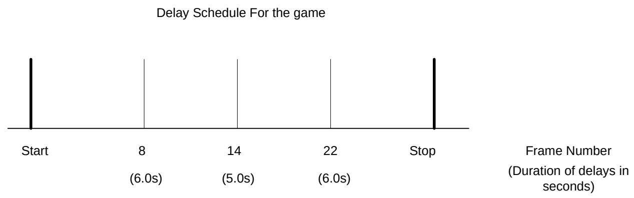

Figure 3: Delay Schedule for the computer game

Instrumentation Hardware

The complete instrumental setup used is illustrated in Figure 4. The stimulus program described above ran in a laptop PC. While solving the visual puzzles, the subject had the GSR and BVP sensors described attached to his/her left hand (Figure 1). Additionally, the eye gaze tracking system described had been calibrated and recorded PD data to a file on its own interface PC, at rate of 60 samples/second.

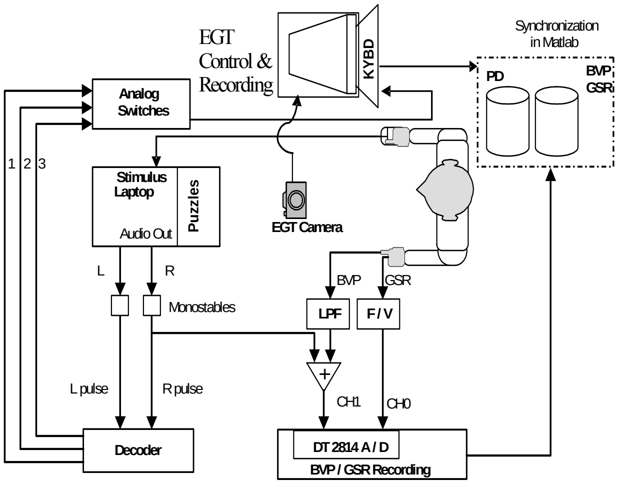

Figure 4: Instrumental setup

Conditioning and A/D Conversion of the BVP and GSR Signals

Both the GSR and BVP signals were converted, after appropriate conditioning described in the following paragraphs, using a multi-channel Data Translation DT2814 Analog-to-Digital Conversion board, as two independent channels, each at a sampling rate of 360 Hz. The acquisition of these signals to disk was initiated manually by the experimenter, a few seconds before the subject would effectively start to solve the visual puzzles, clicking the "start" button on the laptop program, which would send the first timing burst through its audio output. Similarly, the acquisition to disk would continue until it was stopped manually by the operator, after the subject had completed all the visual puzzles and terminated the session by pressing the "stop" button in the laptop program, which issued the corresponding timing burst through its audio output.

As mentioned before, the output of the photoplethysmographic BVP sensor was filtered by a 2^{nd} order low-pass Butterworth active filter, with a corner frequency at 10 Hz. Moreover, the filtered BVP signal was not connected directly to the A/D board input. Instead, it was mixed with the TTL-level signal obtained from the left audio channel of the stimulus laptop, using a non-inverting, unity gain summing amplifier arrangement.
Accordingly, when the user started or concluded the experimental session a large offset pulse would be added to the BVP signal, which allowed the easy identification of these timing marks in the digital record. Since these events (Start = event 1 = 01_b, and Stop = event 3 = 11_b) were also simultaneously recorded (as keystrokes "1" and "3") in the eye gaze and pupil diameter file, captured by the eye gaze tracking system's computer, it is possible, after the fact, to

align both files. The need to fuse both files together is also the reason for our choice of sampling rate for the BVP and GSR signals: 360 samples/sec.

The eye gaze tracking system records the pupil diameter 60 times per second. It should be noted that this "destructive" tagging of the BVP channel, to indicate the start and stop events, so that the BVP-GSR file could be aligned with the PD file, briefly distorts the BVP signal at the very beginning and at the very end of the session. However, these segments of the BVP signal need not be analyzed, anyway.

Pupil Diameter (PD) Recording

To assure reliability of the changes of PD, the lighting of the environment was kept at a constant level during the experiment and the illumination of the system was the same for all the subjects. The values of the pupil diameter of the subjects throughout the experiment were collected at 60 Hz by the ASL-504 eye gaze tracking system. The software for this system allows the extraction of selected variables (in this case the pupil diameter and the marker channel) to a smaller file, which in turn can be read into Matlab ®, where it can be aligned with the BVP and GSR obtained through the A/D board, thanks to their common timing marks for events 1 (start) and 3 (stop). At this point the pupil diameter data can be upsampled (interpolated) by six, to achieve a common sampling rate of 360 Hz, for all three measured signals.

