Toni Vanhala

Towards Computer-Assisted Regulation of Emotions

ACADEMIC DISSERTATION

To be presented with the permission of the School of Information Sciences of the University of Tampere, for public discussion in Pinni auditorium B1097 on December 9^{th}, 2011, at noon.

School of Information Sciences
University of Tampere

Dissertations in Interactive Technology, Number 13
TAMPERE 2011

ACADEMIC DISSERTATION IN INTERACTIVE TECHNOLOGY

Supervisor:	Professor Veikko Surakka, Ph.D. School of Information Sciences University of Tampere Finland
Opponent:	Professor Manfred Thüring, Ph.D. Department of Cognitive Psychology and Cognitive Ergonomics Technical University of Berlin Germany
Reviewers:	Professor Gary Bente, Ph.D. Department of Social Psychology and Media Psychology University of Cologne Germany Professor Tapio Takala, Ph.D. Department of Media Technology Aalto University Finland

Dissertations in Interactive Technology, Number 13

School of Information Sciences
FIN-33014 University of Tampere
FINLAND

ISBN 978-951-44-8612-8
ISSN 1795-9489

ISBN 978-951-44-8639-5 (pdf)
ISSN 1456-954X
http://acta.uta.fi

Tampereen yliopistopaino Oy
Tampere 2011

“Computer, tell me why I am
Not content or in love with a simple feeling.
I ask for I have lost my sense
Of peace and you are extraordinarily clever.”
To My Boy: Tell Me, Computer.
Messages (2007), XL Recordings.

Abstract

Emotions are an important part of all motivated human behaviour. They are intimately connected with our memory, thinking, mental and physical health, and they have an important role in coordinating our interactions with other people. Consequently, the level of competence in regulating the content and intensity of emotions can have effects far and wide, ranging from one's daily well-being to general success in life. Similar to the deep-seated meaning of emotions for our daily functioning, technology has recently become a pervasive and unavoidable aspect of our modern lives. We interact daily with some kind of technology. These ubiquitous interactions could provide an opportunity to affect and regulate emotions virtually anywhere and at anytime.

The aim of the present thesis was to create a theoretical and empirical basis for constructing systems for computer-assisted emotion regulation. First, a theoretical framework for studying and developing such systems was defined. This framework identified artificial perceptual and expressive intelligence as essential capabilities for technology that aims to support the regulation of emotions. Then, the practicality of developing and more widely applying perceptual and expressive technology was studied using constructive and empirical methods. For this purpose, an unobtrusive method for perceptual intelligence was developed and experimentally validated by using a special office chair to measure body movement responses to artificial social and emotional cues. Feasible tools for expressive intelligence were developed in experiments investigating the experiential and physiological effects of virtual bodily distance (i.e., proximity) and facial expressions of humanlike computer characters. Finally, a platform was constructed for studying the effectiveness of voluntary facial activations in regulating emotion related experiences and physiological processes during human-technology interaction.

The empirical results suggest that perceptual and expressive intelligence can provide practical methods for regulating human emotional responding. The developed unobtrusive method for body movement analysis was able to detect significant differences between body movement responses to virtual stimulation with unpleasant, neutral, or pleasant emotional content. Virtual proximity and facial expressions of humanlike computer characters were found to significantly affect human body movement, physiology, and subjective experiences of emotion and attention. Voluntary facial activations were found to significantly regulate subjective experiences of pleasantness, arousal, and dominance as well as physiological activity during virtual social communication with humanlike characters. Further, the constructed platform was found to be usable with relatively little training and preparation, which suggests that it could be a practical starting point for continuing the study of computer-assisted regulation of emotions. In sum, based on the empirical results, the theoretical and constructive work of the present thesis is a solid basis for developing the first generation of systems for computer-assisted emotion regulation.

Acknowledgments

Emotions are important. Without them, life would be dull and bland. Fortunately, my road towards the Ph.D. was filled with lots of emotions.

I sincerely and greatly thank my supervisor, Professor Veikko Surakka. He was very generous with his time and advice about scientific, academic, and more mundane matters of life. He truly cares about initiating his advisees to academia. I trust that many more students will see this as he heads the UCIT graduate school.

Other professors in Interactive Technology – especially Prof. Kari-Jouko Räihä, Prof. Roope Raisamo, and Prof. Poika Isokoski have greatly helped this thesis by providing high-quality teaching and administration at the TAUCHI research unit. The UCIT graduate school, with its the former head Prof. Räihä and ex-coordinator Prof. Isokoski, provided many resources for this work. They did a great job in arranging relevant, useful, but also fun activities for the graduate students.

From my former colleagues at TAUCHI, the present coordinator of UCIT, M.Sc. Jenni Anttonen deserves separate thanks for presenting our work at NordiCHI 2008. Kudos to Harri Siirtola for the LaTeX template. My former research group at TAUCHI was a great place to work. Thanks to all of you nice people and other friends at the former department of Computer Sciences. At the university cafeteria, Dr. Järvelin, Dr. Talvensaari, M.Sc. Saarinen, and Mr. Anttalainen kept my cup full of coffee and my hovercraft full of eels.

I thank the reviewers Prof. Bente and Prof. Takala for their effort.

My parents Suvi and Esa and my brother Tatu encouraged me to learn about computers and the human condition. Loving thanks to Katriina, Maija, and Aliisa for taking part in this human experiment. You were a great help in regulating my emotions towards the positive.

Tampere, 23rd of October, 2011
Toni Vanhala

Contents

List of Publications

This thesis consists of a summary and the following publications, which are reproduced here by permission of their publishers. The papers are referred to in the text by the correspoding Roman numerals.

Author's Research Contributions

Each publication included in this thesis was a joined effort of multiple contributors. Although the research involved collaboration between the authors at every stage, the personal contributions of the present author (Vanhala) can be briefly identified as follows.

Publication I was first drafted by Surakka, then significantly updated by Vanhala, and finally revised in collaboration between Surakka and Vanhala. For Publications II–VI, Vanhala was a central contributor for the design of the experiments, the implementation of the hardware and software setups, the conduction of the empirical work, the analysis of collected data, and writing the publications. Vanhala's contributions to the empirical work included working as the experimenter (i.e., performing the laboratory work) in each study as well as the development of original research software for the implementation of the experimental setups and for conducting data analyses. Vanhala wrote the first draft of each publication.

1 Introduction

Imagine that you are driving to home after a hard day at work. You had a heated argument with a colleague and can't help but to still brood angry thoughts. While you are driving, you notice that the in-car navigation system suggests an alternative route to home, which is slightly longer than your usual route but has much more pleasant scenery. You decide to break your habit and enjoy a new road for a change. Approaching home, you realize that you need to stop at a convenience store to pick up some essentials. While you are standing at the queue to the cashier, someone cuts in line to the front of you. Your heart starts to pound and you feel tense. Just when you are about to shout to the offender, you notice soothing music playing in the background. It appears that your favourite song is playing on the hands-free headset that you are wearing. You manage to calm down and politely point out that you were first in line. The other customer seems to be truly sorry and lets you pass without a conflict that could have ensued if you had remained agitated.

The above example shows how changing the course of our emotions is a natural part of our everyday lives. For example, unpleasant events do not necessarily lead to angry outbursts, if we are able to focus on more positive aspects of a situation. Otherwise, it would be difficult to move beyond impulses (e.g., starting an angry argument) and towards more useful purposes (e.g., finishing shopping and getting home). Thus, the regulation of emotions has a fundamental role in leading a normal life. Recently also scientific interest for studying emotion regulation has grown and the topic can now be considered an established field of study with conferences, special issues in scientific journals, and books dealing with the topic (Tamir, 2011). However, the work in this field has been split into several approaches depending on, for example, different views about the functions of emotion regulation.

Regulation of emotions is indeed a multi-faceted task. First, it is not always clear what the best emotional state would be at a particular moment. Maximizing positive emotions might seem like an attractive option at first hand, but negative emotions also serve important functions in life (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001). For example, feeling bad about immoral behaviour motivates one to be a better person in the future. Thus, simple rules like maximizing positive and minimizing negative emotion would oversimplify the role of emotion in human behaviour.

Second, the extent to which the regulation of emotions should be supported may depend on context. In some cases it would rather be preferable to train the person's own competence in emotion regulation instead of taking over most of the responsibility. For example, increasing the competence to handle emotions is a developmental task for children who need to learn how to deal with anger elicited by disappointments. Third, emotions as such are multi-component phenomena that consist of changes in subjective experience (i.e., feelings), physiology, and behaviour (Mauss, Levenson, McCarter, Wilhelm, & Gross, 2005).

The multi-faceted nature of emotion both poses challenges and offers possibilities for regulating emotions. For example, focusing on suppressing behaviour that expresses the emotion (e.g., cursing and frowning) may not be sufficient alone in modifying the core emotion if feelings of anger and physiological responses are not similarly controlled. Quite the contrary, there is evidence that suppressing negative expressions may even lead to heightened physiological responsiveness (Gross & Levenson, 1997). On the other hand, the multi-component nature of emotions can also facilitate regulation by providing multiple channels for expressing, recognizing, and influencing emotions.

Sometimes one's own efforts and competences will not be sufficient for affecting emotions as intended. In these cases, external support could be helpful. Now, as technology pervades more and more of our modern environment, technological support for emotion regulation could be made available virtually all the time and everywhere. For example, the performance and connectivity of a modern mobile phone could already provide a feasible platform for quite sophisticated applications.

On the other hand, as we interact with various technological devices throughout the day, the emotions evoked by these interactions are also a significant part of our lives. For example, one study of student and workplace computer users found that computers regularly elicit high levels of frustration (Lazar, Jones, Hackley, & Shneiderman, 2006). In another survey, several participants reported that they frequently have violent thoughts and even cause damage to their computers: one respondent wrote "I often show my PC the middle finger!", while another reported having "[p]unched the computer hard enough to leave a dent..." (*Computer Rage: Reported Acts of Rage Against Computers*, 2005). Even if these are extreme examples, it is clear that even milder emotions can significantly affect the perceived quality of these daily interactions with technology (Hassenzahl & Tractinsky, 2006). So far, these emotional responses have been largely ignored by technology, but they could be regulated in order to facilitate the quality and effectiveness of human-technology interaction.

In the opening example of the way home from work, the technological assistance for regulating emotional responses blended in to everyday activities in the form of an emotionally intelligent car navigation system and a music player that sensed and guided emotions. The car navigation system was able to detect the negative mood of the driver and suggest a more pleasant route to home. The intelligent music player was able to sense when anger was elicited and automatically select a song to counter the emotion. The exact technology that was used in detecting the emotional state of the person was not specified and could be hidden from plain sight. Perhaps the hands-free head set worn by the person was equipped with physiological sensors that could detect response patterns associated with the negative emotions. In both cases, the technology in the example took its own initiative by suggesting alternative driving routes and choosing songs to play without explicitly commanded to do so.

Recent technological advances in intelligent sensing of the person and the environment suggest that such implicit interaction (i.e., without explicit commands) can become mainstream sooner than many would expect. For example, a smart system installed at AT&T Laboratories Cambridge moved the personal computer desktop automatically from one monitor to another when one moved away from the first one and went near to the second one (Addlesee et al., 2001). In this case, the smart space removed the need for logging out from one computer and logging in to the second one. A similar system could be made to support the regulation of emotions, for example, by projecting the photographs of loved ones on the work desk beside the monitor.

In addition to sensing the physical environment, another way in which technology could provide new tools for emotion regulation is through physiological computing (Allanson & Fairclough, 2004). This approach can be seen as a continuation of the long tradition in psychophysiological research which is a way of looking at functions of the mind by studying functions of the body. In psychophysiology, physiological signals that reflect bodily processes are measured and interpreted as reflections of psychological processes. Such measurements can be quite intimate in many ways, for example, they reflect internal processes that are normally not available to other people, and they can be acquired covertly and recorded continuously by wearing sensors underneath one's clothes. This way, technology could offer new tools for supporting the regulation of core emotional responses that otherwise would be difficult to observe.

The above examples illustrate the potential that technology has for supporting emotion regulation. The potential advantages of technology include the possibility for extensive, detailed, and tireless monitoring of emotion related processes throughout the day, potentially rapid adaptation to detected changes in order to actively guide emotional responding, and the resulting implicit mode of interaction that could allow emotion regulation to be facilitated without explicit effort from the person her- or himself. However, although research on emotion regulation is no longer entirely fresh, a general consensus about the basic principles of the field has not yet emerged (Tamir, 2011). Further, there has so far been very little work that specifically considers emotion regulation in the context of technology. Most of such work has focused on reducing the intensity of excessive emotional responses (e.g., symptoms of phobia), while little attention has been paid to more common everyday scenarios (Marks, Cavanagh, & Gega, 2007, provide clinical examples). Thus, there is a lack of a well-structured basis for building the first generation of systems that would specifically target the regulation of emotions during daily activities.

The aim of the present thesis was to provide a theoretical and empirical basis for the study of computer-assisted emotion regulation. First, a theoretical framework for unifying the existing research and directing future work was defined. This work was divided into two subgoals: establishing a practical approach for integrating emotions into human-technology interaction in Publication I and providing a basis for applying this approach in computer-assisted emotion regulation in Publication II. Then, the validity and practicality of implementing systems with the proposed framework was established with constructive and empirical research in Publications III and IV. Finally, a platform to be used as a starting point for more detailed studies of computer-assisted regulation of emotions was created and experimentally validated in Publications V and VI.

2 Human Emotion

2.1 The Structure of Emotion

2.1.1 Discrete Emotions

One influential branch of theories views emotions as discrete states, such as, happiness, anger, disgust, fear, and sadness (Ekman, 1992; Johnson-Laird & Oatley, 1989; Panksepp, 1998; Tomkins, 1962). In their simplest form, these approaches list prototypical examples of emotion without defining their exact meaning, or criteria (i.e., sufficient or necessary conditions) for something to be considered an emotion (Gross & Thompson, 2007; Oatley & Jenkins, 1996). In some cases, these kinds of emotion lists are made to meet the demands of a particular field of study or application, for example, to infer periods of frustration, curiosity, and boredom that occur while learning (e.g., D'Mello, Graesser, & Picard, 2007). This will inevitably lead to the use of diverse sets of emotions by different researchers as well as in common language.

For example, after asking 200 undergraduate students to write down examples of emotion and merging syntactic variants of the same emotion (e.g., happy, happiness, happily), Fehr and Russell (1984) were left with a list of 383 different emotions. Such loosely described prototypical examples of emotion avoid the problem of defining the exact nature of emotions, but may lead to further problems when trying to study emotions using empirical methods. For example, it is not clear how much different emotions (e.g., sadness and grief) have in common or if they are in fact instances of the same emotion.

In more structured approaches to defining a discrete set of emotions, typically hierarchies of emotion have been formulated (e.g., Ekman, 1992; Parrott, 2001; Plutchik, 1991). It follows from this approach that some emotions can be considered to be more fundamental than others, while others are mixtures or subcategories of the more basic ones. Ekman (1992) suggested nine characteristics that allow a set of more basic emotions to be recognized and separated from each other as well as other affective phenomena (e.g., moods and more complex emotions). In addition to studying which emotions have clearly separate expressions (that other people can also perceive as being different from each other), Ekman's (1992) criteria require further evidence from cross-cultural studies, human physiology, and animal behaviour to state that an emotion is in fact basic by nature.

Based on these criteria, Ekman (1992, 2004) has suggested that there is strong evidence for a limited set of basic emotions, which may include at least anger, fear, disgust, contempt, surprise, sadness, and enjoyment. For example, although facial cues may express over 60 different variations of anger, the common pattern of facial activity in angry expressions, consisting of furrowing and lowering of the brow, raising of the upper eye lid, and tightening of the muscle around the mouth, is remarkably consistent between different cultures (Ekman, 1979, 1992). Further, human and animal expressions of certain emotions share similarities. These similarities were noted as evidence for the evolutionary basis of our emotions (and species) by Darwin and published in his book *The expression of the emotions in man and animals* in (1872) (see Figure 2.1). Similarities of some human and animal emotion would indicate a strong genetic basis for such emotions that can be traced back to the past evolution of our species (Schmidt & Cohn, 2001).

There is also evidence that each basic emotion (e.g., anger) is accompanied by certain physiological changes that separates it from other basic emotions (e.g., fear, sadness, and disgust) (Levenson, Ekman, & Friesen, 1990). Further, such specific facial expressions and physiological patterns for basic emotions have been found when studying members of both Western and non-Western cultures (Ekman, 2004; Levenson, Ekman, Heider, & Friesen, 1992). Taken together, it is hard to deny that these different forms of evidence do make a compelling argument for the existence of a few discrete and universal emotions (or related families of emotion, as discussed by Ekman, 1992).

However, some researchers have challenged the labeling of these patterns of behaviour and physiology as basic, arguing that the so-called basic emotions may be the result of learning, that is, not fixed and genetically pre-determined at birth (e.g., Posner, Russell, & Peterson, 2005). In this view, discrete basic emotions would rather be the result of cognitive

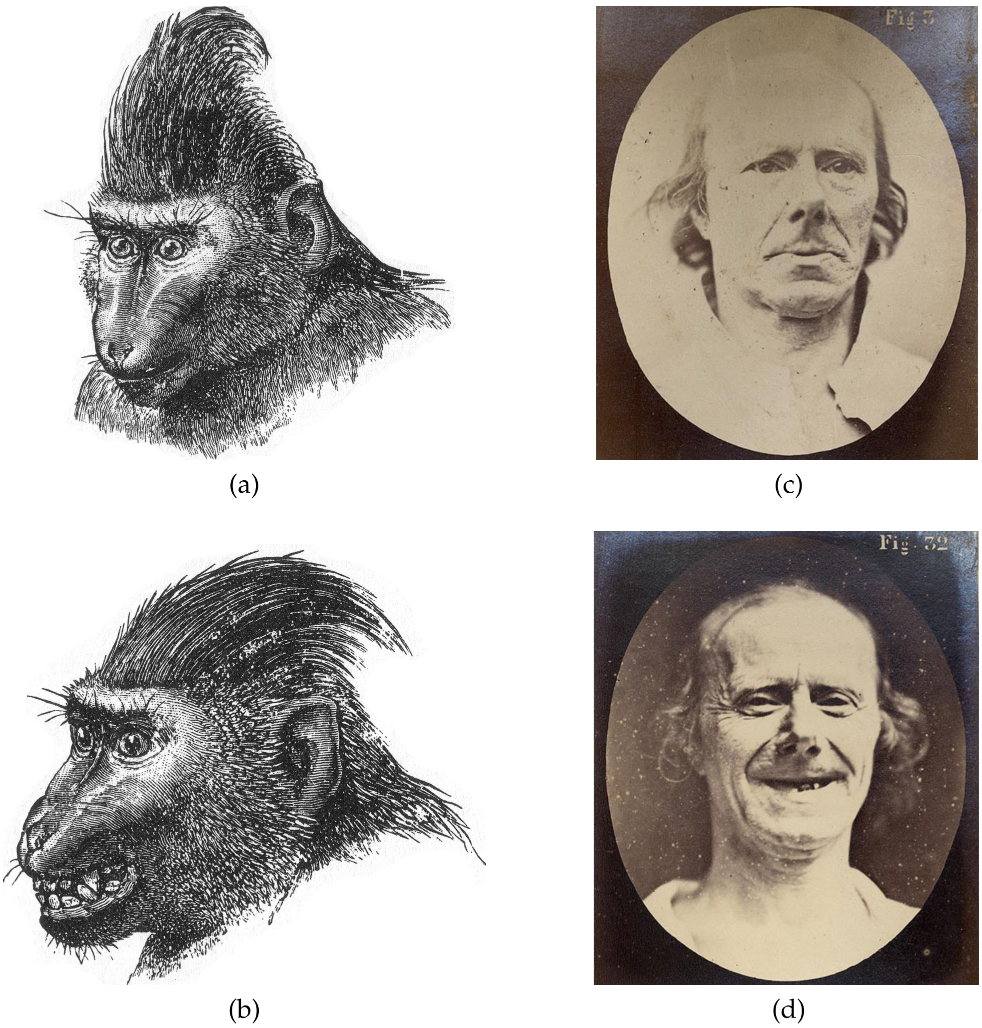


Figure 2.1: Bare-teeth smile of (b) a Celebes Crested Macaque monkey (*macaca nigra*) and (d) a male person. Examples from Darwin (1872) who described the monkey as (a) in a "placid condition", (b) "pleased by being caressed", and the man being (c) in "passive condition", and (d) "naturally smiling". (a & b) drawn by Mr. Wolf and reprinted from Darwin (1872). (c & d) reprinted from Duchenne (1876). Public domain.

constructions that have been later formed to interpret and extend more basic emotional processes (e.g., experience of hedonic valence: pleasure or displeasure). These two views (i.e., "basic" emotions and "learned, constructed" emotions) can also be viewed as compatible and complementary.

For example, Ekman (1992) has suggested that: "To identify separate discrete emotions does not necessarily require that one also take an evolutionary view of emotions. . . . [O]ne can attribute universals to . . . social learning which will usually occur for all members of the species regardless of culture ...". In other words, humans may share a set of basic emotions due to, for example, common life challenges that require us all to learn the same emotional responses.

Taking the argument further, even without a strong universal basis, a set of discrete emotions could still provide a meaningful unit of study and a basis for creating practical applications (e.g., to build computers that respond to emotions). This kind of a practical approach might cause some concerns like how far the research can be generalized (e.g., between cultures), but would allow more advanced applications to be investigated (e.g., automatic recognition of emotions while driving a car; Nasoz, Lisetti, & Vasilakos, 2010).

2.1.2 Dimensions of Emotion

The dimensional theory views emotions as multi-component phenomena (Russell & Mehrabian, 1977; Schlosberg, 1954; Wundt, 1896). In this view, emotional responses can be organized using a combination of certain dimensions, such as, emotional valence (i.e., from unpleasant to pleasant), arousal (i.e., from calm to excited), and dominance (i.e., from being in control of to being controlled by a situation). Several different emotions can be equal on one dimension (e.g., highly arousing), while being completely different on another (e.g., negatively arousing versus positively arousing). In other words, this framework uses a certain point in a multidimensional space of emotions to represent an individual state of emotion.

The basis for the contemporary research on dimensional emotions was laid by Osgood (1952) who found using factor analysis that several types of verbal emotion assessments could be structured according to emotional valence, arousal, and dominance. However, a similar structure of emotion, including dimensions of valence and arousal, had been inferred using introspective methods (i.e., through reflection of own feelings and thoughts) by Wilhelm Wundt almost hundred years earlier (Wundt, 1896). More recently, further evidence for the dimensional structure of emotional responses has been acquired from neurophysiological and psychophysiological experiments (Bradley, 2000; Bradley, Codispoti, Cuthbert, & Lang, 2001; Colibazzi et al., 2010). In general, the current state of research suggests that a major part of emotional responses can be described using the dimensions of valence and arousal, while the dominance dimension has a less significant relation to emotional experience, physiology, and behaviour.

Some researchers have described the dimensions of valence and arousal as *core affect*, that is, the basis for discrete emotional states (Barrett, 2006; Posner et al., 2005, 2009). This view is supported by neurophysiological evidence suggesting that these two dimensions may correspond to two basic brain mechanisms (Colibazzi et al., 2010; Heilman, 1997). According to this view, a certain pattern of neurophysiological activation that varies according to the two dimensions (e.g., positive arousal) would be interpreted by the person as a specific emotional state (e.g., joy), that is, a discrete emotion.

An earlier approach to combining the discrete and dimensional theories of emotion is based on the circumplex model of affect (e.g., Russell, 1980; Watson & Tellegen, 1985). The circumplex of affect is formed when different categories of emotion (e.g., happy, excited, and alert) are placed in the emotional space of valence and arousal. Figure 2.2 illustrates how different, more or less discrete, emotional states could be organized in the space of valence and arousal dimensions. A similar circular pattern has been found in several types of emotion ratings (e.g., ratings of subjective experiences or evaluations of emotion related words) using several different empirical procedures (e.g., sorting done by the participants or factor

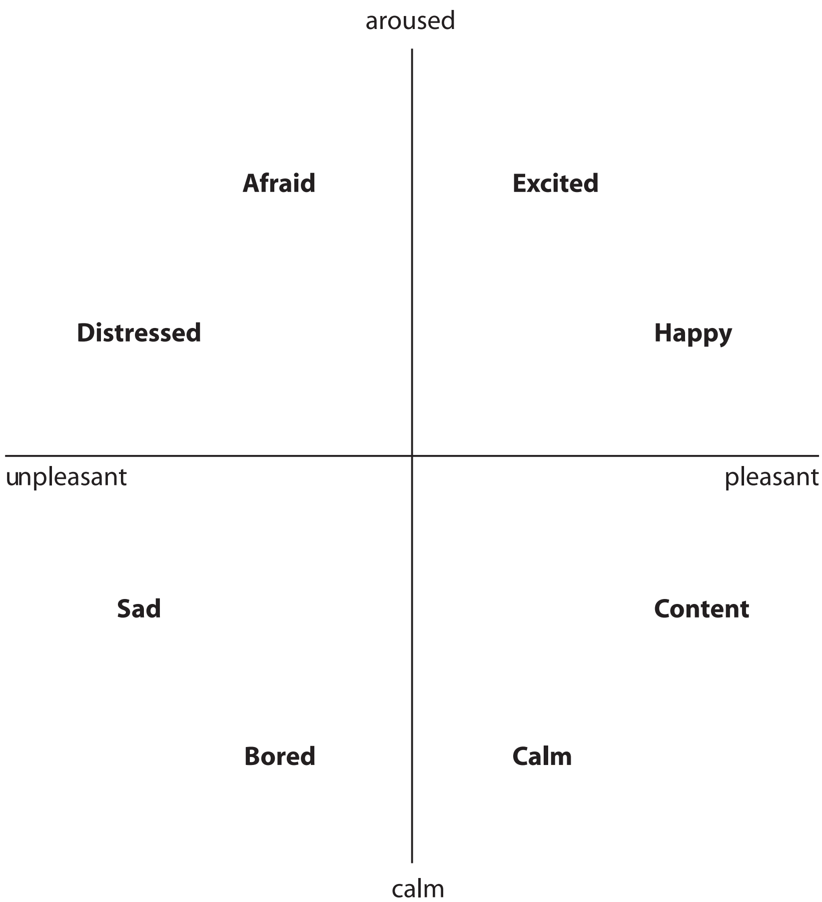


Figure 2.2: A possible organization of emotional concepts in a circumplex of affect. Redrawn and adapted from Russell (1980).

analysis). These results suggest that more basic psychological constructs of valence and arousal could underlie also self-reported experiences of discrete emotional states (Barrett, 2006).

