ELSEVIER

Available online at www.sciencedirect.com
SCIENCE @ DIRECT®

Int. J. Human-Computer Studies 59 (2003) 185–198

International Journal of
Human-Computer
Studies

www.elsevier.com/locate/ijhcs

# Pupil size variation as an indication of affective processing

Timo Partala[a], Veikko Surakka[a,b,*]

[a] Tampere Unit for Computer-Human Interaction, Department of Computer and Information Sciences, University of Tampere, FIN-33014 Tampere, Finland
[b] Tampere University Hospital, Department of Clinical Neurophysiology, P.O. Box 2000, FIN-33521 Tampere, Finland

Received 25 September 2002; received in revised form 20 November 2002; accepted 30 November 2002

## Abstract

The present objective was to investigate pupil size variation during and after auditory emotional stimulation. Thirty subjects' (15 females and 15 males) pupil responses were measured while listening to 10 negative and 10 positive highly arousing sounds (e.g. a baby crying and laughing), and 10 emotionally neutral sounds (e.g. regular office noise). The subjects also rated their subjective experiences related to the stimuli. The results showed that the pupil size was significantly larger during both emotionally negative and positive stimuli than during neutral stimuli. The results for the time period of 2 s following the stimulus offset showed that pupil size was significantly larger after both negative and positive than neutral stimulation. These results suggest that the autonomic nervous system is sensitive to highly arousing emotional stimulation. The subjective ratings confirmed that the stimuli influenced the subjects' emotional experiences as expected. Further analyses showed that female subjects had significantly larger pupil responses than males only to neutral stimuli and only during the auditory stimulation. In sum, our results showed that systematically chosen stimuli significantly affected the subjects' physiological reactions and subjective experiences. It could be possible to use pupil size variation as a computer input signal, for example, in affective computing. Auditory emotion-related cues could also be utilized to modulate the user's emotional reactions.

*Keywords:* Emotions; Computers; Pupil size; Autonomic nervous system; Affective computing

---

*Corresponding author. Tampere Unit for Computer–Human Interaction, Department of Computer and Information Sciences, University of Tampere, FIN-33014 Tampere, Finland. Tel.: +358-3-215-8551, fax: +358-3-215-8557.
*E-mail address:* veikko.surakka@uta.fi (V. Surakka).

1071-5819/03/$ - see front matter © 2003 Elsevier Science Ltd. All rights reserved.
doi:10.1016/S1071-5819(03)00017-X

## 1. Introduction

A natural approach in the search for alternative and better means for human–computer interaction (HCI) is the understanding of centrally important factors in intra- and interindividual human communication and behavior. Concepts such as perceptually intelligent machines, neuroinformatics, affective computing, and emotion- and interest-sensitive media refer to a future generation of flexible, learnable, and emotionally responsive technologies. These technologies utilize, for example, video cameras and microphones for perceiving the users' behaviors. Various types of physiological indicators, such as the correlates of the autonomic nervous system activation (e.g., pupil size[1] variation), the responses of the central nervous system (e.g. electric brain potentials) and the electrical activity of facial muscles can also be utilized in connecting the user with the computer (Allanson et al., 1999; Kübler et al., 1999; Laakso et al., 2001; Partala et al., 2001; Pentland, 2000; Picard, 1997).

The reason for the growing interest in these types of technologies is the fact that currently, it has begun to be more widely accepted that emotional reactions (physiological, experiential, and expressive) are central factors in rational human behavior (Damasio, 1994; Lang, 1995; Lang et al., 1993; LeDoux, 1998). For example, it has been shown that decision-making can be seriously disturbed after brain damage to areas important for emotional processing (Damasio, 1994). Our recent studies have also revealed that automatic change detection from an auditory stimulus stream in the auditory cortex is significantly attenuated while processing emotionally positive visual information (Surakka et al., 1998). In a classic paper Zajonc (1980) showed convincing evidence that emotional processing can have primacy over cognition. Thus, for example, perceiving others' facial emotional expressions can significantly affect the interpretation of non-emotional information (Surakka et al., 1999; Zajonc, 1980). In interindividual behavior emotions and emotional expressions significantly affect the quality of interaction. There is evidence that other peoples' emotional expressions (facial and vocal) can evoke parallel emotional reactions in the observer in terms of electrical facial muscle activity and emotional experiences. In other words, emotions can be contagious (Hietanen et al., 1998; Lundqvist and Dimberg, 1995; Surakka and Hietanen, 1998).

Clearly, emotions are of central importance in human behavior and, thus, very much needed in the development of interaction with technology. Emotion-related factors are important in developing new alternative interaction techniques and in regulating the quality of human–technology interaction. A lot of basic as well as applied research is, however, needed on human emotions. This research extends from studies on the association of human physiology and emotions to the understanding of what kind of emotion-related signals are applicable to HCI.

---

[1] Preliminary results (i.e. pupil size data with no baseline corrections) of this study were presented in Proceedings of ETRA 2000, Symposium on Eye Tracking Research and Applications, ACM Press, New York, pp. 123–129.