Affect Detection

In this study we have monitored the physiological response of a person to detect that person's level of frustration. Figure 4 shows an example of the three synchronized signal traces, displayed in Matlab® for one of the subjects. The order of the signals is (from top to bottom): BVP, GSR and PD. The three stem marks, with circles at their tops, in each trace, indicate the occurrence of events of type 2 (temporarily unresponsive cursor).

Cardiac response:

HRV is a measure of the oscillation of the interval between consecutive heartbeats. It has a strong potential to determine the role of autonomic nervous system (ANS) fluctuations in the human body. Increased activity of the sympathetic branch causes an increase in the heart rate while an increase in the parasympathetic branch results in a slowing down of the heart rate. Here we use HRV to detect the rapid transient occurrence of a frustration state.

Measuring the intervals between peaks of adjacent BVP pulses, the instantaneous Heart Rate (HR) can be calculated (shown in Fig. 6). The HR is typically 70 – 80 beats per minute. The HRV can be approximated as the standard deviation in a sliding window, normalized by the window mean. Then, defining an interval of suspected frustration around HRV increases that surpass a given threshold, the "BVP suspected frustration intervals" shown in Fig. 6 are found.

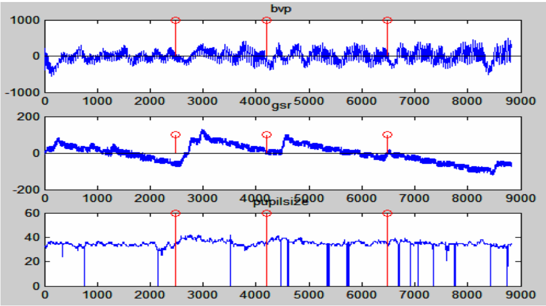

Figure 5: Synchronized BVP, GSR and PD data collected from subject #11.

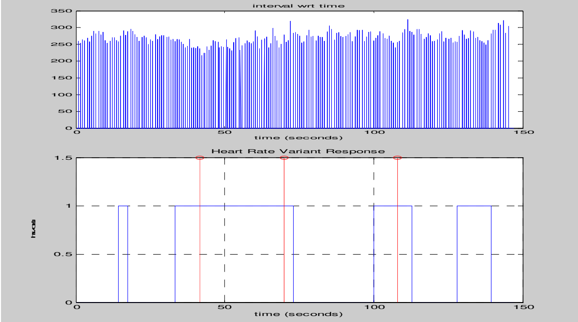

Figure 6: HR from BVP and BVP suspect intervals

Electrodermal response:

Electrodermal response, also called Galvanic Skin Response, is one of the fastest responding measures affected by anxiety or stress. It has been found to be one of the most robust and non-invasive physiological measures of sympathetic activity. We remove the base trend and denoise the GSR signal with a moving median filter. The resulting signal is shown in Fig 7, along with the "GSR suspected frustration intervals" that are defined from it, by thresholding.

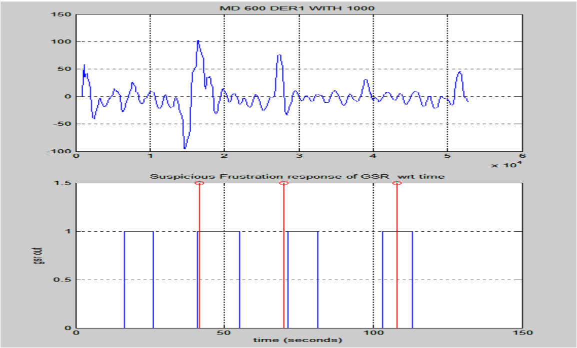

Figure7: Processed GSR and suspect intervals

Pupil diameter response:

In the case of the PD signal, gaps due to blinking have to be filled by interpolation. A threshold is then applied to the amended sequence of PD values, to isolate increases in pupil diameter that may signal an affective shift, perhaps due to frustration. The resulting “PD suspected frustration intervals” are shown in Figure 8.