2.1.3 Summary

Based on previous research, both the discrete and the dimensional approaches to emotions have a solid empirical basis. On one hand, the dimensional structure has been found in both verbal assessments of emotion (e.g., self-reported emotional experiences) and bodily (i.e., neurophysiological and psychophysiological) emotional responses. On the other hand, the cross-cultural, physiological, and behavioral evidence for the existence of a limited set of basic discrete emotions is also convincing. Thus, using these two theories as complementary approaches, instead of choosing one over the other, seems like an attractive option.

However, in practice it may be beneficial to use one of the approaches as a working hypothesis. In the empirical research for the present thesis, the dimensional theory was chosen as a starting point. This approach allowed the experiments to be designed so that well-established measures of physiological and subjective responses could be matched to the characteristics of the stimulation. For example, virtual computer characters were used in Publications III and IV. In both studies, the virtual stimulation was designed to vary in emotional valence, and the participants' perception of the stimuli was confirmed using subjective ratings of valence. Further, in the research reported in Publication IV, we also measured the electrical activity of certain facial muscles that have been previously well-connected with emotional valence. Thus, the chosen dimensional approach provided us a framework for choosing the appropriate measurements for each experiment.

2.2 The Three Components of Emotion

2.2.1 Experience

Our everyday usage of the term "emotion" seems to stress the significance of subjective experiences of emotion, that is, "feelings" (Gross & Thompson, 2007). In common language, the terms "emotion" and "feeling" are often used interchangeably. However, general scientific view of emotion does not emphasize conscious subjective experience over other components of emotion, but describes all components as fundamental to emotion (Mauss et al., 2005). In fact, there is evidence that responses to emotionally meaningful stimuli can occur without conscious experience of the stimuli. For example, Dimberg, Thunberg, and Elmehed (2000) found that brief (30 ms long) presentations of facial expressions of happiness and anger induced congruent facial reactions in the observer, even though conscious perception of the expressions was prevented using the backward-masking technique (i.e., presentation of a neutral face to mask the preceding stimulation).

The measurement of subjective experience of emotion requires not only that the person is aware (i.e., conscious) of the emotion, but also that she or he translates the experience into a semantic form that allows the emotion to be reported. This may be challenging for several reasons (Russell, 1980). For example, the accuracy of the report will depend on how accurately the person is able to describe her of his feelings. The lay terms for describing an emotion may be insufficient for capturing the essential parts of the experience at a fine-grained level. Further, any individual person will have naïve (i.e., untrained) views of emotion that vary in the degree that they correspond to scientific theories of emotion or to other lay views of emotion. These implicit theories of emotion guide how a person interprets and describes emotion. For example, one participant in a study could describe an escalating response to rude behaviour as a negative emotion that increases in intensity, while another participant might label the initial reaction and the subsequent response as two different emotions, such as, disapproval and anger. Reconciliating such differences in subjective reports of emotional experience could pose significant effort for a researcher.

For the above reasons, structured approaches to measuring emotional experiences may be particularly appealing for empirical work based on established theories of emotion. One prominent approach to measuring emotion uses bipolar scales to measure the emotional dimensions of valence (i.e., from unpleasant to pleasant), arousal (i.e., from calm to excited), and dominance (i.e., from being in control of to being dominated by the stimulus). This three dimensional framework for measuring emotion has been the basis for creating standardized sets of pictures (International Affective Picture System; IAPS), sounds (International Affective Digital Sounds; IADS), English words (Affective Norms for English Words; ANEW), and brief English sentences (Affective Norms for English Text; ANET), along with their normative ratings (Bradley & Lang, 1999a, 1999b, 2007; Lang, Bradley, & Cuthbert, 1999).

Such standard ways of classifying stimuli provide a framework for researchers to build upon previous work and design experiments that address specific questions that remain open. For example, everyday experience and laboratory studies have long suggested that emotional events are remembered in much greater detail than less emotionally rich events and stimuli (Hamann, 2001). Later studies have provided a more detailed view of the psychological and neurophysiological mechanisms behind these effects, suggesting that these enhancements in memory are mainly associated with the physiological arousal elicited by either emotionally negative or positive stimuli (LaBar & Cabeza, 2006). Valence of the stimulus, on the other hand, seems to make a relatively small contribution to the observed memory enhancements (Kensinger & Corkin, 2003). Standard ways of rating the stimuli have allowed researchers to carefully select the appropriate stimuli for such experiments, combining different levels of arousal with different categories of stimulus valence (e.g., from less to more arousing negative stimuli).

However, applied settings may pose additional challenges for measuring subjective experiences. For example, subjective experience is a significant component in the clinical treatment of anxiety disorders. Acquiring a detailed measure (e.g., several ratings) would be inconvenient due to the required effort that distracts the person from the therapeutic tasks. In practice, a measure of Subjective Units of Distress (SUD) is often applied in this setting (Krijn, Emmelkamp, Olafsson, & Biemond, 2004; Wiederhold & Wiederhold, 2003). The measure consists of a single rating of the current experience, which is typically given verbally on a scale of 0 (i.e., no distress) to 100 (i.e., extreme distress). The relatively small effort imposed by the measurement allows the rating to be queried repeatedly throughout a clinical session, which can then provide information for guiding the treatment.

On the other hand, the area of application may also serve as a context that directs the inquiry to employ the most relevant measurements. For example, SUD is a feasible measure for supporting the clinical treatment of anxiety because negative emotion, that is, distress, and its regulation are the most relevant aspects of emotional experience to consider during the course of treatment. In less intensive contexts, it is feasible to employ measures that collect more fine-grained measures that address the specific area of application. For example, Craig, D'Mello, Witherspoon, and Graesser (2008) devised an emote-aloud procedure that was applied to studying emotional responses during computer-assisted learning. The procedure was based on eight experiential states that arguably relate especially to learning, such as, boredom, confusion, and curiosity. Whenever participants experienced an emotion, they were to choose one of these eight pre-defined states and verbally report it.

In general, when the aim is to measure discrete (e.g., basic) emotions, typically either forced-choice ratings or ratings of intensity are used. In the former option, a person is presented with a set of discrete emotions that are effectively treated as exclusive. The task is to choose the emotion that best matches the subjective experience elicited by, for example, a visual stimulus. In addition to the emote-aloud procedure as discussed above, this approach has been employed in several studies of human perception of emotions (e.g., Juslin & Laukka, 2003; Coulson, 2004). The second option is to have the person rate the felt intensity of several discrete emotions at the same time. This way, the set of emotions may be extended to cover more complex mixtures of emotions. For example, a concurrent experience of both happiness and sadness could be called bittersweet (Larsen, McGraw, & Cacioppo, 2001).

2.2.2 Behaviour

External behaviour can be observed without asking the person to voluntarily express an emotion. Thus, as compared to measures of subjective experience, behavioural measures may allow more spontaneous and involuntary emotional responses to be investigated. Further, several technological (e.g., hidden cameras) and procedural (e.g., the use of a cover story) arrangements allow behaviour to be observed without the participant's awareness, provided this is ethically reasonable. Such hidden recordings can avoid the influence from knowing that others are observing one's behaviour, and thus capture more spontaneous behaviour than other, more obvious measurements (Picard, Vyzas, & Healey, 2001).

The face is in many ways central to emotional behaviour. First, the face allows a wide variety of expressions to be displayed, as facial musculature is fine-grained and well-innervated (Rinn, 1991; Schieber, 2001). Second, the face is normally always visible and readily providing some information (Cohn & Ekman, 2005). Third, in addition to facial expressions, the face is the source of several other kinds of messages, for example, speech and personal characteristics (e.g., identity, gender, and age). Thus, we can be expected to pay significant attention to other people's faces.

There are several techniques for systematically describing (i.e., coding) human facial behaviour (see Cohn & Ekman, 2005, for a review). These systems are based on detailed observation and encoding of individual facial actions, that is, observable changes in the face (e.g., wrinkles in skin and movements of the eye brow). As such, most systems do not impose a certain framework for judging which emotions the individual movements convey (i.e., what is their significance "as a whole"). However, many of them are based on a particular theory of emotional responding, which limits the extent that the systems can be used to code expressions not predicted by the theory in question (Cohn & Ekman, 2005). In contrast, the most widely used system called Facial Action Coding System (FACS; Ekman, Friesen, & Hager, 2002) covers all anatomically feasible actions, being blind to theory in this sense.

In any case, after facial behaviour is observed and encoded, it needs to be interpreted and this process will be largely dependent on the theory that guides the research. For example, Posner et al. (2005) have argued that although observational evidence from new-borns seems to show behaviour consistent with certain discrete basic emotions shortly after birth, these observations may be confounded with the researchers' interpretation of the actual behaviour. For example, although infants display behaviour (e.g., smiling) that can be associated with certain basic emotions (e.g., joy), these behaviours could also be taken to reflect different patterns of activity according to dimensional theories of emotion (e.g., a highly pleasant and aroused state). In line with this, some researchers have argued that dimensional theories adequately explain most of the variance in facial behaviour (Mauss & Robinson, 2009).

However, facial coding systems have — at least so far — been rarely used to investigate emotional responses according to dimensional theories. On the other hand, the connections between other channels of expression and specific dimensions of emotion have been quite thoroughly studied. For example, studies of vocal expression of emotions suggest that parameters of voice and speech are mainly affected by the current state of arousal, while the valence of the experienced emotion is more difficult to observe, at least from individual parameters (see Cowie et al., 2001; Mauss & Robinson, 2009; Murray & Arnott, 1993, for reviews). For example, anger (i.e., negative arousal) and happiness (i.e., positive arousal) have both been found to increase pitch, intensity, and rate of speech. On the other hand, there is some evidence that certain categories of emotion — including anger and happiness — could be discriminated from each other by comparing complex patterns of several acoustic parameters (Banse & Scherer, 1996). Further, other people (i.e., listeners) are able to recognize different categories of vocalized emotion (e.g., angry and sad speech) remarkably well (Juslin & Laukka, 2003).

In addition to affecting behaviour that may serve in conveying emotions to others (e.g., facial expressions and speech), emotions can also have effects on basic behavioural tendencies that serve to facilitate an individual's own intentions. Perhaps the most direct link between emotions and motivated behaviour is provided by the basic tendencies to approach positively valenced objects (e.g., nourishment) and avoid negatively valenced objects (e.g., punishment) (Bradley et al., 2001; Elliot, 2006; Lang, 1995). These tendencies are present in virtually all animals and they have been suggested to serve as the basis for more complex human emotions (Davidson, 1993; Lang, Bradley, & Cuthbert, 1992).

An intuitive approach to studying such tendencies is to observe body movement responses to emotionally appealing and aversive stimuli. An often used method for accurately measuring the body posture and sway (i.e., amount and magnitude of movements) is to have the person stand on a platform that has embedded force sensors. Using this technique several researchers have observed a reduction in body movements, that is, freezing-like behaviour or behavioural inhibition in response to unpleasant picture stimuli (Azevedo et al., 2005; Facchinetti, Imbiriba, Azevedo, Vargas, & Volchan, 2006; Stins & Beek, 2007). Further, Hillman, Rosengren, and Smith (2004) found that unpleasant, neutral, and pleasant pictures all induced posterior movement, that is, movement away from the stimuli. These results seem to be against a direct connection between emotional valence and approach-avoidance behaviour. On the other hand, when Miles (2009) made the association to approach behaviour very clear by using neutral and positive facial images that appeared to move towards each participant, the approach behaviour of participants was indeed enhanced towards the smiling (i.e., positively valenced) images as compared to neutral images.

In addition to measuring body movements, approach-avoidance tendencies have been investigated using reaction times to the perceived behaviour of others. The underlying premise of this approach is that movement which is congruent with the approach-withdraw tendency conveyed by other cues (e.g., facial expression) should facilitate reaction times in a variety of tasks. In studies using approaching and withdrawing facial images as stimuli, such advantages have been found for both approaching and withdrawing angry faces as compared to other facial expressions (Adams Jr., Ambady, Macrae, & Kleck, 2006; Van Peer, Rotteveel, Spinhoven, Tollenaar, & Roelofs, 2010). This suggests that the conveyed intention to approach or withdraw may depend on other context besides the facial expression as such (e.g., anger could signal both approach and withdrawal tendencies).

The above results suggest that the connection between approach-avoidance motivation and valence may be less direct than some have assumed (see Amodio, Master, Yee, & Taylor, 2008; Carver, 2006; Harmon-Jones & Allen, 1998; Norris, Gollan, Berntson, & Cacioppo, 2010, for similar results). Thus, there is a basis for viewing approach-avoidance motivation as a separate component which influences behaviour and may be associated with (but not equal to) emotional valence (Mauss & Robinson, 2009).

2.2.3 Physiology

It can be argued that physiological responses are normally under less voluntary control than self-reports of emotion and behaviour. Consequently, some have suggested them to reflect more "true" emotion as compared to subjective and behavioural measures (Knapp, Kim, & André, 2011). However, physiological processes primarily serve many purposes related to basic mechanisms of living (Berntson & Cacioppo, 2000). Thus, physiological activity can reflect a wide variety of other processes (e.g., digestion and homeostasis) besides emotion. Nonetheless, there are several ways in which physiology is significantly affected by emotion, and these effects can be a valuable source of information about the basic mechanisms behind emotional processes.

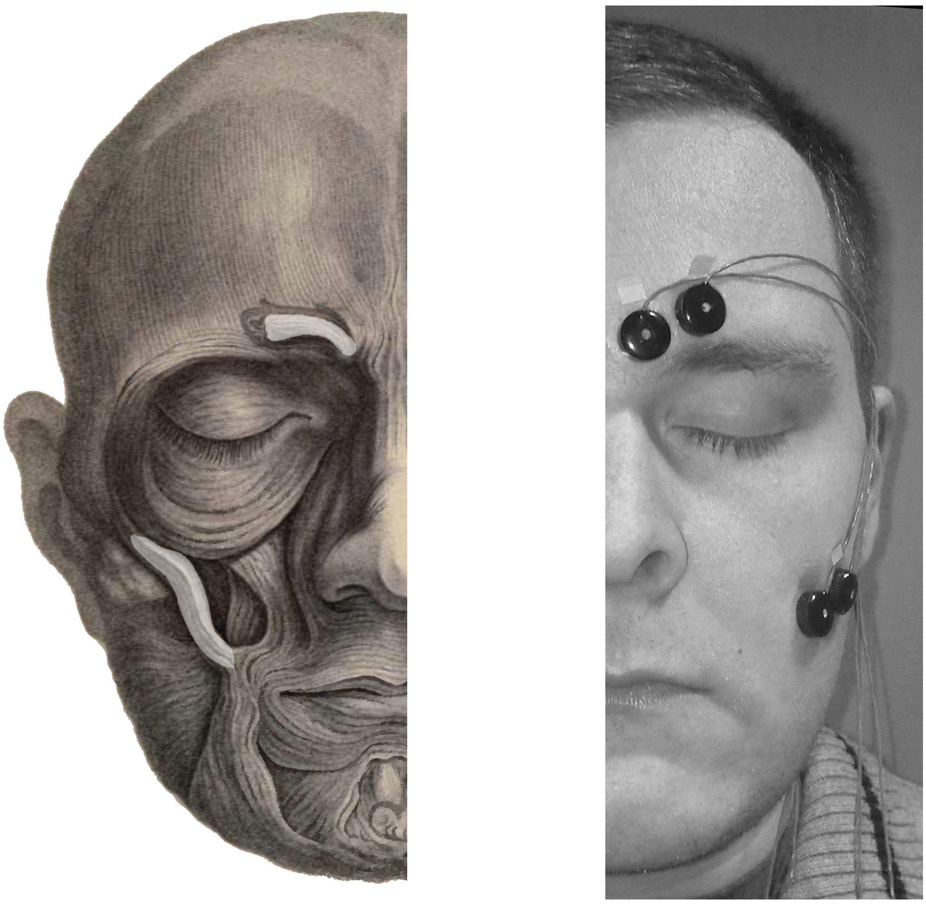
One line of evidence for the association of certain discrete emotions and patterns of physiological activity comes from the studies of Levenson and colleagues (Levenson & Ekman, 2002; Levenson et al., 1990, 1992). In these studies, participants were instructed to produce sets of facial muscle activations that replicated facial configurations associated with more spontaneous emotional expressions (e.g., happy and angry faces). The results showed that certain physiological parameters differentiated among specific pairs of expressions. For example, expressions of fear and happiness were associated with significantly different skin conductance responses, while the change in skin conductance was nearly equal for fear and disgust. However, fear and disgust were separated by significantly different heart rate responses. Similarly, Rainville, Bechara, Naqvi, and Damasio (2006) more recently found relatively complex patterns of physiological and behavioural (i.e., respiration, heart activity, and facial activity) responses that separated fear, anger, sadness, and happiness elicited with affective imagery. Thus, the evidence suggests that some emotions are associated with specific physiological patterns.

On the other hand, such discrete physiological patterns are relatively complex as compared to the more straight-forward connections between physiology and the dimensional view of emotions. Probably the most well-established measures associated with experienced valence (i.e., pleasantness) are heart rate responses and electrical measures of facial muscle activity (Bradley, 2000; Bradley et al., 2001; Tassinary & Cacioppo, 2000). Heart rate responses to emotionally arousing (i.e., unpleasant or pleasant) stimuli typically have shown a deceleration with a varying pattern of accelerations and decelerations following the stimuli (Bradley, 2000; Codispoti & De Cesarei, 2007). Negative stimulation has induced a lower heart rate than neutral and positive stimulation. However, although such typical heart rate patterns have been found, the patterns have also varied considerably between different types of stimulation and tasks (Anttonen, Surakka, & Koivuluoma, 2009). For example, viewing emotion eliciting pictures has generally induced heart rate decelerations, while imagining emotional situations has evoked accelerative responses (Bradley et al., 2001; Witvliet & Vrana, 1995).

There is evidence that electrical measures of facial activity could be a more specific index of experienced emotional valence. Facial electromyography (EMG) can be so sensitive that small muscle activations can be detected without externally observable changes occurring in the face (Gross

& Levenson, 1997; Dimberg et al., 2000; Tassinary & Cacioppo, 2000; Larsen, Norris, & Cacioppo, 2003). Especially two muscles, one in the cheek and one in the forehead, have been well-connected with subjective experiences of emotional valence elicited with, for example, picture stimuli, sound stimuli, and imagery (Lang, Greenwald, Bradley, & Hamm, 1993; Bradley & Lang, 2000; Codispoti, Mazzetti, & Bradley, 2009). Figure 2.3 illustrates the two muscles and the placement of EMG electrodes for measuring their activity according to the guidelines of Fridlund and Cacioppo (1986).

The activity of the *zygomaticus major* muscle in the cheek (activated when smiling) increases during pleasant experiences and decreases during unpleasant experiences. The activity of the *corrugator supercilii* mus-



(a) (b)

Figure 2.3: Muscles of the face. (a) A diagram of the facial muscles on the right side of the face with the two muscles highlighted. (b) Electrode placement for measuring electromyographic activity of the *corrugator supercilii* and the *zygomaticus major* muscles from the left side of the face. (a) adapted from public domain work by Bell (1865).

cle in the forehead (activated when frowning), on the other hand, increases during unpleasant experiences and decreases during pleasant experiences.

Besides emotional valence, other dimensions of emotion have also physiological associations. The specificity of the connection between sweat gland activity and emotional arousal has been long argued (Schlosberg, 1954). Electrodermal measures are acquired using a small voltage electrical current, which is typically applied between two fingers of the same hand (Dawson, Schell, & Filion, 2000). The electrical conductivity of the skin depends on, and can be used to assess, the activity of the sweat glands.

Some have even suggested using such electrodermal measures as direct indices of anxiety related arousal (Fowles, 1988). More generally, short- and long-term changes in electrodermal activity can be consired to be quite specific measures of sympathetic activation of the autonomic nervous system (Dawson et al., 2000). These changes have been used to assess both short-term (i.e., few seconds long) emotional responses to a certain stimulus and longer term (i.e., minutes or hours) changes in the general level of arousal. For example, in one study excessive fear of flying was found to be associated with both a generally higher level of skin conductance and greater short-term changes in skin conductance, as compared to participants without flying phobia (Wilhelm & Roth, 1998).

However, although such connections between physiology and emotional dimensions have been found consistently, recent studies suggest that similarly valenced and arousing emotions can still evoke quite different physiological responses. For example, Kreibig, Wilhelm, Roth, and Gross (2007) found that fear and sadness inducing videos evoked significantly different heart rate responses, although subjective ratings of valence and arousal showed almost no difference between the videos. Stemmler, Aue, and Wacker (2007) instructed their participants to imagine soccer scenarios associated with either fear or anger and a motivational tendency to either approach or withdraw. Their results showed that heart rate and electromyographic response patterns could significantly separate not only the negatively arousing emotions of fear and anger, but also the direction of motivation (i.e., approach or withdraw).

These results suggest that the dimensions of valence and arousal do not necessarily capture the whole variance of physiological responses associated with emotion. Neurophysiological studies provide further evidence that emotional responses are associated with specific approach-withdrawal tendencies that may at least partially be distinct from valence and arousal. One line of evidence comes from studies of electroencelographic (EEG) alpha asymmetry, that is, the relative activation of the left- and right-side of the brain as assessed by electrical measures. Greater left frontal brain activity has been associated with the tendency to approach, while greater right frontal brain activity has been found to reflect the tendency to withdraw (Coan & Allen, 2004).

Coan, Allen, and Harmon-Jones (2001) measured frontal brain activity while participants produced voluntary facial configurations resembling spontaneous emotional expressions, that is, in a procedure similar to the studies of Levenson and colleagues (Levenson et al., 1990, 1992). They found that approach-oriented emotions (e.g., anger) elicited greater left frontal brain activity as compared to withdraw-oriented emotions (e.g., fear). In general, these basic motivational tendencies seem to be central to physiological patterns of emotional responding and they cannot be accounted for by most traditional dimensional models of emotion (e.g., valence/pleasure and arousal/activation) (Carver & Harmon-Jones, 2009).

2.2.4 Summary

Although experience, behaviour, and physiology are all central to emotional responding, they are best viewed as loosely-coupled systems (Mauss et al., 2005; Mauss & Robinson, 2009). In other words, emotions are made of multiple, individually variable components. Thus, the observed effects of emotion may diverge from measure to measure and from one study to another. This variance in measures underlines the benefits of acquiring complementary measures from multiple systems (i.e., experience, behaviour, and physiology) in order to get a more complete picture of an emotional response.

The multi-component nature of emotion in terms of these loosely-coupled responses seems to be analogical to dimensional models of emotion, which view emotions as a combination of individual components, such as, valence and arousal. In line with this notion, some researchers have argued that most of the variance in experience, behaviour, and physiology can be explained using the dimensions of valence and arousal (Posner et al., 2005). However, there is a growing body of both behavioural and physiological evidence which suggests that, in addition to the more established dimensions of valence and arousal, the direction of motivation can be — at least in part — a separate component of emotional responding (Coan & Allen, 2004; Stemmler et al., 2007). For example, although emotional valence and motivation have strong associations, the motivational effects of emotional stimulation may be affected by other factors as well.

2.3 The Functions of Emotion

2.3.1 The Effects of Emotion

Emotions are an important part of human functioning. For each individual person, emotions have essential roles in influencing cognitive processes, such as, attention, memory, problem-solving, and decision making (Bechara, Damasio, & Damasio, 2000; Damasio, 1994; LaBar & Cabeza, 2006; Simon, 1967; Singer & Salovey, 1988; Wieser, Pauli, Reicherts, & Mühlberger, 2010). Further, emotions can also have significant effects on personal health, that is, both mental and physical well-being (Bishop, 2001; Butler, Wilhelm, & Gross, 2006; Malliani, Pagani, Lombardi, & Cerutti, 1991). For example, a meta-analysis of studies investigating the effects of hostility on physical health suggested that dispositional anger is at least as significant risk factor for coronary heart disease as the more generally acknowledged ones like smoking and high blood pressure (Miller, Smith, Turner, Guijarro, & Hallet, 1996).

Positive emotions, on the other hand, may have significant and potentially beneficial physiological and cognitive effects. For example, there is evidence that cardiovascular responses are quicker to normalize after negative stimulation when followed by positive stimulation (Fredrickson & Levenson, 1998). Further, another pair of experiments suggested that positive emotions may facilitate thinking in terms of broadening the scope of attention and the diversity of answers to open-ended questions (Fredrickson & Branigan, 2005).

In line with the above results, the significance of emotion for human behaviour is an integral part of most emotion theories. In Frijda's (1986) view, the central function of emotion is to generate a change in action readiness. This way, emotions can be seen as an adaptive response that allows refocusing of the behavioural system to address issues that arise from the emotion eliciting stimulus or situation (Oatley & Jenkins, 1996). A similar suggestion for the function of emotions as an interruption mechanism for cognition was made by Simon (1967). In his view, emotions could solve problems in coordinating goal-oriented behaviour by prioritizing certain goals, interrupting actions towards less important goals, and shifting focus to more important goals when appropriate.

In addition to being essential for an individual's own functioning, emotions serve also social purposes. Shared emotions may facilitate group cohesion, that is, how well-bonded the members of the group feel to the other members (Barsade, 2002). They also function to coordinate social interaction in a group (e.g., provide information about the environment quickly to other group members) and inform the person about the quality of current on-going interaction (Parkinson, 1996; Spoor & Kelly, 2004).

Emotions can also be used to elicit appropriate social responses from others. For example, making angry remarks to a colleague who misses a deadline may make her or him feel guilty and more willing to take the responsibility for meeting the next deadline.

Leading a successful every-day life requires us to retain from expressing all our emotions (Gross & Thompson, 2007). For example, making angry remarks to one colleague may facilitate the progress of her work, but scathing another may lead to a conflict which drains everyone's energy and efforts. Chastising one's boss is probably never a good career move. On a cultural level, emotional expressions are regulated by *display rules*, that is, social rules that specify which (and when) emotions are appropriate to show to a certain person (Ekman, 1979).

If some emotions cannot be shown at a given time, there occurs a need to suppress either the emotion itself or the emotional expression. There is evidence that suppressing an emotion (e.g., by reappraising the situation) is more effective and may have less negative consequences (e.g., reduce symptoms of depression) than suppressing the expression of the emotion (Gross, 2002). More generally, successful regulation of own emotions, that is, controlling the intensity and tone of emotions, can promote personal well-being and satisfaction in life (Gross & John, 2003). For example, children who are more skillful regulators of emotion in kindergarten perform better academically in the first-grade (Trentacosta & Izard, 2007). Another study found that college students who had better abilities for self-regulating emotions were nominated by their peers significantly more often as prosocial and were more often involved in reciprocal friendships as compared to other students (Lopes, Salovey, Coté, & Beers, 2005).