The association of emotions and physiological reactions controlled by the autonomic nervous system is one area of research that needs to be studied further. There is some evidence that certain specific emotions, in terms of different facial expressions (e.g. fear, happiness, and anger), associate with differential autonomic nervous system (ANS) activity. These ANS differentiations between emotions are complex, but anger, for example, has been associated with higher heart rate than happiness, and on the other hand, anger has been associated with higher finger temperature than fear (Ekman et al., 1983; Levenson, 1992; Levenson et al., 1990, 1992).

One fairly simple indicator of ANS activity is the variation of the pupil size. It is known that pupil dilations and constrictions are governed by the ANS, for example, to adjust the amount of light entering the eye (Andreassi, 1995). Previous studies have suggested that pupil size variation is also related to both cognitive and affective information processing. For example, it has been shown that pupil dilation is positively associated with increased cognitive load (Hyönä et al., 1995; Kahneman and Beatty, 1966). Most of the existing studies on pupil size and affects are from the 1960s or 1970s. Regardless of considerable development in eye-tracking technology (e.g. ease of use, improved accuracy, and enhanced sampling rate), recent studies have been infrequent. This is surprising considering that pupil size variation has been recognized as a potential computer input signal (Jacob, 1996). If emotions and pupil size variation were reliably associated with each other, then current eye-tracking technology would offer a possibility for unobtrusive monitoring of emotion-related reactions during HCI because no sensors need to be attached to the user. In order to be able to evaluate the possibilities of using pupil size measurement in HCI for detecting the users' emotional responses, we need to understand, how emotions and pupil size variation relate to each other.

There is some evidence that pupil size has a communicative meaning in human–human communication of emotion. For example, it has been shown that pupil size affects males' judgments of females on a negative–positive scale. Females with larger pupils evoked more positive feelings in males (e.g. Hess and Petrovich, 1987). It has also been found that pupil size variation interacts with sex. For example, when subjects were shown facial stimuli of the opposite sex, the bigger the pupil of the stimulus the bigger the pupil of the observing subject (Simms, 1967). Early studies also focused on sexual arousal indicating pupil dilation to sexually interesting pictures (Hess and Petrovich, 1987).

Previous research on the relation of affective processing and pupil size variation has also been somewhat controversial. Hess (1972) suggested that there is a continuum ranging from extreme dilation to interesting or pleasing stimuli to extreme constriction to unpleasant or distasteful material. Janisse (1974) challenged this bi-directional view and argued that there is no pupil constriction to negative stimuli, or the constriction response may be limited to a few individuals and a small range of stimuli. Earlier Loewenfeld (1966) had also studied the effects of various sensory and psychological stimuli to pupil size variation and argued that none of them caused pupil constriction except increased light intensities. Janisse (1974) suggested that pupil size is linearly related to the intensity dimension of the stimuli.

According to him, pupil size variation seems to behave curvilinearly on the valence scale. It is largest at the negative and positive ends of the continuum and smallest at the center, which represents neutral affect.

Part of the controversial results and theories concerning the relation of affects and pupil size variation may be due to the various stimulus materials used. Mostly, they have been limited sets of pictures varying in content. The materials used have also suffered from methodological problems with color, luminance, and contrast (Hess and Petrovich, 1987). Clearly, controlled stimulus sets are a necessary precondition for the systematic study of the effects of emotions to pupil size variation.

Recently, Bradley and Lang (1999) published a set of systematically studied affective sound stimuli called the International Affective Digitized Sounds (IADS). This stimulus set consists of 117 affective sound stimuli. These stimuli have been extensively studied using the Self-Assessment Manikin, which is a method to study differences in affective experiences using nine-point bipolar rating scales (Bradley and Lang, 1994). Their subjects listened to the stimuli and rated them on three bipolar dimensions: emotional valence, arousal, and dominance. However, valence and arousal are the most frequently used dimensions to capture the nature of emotional information. The valence dimension varies from negative to positive emotional experience, and the arousal dimension varies from calm to highly excited. Lang and his co-workers have suggested that the valence dimension reflects the presence of appetitive or aversive motive systems linked to behavioral tendencies of approach and withdrawal, respectively. The arousal dimension reflects the intensity of either appetitive or aversive systems (Bradley and Lang, 2000; Lang, 1995; Lang et al., 1993).

The stimulus set of Bradley and Lang (1999) offers a well-grounded starting point for investigating pupil size responses to emotional stimuli. Because the set consists of auditory stimuli, the problems with visual stimuli can be avoided. The present aim was to investigate pupil size variation in response to auditory stimulation using IADS stimulus materials. We wanted to explore the association of pupil size on three types of emotional sound categories. The categories were negative and positive highly arousing sounds, and emotionally neutral sounds. We also measured subjective experiences to the stimuli. As there is some evidence that males and females may differ in their pupil behavior, we also explored the possible sex differences in pupil size responses.