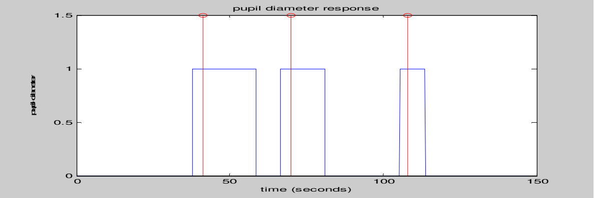

Figure 8: Suspect intervals from PD increase

When these detection signals are combined according to the following logic equation:

$$FR = (GSR \wedge BVP) \vee (GSR \wedge PD) \quad (1)$$

the frustration suspected intervals determined by the overall system (FR) emerge. These are shown in Figure 9.

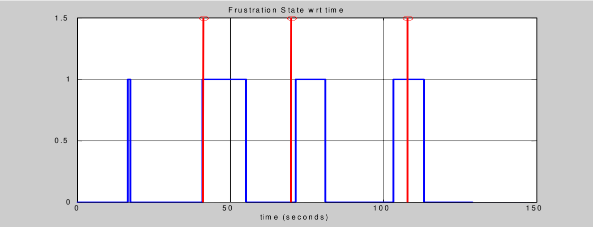

Figure 9: Overall frustration suspect intervals

Conclusion and future work

The results indicating the suspicion of episodes of frustration show a good temporal correlation with the stimulus (mouse unresponsiveness) time markers. Therefore, they have a strong potential of quantifying this kind of relationship. They also support the appropriateness of this instrumental setup for the comprehensive monitoring of the affective state of a computer user.

Other subjects from the pool of 15 that were tested with this instrumentation did not yield results that were as clear cut as the ones shown, with the simple processing used so far. In future work, more powerful approaches will be tried to extract and analyze the useful information effectively. The future directions of this research may include the use of machine learning and fuzzy logic as well as the monitoring of additional physiological signals.

Acknowledgment

This work was sponsored by NSF grants IIS-0308155, HRD-0317692 and CNS-0426125. Ms. Zhai is the recipient of a Florida International University Presidential Fellowship.

References

Cacioppo, J. and Tassinary, L.,1990, "Inferring psychological significance from pshysiological signals," *American Psychologist*, vol. 45, pp. 16-28.

Dawson, M. E., Schell, A. M. and Filion, D. L., "The Electrodermal System," in *Handbook of Psychophysiology*, Cacioppo, J. T., L.G., T. and G.G., B., Eds., 2nd Edition ed. Cambridge,UK: Cambridge University Press, 2000.

Dishman, R. K., Nakamura, Y., Garcia, M. E., Thompson, R. W., Dunn, A. L. and Blair, S. N.,2000, "Heart rate variability, trait anxiety, and perceived stress among physically fit men and women," *International Journal of Psychophysiology*, vol. 37, pp. 121-133.

Giardino, N. D., Lehrer, P. M. and Edelberg, R.,2002, "Comparison of finger plethysmograph to ECG in the measurement of heart rate variability," *Psychophysiology*, vol. 39, pp. 246-253.

Grings, W. W. and Dawson, M. E.,1978, *Emotions and Bodily Responses A psychophysiological Approach*: Academic Press, INC.

Hansen, A. L., Johnsen, B. H. and Thayer, J. F.,2003, "Vagal influence on working memory and attention," *International Journal of Psychophysiology*, vol. 48, pp. 263-274.

Hess, E. H.,1975, *The Tell-Tale Eye*: Litton Educational Publishing, Inc.

Hugdahl, K., 1995, *Psychophysiology The Mind-Body Perspective*: Harvard University Press.

Lawson, R.,1965, *Frustration:The Development of a Scientific Concept.* New York: MacMillan.

Partala, T. and Surakka, V.,2003, "Pupil size variation as an indication of affective processing," *International Journal of Human-Computer Studies*, vol. 59, pp. 185-198.

Picard, R. W.,1997, *Affective Computing*. Cambridge,MA: M.I.T. Press.

Picard, R. W. and Klein, J.,2002, "Computers that recognise and respond to user emotion: theoretical and practical implications," *Interacting with Computers*, vol. 14, pp. 141-169.

Reeves, B. and Nass, C.,1996, *The Media Equation*: Cambridge Univ. Press.

Scheirer, J., Fernandez, R., Klein, J. and Picard, R. W.,2002, "Frustrating the user on purpose: a step toward building an affective computer," *Interacting with Computers*, vol. 14, pp. 93-118.

Wilson, G. F. and Eggemeier, F. T., "Psychophysiological assessment of workload in multitask environments," in *Multiple-Task-Performance*, Damos, D. L., Ed. London: Taylor & Francis, 1991, pp. 329-360.