On the other hand, less skillful regulation of emotion is associated with adverse mental and physical consequences, such as, depression and heightened physiological responses to negative emotion (Gross & John, 2003; Butler et al., 2006). The flip-side of the significance of emotion for human functioning is that excessive emotion can be debilitating. For example, being wary of snakes may be reasonable and even promote survival in a rural environment, but excessive fear of snakes may significantly reduce a person's quality of life. Thus, not only the content but also the intensity of emotions may contribute to their practicality.

2.3.2 The Process of Emotion

Emotions do not happen as sporadic events without any context, but are more appropriately characterized as ongoing processes that relate to other events and situations. Frijda (1986) suggested a model which represents emotions as a process that goes through stages of appraisal, context evaluation, change in action readiness, and changes in physiology, expression, and action (i.e., behavior). Other theorists have also suggested similar information-processing models that are generally in line with Frijda's well-accepted formulation, although several details (e.g., order and timing of different parts of the process) may be open for argument and further investigation (Lanctôt & Hess, 2007; Oatley & Jenkins, 1996).

The first stage of the process, appraisal, determines whether an event is significant for the person and, thus, which emotion (if any) should occur (Ellsworth & Smith, 1988; Oatley & Johnson-Laird, 1987; Scherer, 1993). The actual process of appraisal (i.e., the causes and mechanisms of emotional inferences) can be highly automated and unconscious, that is, outside of awareness (Ekman, 1992; Oatley & Jenkins, 1996). Then, in the stage of context evaluation "[t]he stimulus situation as a whole is appraised in terms of what the subject can or cannot do about it" (Frijda, 1986, p.455). This secondary appraisal, in Frijda's terminology, serves as the basis for preparing and selecting among potential actions in later stages. Finally, previous stages lead to a coordinated change in action readiness that manifests in physiological changes and eventually may lead to overt action.

However, different appraisals and motivational tendencies do not necessarily translate to specific actions. Frijda (1986) emphasized the role of different regulatory mechanisms in determining the resulting actions and suggested that all stages preceding action may be subject to emotion regulation. Gross and Thompson (2007) extended this view even further by pointing out that in addition to regulating own emotions (i.e., intrinsic regulation), one can also also regulate another person's emotions (i.e., extrinsic regulation). However, the means for regulating emotion in self and others are essentially the same.

Figure 2.4 shows the process model of emotion regulation as presented by Gross (1998). As emotions are elicited by situations, the first method for regulating which emotions occur is to select appropriate situations. For example, one could rent the latest comedy to cheer up either oneself (i.e., intrinsic regulation) or a spouse who is having a bad day (i.e., extrinsic regulation). Later, when an emotion eliciting situation occurs, it is possible to modify the situation itself. For example, noticing that others are terrified while you are driving fast on a highway may cause you to slow down. In this case, the expressions of others serve to first modify the emotional tone of the social environment, while the consequent slowing down modifies the physical environment, which can both be seen as a type of extrinsic regulation in this case.

Modifying the external situation may be very effective for determining which emotions occur. However, it is not always feasible to change the environment. On the other hand, internal processes for regulating

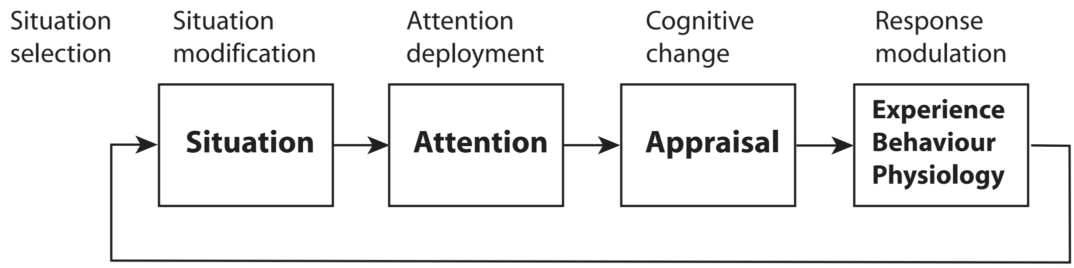


Figure 2.4: The process model of emotion regulation. Top row lists different means for regulating emotions in different stages of emotional responding. Redrawn and adapted from Gross and Thompson (2007)

emotion are always potentially available, as they depend only on the person's own competence. First, a person may affect how she or he directs attention to different aspects of the situation. Second, cognitions relating to the situation may be voluntarily changed to affect how the situation is appraised. For example, one could try recollect a previous positive experience from a public performance before speaking again to a crowd.

Finally, it is possible to regulate emotional responses to the situation as it happens. For example, different cultures impose specific demands for regulating emotional expressions based on which emotions are appropriate in a certain context. Ekman (1979) has argued that some of these *display rules* are learned so well that they operate automatically without conscious effort. For example, in one study Americans rated facial expressions of happiness as more appropriate to show to outgroups (i.e., people outside their own social group) as compared to Japanese, while Japanese rated anger and fear as more appropriate to show to outgroups as compared to Americans (Matsumoto, 1990). However, another study referenced in Ekman (1979) showed that when an outside experimenter was present, the Japanese suppressed negative facial expressions and masked them with smiling more often than Americans. This mismatch of voluntary ratings and spontaneous behaviour seems to support the view that behaviour is largely influenced by automated processes, such as, display rules.

In addition to regulating which emotions can and should be expressed, the intensity of a response can also be regulated. For example, although one might get very angry when someone cuts in front of her or him on a freeway, typically these feelings are suppressed or coped with in such ways as mumbling curses under one's breath. However, in some cases the driver is unable to control her or his temper, which may lead to so-called *road rage* (Wells-Parker et al., 2002). Milder behaviours associated with road rage include flashing the lights and using the horn, while in more extreme cases the driver may deliberately try to hit another car. It is not surprising that such angry behaviours on the road lead to an increased risk of crashing while driving (Wells-Parker et al., 2002).

Clearly, there would be significant benefits in facilitating emotion regulation in these kinds of situations (e.g., driving on a free-way), where hazardous behaviour poses serious risks to oneself as well as other people. One recent line of research aims to develop technologies that would give human-like assistance and support to the driver (Eyben et al., 2010; Nasoz et al., 2010). In a way, commonly used in-car navigation systems provide a contemporary example of this kind of virtual assistant that uses artificial speech to facilitate driving. A recent study by Nass et al. (2005) provides a rousing example of how emotion-sensitivity could be implemented to improve such voice-based systems. The results of this study showed that when the emotional tone of an artificial voice (i.e., happy or upset) in a car simulator matched the emotional state of the driver, the driver attended more to the road and made less accidents during simulated driving, as compared to when the emotions of voice and the driver were mismatched. Thus, an emotion-sensitive in-car system could have significant potential for improving the performance of a driver and, consequently, facilitating road traffic safety.

Technology is becoming increasingly ubiquitous, for example, embedded within our environment, clothes, and even our bodies (Tennenhouse, 2000; J. Rantanen et al., 2002). Mobile technologies already allow us to be reached virtually at any time and place. Thus, as technology pervades more and more of our everyday lives, there is an increasing potential for creating artificial systems that could assist emotion regulation whenever needed. The applications for such systems could range from the more dramatic (e.g., reducing road rage while driving; Nasoz et al., 2010) to the more ordinary (e.g., automatically choosing to play soothing music after a hard day at work; Chung & Vercoe, 2006).

2.3.3 Summary

The significance of emotion for human functioning is evident in its effects on cognition (e.g., memory and problem-solving), physical health, mental well-being, social interaction, and life success in general (Bechara et al., 2000; Damasio, 1994; Gross & Thompson, 2007; LaBar & Cabeza, 2006; Trentacosta & Izard, 2007). Consequently, abilities to regulate emotion also have great significance for leading a successful life.

There are several ways in which emotions can be regulated before, during, and after they have been elicited (Gross & Thompson, 2007). However, all attempts to self-regulate emotion cannot be successful. Sometimes other people can provide assistance (i.e., extrinsic regulation), but it may be that they are not available, capable, or willing to help when one is most in need of support. On the other hand, artificial systems for facilitating emotion regulation could be made available virtually all the time without needing rest or reciprocal assistance. Thus, there seems to be clear potential for facilitating the regulation of emotions with technology.

3 Computer-Assisted Emotion Regulation

3.1 Emotionally Intelligent Technology

3.1.1 Affective Computing

Affective computing is a relatively recent field of study which can be broadly defined as "computing that relates to, arises from, or deliberately influences emotions." (Picard, 1997). The ultimate aim of this approach can be seen as the development of computer systems that have *emotional intelligence*, that is, the ability to recognize, express, and regulate emotions in self and others (Picard et al., 2001). Thus, the field by definition covers a wide range of emotion related phenomena, for example, the development of computational models that can be used to simulate emotions and their regulation in computer software agents (e.g., Bosse & de Lange, 2008; Bosse, Gratch, Hoorn, Portier, & Siddiqui, 2010; Marsella & Gratch, 2003). Perhaps a better way to describe the more specific work towards technology that is sensitive to human emotion is to separate the task into two subgoals: creating perceptually intelligent technology and building expressive intelligence for technology (Ochs, Niewiadomski, Pelachaud, & Sadek, 2005; Pentland, 2000).

3.1.2 Perceptual Intelligence

The aim of perceptual intelligence is to develop capabilities that allow computers to attend to people and surroundings like a human would (Pentland, 2000). In practise, such capabilities can be based on automatic analysis of those channels of expression that humans use to perceive the emotions of others. This is arguably a quite intuitive approach to building technology that perceives emotions and such technologies have already been developed for recognizing the emotional content of practically all human expression, for example, speech, text, facial expressions, and body movements (Castellano, Villabla, & Camurri, 2007; Cohn & Ekman, 2005; Cowie et al., 2001, 2005; Fasel & Luettin, 2003; Wu, Chuang, & Lin, 2006). The accuracy of these systems can be very good. For example, the system of Bernhardt and Robinson (2007) achieved an overall recognition rate of 81% in classifying motion-capture data of a knocking motion (i.e., pounding something like a door) that participants performed in a neutral, happy, angry, or sad style. It is clear that a similar task could be quite challenging for a human perceiver as well. In general, systems that recognize emotion often can match the accuracy of human perceivers, or sometimes even do slightly better than us (Picard et al., 2001).

Although such results are promising, there is a limit to the extent that reliable emotion recognition can be built on individual channels of expression. Voice recognition systems are sensitive to noise in the environment, for example, engine and road noise in the case of in-car systems (Grimm et al., 2007). Computer vision systems for detecting facial expression can be susceptible to changes in lighting and head orientation as well as inaccurate detection of facial features (Cowie et al., 2005). Further, placing a camera that points at the person's face may not be practical for all applications, especially if the system should be mobile and unobtrusive. Thus, a general solution for perceiving emotions should probably use several channels of information, much in the same way as we humans perceive emotion (e.g. Juslin & Laukka, 2003; Van den Stock, Righart, & de Gelder, 2007). Indeed, such emotion classifiers have achieved quite good accuracies (i.e., near or above 90% in some studies) as could be expected based on human performance in similar multi-channel tasks (Bailenson et al., 2008; Busso et al., 2004; Zeng et al., 2004).

There is no reason to restrict the capabilities of automatic perception by limiting them to those cues that humans can readily perceive. For example, Yoshitomi, Sung-Ill, Kawano, and Kilazoe (2000) created classifiers that used thermal images of facial expressions in addition to visible light images and voice parameters. The results showed quite good overall accuracy of 85% for a classifier that combined all three modalities to recognize neutral faces and expressions of happiness, anger, sadness, and surprise. Thermal facial images have the additional advantage that they can be acquired regardless of lighting conditions, which could extend the feasibility of these methods to some contexts that are challenging for traditional computer vision.

Human physiological responses to emotion can also provide information which is not normally accessible to other people. Practically all physiological measures which have been found to significantly reflect emotional processes have been applied also to automatic recognition of emotions, including electrical brain and facial muscle activity, heart rate, respiration, and skin conductivity (Horlings, Datcu, & Rothkrantz, 2008; Picard et al., 2001; Rani, Sarkar, Smith, & Adams, 2003). For example, Partala et al. (2005; 2006) created EMG-based systems for recognizing facial activity. These systems were able to estimate participants' subjective ratings of valence on a 9-point scale with up to a .9 correlation.

It may seem that physiological measurements are feasible for only a limited range of applications due to the complex arrangements and preparations (e.g., electrode and skin preparations) that are required for acquiring them. However, recent advances in wireless and wearable measurement technologies have already markedly facilitated the practicality of such measurements. For example, a wearable headband with embroidered silver thread electrodes for measuring electrical activity of facial muscles has been developed (Nöjd et al., 2005). More recently, a completely contact-free measurement of facial movements has been made possible by a wireless capacitive sensor that is integrated to eye glass frames (V. Rantanen, Niemenlehto, Verho, & Lekkala, 2010). Such novel measurement technologies could provide feasible alternatives to EMG for measuring relatively small movements of muscles and facial skin which are associated with emotional responding.

In addition to these research prototypes, first consumer products with integrated physiological measurements have recently appeared on the market. For example, Textronics offers sports bras and t-shirts with integrated sensors that can be coupled with Polar Electro's watch-like computers for wireless heart rate measurement during sports and exercise training (Textronics Inc., 2011). Emotiv currently (February 2^{nd}, 2011) sells the EPOC headset for EEG measurements at $299 USD which is a fraction of the cost of a clinical EEG system (Emotiv, 2010). In 2009, Mattel released a toy called MindFlex™ that consists of a platform and a headband that is worn to measure brain activity (Mattel, Inc., 2011). MindFlex™ allows a ball floating in air to be controlled with voluntary brain activity. The goal of the game is to pass the ball through different obstacles on the platform. Such systems suggest that physiological measurements will continue to become more convenient through the development of more easily applied measurement technologies. Physiological measurements may also become more common-place and acceptable through the proliferation of these kinds of gadgets.

3.1.3 Expressive Intelligence

In order to affect human social and emotional processes, computers need expressive intelligence, that is, ways of expressing appropriate social and emotional cues that have desirable effects. The basis for building effective expressive cues for computers comes from a line of studies by Clifford Nass and his colleagues (Nass, Steuer, & Tauber, 1994; Reeves & Nass, 1996; Nass, 2004). The results from these studies provide compelling evidence that humans perceive computers as social actors. For example, Partala and Surakka (2004) designed an experiment in which participants performed a sorting task by moving colored squares using a computer mouse. The task was randomly interrupted with pre-programmed delays in mouse movement, which could be followed by a negatively or a positively worded intervention from a speech synthesizer. The results showed that the interventions which used positively worded synthesized speech could significantly improve problem solving performance and increase smiling behaviour.

In general, previous work in this field suggests that people not only perceive computer as social entities, but also treat them as such (e.g., display politeness to computers), and are significantly affected by artificial social and emotional communication. For example, computer-generated text and speech communication is not only perceived to reflect the underlying personality and emotional state of a computer, but textual and prosodic cues also significantly affect human attitudes and behaviour (Nass & Lee, 2001; Brave, Nass, & Hutchinson, 2005). There are several alternative mechanisms which could account for these tendencies (see, e.g., Nass & Lee, 2001, for details). However, it seems likely that they are to a significant extent based on automatic mechanisms for perceiving social and emotional cues in human-human communication. For example, the strong evolutionary basis for the ability to recognize and produce speech causes people to attribute special significance to even nonsense syllables (Nass & Lee, 2001). Further, such automated mechanisms are difficult to extinguish without constant reminders to do so (Nass & Gong, 2000). Thus, computer-generated human-like expressions have the potential to repeatedly activate these same fundamental mechanisms for recognizing social and emotional cues, although on a conscious level one might be aware that the communication is essentially artificial.

One way of approaching the field of expressive intelligence is to consider how much realism or human-likeness is required from speech and other channels of expression in order to (still) evoke significant social and emotional responses in humans (Ilves & Surakka, 2008). Knowing the requirements and limitations of using each channel of expression could assist designers of future technologies to harmonize their potential, for example, optimize the generation of cues according to their computational and other resource requirements and their effectiveness for expressing emotion.

There is evidence that even quite simple artificial cues can have significant social and emotional effects on human emotion and cognition. For example, commonly available speech synthesizers have long provided control over parameters which enable the imitation of human vocal cues of emotion (Cahn, 1990; Murray & Arnott, 1995). Another line of research aims to create visual cues in the form of human-like characters that could enhance human-technology interaction (see Beale & Creed, 2009, for a review). In an early work in this field, Lester et al. (1997) found evidence that the mere presence of an anthropomorphic character could facilitate children's learning experiences in a virtual environment. More expressive agents had a greater impact on learning and students rated them more positively than the less expressive ones, but the effects were significant even when the character was muted, that is, did not provide vocal or animated assistance about the topic of study. The extent and significance of this so-called *persona effect* has been under some dispute since the seminal study (Van Mulken, André, & Müller, 1998; Dehn & Van Mulken, 2000; Moundridou & Virvou, 2002; Prendinger, Mayer, Mori, & Ishizuka, 2003). However, when Yee, Bailenson, and Rickertsen (2007) performed a meta-analysis of empirical studies comparing interfaces with visual characters to those without, the compound analysis confirmed that the presence of a virtual character resulted in more positive interactions with the system as compared to computer interfaces without a virtual representation. Thus, it seems a relatively conservative conclusion that people prefer interfaces that resemble humans to some extent.

Expressive virtual characters enable several social and emotional cues that require embodiement to be conveyed, for example, gestures, posture, facial expressions, gaze direction, and lip-synchronized speech. Empirical studies of such synthetic stimuli have shown that people readily perceive them as emotional cues and that these artificial cues can also significantly affect subjective experiences and physiology (Bailenson, Beall, Loomis, Blascovich, & Turk, 2005; Beale & Creed, 2009; Coulson, 2004; Fukayama, Ohno, Mukawa, Sawaki, & Hagita, 2002; Llobera, Spanlang, Ruffini, & Slater, 2010; Vinayagamoorthy, Brogni, Steed, & Slater, 2006). One relatively challenging line of study concerns the detailed replication of human social cues for each individual channel of expression (see, e.g., Martin, Niewiadomski, Devillers, Buisine, & Pelachaud, 2006; Vinayagamoorthy, Garau, Steed, & Slater, 2004, for detailed approaches).

On the other hand, some cues may be quite simple to generate for one character and also relatively easy to generalize for other virtual characters. For example, proximity to a visual stimulus (e.g., virtual computer character) could be simulated simply by varying its size (Loftus & Harley, 2005). Partala, Surakka, and Lahti (2004) found in one of the first studies of virtual proximity that a closer simulated distance to a human-like head decreased the experienced subjective dominance, that is, participants felt they were less in control of the stimulus when the head was closer. The study reported in Publication IV of the present thesis obtained similar results using full-body human-like characters.

Although such relatively simple cues may suffice at least in some cases, artificial expressions of emotion that more closely resemble human expressions may enable richer social and emotional communication. For example, Ilves and Surakka (2004) compared two speech synthesizers and found that the emotional content of synthesized speech evoked significant facial muscle responses only when the more human-like voice of the two was used. For the simulation of facial expressions, the development of increasingly sophisticated technologies has enabled the generation of highly detailed real-time facial models (e.g., Courgeon, Buisine, & Martin, 2009; Radovan & Pretorius, 2006). Figure 3.1 presents one such model that uses recently developed software called the Multimodal Affective and Reactive Character (M.A.R.C.). M.A.R.C. uses high resolution textures and 3D models with about 80,000 polygons per virtual character.

The realism of the synthesized facial expressions can be enhanced by more global changes in facial appearance. For example, Figure 3.1a shows that a smile mainly produces local changes in the mouth area, but also increases skin wrinkles around the eyes as compared to the neutral expression of M.A.R.C. in Figure 3.1b. Such fine-grained synthetic facial cues can not only facilitate realism, but also potentially provide access to the utilisation of rapidly occurring strong reactions to the human face, for example, unconscious processing of facial expressions, the recognition and separation of posed and genuinely felt enjoyment (i.e., so-called Duchenne smiles), and the integration of auditory and visual information to facilitate the perception of speech (Cosker, Paddock, Marshall, Rosin, & Rushton, 2005; Dimberg et al., 2000; Ekman, Davidson, & Friesen, 1990). It can be expected that future simulations of human expressive channels will continue to increase in realism and consequently also affect our emotional core more deeply than we can imagine.

3.1.4 Summary

There are several channels which can be used to perceive and express emotions in human-technology interaction. In addition to the expressive

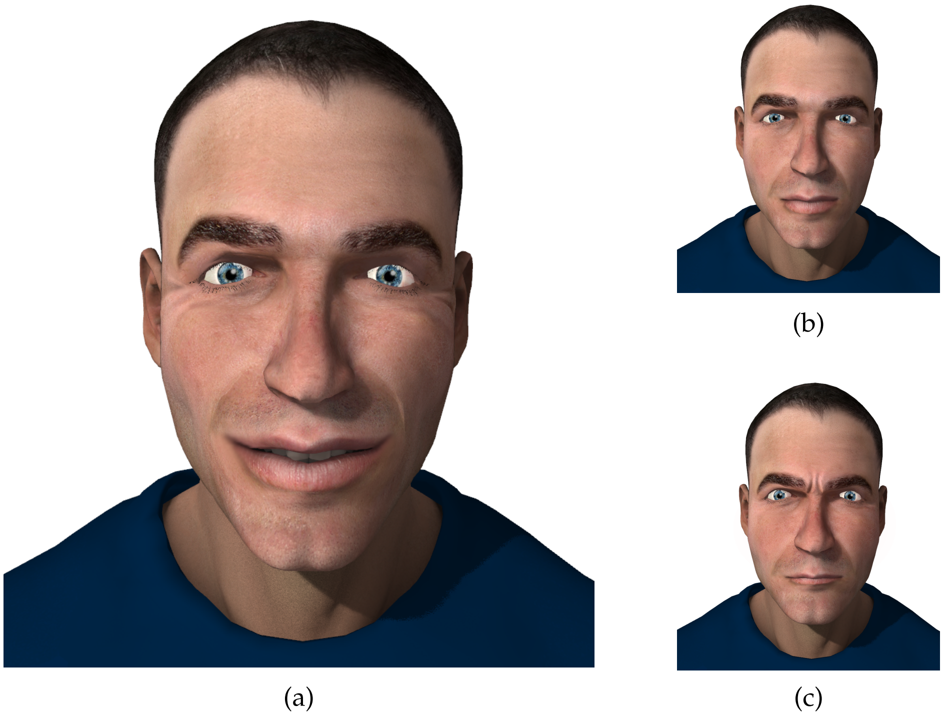


Figure 3.1: M.A.R.C. software showing a real-time simulation of a human-like face with (a) a smile, (b) a neutral expression, and (c) a frown. Images courtesy of Matthieu Courgeon, Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur, Centre national de la recherche scientifique, Université Paris-Sud 11, France

channels that are used in human perception of the emotions of others (e.g., speech and facial expressions), physiological signals provide an attractive source of information for future computer systems that recognize emotion. In particular, new prototype technologies and consumer products — even toys — provide a glimpse into upcoming wearable and wireless physiological measurement devices that can be readily used with little or no expertise. On the other hand, the most effective expressive channels for significantly influencing human emotion are likely those that resemble human expression of emotion. Such modes of expression can take advantage of the automatic nature of the human responses that are elicited by emotional and social cues.

Although the aim of affective computing — at least to some extent — is to build emotionally intelligent technology, the research so far has focused on these two separate tracks of automatically perceiving and synthetically expressing emotions (Picard et al., 2001). It may be that combining the two in a relatively straight-forward manner, for example, by perceiving emotions and expressing/projecting them on an avatar, can leverage the quality of human-technology interaction as such (see, e.g., Höök, 2008).

On the other hand, several researchers have warned against using impressive human-like expressions that may create false expectations about the capabilities of a computer with no underlying real intelligence (e.g., Dehn & Van Mulken, 2000). Further, it seems likely that some applications will require a more proactive approach to regulating emotions. For example, it may be insufficient to raise a person's own awareness of her or his excessive emotional responses, if the person does not have sufficient abilities for regulating those responses. In such cases, more proactively intelligent technologies would not just react to perceived emotions, but could also facilitate the regulation of those emotions, for example, by guiding the person's own reactions appropriately.

3.2 Computer Systems for Emotion Regulation

3.2.1 Approaches to Computer-Assisted Emotion Regulation

The process model of Gross (1998) provides a starting point for investigating how technology could help in different stages of emotion regulation. Figure 3.2 exemplifies some ways in which current and future technology could facilitate and possibly extend non-technological methods of emotion regulation.

Starting at the first stages of the process model of Gross (1998), computer systems could assist in selecting appropriate situations or modify-

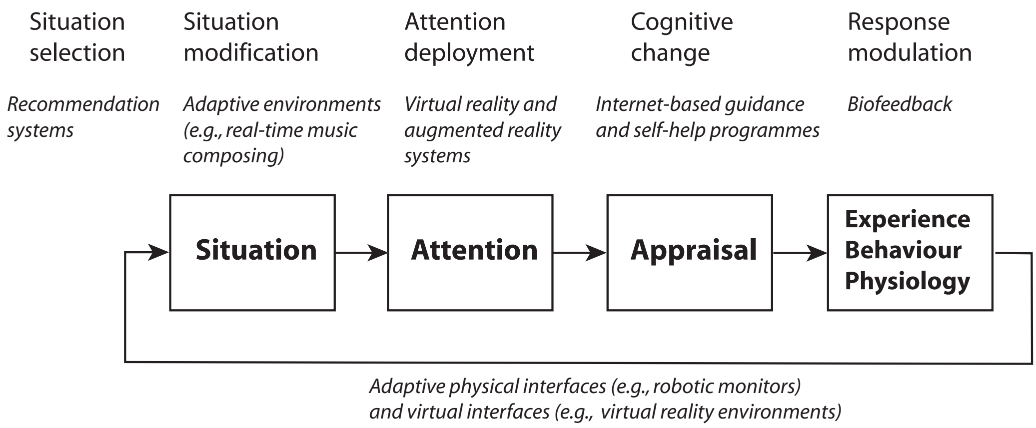


Figure 3.2: Some examples of technological applications that could provide support for different stages of the process of emotion regulation.

ing existing ones. Different *recommendation systems* could offer computerized support for these tasks. For example, Last.fm (Last.fm Ltd., 2011) offers a popular music recommendation service that allows people to listen to songs and mark those that they "Love". Based on this simple expression of emotion — together with other data (e.g., the social network of other listeners) — the service is able to automatically choose and play other songs that should elicit similar feelings for the particular person. In effect, this application can be seen as a relatively simple example of a system that facilitates the selection of an emotionally positive situation (i.e., background music) by recommending songs based on expressive intelligence about music.