## 2. Methods

### 2.1. Subjects

Thirty-one voluntary students were recruited from an elementary programming course. Data from one subject had to be discarded because of technical problems. Thus, the results are based on the data collected from 30 subjects (15 females and 15 males, mean age 23.5 years, range 18–40 years). All subjects had normal or corrected-to-normal vision, normal hearing by their own report, and they were not informed of the purpose of the study.

### 2.2. Stimuli

Thirty stimuli, 10 per category were selected from the IADS. The IADS mean rating values of two dimensions, valence and arousal, were used as criteria for stimulus selection (Bradley and Lang, 1999). The categories were negative highly arousing (e.g. a couple fighting), neutral (e.g. background office noise), and positive highly arousing (e.g. a baby laughing) stimuli. Erotic stimuli were not used. The respective means for valence and arousal for the different categories were for negative (2.6 and 6.8), neutral (5.1 and 4.2), and positive (7.4 and 6.6). All the stimuli were about 6 s long.

The stimulus presentation was controlled by the PsyScope program (Cohen et al., 1993) running in a MacIntosh PowerPC 7500/100 computer. The stimuli were delivered via headphones to both of the subject's ears at a comfortable constant volume level. The stimulus presentation was fully randomized.

### 2.3. Procedure

First, the laboratory was introduced to the subject and then she/he was comfortably seated in an adjustable chair in front of a computer screen. The subject was told that involuntary characteristics of her/his eye movements would be measured during auditory information processing. The headphones were put on and an adjustable headrest was used to keep the eyes at a distance of 60 cm from the center of the computer screen.

The eye tracker was calibrated and the subject was instructed to look at a fixation cross at the center of the screen. Twenty seconds from the onset of the fixation point the first stimulus was delivered. Ten seconds after the stimulus offset the subject heard a beep and the fixation point disappeared, which indicated that the subject had an opportunity to move her/his eyes to prevent eye fatigue. Ten seconds later, the subject heard the beep again, and the fixation cross reappeared to signal the end of the resting period. Following this there was a randomized pause of $6 \pm 1$ seconds before the next stimulus was delivered.

The pupil of the subject's left eye was monitored with an Applied Science Laboratories series 4000 eye tracking system running on a PC computer. The distance from the mirrors of the eye tracker to the eye was 85 cm. The sampling rate of the system was 50 Hz. The lighting of the laboratory was kept at a constant level for all subjects.

Finally, the subject rated her/his experiences evoked by the stimuli on two dimensions: valence and arousal (Bradley and Lang, 1999). On both scales, the subject gave a rating on a nine-point (1–9) scale using the computer keyboard. On the valence scale, the lower end represented a very negative emotional experience, and the upper end represented a very positive emotional experience. The center of the scale represented a neutral experience. The subjects also rated their experienced arousal from calm (the lower end) to highly aroused (the upper end). The PsyScope program was used to control the rating phase and to record the responses. In the beginning there were two neutral exercise sounds (different from the actual stimuli) in order to make sure that the subject was comfortable with giving the ratings. The stimuli were presented randomly and the ratings were given using the keyboard. First, the subject heard a stimulus and gave a valence rating. Then the stimulus was repeated and she/he gave an arousal rating. After each rating a beep was delivered to indicate that the rating had been registered. After each rating there was a 1-s pause before the next sound.

### *2.4. Data analysis*

First, for a smoothed visual timeline presentation blinks were removed according to Bernhardt et al. (1996) using the last valid pupil diameter value before each blink as a replacement for the blink.

Second, for the statistical analyses the blinks and artefacts were removed. In addition to blinks sudden brief increases or decreases of at least 0.375 mm within a 20 ms interval were judged as artefacts. The pupil size data were analysed using a 1000 ms prestimulus baseline correction. Then the data were categorized according to the stimulus categories to two sets of data. The first set consisted of baseline corrected data averaged over the whole 6-s auditory stimulus interval. The second set consisted of baseline corrected data averaged over a 2-s time interval after each stimulus offset.

All the data were analysed with repeated measures analyses of variance. Greenhouse-Geisser corrected degrees of freedoms were used in reported significance levels. Pairwise Bonferroni corrected paired *t*-tests were used for post hoc analyses.

### *2.5. Results*

The results are presented so that first a descriptive part presenting the timelines during and 2 s after the stimulus presentation are given. Following this the statistical analyses are presented.

The timeline analysis of the pupil size data revealed a common pattern to all stimulus categories (Fig. 1). First, there was a period of about 400 ms without any pupil dilation, followed by a steep increase in the pupil size. The peaks were reached at about 2–3 s from the stimulus onset. Then a smooth constriction started and continued until the end of the stimuli. Just following the stimulus offset there seemed to be a small dilation, after which the constriction continued until the end of our analysis window. It is worth noting that at about 1 s from the onset of the stimuli, the lines were separated by emotional vs. neutral stimuli. As the responses to the highly arousing positive and negative stimuli provoked roughly the same amount of pupil dilation, the neutral stimuli provoked clearly less pupil dilation than the emotional stimuli.