Taking the personalization of music even further, Chung and Vercoe (2006) created a prototype system based on both perceptual intelligence and expressive intelligence about musical experiences. Their system could modify songs by re-arranging (e.g., composing) music in real-time based on affective responses that were automatically perceived from physiological and behavioural data. The expressive intelligence of the system aimed to select music in order to induce a specific emotional state, such as, positive arousal. It would be relatively straight-forward to apply such a system as a tool for regulating emotions in self or others, for example, in order to cheer up a dispirited friend. Although such systems may seem like novelty applications at first hand, there is evidence that music can in fact significantly affect physiology, for example, speed the recovery from the neurophysiological and physiological effects of negative emotions (Sokhadze, 2007). Thus, automatically arranged music could offer a tool for significantly suppressing the effects of negative emotions and perhaps even alleviating some of their adverse health consequences (Bishop, 2001).

The music recommendation and re-arranging systems discussed above illustrate basic approaches that could be used to assist a person to select emotionally appropriate situations (e.g., a song to play or a social event to attend) or automatically modify situations in emotionally intelligent ways (e.g., modify the tempo of music, change the tone of lighting, or adjust room temperature). However, it is practically impossible to assess and anticipate all aspects of every situation beforehand, and a person may need to take part in emotionally taxing situations in any case. For example, the success of a business and one's job may depend on attending an excruciating meeting with a difficult client. Thus, there is also a need to facilitate the regulation of emotions in situations once they occur. In these cases, the focus of technological support shifts from the modification of external environment to influencing the person's internal processes.

According to the process model of emotion regulation (Figs. 2.4&3.2),

the first opportunity to affect the regulation of emotions once a situation occurs is to facilitate directing attention to appropriate aspects of the situation. Perhaps the most intuitive approach would be to encourage a person to focus on more pleasant aspects, while avoiding the negative aspects of the situation. As an example, Amir et al. (2009) devised a computer system for training participants with generalized social phobia to direct attention away from threatening faces. The training significantly reduced self- and clinician-assessed symptoms of phobia, and half of the trained participants no longer met the diagnostic criteria for generalized social phobia after 4 weeks. Another quite dramatic example of this approach comes from burn-injury treatment, where virtual reality technology has effectively reduced anxiety by directing attention towards a computer-generated environment (e.g., a world of snow) and away from physical pain induced by wound care (see Morris, Louw, & Grimmer-Somers, 2009, for a review).

However, sometimes it may be essential to direct attention to the more negative aspects of the situation, that is, those that one would otherwise tend to avoid. For example, one method of treatment for certain anxiety disorders involves exposing a person to the object of her or his fear. There is evidence that this kind of exposure has a greater effect when the person is focused on the distressing stimulation (Foa & Kozak, 1986; Grayson, Foa, & Steketee, 1982). However, a person who is anxious of something tends to — quite naturally — look away from the object of the fear, which may relieve short-term anxiety but upkeep the long-term negative consequences of phobia (Tolin, Lohr, Lee, & Sawchuk, 1999). In such cases, technological tools could be used to facilitate attention towards distressing stimuli. For example, augmented reality technologies allow virtual information and objects to be added over images of the real physical world (e.g., by projecting images to the lenses of regular see-through eye glasses) (Azuma et al., 2001). Together with the tracking of eye gaze, such technologies could allow more control over the exposure to the stimuli (e.g., time and length of exposure), for example, by always having the object appear where the participant turns her or his gaze at.

The next stage of the model of Gross (1998) involves cognitive change, that is, changing the meaning that the situation has for the person. According to Gross and Thompson (2007), cognitive change can be directed to either appraisal of the situation as such (i.e., attitudes towards the situation) or the person's own perceived capabilities for managing the situation. Changing appraisals as they happen (i.e., in real-time) can be quite challenging, so perhaps the greatest potential for technology to change cognitions may lie in longer-term assistance. As an example of such longer-term computer-assisted cognitive change, an Internet-based guided exposure system called FearFighter (CCBT Ltd., 2011) has been implemented for self-help therapy of anxiety disorders (Marks et al., 2007). This system coaches a person through a step-wise programme that aims to help in changing fear and anxiety related cognitions through controlled self-exposure. However, although the computer system provides the structure for the therapy and tools for documenting it, this system still relies on human supervision. A completely automated system that also monitors and responds to the person's reported progress would require more sophisticated artificial intelligence. Perhaps such methods can be implemented in the longer run.

Finally, the elicited experiential, behavioural, and physiological responses may be regulated. Technological support for the regulation of physiological responses has a long history in the field of *biofeedback*. In biofeedback, the person receives information (e.g., visual or auditory computer feedback) about her or his physiological processes. The aim is that the person will become more aware of involuntary and unconscious processes and ultimately learn to control them. Traditional applications of biofeedback include the reduction of muscle tension in order to treat headache and motor rehabilitation after a stroke (Nestoriuc, Martin, Rief, & Andrasik, 2008; Tassinary & Cacioppo, 2000). However, biofeedback can also be effective for modulating emotional responses, as physiological responses are coupled with other components of emotional responding. For example, there is evidence that biofeedback can facilitate the reduction of fear of flying as well as reduce the number of remissions (i.e., re-occurring symptoms after treatment) (Wiederhold & Wiederhold, 2003).

There is evidence that the regulation of behaviour can also have significant effects on how a person feels or perceives emotional information (Cacioppo, Priester, & Berntson, 1993; Flack Jr., 2006; Marsh, Ambady, & Kleck, 2005). For example, Ahn, Teeters, Wang, Breazeal, and Picard (2007) implemented a robotic monitor that could regulate its user's posture by moving to a different position. The results showed that when the monitor promoted a posture that was congruent with the user's emotional state, their persistence in a problem solving task was facilitated. This kind of a system could take a more proactive approach to regulating emotions by directly modifying the situation itself (e.g., monitor position) instead of providing feedback (e.g., visualization of physiological signals) to promote changes in emotional responding.

In addition to influencing the physical environment (e.g., position of robotic monitors), computer systems can also provide more direct influence over the antecedents of emotion (i.e., situation and stimulation) by the adaptation of virtual stimulation. For example, using virtual reality it is possible to create immersive computer-generated environments which potentially allow control over both fine-grained details (e.g., the ways in which bugs move on the ground) and more global aspects (e.g., the weather). Such environments have been successfully used as a replacement for real-world exposure to distressing stimuli in the treatment of different anxiety disorders, including fears of flying and spiders (Gerardi, Cukor, Difede, Rizzo, & Rothbaum, 2010; Powers & Emmelkamp, 2008). Further, as computers are becoming an increasingly permanent part of our work and living (e.g., ranging from office work with computers to buying gasoline from an automated machine), the significance of taking the user's emotions into account in these everyday human-technology interactions is also growing. Although such interactions normally happen through less immersive channels of communication (i.e., compared to virtual reality), effective social and emotional cues could be designed for such simpler interfaces as well.

This way, technology could offer a way to modulate how emotional responses affect the (real or virtual) situation, that is, affect the link from the emotional response to the new emerging situation in the model of Gross (1998). Further, through the development of perceptually and expressively intelligent technologies, natural human responses could be exploited in these affective loops by adapting situations intelligently based on automatically perceived responses. Such *implicit interaction* between the system and the person could provide means to construct truly cybernetic systems for emotion regulation. A central principle of cybernetics is that feedback from a process enables it to be controlled (Wiener, 1948). In the case of computer-assisted emotion regulation, a cybernetic system could achieve a kind of an emotional equilibrium through tight coupling (i.e., reciprocal feedback) between the system and human emotional responding (Rani et al., 2003). A more detailed model for such systems was presented in Publication II of the present thesis.

3.2.2 Applications for Computer-Assisted Emotion Regulation

The field of affective computing covers many areas of applications that utilize perceptual and expressive capabilities but are not aimed at regulating emotion per se (Picard, 1997; Picard et al., 2001; Picard & Klein, 2002). For example, expressive intelligence could be used to facilitate the acceptability of domestic robots by making them more human-like (Nomura, Kanda, Suzuki, Yamada, & Kato, 2009). Similarly, perceptual intelligence could be used for many purposes, such as, to facilitate the mediation of emotional information from a patient to a health-care provider, the evaluation of positive and negative responses to computer games, and the assessment of Web page usability (Hazlett, 2003, 2006; Lisetti, Nasoz, LeRouge, Ozyer, & Alvarez, 2003; Ward & Marsden, 2003). However, some recent prototype systems can be seen to provide suggestions about the potential of applying technology to the more specific area of computer-assisted emotion regulation.

A contemporary step towards computer-assisted emotion regulation is exemplified in computer games that adapt their own functionality based on physiological responses (Kuikkaniemi, Kosunen, Turpeinen, Laitinen, & Lievonen, 2010; Sakurazawa, Yoshida, & Munekata, 2004). The aim of this adaptation is to provide a suitable level of challenge (i.e., not too easy or hard) for each individual person and thus make the games more enjoyable. The ultimate aim for such a system could be achieving and maintaining a *flow experience*, that is, immersion to the task through complete alignment of emotions with the task (Csikszentmihalyi, 1990). The goal of reaching an experience of flow (i.e., maximizing positive emotion) has been explicitly stated for some computer-assisted learning systems, which aim to adjust their functioning according to perceived emotion, for example, based on skin conductivity and facial expressions (e.g., Burleson & Picard, 2007; D'Mello et al., 2007). In-car systems that aim to avoid extremely negative emotions (e.g., road rage) while driving a car provide a somewhat opposite (i.e., reducing the intensity of negative emotions) and a more dramatic example of potential real-time systems for emotion regulation (Eyben et al., 2010; Nasoz et al., 2010).

Failure to regulate the intensity of emotional responses may also contribute to the development and sustenance of different anxiety disorders, for example, specific phobias (Amstadter, 2008). Thus, explicit facilitation of emotion regulation could be beneficial for the treatment of these disorders as well. For example, there is some evidence that the use of biofeedback to facilitate awareness of physiological responses can further promote the effectiveness of virtual exposure therapy in treating anxiety disorders (Wiederhold & Wiederhold, 2003). However, such systems so far rely on human operation in monitoring and explicitly adjusting own behaviour according to the measured signals, while truly automated cybernetic adaptation (i.e., a loop of perceptual and expressive intelligence) is yet to be investigated in this context.

To reduce milder forms of anxiety, Prendinger and Ishizuka (2005) presented a system that could perceive emotional reactions using physiological sensors and provide empathetic feedback for the person during a potentially stressing virtual job interview. In another setup, Prendinger et al. (2003) used the same system to implement an *Emotion Mirror* that directly reflected the automatically perceived emotions back to the person using virtual characters that showed both verbal and non-verbal emotional cues. In both setups, the aim was to facilitate a person's own competences in emotion regulation by allowing virtual training in real time. Such systems could benefit, for example, job seekers in preparing for interviews. However, the effectiveness of these setups is yet to be demonstrated as the results of Prendinger and Ishizuka (2005) did not reveal any significant positive effects for receiving the empathetic as compared to non-empathetic feedback.

In addition to facilitating real-time interaction, the capture and storing of data about emotional responses could also provide opportunity for later reflection, that is, offer a kind of an *affective diary* (Lindström et al., 2006). This kind of a system could be used to facilitate awareness of own emotion related responses in the long-term, augmenting systems that support awareness of short-term responses (e.g., biofeedback). Exercise and fitness management systems that engage a person through personal feedback (e.g., heart rate monitoring) and social involvement provide current examples of widely adopted systems for physiological awareness (Sirkiä, 2010). As health and weight may be emotionally sensitive issues, it is evident that such systems could gain significant benefits from being emotionally intelligent and supportive. In addition to facilitating more common aims of health management and promotion, this kind of support for regulation of exercise and weight related emotions could also potentially benefit more extreme cases, such as, the mitigation of eating disorders that are related to emotion dysregulation (Gilboa-Schechtman, Avnon, Zubery, & Jeczmien, 2006).

Based on the examples discussed above, there are several promising application areas that could significantly benefit from technology-assisted emotion regulation. These applications cover diverse scenarios from automatic adaptation based on short-term activity (e.g., games) to longer-term reflection of emotional responding (e.g., affective diary), and from implicit interaction using spontaneous emotional reactions (e.g., empathetic agents) to facilitating voluntary changes through explicit effort (e.g., Emotion Mirror). However, perhaps greatest potential for breakthrough technologies is in application areas that are difficult to anticipate. Thus, the full extent of benefits from computer-assisted emotion regulation may only be realized after the first generation of prototypes demonstrates the feasibility of these technologies and seeds future innovations.

3.2.3 Summary

A review of the most recent technologies and computer systems suggested that there is significant potential for technology to facilitate emotion regulation. There are several existing end-user products and prototype solutions that could already provide tools for facilitating every stage of emotion regulation from antecedents of emotion to the consequent emotional responses. Technology could also provide novel opportunities to extend the means of regulating emotion by allowing the environment to be directly influenced (e.g., background music or facial expressions of virtual computer characters) in order to adapt to naturally occurring human responses. This is in contrast to more traditional approaches, for example, biofeedback, which rely on an individual person's competences to adapt her or his functioning according to computer feedback. More proactive approaches, on the other hand, would have significant potential for achieving truly natural emotional human-technology interaction through the development of more intelligent computer perception and expression of emotional cues (Tennenhouse, 2000). Further, the wide variety of potential applications suggests that there are clear needs and possibilities for improving the quality of life by facilitating emotion regulation in a variety of common every-day situations as well as in special scenarios, such as, the treatment of anxiety and eating disorders.

However, several challenges remain to be solved for the first generation of computer-assisted emotion regulation systems, especially if they are to be widely adopted for distinct application domains. First, previous work has not yet led to widely accepted and feasible models for implementing robust systems for computer-assisted emotion regulation. For example, although studies in laboratory conditions have shown that computers can even exceed the accuracy of humans in recognizing some emotional cues, the performance in diverse real-world settings is likely to suffer and may never reach complete accuracy. Existing system architectures (e.g., Lisetti et al., 2003; Rani et al., 2003) often do not explicitly offer practical ways to compensate and account for these inaccuracies in emotion recognition. Second, there is a need for practical, non-invasive ways to measure and influence social and emotional responses in order to promote wider applicability of emotion regulation systems. For example, traditional ways to measure human physiology using wired electrodes may be inconvenient for daily use in domains such as car driver monitoring. Third, the functionality of systems that facilitate emotion regulation in real-time is yet to be demonstrated. These challenges provided a background and a research agenda for the present thesis.

4 The Contributions of the Publications

4.1 Theoretical Framework for Computer-Assisted Emotion Regulation

4.1.1 Publication I: Emotions in Human-Computer Interaction

Publication I aimed to establish a theoretical framework for integrating emotions into the development of functional and natural human-technology interaction by reviewing and structuring previous work related to this area. In the course of this work, studies of expressive intelligence were established as a first priority in order to establish that humans actually respond to computer-generated stimulation in a social and emotional way. Then, methods for intelligently perceiving and reacting to these elicited responses could be created. Finally, complete affective loops could be formed by continuously perceiving user's reactions to synthetic stimuli and adapting these computer-generated expressions accordingly.

The previous work reviewed in Publication I highlighted the value of emotionally intelligent computers in many ways. The reviewed studies suggested that emotional communication could be used to facilitate human-technology interaction by tapping automatic and rapid human responses to such information. More generally, studies of emotions in human-computer interaction were argued to represent the continuation of a well-established research tradition on human emotions in at least two ways. On one hand, since humans seem to respond to computers in a social and emotional way, technological tools may allow more fine-grained and controlled investigations of the characteristics of human-to-human social and emotional communication. On the other hand, technology is becoming increasingly pervasive in our everyday lives, which means that human-technology interaction already forms a major part of our everyday lives. These interactions will in any case evoke emotions that can either be ignored or studied.

Synthetic human-like characters were identified as one of the most promising venues for research on computer-generated emotional expressions, because such simulations allow a wide range of social cues to be used. However, it was also noted that much work remains to be done before such technologies can extensively replicate human appearance and behaviour. In general, there are several potentially powerful social and emotional cues that could be used in human-technology interaction, but we are still far from understanding them in detail and using them effectively. For example, although both positive and negative information seems to have significant effects on human cognitive and emotional processes, especially the use of negatively toned messages may require indepth knowledge and sensitivity (e.g, to appropriate timing) in order to avoid being rude.

The reviewed approaches to perceptual intelligence suggested that especially physiology-based systems show potential towards supporting implicit modes of interaction. Recent developments in wireless, wearable, and unobtrusive technologies were seen to suggest that physiological measures will soon become widely used. Such technologies will allow monitoring of core emotional responses, which are based on neurophysiological and physiological changes that are normally inaccessible to human perception of emotion. Thus, it was argued that through technological advances computers could gain better access to our emotional processes than other people. However, it was also noted that development of such technology will require several iterations of hardware and software (e.g., signal processing algorithms for perceptual intelligence).

In sum, Publication I provided a framework for structuring research on emotions in human-computer interaction. The development of emotionally intelligent systems will require work on both perceptual intelligence and expressive intelligence. Finally, the merging of these two fields will allow functional loops of implicit and smooth emotional interaction to be created. Several topics for further research were identified in the process of this work, including further refinement of synthetic social and emotional cues, more detailed studies of their effects and how they would fit into systems with complete affective loops, and developing more accurate and practical (e.g., wearable and wireless) methods for perceptual intelligence.

4.1.2 Publication II: Computer-Assisted Regulation of Emotional and Social Processes

Publication II aimed first to extend the framework of Publication I by establishing virtual exposure therapy, that is, treatment of anxiety disorders using controlled exposure to computer-generated stimuli, as one of the most promising application domains that would benefit from computer-assisted emotion regulation. The second task was to design a robust model for the first generation of computer-assisted emotion regulation systems. The first priority for this task was to design a model that would be practical to use in the course of virtual exposure treatments.

The field of exposure treatments of anxiety disorders was found to be suitable as a test-bed for computer-assisted regulation of emotions based on several different aspects. First, anxiety disorders are a significant public health issue that could benefit from more effective and accessible (e.g., over-distance, Internet-mediated) methods of treatment. Second, exposure therapy sessions are performed by perceiving emotional reactions and adapting the stimulation in real-time, which would be a suitably challenging setting for testing the efficiency of computerized methods. Third, some preliminary evidence was found to suggest that by incorporating computers into the treatment procedures (e.g., by using virtual reality and biofeedback) the process and results of treatment could be facilitated. Especially the use of physiological measures of emotional responses was found to be promising for the purposes of finding exact timing of responses as well as for reducing the distractions caused as compared to monitoring the level of anxiety using conventional methods (e.g., self-reports).

However, the need for defining a more detailed and robust model for computer-assisted regulation of emotions became clear when reviewing the existing methods of perceptual and expressive intelligence in the context of virtual exposure therapy. In particular, it became evident that a critical aspect of exposure therapy would be to facilitate the training of emotion regulation in order to achieve greater competence in self-regulation after the exposure sessions. Thus, the emphasis should be somewhat different from the more distant goal that was formulated in Publication I, that is, a computer system that could regulate the person's emotions with little intervention. Further, although computers could potentially take much responsibility for effective administration of therapy in the future, the first plausible step in integrating computer systems to present therapy sessions would involve the creation of supportive tools for the patient and the therapist (e.g., to facilitate human perception of emotional responses). Human supervision would likely be needed, for example, in order to compensate for the imperfect real world performance of automatic emotion recognition and analysis methods.

The above considerations led to the development of the model presented in Publication II (Fig. 4.1). The model includes perceptual intelligence for extracting emotional responses from several measures, proactive reasoning that aims to adapt the virtual stimulation intelligently, and a human supervisor for the system. Proactive reasoning in the model covers both the adaptation of the stimuli according to knowledge about human social and emotional responses (i.e., expressive intelligence) and explaining the current state of the system (e.g., including perceived emotions of the person and the reasoning behind stimulus adaptation) to a human supervisor.

It was suggested that systems that are built according to the model presented in Publication II could prevent a human supervisor from being overwhelmed by the wide variety of detailed measures of emotion. The model would also maintain the supervisor in control of the system by allowing direct intervention to the stimulation. On the other hand, this kind of a system could normally function without human supervision by using an implicit cybernetic loop from automatic perception to synthetic expression of emotions (see Figure 4.1), while both the supervisor (e.g., therapist) and the person being treated could mainly focus on human-human interaction (e.g., coaching the person to relax and breathe calmly or otherwise support the regulation of emotions).

Although the work presented in Publication II revealed that computer systems have great promise for improving regulation of emotions during virtual exposure, several open questions in the development of these kinds of systems were also identified. For example, previous work on automatic analysis methods for emotional responses has focused on classifying emotions according to categories or emotional tone (see Bailenson et al., 2008, for an exception), while exposure therapy is based on adapt-

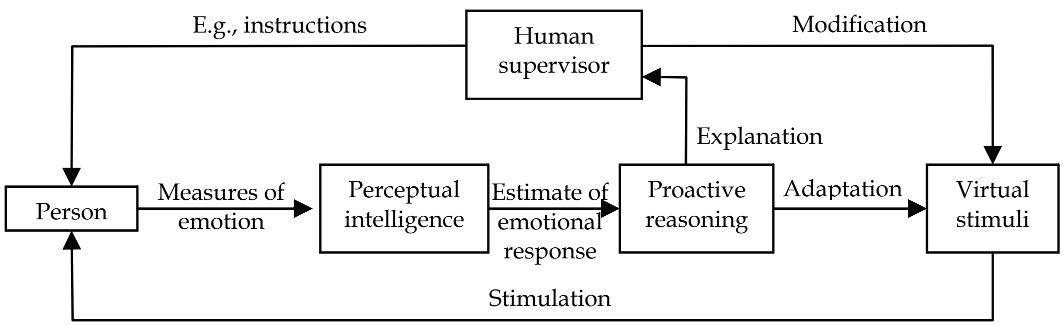


Figure 4.1: A proactive model for virtual exposure therapy. Reprinted from Publication II.

ing the stimulation according to the intensity of the emotion (e.g., level of distress). Thus, there is a need to create automatic analysis for emotional intensity in order to create perceptual intelligence for this purpose. Further, expressive intelligence for this field will require a more in-depth analysis of the effects of different kinds of adaptive computer-generated stimulations.

4.2 Perceptual and Expressive Intelligence

4.2.1 Publication III: Measuring Bodily Responses to Virtual Faces with a Pressure Sensitive Chair

The aims of the study reported in Publication III were twofold. First, while established behavioural and physiological measures of emotion (e.g., facial expressions and heart rate) have already been extensively studied for computer perception of emotion, there is still a need for less invasive and potentially ubiquitous measurements. Such measures could promote wider use of these technologies in various contexts where conventional (e.g., wired) measurements are not practical. Second, the study was one of the first to investigate basic body movement responses to virtual computer characters (see Alatalo, Juhola, Surakka, & Tossavainen, 2009, for related work). A special office chair was used for covert (i.e., unnoticed by the participants) measurement of body movement responses during computer-generated stimulation with virtual characters displaying unpleasant, neutral, or pleasant facial expressions.

The results showed that covert pressure sensors (i.e., unnoticed by each participant) were able to measure significant body movement responses elicited by the virtual computer characters. Participants leaned forward in response to each stimuli and their body sway in the forward-backward axis was reduced as compared to before onset of stimulation. Participants' postures remained more forward leaning during emotionally engaging (i.e., unpleasant and pleasant) stimuli as compared to neutral stimuli. Longest forward leaning response was elicited by the unpleasant stimuli.

Thus, the results showed that different facial expressions of virtual computer characters elicited significantly different body movement responses. More generally, the results suggested that novel body movement measurement technologies could offer a potential tool for recognizing behavioural responses to synthetic social and emotional cues. In particular, embedded pressure sensors could provide one convenient and unobtrusive way of measuring behavioural responses in HCI. Such easily applied technologies could help to adopt similar measurements more widely as well as promote the measurement of more spontaneous emotional responses (e.g., reduce awareness of the measurement and thus voluntary regulation of body movements). On the other hand, these kinds of technologies could support the regulation of body postures by allowing more convenient measurement of voluntary movements as well.

4.2.2 Publication IV: Virtual Proximity and Facial Expressions of Computer Agents Regulate Human Emotions and Attention

The study reported in Publication IV aimed to investigate simulated proximity and facial expressions of computer characters as potential social and emotional cues for regulating human-computer interaction. For the purposes of expressive intelligence, proximity could be an attractive cue for regulating emotions as proximity could be easily combined to different stimulations (e.g., several different characters). On the other hand, intuitively it seems clear that proximity is a strong cue for attention, for example, closer (i.e., larger) stimuli could also be expected to grab the attention more effectively than further away (i.e., smaller) ones. Such attentional effects could be a potentially powerful tool for expressive intelligence to affect human behaviour. On the simplest level, it is quite clear that some attention must be paid to a stimulus for it to have any effects. For example, there is evidence that attentional engagement during exposure to distressing stimuli may facilitate the habituation of fear (Foa & Kozak, 1986). Thus, in order to provide an empirical basis for utilizing these effects in computer-assisted regulation of emotions, the attentional effects of computer characters were investigated in the present study using a tentative set of subjective rating scales.

Similar to the work of Partala et al. (2004) where a virtual humanlike head was used as the stimulus, the size of computer characters was varied in order to simulate different proximities, that is, bodily distances between the character and the participant. As an extension to this previous work, in the present study full-body agents displaying unpleasant, neutral, and pleasant facial expressions were used, physiological responses (i.e., heart rate and facial EMG) were measured, and ratings of subjective experiences of attention (i.e., conspicuousness, concentration, and interestingness) in addition to emotional ratings (i.e., valence, arousal, and dominance) were collected. Principal component analyses were used to explore the patterns of subjective ratings in relation to theoretical expectations.

The analyses revealed that the variation of attentional ratings was adequately represented by two components. The components seemed to closely match the hypothesized factors of *stimulus-driven attention*, which was associated with high conspicuousness, interestingness, and distraction (i.e., low concentration), and *voluntary attention*, which was associated with high concentration and interestingness ratings. Emotional ratings mostly varied according to one component that was labelled as *comfort*, ranging from an alerting (i.e., unpleasant, aroused, and dominating) to a comfortable (i.e., pleasant, calm, and controlled) experience. Figure 4.2 shows the results of this analysis using the data published in Publication IV but with a new visualization to illustrate the associations between the ratings of emotional and attentional experience. In this figure, mean responses to different types of stimuli are represented by points in a two-dimensional space of comfort and stimulus-driven attention.

The related but distinct attentional and emotional effects of proximity and facial expressions are illustrated in Figure 4.2 in at least three ways. First, the different stimulus categories (i.e., points) are clearly separated from each other, that is, they each occupy their own space and do not overlap in the visualization. This suggests that different combinations of proximity and facial expression could be used (e.g., by expressive intel-

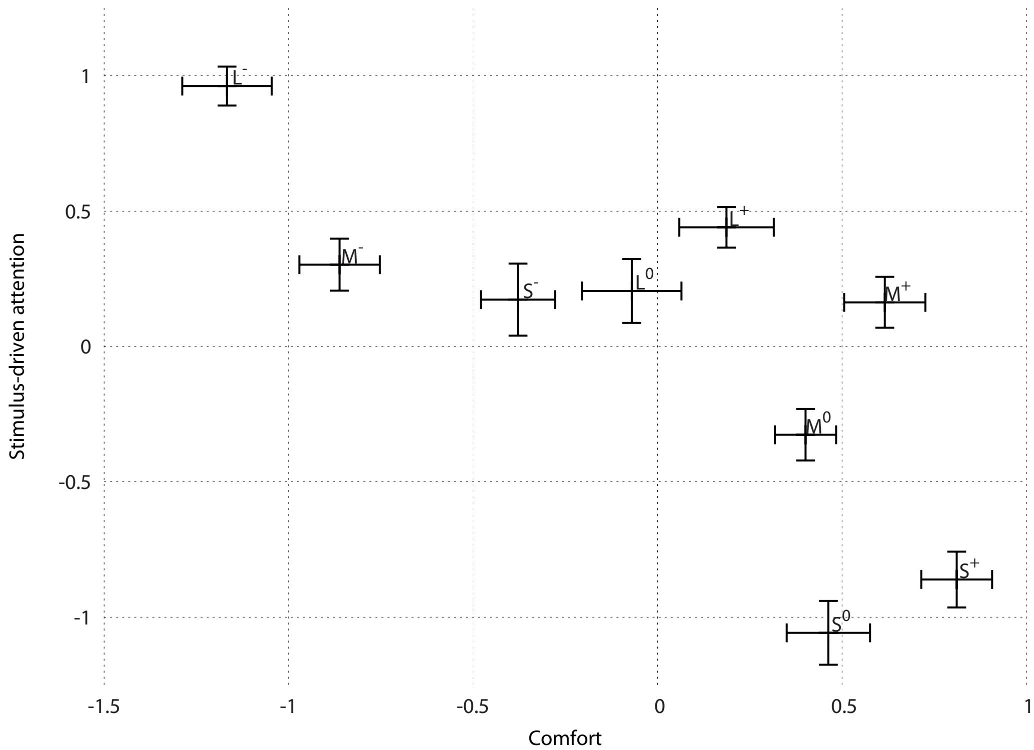


Figure 4.2: Mean comfort and stimulus-driven attention scores for different facial expressions and sizes of the computer characters. The error bars correspond to the standard errors of each mean. S = small, M = medium, L = large size character. - = unpleasant, 0 = neutral, + = pleasant facial expression. Based on data reported in Publication IV.

ligence) to elicit distinct subjective experiences of attention and emotion. Second, there however appeared to be a direct connection between emotion and attention, since there was a clear trend in that alerting stimuli received more stimulus-driven attention (e.g., unpleasant large characters, L^-) and comfortable stimuli received less stimulus-driven attention (e.g., small pleasant characters, S^+). Third, the effects of proximity and facial expression were dependent on each other. For example, if we consider that the facial expression of the character is fixed to neutral, then increasing its size from small (i.e., S^0) to medium (i.e., M^0) leads to a higher level of stimulus-driven attention. However, if the expression was fixed to unpleasant, then similar increase of size from small (i.e., S^-) to medium (i.e., M^-) would have little effect on attentional ratings, but a notable effect on emotional ratings (i.e., decreasing comfort).

The results also showed significant physiological responses to both the simulated proximity and the facial expressions of the computer characters. Proximity affected heart rate which was higher when viewing large characters as compared to medium sized characters. The acceleration of heart rate may have reflected preparation for action, as closer or larger stimuli need to be resolved more quickly than smaller and farther away ones. The facial expressions of the characters affected frowning related *corrugator supercilii* EMG activity. *Corrugator supercilii* EMG responses were suppressed during unpleasant stimulation and increased during pleasant stimulation, which was a somewhat surprising effect as it was opposite to the emotional tone of the characters. However, there is evidence that reactions to human facial expressions may also vary between different contexts, for example, depending on affiliation with the expresser (Bourgeois & Hess, 2008).

In sum, the results reported in Publication IV showed that simulated proximity and facial expressions of computer characters could elicit significant subjective and physiological effects. Thus, both artificial cues could be effective as methods of expressive intelligence. In particular, virtual proximity was found to be an effective social and emotional cue which can be hypothesized as easy to simulate in different characters. Emotional and attentional responses to facial expressions, on the other hand, were interacting with size as well as with each other. This result underlines that expressive intelligence needs to account for both the individual and the joined effects of concurrently presented social and emotional cues. More generally, the results highlighted that both perceptual and expressive intelligence will need to be sensitive to other context (e.g., attention and social aspects like affiliation) besides the effects of individual social and emotional cues.

4.3 Computer-Assisted Emotion Regulation

4.3.1 Publication V: Facial Activation Control Effect (FACE)

The aim of the study reported in Publication V was to investigate voluntary facial activations as a method for controlling emotion related subjective and physiological responses during human-technology interaction. The focus was especially on the effectiveness of simple computer-instructed activations of single facial muscles in regulating more spontaneous physiological processes. In each task, participants activated either the *corrugator supercilii* or the *zygomaticus major* muscle according to visual feedback and text instructions given by the computer. Heart activity was wirelessly measured during the tasks and subjective ratings of emotional valence and the difficulty of the tasks were collected following the tasks.

The results showed that different muscle activations produced both task-specific emotional experiences and significant changes in heart activity. Heart rate decelerated during most activations, with the exception of high intensity (i.e., above 60% of maximal EMG activity) level activations of either muscle during which the mean heart rate remained close to a pre-task baseline. Both low and high frequency heart rate variability were suppressed during every task, which suggests that the general level of autonomic arousal (i.e., both sympathetic and parasymphatetic activity) was decreased. Low intensity activations of either muscle elicited larger heart rate changes and were rated as easier to perform and as more pleasant as compared to medium and high intensity activations.

The above results suggest that voluntary facial activations could be an effective method for influencing more spontaneous physiological processes. Although the presently used tasks were quite simple (c.f., Coan et al., 2001; Levenson et al., 1990), they had significant effects on heart rate deceleration and autonomic relaxation. The present setup that allowed participants to control their facial activity according to computer instructions and feedback was also quite simple in terms of required preparation and training. Thus, a similar setup could be feasible for use in other contexts as well. In sum, computer-assisted moderate intensity facial activations could be an effective, a pleasant, and a practical method for regulating emotion related physiological activity during human-technology interaction.

4.3.2 Publication VI: Voluntary Facial Activations Regulate Physiological Arousal and Subjective Experiences during Virtual Social Stimulation

The study reported in Publication VI aimed to investigate voluntary facial activations as a method for regulating physiological and subjective responses to computer-generated social stimulation. Participants with either low or high level of social anxiety performed trials in which they activated either the *corrugator supercilii* or the *zygomaticus major* muscle in order to keep a female or a male computer character walking towards them. Once the character reached a pre-defined distance, it used speech synthesis to deliver an arithmetic task for the participant to answer. Electrodermal activity (i.e., changes in electrical conductivity of the skin) was recorded during the trials and subjective ratings of valence, arousal, and dominance were collected before and after each trial.

The results showed that both *corrugator supercilii* and *zygomaticus major* activations had significant effects on physiology and emotional experiences. The long-term level of skin conductance decreased during all tasks. However, *corrugator supercilii* activations significantly facilitated the rate of decrease, as the level of skin conductance was lowest after these tasks. Further, both *corrugator supercilii* and *zygomaticus major* activations enhanced the magnitude of short-term changes in skin conductance as compared to tasks without facial activation. Subjective ratings given by the less and the more socially anxious participants showed completely opposite patterns. Ratings of the more socially anxious participants decreased in pleasantness, increased in arousal, and decreased in subjective dominance, while the less socially anxious rated their experience as more pleasant, less aroused, and more in control of the situation after *zygomaticus major activations*.

The electrodermal findings suggest that voluntary facial activations were effective in regulating more spontaneous physiological activity during simulated social communication. In the short-term, the magnitude of electrodermal responses was enhanced during both types of facial activations, which suggests that facial activity increased sympathetic arousal. On the other hand, longer term changes in electrodermal activity suggested that sympathetic arousal decreased at a higher rate after *corrugator supercilii* activations. These effects may seem contradictory at first hand. However, similar results have been reported when relaxation has been facilitated with other methods (e.g., controlled breathing), for example, during exposure therapy (Foa & Kozak, 1986). Thus, it may be that physiological activation in the short-term will enhance long-term habituation to the stimulation.

Subjective ratings after *zygomaticus major* activations showed that socially anxious participants did not feel comfortable in smiling to the computer characters. This suggests that virtual social communication similar to the present one could be used in eliciting social anxiety related responses. Further, the present setup required relatively little preparation in terms of technical setup and training. In sum, Publication VI presented a platform that could be used to implement a potentially effective and practical setup for training the regulation of social anxiety also in a less controlled setting, for example, during clinical sessions.

5 Discussion

The present thesis showed that computer-assisted emotion regulation can be feasible and potentially very effective. The theoretical basis for implementing such systems was laid out in Publications I and II. Publication I discussed the potential benefits and methods of regulating emotions in the more general field of human-computer interaction. Publication II extended these conclusions by presenting a framework for building the first generation of human-technology systems for emotion regulation.

The resulting model for computer-assisted emotion regulation (see Figure 4.1) can be compared with the more general model of human regulation of emotion as presented by Gross (1998) (see Figure 2.4). The presently suggested model supports emotion regulation by applying two general technologies: perception of emotional responses (i.e., perceptual intelligence) and adaptation of virtual stimulation (i.e., expressive intelligence). It is quite straight-forward to apply the former to support the person's awareness of her or his own responses (e.g., using biofeedback), which is clearly in line with the voluntary regulation of those responses (i.e., *response modulation* in the model of Gross, 1998). Adaptation of stimulation, on the other hand, can be seen as a way of affecting the antecedents of emotion, that is, as *situation modification*. Truly intelligent expression could also affect the following stage of Gross's model (i.e., *attention deployment*) by expressing (virtual) cues that regulate attention appropriately.

The empirical part of the thesis established virtual proximity and facial expressions of computer characters as relatively simple and effective emotional and attentional cues for expressive intelligence. Both proximity and facial expressions were found to elicit significant subjective, behavioural, and physiological responses. Facial expressions were found to affect body movement responses in Publication III and subjective ratings of emotion and attention as well as electrical facial muscle activity in Publication IV. It was also reported in Publication IV that, similar to facial expressions, virtual proximity significantly affected both emotional and attentional ratings as well as physiology in terms of changes in heart rate.

Thus, it seems that virtual proximity and facial expressions could be used quite directly to create stimulations, or more generally situations, that elicit desirable (e.g., goal-congruent) emotions. This way, virtual stimulation enables relatively effortless *situation modification* using easily synthesized social and emotional cues. Proximity and facial cues were also found to have potential for significantly affecting the next stage of emotional responding, that is, *attention deployment*, in terms of self-reported levels of attention. In addition to the present promising results, it is likely that such cues could be further enhanced by creating more detailed and vivid simulations of human expression. For example, contemporary computer models are capable of displaying not only facial movements but also other small changes in the face, such as, skin wrinkles that result from these movements (Courgeon et al., 2009). Such fine-grained synthetic facial cues could offer a way to evoke rapid and strong reactions to the human face, such as, different responses to smiles that signal genuine enjoyment as compared to other types of smiles (e.g., fake enjoyment or social agreement) (Ekman et al., 1990). On the other hand, the present results showed that effective — socially and emotionally meaningful — virtual cues can already be created using contemporary technology.

Publications V and VI showed that computer-assisted regulation of facial behaviour could affect more spontaneous physiological processes and, thus, provide a potential method for regulating emotion related physiological responding. The setup of these experiments can be seen as an example of an effective method of *response modulation*, that is, the final stage in the model of Gross (1998). In particular, publication VI demonstrated a setup where virtual social stimulation was controlled using voluntary facial activity. The results showed that facial activity had significant effects on autonomic arousal as well as on subjective ratings which were opposite for the less and the more socially anxious participants. Thus, this kind of a setup could be used for eliciting and regulating both experiential and physiological responses related to emotions in general and social anxiety in particular. The platform constructed for Publication VI could as such provide a starting point for implementing exposure sessions for training the regulation of social anxiety.

In addition to the regulation of facial behaviour, computer systems could be used to assist in regulating other emotion related behaviours as demonstrated in Publication III. In this study, a pressure sensitive chair was presented as a potential method for unobtrusive measurement of body movements. Such easily applied measurement technologies could facilitate wide use of systems for monitoring and regulating bodily behaviour. The regulation of body movements is a particularly promising approach, as body posture has been found to significantly affect emotional experiences, influence judgements, and facilitate problem solving (Ahn et al., 2007; Cacioppo et al., 1993; Flack Jr., 2006). On the other hand, there have been several recent advances in wireless and wearable physiological measurement technologies that monitor, for example, electrical facial muscle activity, movements of facial skin, brain activity, and heart activity (Nöjd et al., 2005; V. Rantanen et al., 2010; Vehkaoja & Lekkala, 2004, 2006). Thus, there is growing potential for wider adoption of perceptually intelligent technology that provide the basis for regulating behavioural and physiological responses.

The present work covered the process of emotional responding from the antecedents of emotion (i.e., the eliciting situation) to the resulting experiential, behavioural, and physiological responses. Computer support for regulating emotions by facilitating the selection of appropriate situations (i.e., first stage in the model of Gross & Levenson, 1997) was not empirically addressed as a part of this work. However, some approaches to provide support for this task, such as, music recommendation and composition systems, were suggested in the literature review of this thesis. Importantly, whether people choose to use computer-assisted systems for emotion regulation in the first place can also be seen as a way of selecting a situation (i.e., with or without computer assistance). These decisions will determine whether such systems become widely adopted or not. In the end, perhaps the most essential question is how people will perceive these computer systems that aim not only to influence but also to change their emotional functioning.

In sum, the present results provide a solid theoretical and empirical basis for computer systems that facilitate emotion regulation. Three central challenges for creating such systems were identified and addressed in this work. First, a theoretical framework for structuring the work in this area was presented in Publications I and II. In particular, Publication II presented a model that could be widely applied to create robust systems for computer-assisted emotion regulation. Second, practical and effective ways for measuring and influencing social and emotional responses were developed in Publications III and IV. Publication III focused on unobtrusively measured body movement responses, while Publication IV studied physiological and subjective responses to virtual cues of proximity and facial expressions. These technologies were shown to be promising approaches for developing unobtrusive perceptual and effective expressive intelligence that could facilitate wider acceptance and adoption of emotion regulation systems. Third, the first generation of easy-to-use setups for facilitating emotion regulation in real-time human-technology interaction were studied in Publications V and VI. In sum, the present empirical and constructive work demonstrated the potential of computer-assisted emotion regulation systems in terms of their easy adoption and effectiveness in influencing emotional responding on experiential, physiological, and behavioural levels.

6 Conclusions

The theoretical and empirical work of the present thesis addressed three themes: the theoretical basis for computer-assisted emotion regulation, perceptual and expressive intelligence, and real-time emotion regulation using computer stimulation and voluntary facial activations. The following conclusions were drawn from this work:

- Affective loops that allow implicit interaction with computers (i.e., without explicit commands or feedback) can be created by merging perceptual and expressive intelligence.
- Perceptual intelligence will become increasingly convenient, pervasive, and accepted through the proliferation of wireless and wearable technologies.
- The measurement of human body movements can be used to create truly unobtrusive perceptual intelligence (e.g., undetected by the person being measured) for monitoring social and emotional responses.
- Virtual proximity and facial expressions provide easy-to-apply, practical cues for expressive intelligence that significantly affects human social and emotional responses in a controlled fashion.
- Voluntary facial activations can be used as a simple but effective method for regulating experiential and physiological processes during socially and emotionally meaningful stimulation.
- Current technology allows real-time cycles of perceptual and ex-

pressive intelligence to be created for easy-to-use computer-assisted regulation of social and emotional processes.

These conclusions can be used for the development of the first generation of systems for computer-assisted emotion regulation. In particular, the present theoretical model has potential in leveraging such systems beyond command-based modes of interaction and towards more implicit affective loops. Further, the development of practical and effective perceptual and expressive intelligence is possible based on the empirical work of the thesis. These results provide a solid platform for utilizing the developed methods in more applied contexts, for example, as tools for learning to regulate excessive anxiety during exposure to distressing virtual stimulation.

BIBLIOGRAPHY

Adams Jr., R. B., Ambady, N., Macrae, C. N., & Kleck, R. E. (2006). Emotional expressions forecast approach-avoidance behavior. *Motivation and Emotion*, *30*, 179–188.

Addlesee, M., Curwen, R., Hodges, S., Newman, J., Steggles, P., Ward, A., et al. (2001). Implementing a sentient computing system. *IEEE Computer*, *34*(8), 50–56.

Ahn, H., Teeters, A., Wang, A., Breazeal, C., & Picard, R. (2007). Stoop to conquer: posture and affect interact to influence computer users' persistence. In A. Paiva, R. Prada, & R. W. Picard (Eds.), *ACII 2007* (Vol. 4738, pp. 582–593).

Alatalo, T., Juhola, M., Surakka, V., & Tossavainen, T. (2009). Creation of a virtual reality stimulus for a psychophysiological body balance measurement. *Psychology Research and Behavior Management*, *2*, 13-22.

Allanson, J., & Fairclough, S. H. (2004). A research agenda for physiological computing. *Interacting with Computers*, *16*, 857-878.

Amir, N., Beard, C., Taylor, C. T., Klumpp, H., Elias, J., Burns, M., et al. (2009). Attention training in individuals with generalized social phobia: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *77*(5), 961–973.

Amodio, D. M., Master, S. L., Yee, C. M., & Taylor, S. E. (2008). Neurocognitive components of the behavioral inhibition and activation systems: Implications for theories of self-regulation. *Psychophysiology*, *45*, 11–19.

Amstadter, A. (2008). Emotion regulation and anxiety disorders. *Journal of Anxiety Disorders*, *22*(2), 211 – 221.

Anttonen, J., Surakka, V., & Koivuluoma, M. (2009). Ballistocardiographic responses to dynamic facial displays of emotion while sitting on the EMFi chair. *Media Psychology*, *21*(2), 69-84.

Azevedo, T. M., Volchan, E., Imbiriba, L. A., Rodrigues, E. C., Oliveira, J. M., Oliveira, L. F., et al. (2005). A freezing-like posture to pictures of mutilation. *Psychophysiology*, *42*(3), 255–260.

Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., & MacIntyre, B. (2001). Recent advances in augmented reality. *Computer Graphics and*

Applications, IEEE, *21*(6), 34 – 47.

Bailenson, J. N., Beall, A. C., Loomis, J., Blascovich, J., & Turk, M. (2005). Transformed social interaction, augmented gaze, and social influence in immersive virtual environments. *Human Communication Research*, *31*(4), 511-537.

Bailenson, J. N., Pontikakis, E. D., Mauss, I. B., Gross, J. J., Jabon, M. E., Hutcherson, C. A., et al. (2008). Real-time classification of evoked emotions using facial feature tracking and physiological responses. *International Journal of Human-Computer Studies*, *66*(5), 303-317.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, *70*(3), 614-636.

Barrett, L. F. (2006). Solving the emotion paradox: categorization and the experience of emotion. *Personality and Social Psychology Review*, *10*(1), 20-46.

Barsade, S. G. (2002). The ripple effect: Emotional contagion and its influence on group behavior. *Administrative Science Quarterly*, *47*(4), 644 - 675.

Baumeister, R., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*(4), 323–370.

Beale, R., & Creed, C. (2009). Affective interaction: How emotional agents affect users. *International Journal of Human-Computer Studies*, *67*, 755 - 776.

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbifrontal cortex. *Cerebral Cortex*, *10*, 295-307.

Bell, C. (1865). *The anatomy and philosophy of expression as connected with the fine arts*. London: Bohn. Available from `http://www.archive.org/details/anatomyphilos00belluoft`

Bernhardt, D., & Robinson, P. (2007). Detecting affect from non-stylised body motions. In *ACII'07, LNCS* (Vol. 4738, p. 59-70).

Berntson, G. G., & Cacioppo, J. T. (2000). From homeostasis to allodynamic regulation. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 459–481). New York: Cambridge University Press.

Bishop, G. D. (2001). Emotions and health. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social & behavioral sciences* (p. 4454 - 4459). Oxford: Pergamon.

Bosse, T., & de Lange, F. P. J. (2008). Development of virtual agents with a theory of emotion regulation. In *IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology (WI-IAT '08)* (Vol. 2, pp. 461–468).

Bosse, T., Gratch, J., Hoorn, J., Portier, M., & Siddiqui, G. (2010). Comparing three computational models of affect. In *Advances in practical applications of agents and multiagent systems* (Vol. 70, pp. 175–184). Berlin / Heidelberg: Springer.

Bourgeois, P., & Hess, U. (2008). The impact of social context on mimicry. *Biological Psychology*, *77*(3), 343 – 352.

Bradley, M. M. (2000). Emotion and motivation. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed., p. 602-642). Cambridge, UK: Cambridge University Press.

Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion*, *1*(3), 276–298.

Bradley, M. M., & Lang, P. J. (1999a). *Affective norms for english words (ANEW): Stimuli, instruction manual and affective ratings* (Tech. Rep. No. C-1). Gainesville, FL, USA: The Center for Research in Psychophysiology, University of Florida.

Bradley, M. M., & Lang, P. J. (1999b). International affective digitized sounds (IADS): Stimuli, instruction manual and affective ratings (tech. rep. no. b-2) [Computer software manual]. Gainesville, FL, USA.

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

Bradley, M. M., & Lang, P. J. (2007). *Affective norms for english text (ANET): Affective ratings of text and instruction manual* (Tech. Rep. No. D-1). Gainesville, FL, USA: University of Florida.

Brave, S., Nass, C., & Hutchinson, K. (2005). Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent. *International Journal of Human-Computer Studies*, *62*, 161-178.

Burleson, W., & Picard, R. W. (2007, July/August). Gender-specific approaches to developing emotionally intelligent learning companions. *IEEE Intelligent Systems*, *22*(4), 62–69.

Busso, C., Deng, Z., Yildirim, S., Bulut, M., Lee, C. M., Kazemzadeh, A., et al. (2004). Analysis of emotion recognition using facial expressions, speech and multimodal information. In *Proceedings of ICMI* (pp. 205–211). ACM Press.

Butler, E. A., Wilhelm, F. H., & Gross, J. J. (2006). Respiratory sinus arrhythmia, emotion, and emotion regulation during social interaction. *Psychophysiology*, *43*, 612–622.

Cacioppo, J. T., Priester, J. R., & Berntson, G. G. (1993). Rudimentary determinants of attitudes: II. arm flexion and extension have differential effects on attitudes. *Journal of Personality and Social Psychology*, *65*(1), 5–17.

Cahn, J. E. (1990). The generation of affect in synthesized speech. *Journal of the American Voice I/O Society*, *8*, 1-19.

Carver, C. S. (2006). Approach, avoidance, and the self-regulation of affect and action. *Motivation and Emotion*, *30*, 105–110.

Carver, C. S., & Harmon-Jones, E. (2009). Anger is an approach-related affect: Evidence and implications. *Psychological Bulletin*, *135*(2), 183-204.

Castellano, G., Villabla, S. D., & Camurri, A. (2007). Recognising human emotions from body movement and gesture dynamics. In A. Paiva, R. Prada, & R. W. Picard (Eds.), *ACII 2007* (Vol. 4738, pp. 71–82).

CCBT Ltd. (2011). *FearFighter — panic and phobia treatment online.*

Chung, J., & Vercoe, G. S. (2006). The affective remixer: personalized music arranging. In *CHI '06 extended abstracts on human factors in computing systems* (pp. 393–398). New York, NY, USA: ACM.

Coan, J. A., & Allen, J. J. B. (2004). Frontal EEG asymmetry as a moderator and mediator of emotion. *Biological Psychology*, *67*(1–2), 7–49.

Coan, J. A., Allen, J. J. B., & Harmon-Jones, E. (2001). Voluntary facial expression and hemispheric asymmetry over the frontal cortex. *Psychophysiology*, *38*, 912-925.

Codispoti, M., & De Cesarei, A. (2007). Arousal and attention: Picture size and emotional reactions. *Psychophysiology*, *44*, 680–686.

Codispoti, M., Mazzetti, M., & Bradley, M. M. (2009). Unmasking emotion: Exposure duration and emotional engagement. *Psychophysiology*, *46*, 731-738.

Cohn, J. F., & Ekman, P. (2005). Measuring facial action. In J. A. Harrigan, R. Rosenthal, & K. R. Scherer (Eds.), *The new handbook of methods in nonverbal behavior research* (p. 9-64). New York: Oxford University Press.

Colibazzi, T., Posner, J., Wang, Z., Gorman, D., Gerber, A., Yu, S., et al. (2010). Neural systems subserving valence and arousal during the experience of induced emotions. *Emotion*, *10*, 377-389.

Computer rage: Reported acts of rage against computers. (2005). College Park, MD. Available from `http://lap.umd.edu/computer_rage/Tech_Report/written_comments/beh.html`

Cosker, D., Paddock, S., Marshall, D., Rosin, P. L., & Rushton, S. (2005). Toward perceptually realistic talking heads: Models, methods, and McGurk. *ACM Transactions on Applied Perception*, *2*, 270–285.

Coulson, M. (2004). Attributing emotion to static body postures: recognition accuracy, confusions, and viewpoint dependence. *Journal of Nonverbal Behavior*, *28*(2), 117–139.

Courgeon, M., Buisine, S., & Martin, J.-C. (2009). Impact of expressive wrinkles on perception of a virtual character's facial expressions of emotions. In Z. Ruttkay, M. Kipp, A. Nijholt, & H. Vilhjálmsson (Eds.), *Intelligent virtual agents* (Vol. 5773, p. 201-214). Springer.