The timelines of the pupil diameter were somewhat different for female and male subjects (Figs. 2 and 3). The positive stimuli provoked the strongest pupil dilations

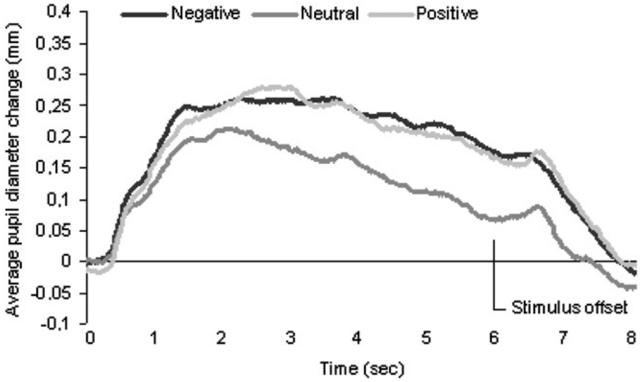

Fig. 1. Averaged pupil diameter timelines for the different stimulus categories.

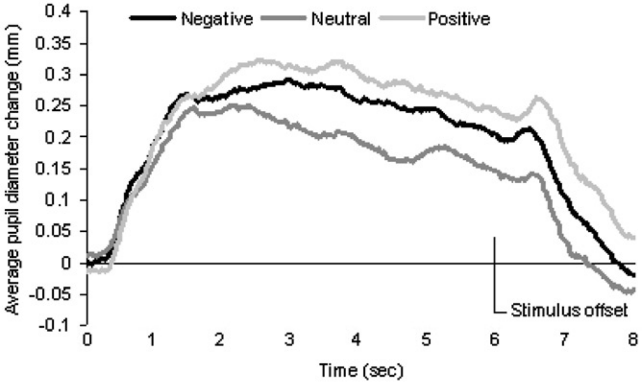

Fig. 2. Averaged pupil diameter timelines for the different stimulus categories for female subjects.

for female subjects, whereas the negative stimuli provoked the strongest dilations for male subjects. However, both female and male subjects showed the smallest dilations to neutral stimuli. At the end of the analysis window both female and male pupil sizes were already near the baseline.

For the data during the stimuli, a $2 \times 3$ (gender × stimulus category) ANOVA showed a significant main effect of gender $F(1, 28) = 6.6$, $p < 0.05$ and a significant main effect of stimulus category $F(2, 56) = 8.6$, $p < 0.01$ on the pupil size. The interaction of the main effects was not significant $F(2, 56) = 1.3$, $p > 0.2$. Independent pairwise comparisons showed that the pupil diameter was significantly larger for female than for male subjects during neutral stimuli $t = 2.9$, $\mathrm{df} = 28$, $p < 0.05$. There were no gender differences during either positive $t = 2.2$, $\mathrm{df} = 28$, $p > 0.1$ or negative $t = 1.2$, $\mathrm{df} = 28$, $p > 0.5$ stimuli.

For the time period following 2 s from the offset of the stimuli, a $2 \times 3$ (gender × stimulus category) ANOVA showed a significant main effect of

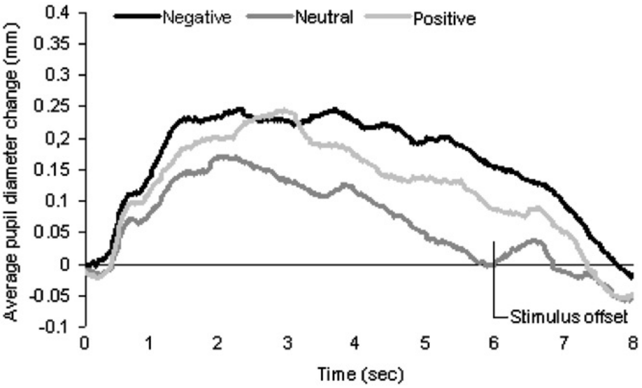

Fig. 3. Averaged pupil diameter timelines for the different stimulus categories for male subjects.

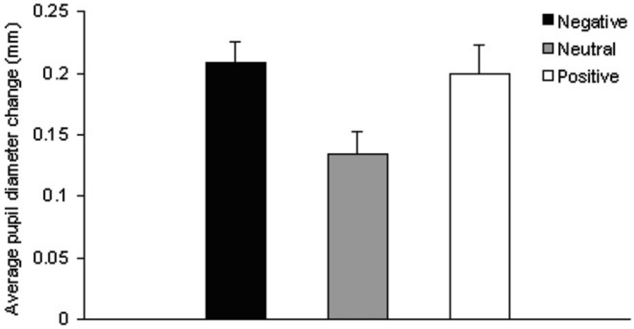

Fig. 4. Averaged pupil diameter deviations from baseline (and S.E.M.) during the different stimulus categories.

stimulus category $F(2, 56) = 4.9$, $p < 0.05$. The main effect of gender $F(1, 28) = 2.6$, $p > 0.1$ and the interaction of the main effects $F(2, 56) = 2.1$, $p > 0.1$ were not significant.