Cowie, R., Douglas-Cowie, E., Taylor, J., Ioannou, S., Wallace, M., & Kollias, S. (2005). An intelligent system for facial emotion recognition. In *Proceedings of IEEE international conference on multimedia and expo* (p. 4 pp).

Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., et al. (2001). Emotion recognition in human-computer interaction. *IEEE Signal Processing Magazine*, *18*(1), 32–80.

Craig, S. D., D'Mello, S., Witherspoon, A., & Graesser, A. (2008). Emote aloud during learning with AutoTutor: Applying the Facial Action Coding System to cognitive-affective states during learning. *Cognition & Emotion*, *22*(5), 777 - 788.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience.* New York, NY, USA: Harper and Row.

Damasio, A. (1994). *Descartes' error: Emotion, reason, and the human brain*. New

York: Putnam.

Darwin, C. (1872). *The expression of the emotions in man and animals*. London: John Murray.

Davidson, R. J. (1993). Parsing affective space: Perspectives from neuropsychology and psychophysiology. *Neuropsychology*, *7*(4), 464–475.

Dawson, M. E., Schell, A. M., & Filion, D. L. (2000). The electrodermal system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (p. 200-223). Cambridge, UK: Cambridge University Press.

Dehn, D. M., & Van Mulken, S. (2000). The impact of animated interface agents: a review of empirical research. *International Journal of Human-Computer Studies*, *52*, 1-22.

Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science*, *11*(1), 86-89.

D'Mello, S., Graesser, A., & Picard, R. W. (2007). Toward an affect-sensitive autotutor. *IEEE Intelligent Systems*, *22*(4), 53-61.

Duchenne, G.-B. (1876). *Mécanisme de la physionomie humaine* (Dèuxieme ed.). Paris: J.-B. Bailliere et Fils.

Ekman, P. (1979). About brows: emotional and conversational signals. In M. von Cranach, K. Foppa, W. Lepenies, & D. Ploog (Eds.), *Human ethology.* Cambridge, UK: Cambridge University Press.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, *6*(3), 169–200.

Ekman, P. (2004). Emotional and conversational nonverbal signals. In J. M. Larrazabal & L. A. P. Miranda (Eds.), *Language, knowledge, and representation* (p. 39-50). Netherlands: Kluwer.

Ekman, P., Davidson, R. J., & Friesen, W. V. (1990). The duchenne smile: Emotional expression and brain physiology II. *Journal of Personality and Social Psychology*, *58*(2), 342-353.

Ekman, P., Friesen, W. V., & Hager, J. C. (2002). Facial action coding system: Investigator's guide [Computer software manual]. Salt Lake City, USA.

Elliot, A. J. (2006). The hierarchical model of approach-avoidance motivation. *Motivation and Emotion*, *30*(2), 111–116.

Ellsworth, P. C., & Smith, C. A. (1988). From appraisal to emotion: Differences among unpleasant feelings. *Motivation and Emotion*, *12*(3), 271-302.

Emotiv. (2010). *Emotiv — brain computer interface technology.* Available from `http://www.emotiv.com`

Eyben, F., Wöllmer, M., Poitschke, T., Schuller, B., Blaschke, C., Färber, B., et al. (2010). Emotion on the road — necessity, acceptance, and feasibility of affective computing in the car. *Advances in Human-Computer Interaction*, *2010*, 17 pages.

Facchinetti, L. D., Imbiriba, L. A., Azevedo, T. M., Vargas, C. D., & Volchan, E. (2006). Postural modulation induced by pictures depicting prosocial or dangerous contexts. *Neuroscience Letters*, *410*(1), 52–56.

Fasel, B., & Luettin, J. (2003). Automatic facial expression analysis: a survey. *Pattern Recognition*, *36*(1), 259 - 275.

Fehr, B., & Russell, J. A. (1984). Concept of emotion viewed from a prototype perspective. *Journal of Experimental Psychology: General*, *113*(3), 464–486.

Flack Jr., W. F. (2006). Peripheral feedback effects of facial expressions, bodily postures, and vocal expressions on emotional feelings. *Cognition and Emotion*, *20*(2), 177–195.

Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: Exposure to corrective information. *Psychological Bulletin*, *99*(1), 20-35.

Fowles, D. C. (1988). Psychophysiology and psychopathology: A motivational approach. *Psychophysiology*, *25*, 373-391.

Fredrickson, B. L., & Branigan, C. (2005). Positive emotions broaden the scope of attention and thought-action repertoires. *Cognition and Emotion*, *19*(3), 313–332.

Fredrickson, B. L., & Levenson, R. W. (1998). Positive emotions speed recovery from the cardiovascular sequelae of negative emotions. *Cognition and Emotion*, *12*(2), 191–220.

Fridlund, A. J., & Cacioppo, J. T. (1986). Guidelines for human electromyographic research. *Psychophysiology*, *23*, 567-589.

Frijda, N. H. (1986). *The emotions*. Cambridge, uk: Cambridge University Press.

Fukayama, A., Ohno, T., Mukawa, N., Sawaki, M., & Hagita, N. (2002). Messages embedded in gaze of interface agents — impression management with agent's gaze. In *Proceedings of the SIGCHI conference on human factors in computing systems: Changing our world, changing ourselves* (pp. 41–48). New York, NY, USA: ACM.

Gerardi, M., Cukor, J., Difede, J., Rizzo, A., & Rothbaum, B. O. (2010). Virtual reality exposure therapy for post-traumatic stress disorder and other anxiety disorders. *Current Psychiatry Reports*, *12*(4), 298-305.

Gilboa-Schechtman, E., Avnon, L., Zubery, E., & Jeczmien, P. (2006). Emotional processing in eating disorders: specific impairment orgeneral distress related deficiency? *Depression and Anxiety*, *23*(6), 331–339.

Grayson, J., Foa, E., & Steketee, G. (1982). Habituation during exposure treatment: Distraction vs attention-focusing. *Behaviour Research and Therapy*, *20*(4), 323-328.

Grimm, M., Kroschel, K., Harris, H., Nass, C., Schuller, B., Rigoll, G., et al. (2007). On the necessity and feasibility of detecting a driver's emotional state while driving. In A. Paiva, R. Prada, & R. Picard (Eds.), *Affective computing and intelligent interaction* (Vol. 4738, p. 126-138). Springer.

Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology*, *2*(3), 271-299.

Gross, J. J. (2002). Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology*, *39*, 281-291.

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being.

Journal of Personality and Social Psychology, *85*(2), 348-362.

Gross, J. J., & Levenson, R. W. (1997). Hiding feelings: The acute effects of inhibiting negative and positive emotion. *Journal of Abnormal Psychology*, *106*(1), 95–103.

Gross, J. J., & Thompson, R. A. (2007). Emotion regulation: Conceptual foundations. In J. J. Gross (Ed.), *Handbook of emotion regulation* (p. 3-24). Guilford Press.

Hamann, S. (2001). Cognitive and neural mechanisms of emotional memory. *Trends in Cognitive Sciences*, *5*(9), 394 - 400.

Harmon-Jones, E., & Allen, J. J. B. (1998). Anger and frontal brain activity: EEG asymmetry consistent with approach motivation despite negative affective valence. *Journal of Personality and Social Psychology*, *74*(5), 1310-1316.

Hassenzahl, M., & Tractinsky, N. (2006). User experience – a research agenda. *Behaviour and Information Technology*, *25*(2), 91–97.

Hazlett, R. L. (2003). Measurement of user frustration: a biologic approach. In *CHI'03 extended abstracts on human factors in computing systems* (p. 734-735).

Hazlett, R. L. (2006). Measuring emotional valence during interactive experiences: boys at video game play. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1023–1026). New York, NY, USA: ACM.

Heilman, K. M. (1997). The neurobiology of emotional experience. *Journal of Neuropsychiatry and Clinical Neurosciences*, *9*, 439-448.

Hillman, C. H., Rosengren, K. S., & Smith, D. P. (2004). Emotion and motivated behavior: postural adjustments to affective picture viewing. *Biological Psychology*, *66*, 51–62.

Höök, K. (2008). Affective loop experiences — what are they? In *Proceedings of the 3rd international conference on persuasive technology* (pp. 1–12). Berlin, Heidelberg: Springer-Verlag.

Horlings, R., Datcu, D., & Rothkrantz, L. J. M. (2008). Emotion recognition using brain activity. In *Proceedings of the 9th international conference on computer systems and technologies and workshop for PhD students in computing* (pp. 6:II.1–6:1). New York, NY, USA: ACM.

Ilves, M., & Surakka, V. (2004). Subjective and physiological responses to emotional content of synthesized speech. In N. Magnenat-Thalmann, C. Joslin, & H. Kim (Eds.), *Proceedings of the 17th international conference on computer animation and social agents, CASA 2004* (p. 19-26). Computer Graphics Society (CGS).

Ilves, M., & Surakka, V. (2008). Emotions, anthropomorphism of speech synthesis, and psychophysiology. In K. Izdebski (Ed.), *Emotions in the human voice. volume III: Culture and perception* (pp. 137–152). San Diego, CA: Plural Publishing Inc.

Johnson-Laird, P. N., & Oatley, K. (1989). The language of emotions: An analysis of a semantic field. *Cognition and Emotion*, *3*, 81–123.

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, *129*(5), 770-814.

Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words:are emotional words more vividly remembered than neutral words? *Memory and Cognition*, *31*(8), 1169-1180.

Knapp, R. B., Kim, J., & André, E. (2011). Physiological signals and their use in augmenting emotion recognition for human-machine interaction. In P. Petta, C. Pelachaud, & R. Cowie (Eds.), *Emotion-oriented systems: The Humaine handbook* (pp. 133–159). Berlin, Heidelberg: Springer.

Kreibig, S. D., Wilhelm, F. H., Roth, W. T., & Gross, J. J. (2007). Cardiovascular, electrodermal, and respiratory response patterns to fear- and sadness-inducing films. *Psychophysiology*, *44*(5), 787–806.

Krijn, M., Emmelkamp, P., Olafsson, R., & Biemond, R. (2004). Virtual reality exposure therapy of anxiety disorders: A review. *Clinical Psychology Review*, *24*, 259-281.

Kuikkaniemi, K., Kosunen, I., Turpeinen, T., M.and Saari, Laitinen, T., & Lievonen, P. (2010). Designing biofeedback for games and playful applications. In *CHI 2010 workshop on brain, body and bytes: Psychophysiological user interaction.* Atlanta, GA, USA. (Retrieved February 5th, 2010, from http://www.eecs.tufts.edu/ agirou01/workshop/papers/index.htm)

LaBar, K. S., & Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nature Reviews Neuroscience*, *7*(1), 54-64.

Lanctôt, N., & Hess, U. (2007). The timing of appraisals. *Emotion*, *7*(1), 207-212.

Lang, P. J. (1995). The emotion probe: Studies of motivation and attention. *American Psychologist*, *50*(5), 372-385.

Lang, P. J., Bradley, M. M., & Cuthbert. (1999). International affective picture system (IAPS): Instruction manual and affective ratings [Computer software manual].

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1992). A motivational analysis of emotion: Reflex-cortex connections. *Psychological Science*, *3*(1), 44–49.

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.

Larsen, J. T., McGraw, A. P., & Cacioppo, J. T. (2001). Can people feel happy and sad at the same time? *Journal of Personality and Social Psychology*, *81*(4), 684–696.

Larsen, J. T., Norris, C. J., & Cacioppo, J. T. (2003). Effects of positive and negative affect on electromyographic activity over the zygomaticus major and corrugator supercilii. *Psychophysiology*, *40*, 776-785.

Last.fm Ltd. (2011). *Last.fm — listen to internet radio and the largest music catalogue online.* Available from `http://www.last.fm/`

Lazar, J., Jones, A., Hackley, M., & Shneiderman, B. (2006). Severity and impact of computer user frustration: A comparison of student and workplace users. *Interacting with Computers*, *18*, 187–207.

Lester, J. C., Converse, S. A., Kahler, S. E., Barlow, S. T., Stone, B. A., & Bhogal, R. S. (1997). The persona effect: Affective impact of animated pedagogical agents. In *Proceedings of CHI'97* (p. 359-366).

Levenson, R. W., & Ekman, P. (2002). Difficulty does not account for emotion-specific heart rate changes in the directed facial action task. *Psychophysiology*, *39*, 397-405.

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.

Lindström, M., Ståhl, A., Höök, K., Sundström, P., Laaksolahti, J., Combetto, M., et al. (2006). Affective diary: designing for bodily expressiveness and self-reflection. In *CHI '06 extended abstracts on human factors in computing systems* (pp. 1037–1042). New York, NY, USA: ACM.

Lisetti, C., Nasoz, F., LeRouge, C., Ozyer, O., & Alvarez, K. (2003). Developing multimodal intelligent affective interfaces for tele-home health care. *International Journal of Human-Computer Studies*, *59*(1–2), 245–255.

Llobera, J., Spanlang, B., Ruffini, G., & Slater, M. (2010). Proxemics with multiple dynamic characters in an immersive virtual environment. *ACM Transactions on Applied Perception*, *8*, 1–12.

Loftus, G. R., & Harley, E. M. (2005). Why is it easier to identify someone close than far away? *Psychonomic Bulletin & Review*, *12*, 43–65.

Lopes, P. N., Salovey, P., Coté, S., & Beers, M. (2005). Emotion regulation abilities and the quality of social interaction. *Emotion*, *5*(1), 113-118.

Malliani, A., Pagani, M., Lombardi, F., & Cerutti, S. (1991). Cardiovascular neural regulation explored in the frequency domain. *Circulation*, *84*, 482-492.

Marks, I. M., Cavanagh, K., & Gega, L. (2007). *Hands-on help: Computer-aided psychotherapy*. Psychology Press.

Marsella, S., & Gratch, J. (2003). Modeling coping behavior in virtual humans: don't worry, be happy. In *Proceedings of the second international joint conference on autonomous agents and multiagent systems (AAMAS '03)* (pp. 313–320). ACM.

Marsh, A. A., Ambady, N., & Kleck, R. E. (2005). The effects of fear and anger facial expressions on approach- and avoidance-related behaviors. *Emotion*, *5*(1), 119–124.

Martin, J.-C., Niewiadomski, R., Devillers, L., Buisine, S., & Pelachaud, C. (2006). Multimodal complex emotions: Gesture expressivity and blended facial expressions. *International Journal of Humanoid Robotics*, *3*(3), 269-291.

Matsumoto, D. (1990). Cultural similarities and differences in display rules.

Motivation and Emotion, *14*, 195–214.
Mattel, Inc. (2011). *Mindflex.* Available from `http://mindflexgames.com`
Mauss, I. B., Levenson, R. W., McCarter, L., Wilhelm, F. H., & Gross, J. J. (2005). The tie that binds? coherence among emotion experience, behavior, and physiology. *Emotion*, *5*(2), 175-190.
Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition and Emotion*, *23*(2), 209-237.
Miles, L. K. (2009). Who is approachable? *Journal of Experimental Social Psychology*, *45*(1), 262 - 266.
Miller, T. Q., Smith, T. W., Turner, C. W., Guijarro, M. W., & Hallet, A. J. (1996). A meta-analytic review of research on hostility and physical health. *Psychological Bulletin*, *119*(2), 322-348.
Morris, L. D., Louw, Q. A., & Grimmer-Somers, K. (2009). The effectiveness of virtual reality on reducing pain and anxiety in burn injury patients: A systematic review. *The Clinical Journal of Pain*, *25*(9), 815–826.
Moundridou, M., & Virvou, M. (2002). Evaluating the persona effect of an interface agent in a tutoring system. *Journal of Computer Assisted Learning*, *18*, 253-261.
Murray, I. R., & Arnott, J. L. (1993). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *Journal of the Acoustical Society of America*, *93*(2), 1097-1108.
Murray, I. R., & Arnott, J. L. (1995). Implementation and testing of a system for producing emotion-by-rule in synthetic speech. *Speech Communication*, *16*, 369-390.
Nasoz, F., Lisetti, C. L., & Vasilakos, A. V. (2010). Affectively intelligent and adaptive car interfaces. *Information Sciences*, *180*(20), 3817 - 3836.
Nass, C. (2004). Etiquette equality: exhibitions and expectations of computer politeness. *Communications of the ACM*, *47*, 35–37.
Nass, C., & Gong, L. (2000). Speech interfaces from an evolutionary perspective. *Communications of the ACM*, *43*, 36–43.
Nass, C., Jonsson, I.-M., Harris, H., Reaves, B., Endo, J., Brave, S., et al. (2005). Improving automotive safety by pairing driver emotion and car voice emotion. In *CHI '05 extended abstracts on human factors in computing systems* (pp. 1973–1976). New York, NY, USA: ACM.
Nass, C., & Lee, K. M. (2001). Does computer-synthesized speech manifest personality? experimental tests of recognition, similarity-attraction, and consistency-attraction. *Journal of Experimental Psychology: Applied*, *7*(3), 171-181.
Nass, C., Steuer, J., & Tauber, E. R. (1994). Computers are social actors. In *Proceedings of CHI'94* (p. 72-78).
Nestoriuc, Y., Martin, A., Rief, W., & Andrasik, F. (2008). Biofeedback treatment for headache disorders: A comprehensive efficacy review. *Applied Psychophysiology and Biofeedback*, *33*, 125–140.
Nöjd, N., Puurtinen, M., Niemenlehto, P., Vehkaoja, A., Verho, J., Vanhala, T.,

et al. (2005). Wireless wearable EMG and EOG measurement system for psychophysiological applications. In *IFMBE proceedings of the 13th nordic baltic conference on biomedical engineering and medical physics* (pp. 144–145).

Nomura, T., Kanda, T., Suzuki, T., Yamada, S., & Kato, K. (2009). Influences of concerns toward emotional interaction into social acceptability of robots. In *HRI '09: Proceedings of the 4th ACM/IEEE international conference on human robot interaction* (pp. 231–232). New York, NY, USA: ACM.

Norris, C. J., Gollan, J., Berntson, G. G., & Cacioppo, J. T. (2010). The current status of research on the structure of evaluative space. *Biological Psychology*, *84*(3), 422-436.

Oatley, K., & Jenkins, J. M. (1996). *Understanding emotions*. Oxford, UK: Blackwell.

Oatley, K., & Johnson-Laird, P. N. (1987). Towards a cognitive theory of emotions. *Cognition and Emotion*, *1*, 29-50.

Ochs, M., Niewiadomski, R., Pelachaud, C., & Sadek, D. (2005). Intelligent expressions of emotions. In J. Tao, T. Tan, & R. Picard (Eds.), *Affective computing and intelligent interaction* (Vol. 3784, pp. 707–714). Berlin, Heidelberg: Springer.

Osgood, C. E. (1952). The nature and measurement of meaning. *Psychological Bulletin*, *49*(3), 197 – 237.

Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. Oxford University Press.

Parkinson, B. (1996). Emotions are social. *British Journal of Psychology*, *87*, 663-683.

Parrott, W. G. (2001). *Emotions in social psychology: Essential readings*. Philadelphia: Psychology Press.

Partala, T., & Surakka, V. (2004). The effects of affective interventions in human-computer interaction. *Interacting with Computers*, *16*(2), 295-309.

Partala, T., Surakka, V., & Lahti, J. (2004). Affective effects of agent proximity in conversational systems. In *Proceedings of NordiCHI 2004* (p. 353-356).

Partala, T., Surakka, V., & Vanhala, T. (2005). Person-independent estimation of emotional experiences from facial expressions. In *Proceedings of IUI 2005* (p. 246-248). ACM Press.

Partala, T., Surakka, V., & Vanhala, T. (2006). Real-time estimation of emotional experiences from facial expressions. *Interacting with Computers*, *18*, 208-226.

Pentland, A. (2000). Perceptual intelligence. *Communications of the ACM*, *43*(3), 35–44.

Picard, R. W. (1997). *Affective computing*. Cambridge, MA, USA: MIT Press.

Picard, R. W., & Klein, J. (2002). Computers that recognise and respond to user emotion: theoretical and practical implications. *Interacting with Computers*, *14*(2), 141 – 169.

Picard, R. W., Vyzas, E., & Healey, J. (2001). Toward machine emotional intelligence: Analysis of affective physiological state. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *23*, 1175-1191.

Plutchik, R. (1991). *The emotions*. University Press of America.

Posner, J., Russell, J. A., Gerber, A., Gorman, D., Colibazzi, T., Yu, S., et al. (2009). The neurophysiological bases of emotion: An fMRI study of the affective circumplex using emotion-denoting words. *Human Brain Mapping*, *30*, 883-895.

Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, *17*, 715-734.

Powers, M. B., & Emmelkamp, P. M. G. (2008). Virtual reality exposure therapy for anxiety disorders: A meta-analysis. *Journal of Anxiety Disorders*, *22*(3), 561-569.

Prendinger, H., & Ishizuka, M. (2005). The empathic companion: A character-based interface that addresses users' affective states. *Applied Artificial Intelligence*, *19*(3/4), 267 – 285.

Prendinger, H., Mayer, S., Mori, J., & Ishizuka, M. (2003). Persona effect revisited: Using bio-signals to measure and reflect the impact of character-based interfaces. In *Proceedings of IVA 2003, lncs* (Vol. 2792, p. 283-291).

Radovan, M., & Pretorius, L. (2006). Facial animation in a nutshell: past, present and future. In *Proceedings of the 2006 annual research conference of the South African institute of computer scientists and information technologists on IT research in developing countries* (pp. 71–79). Republic of South Africa: South African Institute for Computer Scientists and Information Technologists.

Rainville, P., Bechara, A., Naqvi, N., & Damasio, A. R. (2006). Basic emotions are associated with distinct patterns of cardiorespiratory activity. *International Journal of Psychophysiology*, *61*, 5-18.

Rani, P., Sarkar, N., Smith, C., & Adams, J. (2003). Affective communication for implicit human-machine interaction. In *IEEE international conference on systems, man and cybernetics, 2003* (Vol. 5, p. 4896 - 4903).

Rantanen, J., Impiö, J., Karinsalo, T., Malmivaara, M., Reho, A., Tasanen, M., et al. (2002). Smart clothing prototype for the arctic environment. *Personal and Ubiquitous Computing*, *6*, 3-16.

Rantanen, V., Niemenlehto, P.-H., Verho, J., & Lekkala, J. (2010). Capacitive facial movement detection for human-computer interaction to click by frowning and lifting eyebrows. *Medical and Biological Engineering and Computing*, *48*(1), 39–47.

Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. New York, NY, USA: Cambridge University Press.

Rinn, W. E. (1991). Neuropsychology of facial expression. In R. S. Feldman & B. Rimé (Eds.), *Fundamentals of nonverbal behavior* (p. 3-30). Cambridge: Cambridge University Press.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161-1178.

Russell, J. A., & Mehrabian, A. (1977). Evidence for a three-factor theory of emotions. *Journal of Research in Personality*, *11*(3), 273 - 294.

Sakurazawa, S., Yoshida, N., & Munekata, N. (2004). Entertainment feature of a game using skin conductance response. In *Proceedings of the 2004 acm sigchi international conference on advances in computer entertainment technology* (pp. 181–186).

Scherer, K. R. (1993). Studying the emotion-antecedent appraisal process: An expert system approach. *Cognition and Emotion*, *7*(3/4), 325-355.

Schieber, M. H. (2001). Constraints on somatotopic organization in the primary motor cortex. *Journal of Neurophysiology*, *86*, 2125-2143.

Schlosberg, H. (1954). Three dimensions of emotion. *Psychological Review*, *61*(2), 81-88.

Schmidt, K. L., & Cohn, J. F. (2001). Human facial expressions as adaptations: questions in facial expression research. *Yearbook of Physical Anthropology*, *44*, 3-24.

Simon, H. A. (1967). Motivational and emotional controls of cognition. *Psychological review*, *74*(1), 29-39.

Singer, J. A., & Salovey, P. (1988). Mood and memory: Evaluating the network theory of affect. *Clinical Psychology Review*, *8*(2), 211 - 251.

Sirkiä, S. (2010, November). Yksin yhdessä: Internetin sosiaaliset liikuntapalvelut. *Mikrobitti*, 68–70. (In Finnish language)

Sokhadze, E. M. (2007). Effects of music on the recovery of autonomic and electrocortical activity after stress induced by aversive visual stimuli. *Applied Psychophysiology & Biofeedback*, *32*(1), 31–50.

Spoor, J. R., & Kelly, J. R. (2004). The evolutionary significance of affect in groups: Communication and group bonding. *Group Processes & Intergroup Relations*, *7*(4), 398-412.

Stemmler, G., Aue, T., & Wacker, J. (2007). Anger and fear: Separable effects of emotion and motivational direction on somatovisceral responses. *International Journal of Psychophysiology*, *66*, 141–153.

Stins, J. F., & Beek, P. J. (2007). Effects of affective picture viewing on postural control. *BMC Neuroscience*, *8*(83), 7 pp.

Surakka, V., & Vanhala, T. (accepted). Emotions in human-computer interaction. In A. Kappas & N. C. Krämer (Eds.), *Face-to-face communication over the internet.* New York: Cambridge University Press.

Tamir, M. (2011). The maturing field of emotion regulation. *Emotion Review*, *3*(1), 3–7.

Tassinary, L. G., & Cacioppo, J. T. (2000). The skeletomotor system: Surface electromyography. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *The handbook of psychophysiology* (2nd ed., p. 163-199). New York: Cambridge University Press.

Tennenhouse, D. (2000). Proactive computing. *Communications of the ACM*, *43*(5), 43–50.

Textronics Inc. (2011). *Strapless heart rate monitor clothes by NuMetrex.* Available from `http://www.numetrex.com/`

Tolin, D. F., Lohr, J. M., Lee, T. C., & Sawchuk, C. N. (1999). Visual avoidance in specific phobia. *Behaviour Research and Therapy*, *37*(1), 63 – 70.

Tomkins, S. S. (1962). *Affect, imagery, consciousness. volume 1: The positive affects*. New York: Springer.

Trentacosta, C. J., & Izard, C. E. (2007). Kindergarten children's emotion competence as a predictor of their academic competence in first grade. *Emotion*, *7*(1), 77 – 88.

Van den Stock, J., Righart, R., & de Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion*, *7*(3), 487-494.

Van Peer, J. M., Rotteveel, M., Spinhoven, P., Tollenaar, M. S., & Roelofs, K. (2010). Affect-congruent approach and withdrawal movements of happy and angry faces facilitate affective categorisation. *Cognition and Emotion*, *24*(5), 863-875.

Vanhala, T., & Surakka, V. (2007). Facial activation control effect (FACE). In A. Paiva, R. Prada, & R. W. Picard (Eds.), *Affective computing and intelligent interaction* (p. 278-289). Springer-Verlag.