Because pupil size variation was related to gender only when listening to the neutral stimuli we proceeded to analyse the effects of auditory stimulation on a general level (i.e. without the gender). A one-way ANOVA for the data during the different stimuli showed a significant effect of stimulus category on the pupil size $F(2, 58) = 8.5$, $p < 0.01$. Pairwise comparisons showed that the pupil diameter was significantly bigger during the negative than during the neutral stimuli $t = 4.6$, $df = 29$, $p < 0.001$. The pupil diameter was also significantly bigger during positive than neutral stimuli $t = 2.8$, $df = 29$, $p < 0.05$. The difference between negative and positive stimuli was not significant $t = 0.5$, $df = 29$, $p > 0.5$ (see Fig. 4).

A one way ANOVA for the data two seconds from the stimulus offset also showed a significant effect of the stimulus category to the pupil diameter $F(2, 58) = 4.7$, $p < 0.05$. Pairwise comparisons showed that the pupil size was significantly larger

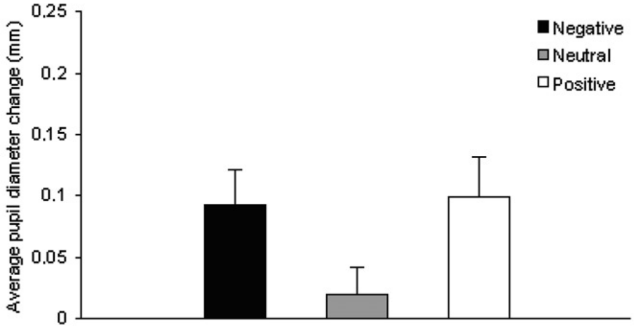

Fig. 5. Averaged pupil diameter deviations from baseline (and S.E.M.) for the period of 2 s following the offset of the different stimulus categories.

after positive stimuli than after neutral stimuli $t = 2.6$, $\mathrm{df} = 29$, $p < 0.05$. The pupil diameter was also significantly larger after negative stimuli than after neutral stimuli $t = 2.7$, $\mathrm{df} = 29$, $p < 0.05$. The difference between positive and negative categories was not significant $t = 0.2$, $\mathrm{df} = 29$, $p > 0.5$ (see Fig. 5).

The overall valence ratings (mean ± S.E.M.) of the different sound categories were as follows: negative $3.8 \pm 0.1$, neutral $4.9 \pm 0.1$, and positive $6.4 \pm 0.1$. The overall arousal ratings (mean ± S.E.M.) of the different sound categories were as follows: negative $6.3 \pm 0.2$, neutral $5.3 \pm 0.1$, and positive $5.9 \pm 0.2$. First, two separate two-way ANOVAs were performed to study the effects of gender on the subjective ratings of valence and arousal. A $2 \times 3$ (gender × ratings of valence) ANOVA showed that the main effect of stimulus category was significant for the ratings of valence $F(2, 56) = 167.3$, $p < 0.001$. The main effect of gender $F(1, 28) = 0.6$, $p > 0.4$ and the interaction of the main effects $F(2, 56) = 0.6$, $p > 0.5$ were not significant. A $2 \times 3$ (gender × ratings of arousal) ANOVA showed that the main effect of stimulus category was significant for the ratings of arousal $F(2, 56) = 18.8$, $p < 0.001$ arousal. The main effect of gender $F(1, 28) = 0.6$, $p > 0.3$ and the interaction of the main effects $F(2, 56) = 0.4$, $p > 0.5$ were not significant.

Because gender had no effect on the ratings of the stimuli one-way ANOVAs were used for further analyses. For valence, a one-way ANOVA showed a significant effect of stimulus category $F(2, 58) = 169.8$, $p < 0.001$. Pairwise comparisons showed that the differences between all stimulus categories were significant: negative vs. neutral $t = 10.0$, $\mathrm{df} = 29$, $p < 0.001$, neutral vs. positive $t = 11.4$, $\mathrm{df} = 29$, $p < 0.001$, negative vs. positive $t = 14.8$, $\mathrm{df} = 29$, $p < 0.001$.

For arousal, a one-way ANOVA showed a significant effect of stimulus category $F(2, 58) = 19.2$, $p < 0.001$. Pairwise comparisons showed that the difference between ratings of negative and neutral categories was significant $t = 8.7$, $\mathrm{df} = 29$, $p < 0.001$. The difference of ratings between neutral and positive categories was also significant $t = 3.7$, $\mathrm{df} = 29$, $p < 0.01$. The difference of ratings between the negative and positive stimulus categories was not significant $t = 1.8$, $\mathrm{df} = 29$, $p > 0.2$.

## 3. Discussion

Our results showed significantly larger pupil size dilation in response to both negative and positive emotional stimuli as compared to the dilation to emotionally neutral auditory stimuli. These significant differences were found both during the stimuli and during the period of 2 s following the offset of the stimuli. Pupil dilation to emotional stimuli was not affected by gender.

The visual inspection of the timeline curves revealed that the averaged pupil dilations coherently followed a similar type of curve. A sudden increase in pupil size at about 400 ms from the stimulus onset was followed by a slower decrease. After the initial increase the subjects' pupil responses were separated from each other at about 1 s from the stimulus onset so that emotional stimuli evoked clearly larger dilation than neutral stimuli. These differences were observable until the end of our analysis window.

Because the size of the pupil is regulated by the autonomic nervous system, the results of this experiment suggest that the autonomic nervous system responds differently to emotionally arousing than to emotionally neutral stimuli. Our results are in line with the recent results that showed higher autonomic activity, in terms of skin conductance responses, to both emotionally positive and negative than to neutral auditory IADS stimuli (Bradley and Lang, 2000). Our results also support those earlier findings that have investigated the association of different emotional responses to differential autonomic nervous system activity. These studies have found, for example, that both during angry and happy facial activity the heart rate speeded up from the baseline. Similarly, finger temperature was higher during both angry and happy facial activity when compared to baseline activity. Furthermore, these studies found significant differences between emotions. The heart accelerated more and the finger temperature was higher during angry than happy facial activity (Ekman et al., 1983; Levenson, 1992; Levenson et al., 1990, 1992). Our results showed that during both positive and negative auditory stimulation the pupil size departed relatively rapidly and steeply from the baseline. However, our data did not show significant pupil size differences between positive and negative auditory stimulation. Although one has to consider the possibility that pupil size may not be sensitive enough to discriminate between either positive or negative emotional responses, there are at least two possible explanations. It is possible that the responses did not differ significantly because we used dimensional stimuli in contrast to those earlier studies that investigated autonomic activity during specific emotions. On the other hand, it is possible that previous studies of discrete emotions may have failed to control for the differences in the amount of arousal generated by the respective emotions. Following this, differences in autonomic responses attributed to particular emotions may have resulted from differences in the amount of arousal generated by the emotions.

In the present experiment, only pupil dilations to different stimuli were found. This finding is in contrast to some earlier findings that indicated pupil size constrictions to negative affective stimuli (Hess, 1972). Thus, our results support the results that have found only pupil dilations to affective stimuli (Janisse, 1974;

Loewenfeld, 1966). As suggested earlier by Janisse (1974), it seemed that for the valence dimension the association with pupil dilation was curvilinear and for the arousal dimension the association was linear.

The ratings of auditory stimuli were in concordance with the earlier findings with IADS stimuli (Bradley and Lang, 1999). In the present study, the ratings of the stimuli showed that for valence dimension all the stimulus categories were experienced as significantly distinct from each other. The ratings of arousal showed that negative and positive stimuli were experienced as equally arousing. Because both the pupil size variations and the arousal ratings of the auditory stimuli did not differ from each other by emotional stimulation, the current results suggest that the magnitude of the pupil response is determined by the amount of emotional arousal.

In the current experiment, we chose to analyse the data by averaging baseline corrected pupil responses during two periods: during the stimuli and during the period of 2 s following the stimulus offset. Although this method missed brief peak responses both during and after the auditory stimulation, by averaging over a longer time period a more stable measure of the user's emotional arousal could be acquired. In respect to HCI applications, it may be important that the analysis of the users' emotional state is based on longer samples of human behavior than single values.

In evaluating the possibilities of using pupil size measurement in human–computer interaction for affective computing one has to be cautious. There are several factors affecting the variation of pupil size. Hess and Petrovich (1987) listed several different sources of pupil size variation, including, for example, the light reflex, different stimulus parameters (e.g. visual and chemical), and information-processing load. In the present experiment, we controlled as many variables as possible. We utilized well-studied auditory stimuli, thus, avoiding the possible problems caused by variations in stimulus luminance while using visual stimuli. The stimuli were of equal length and they were played back at equal volume. Given real world HCI settings controlling factors with such accuracy is not possible. For example, controlling light variation in integrating pupil size measurement with current desktop interfaces is challenging. However, in audio interfaces the monitoring of pupil size variation seems promising. It can also become viable in head mounted displays that cover the eyes from external light and that can incorporate pupil size tracking.

Compared to the measurements of other physiological signals the measurement of pupil size has important advantages. It is an unobtrusive method because no sensors have to be necessarily attached to the user. The use of the technology needed for the measurement of pupil size is relatively simple. Pupil size measurement also has potential over other physiological measures because it may not be as sensitive to artefacts caused by various bodily movements (e.g. in a mobile environment). An important advantage of pupil size measurement is that pupil size variation is an involuntary index of ANS activity. This means that pupil size variation is not easy to control voluntarily and, thus, it indexes real spontaneous activity. It is known from facial expression research that visually observable changes in facial emotional behavior can be and are masked, inhibited, exaggerated, and faked (e.g. Ekman, 1985; Ekman and Friesen, 1982; Surakka and Hietanen, 1998). In this respect, pupil size measurement can also avoid problems that are inherently related, for example, to the analysis of spontaneous vs. acted facial expression monitoring with video cameras. We note, however, that as our data indicates pupil size variation by dimensional stimulation and may not be sensitive to specific emotional reactions (e.g. happiness, sadness, fear, etc.) the use of multiple physiological indices for the analysis of users' emotional states would be an optimal approach. Facial EMG has been, for example, shown to be sensitive to specific emotions (Surakka and Hietanen, 1998).