Vanhala, T., & Surakka, V. (2008). Computer-assisted regulation of emotional and social processes. In J. Or (Ed.), *Affective computing* (p. 405-420). Vienna, Austria: InTech Education and Publishing.

Vanhala, T., Surakka, V., & Anttonen, J. (2008). Measuring bodily responses to virtual faces with a pressure sensitive chair. In *Nordichi '08: Proceedings of the 5th nordic conference on human-computer interaction* (pp. 555–559). New York, NY, USA: ACM.

Vanhala, T., Surakka, V., Courgeon, M., & Martin, J.-C. (2011). Voluntary facial activations regulate physiological arousal and subjective experiences during virtual social stimulation. *ACM Transactions on Applied Perception*.

Vanhala, T., Surakka, V., Siirtola, H., Räihä, K.-J., Morel, B., & Ach, L. (2010). Virtual proximity and facial expressions of computer agents regulate human emotions and attention. *Computer Animation and Virtual Worlds*, *21*(3-4), 215-224.

Van Mulken, S., André, E., & Müller, J. (1998). The persona effect: How substantial is it? In *Proceedings of HCI conference on people and computers XIII* (p. 53-66).

Vehkaoja, A., & Lekkala, J. (2004). Wearable wireless biopotential measurement device. In *Proceedings of the ieee embs 2004* (p. 2177-2179).

Vehkaoja, A., & Lekkala, J. (2006). Wireless measurement band for eeg mismatch negativity registration in mobile activities. In *Imeko xviii world congress and iv brazilian congress of metrology.* Rio de Janeiro, Brazil.

Vinayagamoorthy, V., Brogni, A., Steed, A., & Slater, M. (2006). The role of posture in the communication of affect in an immersive virtual environment. In *VRCIA '06: Proceedings of the 2006 ACM international conference* *on virtual reality continuum and its applications* (pp. 229–236). New York, NY, USA: ACM.

Vinayagamoorthy, V., Garau, M., Steed, A., & Slater, M. (2004). An eye gaze model for dyadic interaction in an immersive virtual environment: Practice and experience. *Computer Graphics Forum*, *23*(1), 1–11.

Ward, R. D., & Marsden, P. H. (2003). Physiological responses to different web page designs. *International Journal of Human-Computer Studies*, *59*, 199-212.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin*, *98*, 219-235.

Wells-Parker, E., Ceminsky, J., Hallberg, V., Snow, R. W., Dunaway, G., Guiling, S., et al. (2002). An exploratory study of the relationship between road rage and crash experience in a representative sample of US drivers. *Accident Analysis & Prevention*, *34*(3), 271 - 278.

Wiederhold, B. K., & Wiederhold, M. D. (2003). Three-year follow-up for virtual reality exposure for fear of flying. *CyberPsychology and Behavior*, *6*, 441-445.

Wiener, N. (1948). *Cybernetics, or control and communication in the animal and the machine*. New York: Wiley.

Wieser, M. J., Pauli, P., Reicherts, P., & Mühlberger, A. (2010). Don't look at me in anger! enhanced processing of angry faces in anticipation of public speaking. *Psychophysiology*, *47*, 271-280.

Wilhelm, F. H., & Roth, W. T. (1998). Taking the laboratory to the skies: Ambulatory assessment of self-report, autonomic, and respiratory responses in flying phobia. *Psychophysiology*, *35*(5), 596-606.

Witvliet, C. V., & Vrana, S. R. (1995). Psychophysiological responses as indices of affective dimensions. *Psychophysiology*, *5*, 436–443.

Wu, C., Chuang, Z., & Lin, Y. (2006, June). Emotion recognition from text using semantic labels and separable mixture models. In (Vol. 5, pp. 165–183). New York, NY, USA: ACM.

Wundt, W. (1896). *Grundriss der psychologie*. Leipzig, Germany: Verlag von Wilhelm Engelmann.

Yee, N., Bailenson, J. N., & Rickertsen, K. (2007). A meta-analysis of the impact of the inclusion and realism of human-like faces on user experiences in interfaces. In *CHI '07: Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1–10). New York, NY, USA: ACM.

Yoshitomi, Y., Sung-Ill, K., Kawano, T., & Kilazoe, T. (2000). Effect of sensor fusion for recognition of emotional states using voice, face image and thermal image of face. In *Proceedings of the 9th IEEE international workshop on robot and human interactive communication* (p. 178 - 183).

Zeng, Z., Tu, J., Zhang, T., Rizzolo, N., Zhang, Z., Huang, T. S., et al. (2004). Bimodal HCI-related affect recognition. In *Proceedings of ICMI* (pp. 205–211). ACM Press.

Appendix A

Publication I

Surakka, V. and Vanhala, T. (2011). Emotions in human-computer interaction. In Kappas, A. and Krämer, N.C. (Eds.) *Face-to-Face Communication over the Internet*, pp. 213–236. Cambridge, UK: Cambridge University Press.

CHAPTER 9

Emotions in human–computer interaction

Veikko Surakka and Toni Vanhala

Overview: Human–computer interaction (HCI) may be significantly improved by incorporating social and emotional processes. Developing appropriate technologies is only one side of the problem. It is also vital to investigate how synthesized emotional information might affect human behavior in the context of information technology. Despite previous suggestions that people treat computers as social actors, we still know relatively little about the possible and supposedly positive effects of utilizing any kind of emotional cues or messages in human–technology interaction. The aim of the present chapter is to provide a theoretical and empirical basis for integrating emotions into the study of HCI. We will first argue and show evidence in favor of the use of virtual emotions in HCI. We will then proceed by studying the possibilities of a computer for analyzing human emotion-related processes and consider some physiological measures used for this purpose in more detail. In this context, we will also briefly describe some new technological prototypes for measuring computer users' behavior. The chapter ends with a discussion summarizing the findings and addressing the advantages of studying emotions in the context of present-day technology.

Introduction

The qualitative improvement and facilitation of human–computer interaction (HCI) has become a central research issue in computer science. Traditionally, attempts to improve HCI have centered on making computers more user-friendly along technical dimensions. In this line of work, perhaps still the most visible milestone for an ordinary user has been the development of a graphical, mouse-driven user interface. This invention dramatically enhanced computer use and raised HCI to a new level. The most important consequences are that the user receives feedback about his or her actions when operating a desktop computer or another kind of information technology, such as handheld devices, televisions, and mobile telephones.

These developments in turn enable the development of various non-traditional input and output channels. For example, computers might be used by touching, speaking, and gazing in their direction. On the other hand, computers themselves might provide rich feedback or output using either the common auditory, visual, or audiovisual channels, or less commonly applied modalities, such as the sense of touch.

Although current computers and user interfaces are reasonably functional (e.g., efficient in computing and relatively easy to use), new problems and lines of research have developed in parallel with the technology itself. Indeed, one of the recently recognized problems relates to the emotional state that working with computers evokes in users. At least two studies have succeeded in bringing attention to the emotional aspects of human–technology relations.

A study by Concord Communications (1999) found that users often admit to help desk managers that they have had violent thoughts and/or engaged in abusive behavior toward their computers. A study called "Rage Against the Machine," conducted by MORI on behalf of Compaq, found that users swear, consider causing damage, deliberately pull out the plug, and even kick their computers.[1] These examples show that working with computers evokes intense negative emotional reactions. Looking a little deeper, these examples also imply that people actually treat computers as if they were social entities, which they would like to threaten and punish in order to make them operate better in the future.

Curiously, although it is evident on a conscious cognitive level that computers do not share our communicative signals, we seem to forget this when interacting with them. For instance, people really seem to appreciate sharing their feelings with computers as well as virtual expressions of empathy, irrespective of the fact that machines cannot really feel for us (Brave *et al.*, 2005; Klein *et al.*, 2002). Moreover, other social signals, such as simulated emotional touch (i.e., haptics and emotions), simulated proximity, and virtual facial expressions of an embodied agent character may affect our subjective feelings and physiological arousal (Partala *et al.*, 2004; Salminen *et al.*, 2008; Vanhala *et al.*, 2010). Generally speaking, people react to virtual cues of emotion and sociality in much the same manner as they do to those expressed by other humans. According to Nass *et al.* (1994), this is due to our strong tendency for social behavior.

[1] www.ipsos-mori.com/researchpublications/researcharchive/poll.aspx?oItemId=1900.

However, we are still left with the fact that computers have limited access to the social signals that we give out, and therefore cannot respond properly to their users. Building *perceptually intelligent* computers is one way to try to construct more socially and emotionally intelligent computers (Pentland, 2000). Concepts such as *perceptually intelligent machines*, *affective computing*, and *emotion and interest sensitive media* refer to a future generation of flexible, trainable, adaptive, and emotionally responsive user interfaces. The ambitious aim of perceptual intelligence is to build machines or interfaces that are as intelligent observers of social and emotional cues as we are. These interfaces utilize, for example, video cameras, microphones, and eye trackers for perceiving users' behaviors and translating them into information that a computer can process.

In addition, various types of physiological indicators (i.e., physiological computing), including correlates of autonomic nervous system activation (e.g., pupil size variation and heart rate), responses of central nervous activation, and electrical activity of facial muscles (Figure 9.1), can be utilized in connecting the user with the computer (Allanson *et al.*, 1999; Anttonen and Surakka, 2005; Barreto *et al.*, 2000; Jacob, 1991; Kübler *et al.*, 1999; Partala *et al.*, 2001, 2005, 2006; Pentland, 2000; Picard, 1997; Surakka *et al.*, 2004; Wolpaw *et al.*, 2002). Although these physiological measures are not readily available for our human communication partners, they can provide rich information for machines that sense them. As physiological activity is a central, underlying factor of all human behavior, it can be argued that computers might potentially have better access to our emotional processes than other people.

Enhancements of the perceptual intelligence of computers need to be matched by corresponding increases in *expressive intelligence*. The ambitious aim in this respect is that computers should be able to express socially and emotionally meaningful information in a human-like manner. Important aspects of communication are sensitivity to the user and appropriate responses to his or her cognitions and emotions. This might be achieved, for example, by the use of friendly visual expressions and even-tempered vocal intonations, while taking into account how users react to them. Work in this field is very much in its infancy. Speech synthesizers, for instance, are slowly beginning to make better use of prosodic features (Ilves and Surakka, 2004). There is also emerging knowledge about the types of virtual cues that are effective in influencing social perceptions of the human perceiver. Recent studies using virtual agents suggest that their gaze direction and other meaningful facial expressions are effective in communicating that they are trustworthy, intelligent social actors (Berry *et al.*, 2005; Schilbach *et al.*,

Figure 9.1 A person interacting with a software agent. This agent records the changes in the level of electrical facial muscle activities, and interprets these activations with the help of signal-processing algorithms. In the prototype, a software agent can be called on for help by raising the eyebrows (a common gesture when one is faced with problems or surprises). This agent can be forced to disappear by lowering the eyebrows (a common gesture used for expressing a negative attitude).

2006). However, these channels have not yet been used to any large extent in actual or real-time human–machine interaction.

Real-time expressive intelligence requires systems that are able to smoothly monitor and interpret users' behaviors via many channels so that the user's reactions to these intelligent expressions can also be taken into account. Computers that utilize video-, bio-, and behavioral signals need efficient signal analysis methods in order to interpret the various sensory signals with which they are faced. Intelligent perception also involves the capability to harness the potential of all available information. In the age of wireless and wearable computing, this means that systems should be able to dynamically integrate new devices (if and when they become available) into their analytical and decision-making processes. Further, they need to drop devices and measurements from their analysis when devices become unavailable (i.e., out of range), and still be able to interpret remaining data properly. These kinds of new developments have been enabled by the recent approaches in software development which emphasize loosely coupled, agent-based architectures (e.g., Vanhala and Surakka, 2005). Thus, complete communicative links can be created between the user and the computer.

An example of a complete communication link between person and machine is a demonstration of the Real Estate Agent (REA) software developed at the Massachusetts Institute of Technology. For example, REA can sense the appearance of a customer or user in front of her with the help of machine vision technology. She (or it) can then make a greeting by speech and nonverbal gestures (i.e., it has both a speech and a gesture database). It is also capable of taking turns during conversation (Cassell *et al*., 2000). Recently, D'Mello *et al*. (2007) argued that such full communication cycles are needed in intelligent tutoring systems, such as their AutoTutor system. The emotions evoked by AutoTutor were explored with self-reports and observations. Perceptual intelligence was built into the system in the form of several automatic emotion classifiers for conversational cues, and for analyzing postural and facial expression. However, the integration of these separate channels and the intelligent use of this knowledge in an affective loop between the student and AutoTutor is still a work in progress.

We have designed an architecture for building functional prototype software that regulates (i.e., both observes and responds to) users' social and emotional responses (Figure 9.2). The software architecture enables the designer of the system to script the responses of an agent to different predefined scenarios of psychophysiological or psychobehavioral activity. For example, one prototypic system was developed for relaxing the respiration intensity of a person with labored breathing by automatic analysis of his or her heart activity patterns (Vanhala and Surakka, 2005). Depending on the automated higher-level interpretation of low-level electrocardiographic signals extracted from the "user," a virtual agent character would appear and calm the person by instructing a more relaxed breathing pattern in synthesized speech. A second system designed to monitor noninvasively the patterns of mouse and keyboard activity was developed to calm people who become negatively aroused when browsing the web with a mouse and keyboard. Intense levels of arousal while using the computer mouse has also been termed "mouse rage." This "rage" has been shown to exist in day-to-day computer use and can potentially lead to harmful physiological consequences (Marsh and Khor, 2006). When the user becomes negatively aroused, he or she may interpret the lack of feedback from the computer as negative information. This in turn

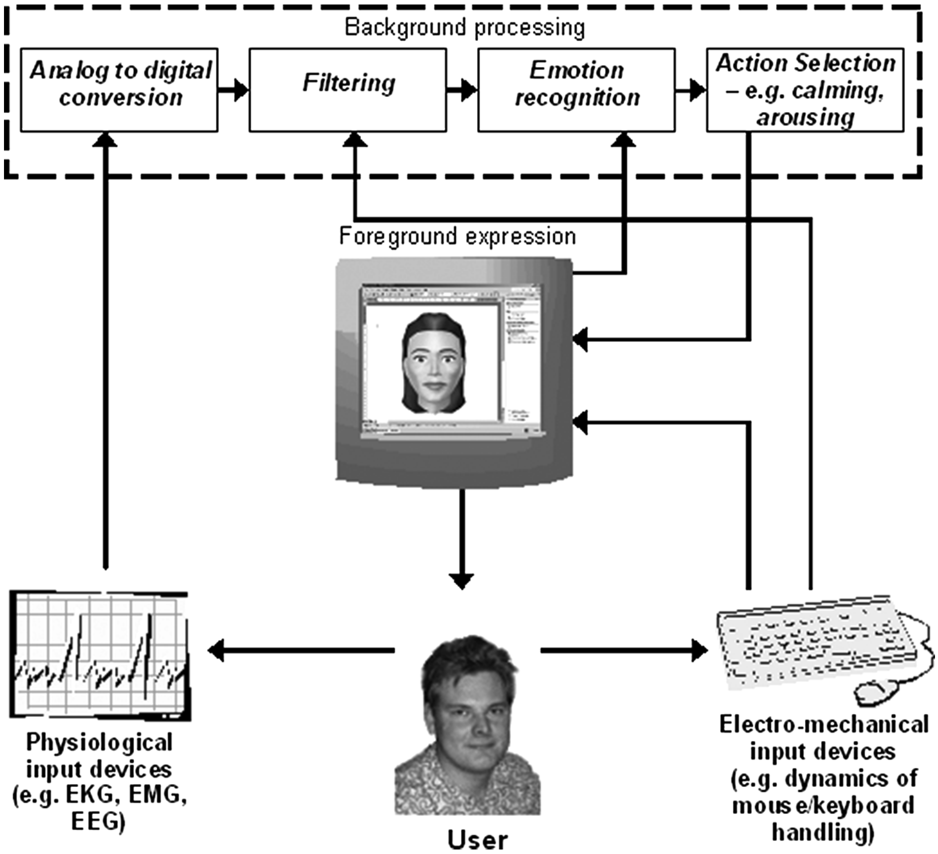


Figure 9.2 A schematic overview of a prototype software developed for detecting users' probable emotion-related responses. For example, significant changes in bioelectrical signals of facial muscles can be recognized with signal-processing algorithms (i.e., psychophysiological computing). In addition, dynamic behavioral changes (i.e., keyboard and mouse manipulation) can be monitored with signals from traditional input devices (i.e., psychobehavioral computing). When predefined algorithmic thresholds are exceeded, for example, an audiovisual agent can be made to appear and communicate in a predefined manner.

may lead to an even greater increase in negative arousal. In this case, our system was designed first to monitor noninvasively the patterns of mouse and keyboard activity, and then, if needed, to invoke a virtual character in order to break the vicious circle.

Although previous examples have shown that contemporary audiovisual agents can be programmed to display socially meaningful signals, this still requires a great deal of effort in preprogramming (i.e., scripting) the behavior of these agents. Even though we are beginning to understand the effectiveness of social and emotional cues in human–computer communication, we are still far from coming up with machines that can independently perform intelligent social emotional communication.

The aim of the present chapter is to provide a basis for a new generation of emotionally intelligent systems. Thus, in the following sections, we will first discuss the advantages of having a technology that is sensitive to emotional and social information. Next, we will present studies that clearly show the potential of emotionally expressive intelligence. Following this, we will present findings on how perceptual and expressive intelligence can be merged to create functional loops of perceptual and expressive intelligence. Finally, we will discuss the findings by adopting a positive perspective for the future.

Emotions and sociality in HCI

A natural approach to the development of social and emotional HCI involves considering which factors are central in human communication and behavior. In addition to the questionnaire and interview findings reported in the introduction, there is other evidence that people use the same kind of social rules in HCI that they use in human–human social communication (Fogg and Nass, 1997; Reeves and Nass, 1996). For example, research on sociality and computers has shown that people automatically interpret a computer agent as extroverted or introverted according to its programmed behavior. It has also been shown that people prefer to interact with agents whose behavior seems to reflect a consistent trait or disposition. For example, they prefer agents that are extroverted and use more courageous language to agents that are extroverted but use more shy language (Nass *et al.*, 2000). Furthermore, there is evidence that people accept and are responsive to flattery and feedback from computers (Aula and Surakka, 2002; Fogg and Nass, 1997). In fact, a research paradigm called "computers are social actors" (CASA) suggests that theories and methods of social psychology are directly applicable to HCI (Nass *et al.*, 1994, 2000; Nass and Moon, 2000; Nass and Steuer, 1993; Reeves and Nass, 1996).

According to Reeves and Nass (1996), our strong tendency to sociality results in people unavoidably reacting to computers' signals as if they had originated from another human being. Although it has not been explicitly stated that the CASA paradigm includes emotions and emotional behavior, it is clear that emotions are one of the most significant factors in social behavior. For example, people like and trust embodied computer agents that display empathic emotions more than agents with self-oriented emotional expressions (Brave *et al.*, 2005). On the other hand, it may also be that inanimate objects, and especially virtual creatures with or without embodiment, automatically evoke our tendency for emotional behavior. This evoked emotional behavior then results in the activation of other social processes. In any case, it is clear that emotional factors are an important component in HCI.

The acknowledged importance of emotions in terms of intra- and interindividual performance and behavior means that the challenge of incorporating emotions into HCI must be faced. For some time, there has been a growing consensus that human emotion systems (physiological, experiential, and expressive) are central factors for rational human behavior (Damasio, 1994; Lang, 1995; Lang *et al.*, 1993; LeDoux, 1994). Affective processes are a necessity for cognitive processing. For example, there is evidence that decision making can be seriously disturbed after brain damage to areas that are important for emotional processing (Damasio, 1994). In a classic paper, Zajonc (1980) suggested that emotional processing can even have primacy over cognition. Because working with computers constantly requires and evokes both affective and cognitive processing, there is no doubt that affective information needs to be considered in the development of interaction with technology. New anthropomorphic interfaces, such as embodied agents, have the potential to enhance the social reactions to computers even further. For example, there is evidence that people react similarly to both virtual and human faces (Schilbach *et al.*, 2005, 2006; Vanhala *et al.*, 2010; Weyers *et al.*, 2006). This may have considerable implications, as, for instance, perception of emotional cues can significantly affect the processing of nonemotional auditory information at a very basic neuronal level (Surakka *et al.*, 1998). Thus, it is becoming increasingly important to take emotions into account in developing alternative new interaction techniques per se and in regulating the quality of human–technology interaction (Aula and Surakka, 2002; Partala and Surakka, 2003, 2004; Salminen *et al.*, 2008; Surakka *et al.*, 2003, 2004).

The argument developed so far implies that fully functional systems need to analyze the aims of the user and the context of use, and be perceptually and expressively intelligent. Clearly, the fulfillment of these aims will require a lot of work. Developing system architectures and programs (as, for example, in Figure 9.2) that are able to process, interpret, and act upon social and emotional information represents a research endeavor in its own right. Although our research team (among others) is ultimately aiming to develop these types of fully functional systems, our recent research efforts have been mainly dealing with the following subtasks.

First, studies that investigate the potential advantages of having expressive intelligence are required. The key question in terms of emotions and human–technology interaction is whether we can evoke reasonable, useful effects with emotional feedback and interventions from technology. Provided that we find useful emotional and social effects, we will then have a well-grounded basis for further developments. Proceeding in this order may initially seem counterintuitive, as perceptual intelligence is actually a precondition for expressive intelligence in fully functional systems. However, if computers were found not to evoke emotional and social responses, further studies would be misguided. In any case, our own studies have shown the clear promise of the potential of emotionally intelligent expression. For this reason, we have proceeded to develop perceptual intelligence. These activities have consisted of both the development of new or alternative hardware-sensing technologies and the analysis of the possibilities of different signal-processing algorithms in interpreting the data flow from various sensing technologies.

Potential of emotionally intelligent expression in HCI

Although people seem to be driven to act socially and emotionally with computers, as yet there is little knowledge regarding what kind of effects (emotional or cognitive) computers expressing this type of information might have on human behavior. The usual and often implicit assumption is that computers that respond to users' stress, frustration, or more specific emotions might significantly improve and facilitate HCI. A related challenge is to identify what measures could be used for detecting these effects automatically in order to build perceptual intelligence into machines. In a sense, these types of investigations can be implemented in a highly similar fashion to investigations in basic emotion research. Thus, we first need to establish, for example, whether pupil size behavior is affected by emotional stimulation. If this proves to be the case, then it is clear that the special nature of HCI needs to be considered at some point by performing pupil size measurements while the user is interacting with a computer.

When we first began to study the potential advantages of emotions in HCI, one of the earlier notions was that tracking the users' gaze direction and pupil size variation could provide an input signal indicating users' interests and emotional states to the computer (e.g., Jacob, 1996). The early results and theories concerning this issue were rather mixed. For example, some studies found pupil constriction to negative emotional stimulation, while others did not find such evidence (e.g., Hess, 1972; Janisse, 1974; Loewenfeld, 1966). Furthermore, gender differences were suggested in some of the early studies (for a review, see Partala and Surakka, 2003). As it was unclear how emotional stimuli would affect the pupil size, it was also unclear whether pupil size variation could be used as an index of users' emotional states. Thus, we conducted studies to address this question, using a set of carefully developed acoustic emotional stimuli called international affective digitized sounds (IADS) by Bradley and Lang (1999, 2000).

In one experiment (Partala *et al.*, 2000; Partala and Surakka, 2003), participants' pupil size was monitored while they were exposed to 30 (10 emotionally negative, 10 neutral, and 10 emotionally positive) stimuli. We found that both negative and positive auditory stimuli evoked significantly greater pupil enlargement than neutral auditory stimuli at the time of presentation. Pupil sizes began to slowly approach the baseline at about 3 s after the stimulus onset, but the statistical differences remained the same even after 2 s from the stimulus offset. We found no evidence of either pupil constrictions or gender differences. The ratings of the stimuli showed that the stimuli elicited the intended emotional responses, and the ratings were in accordance with those of Bradley and Lang (1999). In a further experiment (Experiment 2 in Partala *et al.*, 2000), we found similar results.

The above findings were encouraging in that they showed that auditory emotional stimulation regulates the functioning of the autonomic nervous system as indexed by pupil size variation. The results also showed that subjective experiences could be significantly influenced by auditory stimuli. These findings therefore demonstrate that the measurement of pupil size can be used to reveal possible effects of emotions in the context of HCI.

In a subsequent experiment, we utilized pupil size measurements and ratings of emotional experiences by indexing the users' emotions in the context of HCI. Because pupil size is very sensitive to variations in lighting, we used synthesized speech as a feedback channel during computer interaction. The use of speech synthesis parallels the use of computer agents, but of course without embodiment in this case. In the experiment, we studied the effects of synthesized emotional feedback on cognitive performance and on psychological and physiological reactions during a fully computerized, problem-solving task (Aula and Surakka, 2002). Participants were required to solve series of relatively simple computations (e.g., 3 (?) 3 = 6) in which they had to decide upon the missing (+ or –) operator by pressing a corresponding key. After each series, a Finnish speech synthesizer called Mikropuhe gave *random* negative, neutral, or positive feedback (i.e., feedback that was independent of the participants' performance) with emotional content (e.g., "I am disappointed with your result," "Your result was average," and "Your performance makes me glad," respectively). The feedback messages were controlled so that only the content of the messages varied. Prosodic cues were kept constant across all feedback messages so that there was no variation in intonation, volume, tone, etc. (for more information about the association between vocal cues and emotions, see Scherer *et al.*, 1984). At the end of the experiment, the participants rated their emotional responses to the different feedback messages on two 9-point rating scales assessing emotional valence and arousal (for more information on dimensions of emotions, see, e.g., Lang *et al.*, 1993; Schlosberg, 1954).

In brief, the results from the valence ratings showed that the participants rated different feedback categories as differing significantly in valence (negative as the most unpleasant, positive as the most pleasant). Response times to mathematical tasks were significantly shorter after positive emotional feedback than after negative emotional feedback. Pupil size increased in response to all feedback categories, meaning that autonomic nervous system activity was accentuated as a result of computerized synthetic communication. Interestingly, when tracked over time (5 s from the stimulus offset), the pupil size behavior revealed significant differential pupil responses after the different feedback categories. After positive feedback, there was a greater and faster recovery toward baseline than after neutral feedback.