In sum, the present study gives some promise for the use of pupil size variation in the context of HCI. It might be possible to develop signal analysis methods for automatically detecting significant deviations of pupil size from baseline activity. The success of pupil size variation as an input signal will depend on the effectiveness of associated digital signal processing methods. For example, it has been possible to develop methods to detect successfully emotion related voluntarily produced changes in electrical activity of facial muscles (Laakso et al., 2001, 2002). These studies show that it is possible to extend the ways computers can analyse the users' behavior. In respect to pupil size measurement it might even be possible to distinguish between different kinds of reactions (e.g. negative vs. positive). All in all, we believe that meaningful changes in pupil size can be automatically recognized provided that constant baseline monitoring is applied. This would extend the perceptually intelligent abilities of a computer to perceive and analyse the user's reactions (e.g. Pentland, 2000; Prendinger and Ishizuka, 2001).

In addition to the development of perceptual intelligence with digital signal processing methods for emotion recognition, the present results suggest that by emotional stimulation it is possible to regulate the users' autonomic activity. Clearly, it is important to develop methods to extend the abilities of expressive emotional intelligence of the computer, too. There is evidence that positive emotions can be beneficial both in human behavior and in HCI. In human behavior positive emotions have been argued to increase creativity, help in creating more rich associations for memorized material, and make decision-making more efficient and thorough (Isen, 1993; Zhou, 1998). We have found that when compared to negative auditory feedback, positive auditory feedback given with a speech synthesizer (i.e. brief spoken emotional sentences with negative, neutral, and positive contents) during HCI significantly facilitates cognitive performance and regulates autonomic arousal (Aula and Surakka, 2002). As the current experiment also used auditory stimulation these findings support each other and point out that auditory emotional stimulation can be used for meaningful emotion-cognition regulation in HCI. Well-studied auditory affective feedback can have significant influences on the user, just like the carefully selected stimuli had in this experiment.

## Acknowledgements

The authors would like to thank all the voluntary test subjects. We also want to thank the comments and suggestions of two anonymous reviewers. This research was supported by the Academy of Finland (Project 167491), Tampere Graduate School in Information Science and Engineering, The Information Society Institute of the eTampere programme (Project number 56-01) and University of Tampere.

## References

Allanson, J., Rodden, T., Mariani, J., 1999. A toolkit for exploring electro-physiological human–computer interaction. In: Sasse, M.A., Johnson, C. (Eds.), Proceedings of the INTERACT '99, Seventh IFIP TC.13 Conference on Human-Computer Interaction. IOS Press, Amsterdam, pp. 231–237.

Andreassi, J.L., 1995. Psychophysiology: human behavior and physiological response. Erlbaum, Hillsdale, NJ.

Aula, A., Surakka, V., 2002. Auditory emotional feedback facilitates human–computer interaction. Proceedings of the HCI 2002. Springer, London, pp. 337–349.

Bernhardt, P.C., Dabbs, J.M., Riad, J.K., 1996. Pupillometry system for use in social psychology. Behavior Research Methods, Instruments, and Computers 28, 61–66.

Bradley, M.M., Lang, P.J., 1994. Measuring emotion: the self-assessment manikin and the semantic differential. Journal of Behavioral Therapy and Experimental Psychiatry 25, 49–59.

Bradley, M.M., Lang, P.J., 1999. International affective digitized sounds (IADS): stimuli, instruction manual and affective ratings. Technical Report B-2, University of Florida, The Center for Research in Psychophysiology, FL.

Bradley, M.M., Lang, P.J., 2000. Affective reactions to acoustic stimuli. Psychophysiology 37, 204–215.

Cohen, J.D., MaCwhinney, B., Flatt, M., Provost, J., 1993. Psyscope: a new graphic interactive environment for designing psychology experiments. Behavior Research Methods, Instruments, and Computers 25, 257–271.

Damasio, A., 1994. Descartes' Error: Emotion, Reason, and the Human Brain. Putnam, New York, NY.

Ekman, P., 1985. Telling Lies. W. W. Norton, New York.

Ekman, P., Friesen, W.V., 1982. Felt, false, and miserable smiles. Journal of Nonverbal Behavior 6, 238–252.

Ekman, P., Levenson, R.W., Friesen, W.V., 1983. Autonomic nervous system activity distinguishes among emotions. Science 221, 1208–1210.

Hess, E.H., 1972. Pupillometrics. In: Greenfield, N.S., Sternbach, R.A. (Eds.), Handbook of Psychophysiology. Holt, Richard & Winston, New York, NY, pp. 491–531.

Hess, E.H., Petrovich, S.B., 1987. Pupillary behavior in communication. In: Siegman, A.W., Feldstein, S. (Eds.), Nonverbal Behavior and Communication. Erlbaum, Hillsdale, NJ, pp. 327–348.

Hietanen, J.K., Surakka, V., Linnankoski, I., 1998. Facial electromyographic responses to vocal affect expressions. Psychophysiology 35, 530–536.

Hyönä, J., Tommola, J., Alaja, A.-M., 1995. Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks. The Quarterly Journal of Experimental Psychology 48A, 598–612.

Isen, A.M., 1993. Positive affect and decision making. In: Lewis, M., Haviland, J.M. (Eds.), Handbook of Emotions. The Guilford Press, New York, pp. 261–277.

Jacob, R.J.K., 1996. The future of input devices. ACM Computing Surveys 28(4). Available from: http://www.acm.org/pubs/citations/journals/surveys/1996-28-4es/a138-jacob/.

Janisse, M.P., 1974. Pupil size, affect and exposure frequency. Social Behavior and Personality 2, 125–146.

Kahneman, D., Beatty, J., 1966. Pupillary diameter and load on memory. Science 154, 1583–1585.

Kübler, A., Kotchoubey, B., Hinterberger, T., Ghanayim, N., Perelmouter, J., Schauer, M., Fritsch, C., Taub, E., Birbaumer, N., 1999. The thought translation device: a neurophysiological approach to communication in total paralysis. Experimental Brain Research 124, 223–232.

Laakso, J., Juhola, M., Surakka, V., Aula, A., Partala, T., 2001. Detection of electromyographic events with neural networks. In: Patel, V., et al. (Eds.), Proceedings of MEDINFO, 2001. IOS Press, Amsterdam, pp. 489–492.

Laakso, J., Juhola, M., Surakka, V., 2002. Neural network recognition of electromyographic signals of two facial muscle sites. Proceedings of the Medical Informatics Europe. IOS Press, Amsterdam, 2002, pp. 83–87.

Lang, P.J., 1995. The emotion probe: studies of motivation and attention. American Psychologist 50, 372–385.

Lang, P.J., Greenwald, M.K., Bradley, M.M., Hamm, A.O., 1993. Looking at pictures: affective, facial, visceral, and behavioral reactions. Psychophysiology 30, 261–273.

Ledoux, J., 1998. The Emotional Brain. Simon & Schuster, New York, NY.

Levenson, R.W., 1992. Autonomic nervous system differences among emotions. Psychological Science 3, 23–27.

Levenson, R.W., Ekman, P., Friesen, W.V., 1990. Voluntary facial action generates emotion-specific autonomic nervous system activity. Psychophysiology 27, 363–384.

Levenson, R.W., Ekman, P., Heider, K., Friesen, W.V., 1992. Emotion and autonomic nervous system activity in the Minangkabau of West Sumatra. Journal of Personality and Social Psychology 62, 972–988.

Loewenfeld, I.E., 1966. Pupil size. Survey of Opthalmology 11, 291–294.

Lundqvist, L.-O., Dimberg, U., 1995. Facial expressions are contagious. Journal of Psychophysiology 9, 203–211.

Partala, T., Aula, A., Surakka, V., 2001. Combined voluntary gaze direction and facial muscle activity as a new pointing technique. Proceedings of the INTERACT 2001, Eight IFIP TC.13 Conference on Human–Computer Interaction. IOS Press, Amsterdam, pp. 100–107.

Pentland, A., 2000. Perceptual intelligence. Communications of the ACM 43, 35–44.

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

Prendinger, H., Ishizuka, M., 2001. Simulating affective communication with animated agents. Proceedings of the INTERACT 2001, Eight IFIP TC.13 Conference on Human-Computer Interaction. Amsterdam: IOS Press, pp. 182–189.

Simms, T.M., 1967. Pupillary response of male and female subjects to pupillary differences in male and female picture stimuli. Perception and Psychophysics 6, 115–121.

Surakka, V., Hietanen, J.K., 1998. Facial and emotional reactions to duchenne and non-duchenne smiles. International Journal of Psychophysiology 29, 23–33.

Surakka, V., Sams, M., Hietanen, J.K., 1999. Modulation of neutral face evaluation by laterally presented emotional expressions. Perceptual and Motor Skills 88, 595–606.

Surakka, V., Tenhunen-Eskelinen, M., Hietanen, J.K., Sams, M., 1998. Modulation of human auditory information processing by visual emotional stimuli. Cognitive Brain Research 7, 159–163.

Zajonc, R.B., 1980. Feeling and thinking: preferences need no inferences. American Psychologist 35, 151–175.

Zhou, J., 1998. Feedback valence, feedback style, task autonomy, and achievement orientation: interactive effects on creative performance. Journal of Applied Psychology 83, 261–276.