In sum, the findings suggested that synthetic speech with emotional content regulates emotional responses. Participants experienced feedback messages as emotionally negative, neutral, or positive, and their physiology was consequently affected. The findings also suggest that positive emotional feedback results in improved cognitive performance and faster recovery from physiological arousal than nonemotional feedback. It is noteworthy that the results suggested that positive feedback was effective regardless of the actual performance of the person. This implies that emotional feedback will be effective even with a less than perfect analysis of human performance. This is encouraging for the development of complex emotional technologies. On a more general level, our results clearly showed that synthesized emotional feedback may facilitate the interaction between humans and machines.

Another type of evidence in favor of the use of emotions in human–technology interaction comes from studies using facial EMG and subjective ratings for verifying the evocation of emotional responses in the context of HCI. Unlike earlier findings concerning pupil size, there is an abundance of evidence that facial muscle activity (as measured by EMG) is related to both vocal and facial stimuli that communicate emotions (e.g., Cacioppo and Gardner, 1999; Dimberg, 1990; Hietanen *et al*., 1998; Larsen *et al*., 2003; Surakka and Hietanen, 1998). Consequently, there was no particular need to verify the use of facial EMG responses in emotion research, and we were able to proceed more directly in this study. Thus, in another of our experiments (Partala and Surakka, 2004), we assessed facial EMG responses and ratings of emotional experiences in response to synthetic emotions. We also studied the effects of these synthetic emotions in terms of cognitive performance during a computerized problem-solving task.

In this study, participants relocated different color bars with a mouse according to specific instructions. Participants were exposed to preprogrammed mouse delays in this interactive problem-solving task. Following the mouse delays, emotionally worded positive or negative interventions were given via a speech synthesizer (i.e., Mikropuhe), or in the control condition, no intervention was given. For example, one of the positive interventions involved the computer saying: "The functions of the computer were suspended. The problem will happily soon be over." One of the negative interventions was: "The execution of the program was interrupted. This is annoying." As in our previous speech synthesis studies, prosodic cues were kept constant across all interventions, and only the contents of the interventions varied. Facial EMG responses were recorded from above corrugator supercilii (activated when frowning) and zygomaticus major (activated when smiling) muscle sites (e.g., Fridlund and Cacioppo, 1986). We also measured, among other things, participants' task performance after the different interventions. Finally, participants were asked to rate the different intervention categories on 9-point scales on the dimensions of valence and arousal.

The results showed that subjects rated the interventions as intended, meaning that positive interventions were rated significantly more positively than negative ones. Analysis of problem-solving performance showed better performance following positive than negative interventions. EMG activity was analyzed during and after the different interventions. These results showed that smiling activity was significantly higher during the positive interventions than during the other interventions. It was also significantly higher after the positive than the control condition. Interestingly, frowning activity generally decreased from the baseline during both positive and negative interventions. EMG activity was still below baseline 3 s after the interventions. Frowning activity was significantly lower after the positive interventions than in the control condition. To summarize, the results showed that synthetic speech with emotional content regulated both emotional and cognitive processes. Thus, the results again emphasized the utility of emotions in general and positive emotions in particular (Partala and Surakka, 2004).

It should not be interpreted from the above findings that only positive feedback or intervention is of value in HCI. On the contrary, the results speak in favor of the use of emotions in general because, for example, frowning activity relaxed in response to both negative and positive interventions. Furthermore, negative emotional cues may serve important functions in HCI as well. In a very recent study (Vanhala *et al.*, 2010), we explored how embodied emotionally expressive agents would affect ratings of emotion and attention-related dimensions and physiological responses. Participants were shown a set of female and male embodied agents displaying a negative, a neutral, or a positive facial expression. In addition to the facial cues, the apparent proximity level of the agent was varied by displaying the agent in three different sizes.

The results showed that negative virtual facial cues were effective in capturing attention. Participants rated frowning agents as significantly less pleasant, less relaxing, and more arousing, dominating, conspicuous, and distracting than neutral and smiling agents. The physiological measures provided preliminary evidence suggesting that negative expressions required less mental effort to process. For example, mental effort has been linked with sympathetic activation of the autonomous nervous system (e.g., Chen and Vertegaal, 2004) Sympathetic activity was decreased during negative expressions of the male agent according to our low-frequency heart-rate variability analysis. Thus, the reactions to negative, threatening facial cues seem to be special in terms of the subjective experiences and physiological reactions that they evoke. Regardless of the fact that they are emotionally negative, their utility lies in the fact that they can more efficiently catch one's attention, and that can be very useful for many purposes.

Somewhat similar findings that demonstrate an *anger pop-out* or *threat superiority* effect in responses to human faces have been reported by other researchers (e.g., Fox and Damjanovic, 2006; Hansen and Hansen, 1988). Our study also replicated the earlier results showing that simulated proximity to a computer agent significantly affects the level of the feeling of being in control, so that the bigger the agent, the lower is the experience of being in control (Partala *et al.*, 2004). In both of these studies, agents that appeared to be closer were rated as more dominating than agents that seemed to be further away. Overall, findings from our studies with embodied virtual characters parallel the implications of CASA; that is, similar social and emotional cues (e.g., proximity and facial expressions) seem to be active both in HCI and in human–human interaction. Taking advantage and conscious control of these cues has the potential to significantly facilitate HCI by effecting relatively fast and automatic processes; for example, the rapid allocation of attention to emotional cues in virtual faces.

Developments in emotional perceptual intelligence

Methods for perceptual intelligence require quite complex efforts in different research areas. First, we need engineers to develop new hardware prototypes for unobtrusive measurement technology. Then, the prototypes need to be tested; for example, in the respect that they are able to reliably measure emotion-related responses. Typically, these developments require several iterations before they become functional. Signal-processing methods need to be developed in a similar iterative fashion in order to give computers capacity for lower- and higher-level interpretations from the measured signals. Finally, everything should be coupled together. In the next section, we aim to describe investigations that have dealt with both hardware and software developments in the context of emotions and HCI.

With respect to other emotion-related autonomic nervous system activation correlates, we have recently begun to study different heart-activity measures. Similar to pupil size measurements, recent technological advances have made these measures relatively noninvasive, wireless, and convenient. For example, in one of the earlier studies investigating the relation between emotions and heart rate, heart-rate responses were measured with a regular-looking office chair (Anttonen and Surakka, 2005). This chair contains embedded sensors that are able to detect the heart rate of the person sitting on it. While validating the concept of a nonintrusive measurement chair, we discovered that the chair was able to detect significant changes in heart rate during auditory, visual, and audiovisual stimuli with emotional content. The mean heart rate decelerated in response to emotional stimulation in line with previous findings (e.g., Levenson and Ekman, 2002; Rainville *et al.*, 2006). Further results revealed that heart rate decelerated the most in response to negative stimuli as compared with responses to positive and neutral stimuli.

As both facial activity and heart rate changes have been found to be associated with emotional responding, we performed a further study investigating what kind of heart rate responses could be evoked by computer-guided voluntary facial muscle activations (Vanhala and Surakka, 2007b). Participants were required to activate their facial muscles according to different guidelines as visualized by the computer. Paralleling the methods of biofeedback, participants received real-time visual feedback about the intensity level of their corrugator supercilii and zygomaticus major muscle activations as measured by EMG. Heart activity was registered with a prototype of a wireless electrocardiogram (ECG) measurement patch (Vehkaoja and Lekkala, 2004). We found that activations of both muscles resulted in a deceleration of the mean heart rate, which was extracted from the wireless ECG. This effect was strongest for the moderate intensity level of activations, which were also rated as the most pleasant and easy to perform. The results were encouraging, as they demonstrated how computers can effectively help in regulating emotional and physiological processes.

In a subsequent study, we investigated whether the heart-rate effects could be automatically detected and classified (Vanhala and Surakka, 2007a). The results showed that heart-rate responses to voluntary facial activations could be detected in real time with less than half a second of heart-rate data. In summary, these studies utilizing heart activity have suggested that physiological responses to emotional stimulation can be identified from unobtrusively acquired heart-rate data and that computers can really help in regulating these responses.

Because the facial musculature system is richly developed and well represented in the brain's motor cortex, human facial activity is one promising channel for taking measures to distinguish emotional reactions (e.g., Ekman, 1994; Rinn, 1991). In fact, computing systems built for classifying emotional states into negative, neutral, and positive from continuous facial EMG recordings have shown some promise (Partala *et al.*, 2005, 2006). The setup was as follows. Participants were shown emotionally arousing pictures and videos, and their task was to rate their emotional experiences on a dimensional scale following each stimulation. At the same time, the computer, using various computational regression models, estimated the participants' ratings on the basis of their facial EMG activity. Next, the subjective ratings and the computer's estimations were compared with one another. The comparison showed that the best models were able to estimate the participants' positive and negative ratings of pictorial and video stimuli with average accuracies of over 70 and 80 percent, respectively. Thus, facial EMG is yet another promising method upon which to build perceptual intelligence.

Following these promising results, the development of new wireless electrode measurement technologies has begun (see www.cs.uta.fi/~wtpc/). One of these is an elastic, comfortable-to-wear, facial EMG headband that contains embroidered silver thread electrodes that are "attached" simply by wearing the band. The band includes signal amplifying and radio technology to send the amplified signals wirelessly to machines with computing power. First studies with this prototype have been successful (Nöjd *et al*., 2005; Vehkaoja *et al*., 2005). It is firmly believed that these kinds of advances in technology will bring physiological measurement devices into a wider use in the future.

As a whole, the above research and development has shown that there are several measurements and technologies that can be integrated into computer systems in order to create perceptually intelligent machines. Importantly, many of the measures are becoming increasingly noninvasive, ubiquitous, and convenient. In many cases, the process of carefully attaching sensors according to clinical and experimental standards could be eliminated for the end user of the system. The user would instead merely wear a common accessory (e.g., headband or hat) or be seated on a regular chair that contains unobtrusive sensors.

Discussion

We began our analysis from descriptive findings about surprisingly strong emotional reactions while working with computers. Our further analysis from more controlled experimental studies showed that computers are able to evoke and regulate human emotions while people are working with them. These reactions were more modest than those reported in the public media. However, the findings were clear. It may be that our strong tendency to interpret information socially is the central cause for emotional reactions in HCI, as argued earlier (Brave *et al*., 2005; Reeves and Nass, 1996). On the other hand, it might also be that inanimate objects and especially virtual creatures with or without embodiment automatically evoke our tendency for emotional behavior as a first priority. Following emotions, other social processes like politeness, for example, are activated. In either case, emotions are very much operational and effective in HCI.

Emotions evoked by technology were also found to be otherwise beneficial, as clear proof was found of their effects on participants' emotions and cognitive performance. Positive emotional feedback was found to enhance human mathematical computations. This was the case regardless of the fact that the feedback was randomized, meaning that it had nothing to do with actual performance (Aula and Surakka, 2002). Thus, it may be that positive emotional feedback in particular will be effective even with less than perfect timing and analysis of human performance. Research on emotional interventions (Partala and Surakka, 2004) again showed that emotional messages given by a computer operated in a similar fashion to that in the above study.

The studies reviewed above concerning expressive emotional intelligence showed that even relatively mild, emotion-related social cues can have significant experiential, behavioral, and physiological effects. The use of positive emotional cues seemed to have especially beneficial effects. However, negative cues also showed promise as a method for evoking reactions; for example, in order to draw attention to critical information (Vanhala *et al.*, 2010). The use of negatively toned emotional cues may, in fact, require much more fine-grained analysis of the timing for delivering such cues and more fine-grained analysis of the user performance. Of course, in both cases (i.e., positive or negative emotional cues), more research is needed on the requirements for user performance analysis; for example, with respect to optimal timing for launching emotional interaction between humans and computers. We have shown that it is possible to unobtrusively monitor heart rate and, moreover, to automatically classify changes in the heart rate (Anttonen and Surakka, 2005; Vanhala and Surakka, 2007a). Thus, in this light, automatic analysis of the levels of various physiological processes (e.g., heart rate) is one possibility for detecting the appropriate moments for intervening in users' behaviors. It is possible that people will become annoyed over wrongly timed interventions and feedback, although we have not observed anything to suggest this in our own studies. It is even possible that people will become annoyed over more properly timed communication. There really is much work to do in investigating the synchrony of communication between humans and technology. At this point, we would still like to highlight our finding that people are affected by artificial emotions on many levels (affective, cognitive, social). Perhaps it is well worth risking people's annoyance about computer-generated communication if, on the whole, the benefits clearly outweigh the possibly quite infrequent annoyances.

Moreover, in real life, people tend to be forgiving and tolerant of all kinds of interactions with asynchronies. In this vein, it can be argued that because we tend to treat technology in a human-like manner, we will also be tolerant and forgiving in HCI.

Findings showing promise and potential for expressive intelligence were argued to act as a precondition for further developments in the area of perceptual intelligence. As there is evidence of this potential, there are grounds for developing perceptual intelligence as well. At the moment, there is ongoing work that has already shown that it is possible to develop less obtrusive hardware technologies than, for example, traditional electrode measurement technology. Further technological breakthroughs of this kind are required in the future if we wish to achieve wider acceptance of emotional computing. Our ongoing work has also proved that at least up to a point, software (i.e., signal-processing methods) is capable of continuously analyzing higher-level user processes such as emotional experiences. Of course, there is still a lot to do even regarding these types of inferring algorithms. Moreover, the inclusion of hardware and software technology in the context of use is very much in its early stages.

Studying the effects of emotions and social information in the context of information technology (e.g., HCI, CMC, videoconferencing) clearly reveals the constrained nature of communicative cues available in this context. Due to the limitations of technology, it is unavoidable that, for example, only a face and voice (virtual or real) can be made perceptible. If synthetic characters are used, there is still much to do before they can perfectly mimic human performance, although they do enable the use of a wider range of social cues, such as virtual proximity and facial expressions. At first glance, these shortcomings can be seen as critical barriers to trying to study the possibilities of using this kind of information. However, in some ways, much of the past and current research on human–human emotion has also dealt with virtual or artificial stimuli. For example, studies have widely utilized only facial stimuli (i.e., without voice and body), digitized auditory stimuli (without face and body), etc., yet significant scientific advances concerning human emotions have been made with these stimuli of limited ecological validity. Thus, upon closer inspection, one might argue that studying the effects of emotions in the context of HCI represents the continuation of a relatively well-established research tradition. Furthermore, one could also argue that virtual agents like speech synthesizers, virtual faces, and artificial persons really do offer useful tools for controlled experimentation on the effects of communicative signals on human performance (e.g., Massaro and Egan, 1996). If this is the case, it is reasonable to assume that findings from HCI can be applied directly to human–human interaction.

It is clear that working with computers requires both cognitive and emotional resources even in the absence of dramatic reactions such as swearing at and kicking computers. Perhaps working with computers imposes a continuous cognitive-emotional load, leading to accentuated physiological and psychological arousal. If this arousal is negatively valenced, in the long run there may be serious side effects on our mental and physical health. For example, sympathetic arousal similar to what we have found in response to synthesized emotional cues is known to result, in the long term, in elevated risk of cardiovascular diseases (Malliani *et al*., 1991). One self-evident argument for the necessity of emotionally intelligent technology is that, as we have now seen, computers evoke social and emotional responses in any case. Thus, the question that arises is whether we neglect this or whether we are willing to use this phenomenon in our own favor.

This chapter has presented encouraging findings in favor of using emotional information in HCI. The studies have suggested that emotions can be systematically evoked and regulated by technology in the context of information technology. This kind of emotional HCI can facilitate cognitive performance, regulate attentive processing, and reduce negative physiological arousal. Technology is, and will continue to be, pervasive in everyday life to such an extent that we cannot afford to neglect its effects on emotions. Emotions themselves are equally pervasive factors for all human performance in any context, including HCI.

Acknowledgments

This research has been supported by the Academy of Finland (project numbers 1202183, 177857, 1115997), and the Graduate School in User-Centered Information Technology, Tampere, Finland. Parts of this chapter were published by Surakka (2004) in *Psykologia*, a journal of the Finnish Psychological Association.

References

Allanson, J., Rodden, T., and Mariani, J. (1999). A toolkit exploring electrophysiological human–computer interaction. In M. A. Sasse and C. Johnson (eds), *Human–Computer Interaction – Interact '99* (pp. 231–237). Amsterdam: IOS Press.

Anttonen, J. and Surakka, V. (2005). Emotions and heart rate while sitting on a chair. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI* 2005, (pp. 491–499). Portland, OR: ACM Press.

Aula, A. and Surakka, V. (2002). Auditory emotional feedback facilitates human–computer interaction. In X. Faulkner, J. Finlay, and F. Détienne (eds), *People and Computers XVI: Memorable Yet Invisible, Proceedings of the British HCI* (pp. 337–349). London: Springer.

Barreto, A., Scargle, S., and Adjouadi, M. (2000). A practical EMG-based human–computer interface for users with motor disabilities. *Journal of Rehabilitation Research and Development*, ***37***, 53–64.

Berry, D. C., Butler, L. T., and de Rosis, F. (2005). Evaluating a realistic agent in an advice-giving task. *International Journal of Human–Computer Studies*, ***63***, 304–327.

Bradley, M. M. and Lang, P. J. (1999). *International Affective Digitized Sounds (IADS): Stimuli, Instruction Manual and Affective Ratings.* Technical Report B-2. Gainesville, FL: University of Florida: Center for Research in Psychophysiology.

(2000). Affective reactions to acoustic stimuli. *Psychophysiology*, ***37***, 204–215.

Brave, S., Nass, C., and Hutchinson, K. (2005). Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent. *International Journal of Human–Computer Studies*, ***62***, 161–178.

Cacioppo, J. T. and Gardner, W. L. (1999). Emotion. *Annual Review of Psychology* ***50***, 191–214.

Cassell, J., Bickmore, T., Campbell, L., Vilhjálmsson, H., and Yan, H. (2000). Human conversation as a system framework: designing embodied conversational agents. In J. Cassell, J. Sullivan, S. Prevost, and E. Churchill (eds), *Embodied Conversational Agents* (pp. 29–63). Cambridge, MA: MIT Press.

Chen, D. and Vertegaal, R. (2004). Using mental load for managing interruptions in physiologically attentive user interfaces. In E. Dykstra-Erickson and M. Tscheligi (eds), *Extended Abstracts of the 2004 Conference on Human Factors and Computing Systems* (pp. 1513–1516). Vienna: ACM Press.

Concord Communications (1999). Concord network rage survey. www.concord.com/library/network_rage/.

Damasio, A. (1994). *Descartes' Error: Emotion, Reason and the Human Brain*. New York: Grosset.

Dimberg, U. (1990). Facial electromyography and emotional reactions. *Psychophysiology*, ***27***, 481–494.

Ekman, P. (1994). Strong evidence for universals in facial expressions: a reply to Russell's mistaken critique. *Psychological Bulletin*, ***115***, 268–287.

Fogg, B. J. and Nass, C. (1997). Silicon sycophants: the effects of computers that flatter. *International Journal of Human–Computer Studies*, ***46***, 551–561.

Fox, E. and Damjanovic, L. (2006). The eyes are sufficient to produce a threat superiority effect. *Emotion*, ***6***, 534–539.

Fridlund, A. J. and Cacioppo, J. T. (1986). Guidelines for human electromyographic research. *Psychophysiology*, ***23***, 567–589.

Hansen, C. H. and Hansen, R. D. (1988). Finding the face in the crowd: an anger superiority effect. *Journal of Personality and Social Psychology*, ***54***, 917–924.

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

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

Ilves, M. and Surakka, V. (2004). Subjective and physiological responses to emotional content of synthesized speech. In N. Magnenat-Thalmann, C. Joslin, and H. Kim (eds), *Proceedings of the 17th International Conference on Computer Animation and Social Agents, CASA 2004* (pp. 19–26). Geneva: Computer Graphics Society (CGS).

Jacob, R. J. K. (1991). The use of eye movements in human computer interaction techniques: what you look at is what you get. *ACM Transactions on Information Systems*, ***9***, 152–169.

(1996). The future of input devices. *ACM Computing Surveys*, ***28***(A) (Annex), December.

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

Klein, J., Moon, Y., and Picard, R. W. (2002). This computer responds to user frustration: theory, design, and results. *Interacting with Computers*, ***14***, 119–140.

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

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

Lang, P. J., Bradley, M. M., and Cuthbert, B. N. (1995). *International Affective Picture System (IAPS): Photographic Slides*. Gainesville, FL: University of Florida, Center for the Study of Emotion and Attention.

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

Larsen, J. T., Norris, C. J., and Cacioppo, J. T. (2003). Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii. *Psychophysiology*, ***40***, 776–785.

LeDoux, J. E. (1994). Emotion, memory and the brain. *Scientific American*, ***270*** (6), 32–39.

(1998). *The Emotional Brain*. New York: Simon & Schuster.

Levenson, R. W. and Ekman, P. (2002). Difficulty does not account for emotion-specific heart rate changes in the directed facial action task. *Psychophysiology*, ***39***, 397–405.

Loewenfeld, I. E. (1966). Pupil size. *Survey of Ophthalmology*, ***11***, 291–294.

Malliani, A., Pagani, M., Lombardi, F., and Cerutti, S. (1991). Cardiovascular neural regulation explored in the frequency domain. *Circulation*, ***84***, 482–492.

Marsh, P. and Khor, Z. (2006). Life online: mouse rage! Social Issues Research Centre. www.sirc.org/publik/mouse_rage.pdf.

Massaro, D. W. and Egan, P. B. (1996). Perceiving affect from the voice and the face. *Psychonomic Bulletin and Review*, ***3***, 215–221.

D'Mello, S., Graesser, A., and Picard, R. W. (2007). Toward an affect-sensitive AutoTutor. *IEEE Intelligent Systems*, ***22***(4), 53–61.

Nass, C., Isbister, K., and Eun, J. L. (2000). Truth is beauty: researching embodied conversational agents. In J. Cassell, J. Sullivan, S. Prevost, and

E. Churchill (eds), *Embodied Conversational Agents* (pp. 374–402). Cambridge, MA: MIT Press.

Nass, C. and Moon, Y. (2000). Machines and mindlessness: social responses to computers. *Journal of Social Issues*, ***56***, 81–103.

Nass, C. and Steuer, J. (1993). Voices, boxes, and sources of messages: computers and social actors. *Human Communication Research*, ***19***, 504–527.

Nass, C., Steuer, J., and Tauber, E. R. (1994). Computers are social actors. *Proceedings of CHI '94*, (pp. 72–78). Boston: ACM.

Nöjd, N., Puurtinen, M., Niemenlehto, P., Vehkaoja, A., Verho, J., Vanhala, T., Hyttinen, J., Juhola, M., Lekkala, J., and Surakka, V. (2005). Wireless wearable EMG and EOG measurement system for psychophysiological applications. In R. Lundström, B. Andersson, and H. Grip (eds), *Proceedings of the Nordic Baltic Conference on Biomedical Engineering and Medical Physics* (pp. 144–145). Umeå: Swedish Society for Medical Engineering and Medical Physics.

Partala, T., Jokiniemi, M., and Surakka, V. (2000). Pupillary responses to emotionally provocative stimuli. In A. T. Duchowski (ed.), *Proceedings of ETRA 2000, Eye Tracking Research and Applications* (123–129). Palm Beach Gardens, FL: ACM Press.

Partala, T., Aula, A., and Surakka, V. (2001). Combined voluntary gaze direction and facial muscle activity as a new pointing technique. In M. Hirose (ed.), *Proceedings of INTERACT 2001* (pp. 100–107). Tokyo: IOS Press.

Partala, T. and Surakka, V. (2003). Pupil size variation as an indication of affective processing. *International Journal of Human Computer Studies*, ***59***1–2, 185–198.

(2004). The effects of affective interventions in human–computer interaction. *Interacting with Computers*, ***16***, 295–309.

Partala, T. Surakka, V., and Lahti, J. (2004). Affective effects of agent proximity in conversational systems. In *NordiCHI: Proceedings of the Third Nordic Conference on Human–Computer Interaction* (pp. 353–356). New York: ACM Press.

Partala, T., Surakka, V., and Vanhala, T. (2005). Person-independent estimation of emotional experiences from facial expressions. In J. Riedl, A. Jameson, D. Billsus, and T. Lau (eds), *Proceedings of the 10th International Conference on Intelligent User Interfaces* (pp. 246–248). New York: ACM Press.

(2006). Real-time estimation of emotional experiences from facial expressions. *Interacting with Computers*, ***18***, 208–226.

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

Picard, R. (1997). *Affective Computing*. Cambridge, MA: MIT Press.

Rainville, P., Bechara, A., Naqvi, N., and Damasio, A. R. (2006). Basic emotions are associated with distinct patterns of cardiorespiratory activity. *International Journal of Psychophysiology*, ***61***, 5–18.

Reeves, B. and Nass, C. (1996). *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge University Press.

Rinn, W. E. (1991). The neuropsychology of facial expression. In R. S. Feldman and B. Rimé (eds), *Fundamentals of Nonverbal Behavior* (pp. 3–30). Cambridge University Press.

Salminen, K., Surakka, V., Lylykangas, J., Raisamo, J., Saarinen, R., Rantala, J., Raisamo, R., and Evreinov, G. (2008). Emotional and behavioral responses to haptic stimulation. In *Proceedings of SIGCHI Conference on Human Factors in Computing Systems* (pp. 1555–1652). New York: ACM Press.

Scherer, K. R., Ladd, D. R., and Silverman, K. E. A. (1984). Vocal cues to speaker affect: testing two models. *Journal of the Acoustical Society of America*, ***76***, 1346–1356.

Schilbach, L., Helmert, J. R., Mojzisch, A., Pannasch, S., Velichkovsky, B. M., and Vogeley, K. (2005). Neural correlates, visual attention and facial expression during social interaction with virtual others. In *Proceedings of ICCS-2005 Symposium, Toward Social Mechanisms of Android Science*, Stresa, Italy (74–86).

Schilbach, L., Wohlschlaeger, A. M., Kraemer, N. C., Newen, A., Shah, N. J., Fink, G. R., and Vogeley, K. (2006). Being with virtual others: neural correlates of social interaction. *Neuropsychologia*, ***44***, 718–730.

Schlosberg, H. (1954). Three dimensions of emotion. *Psychological Review*, ***61***, 81–88.

Surakka, V. (2004). Tunteet ja sosiaalisuus ihminen-tietokone vuorovaikutuksessa [Emotions and sociality in human–computer interaction]. *Psykologia*, ***39***, 19–28.

Surakka, V. and Hietanen, J. K. (1998). Facial and emotional reactions to Duchenne and non-Duchenne smiles. *International Journal of Psychophysiology*, ***29***(1), 23–33.

Surakka, V., Illi, M., and Isokoski, P. (2003). Voluntary eye movements in human–computer interaction. In J. Hyönä, R. Radach, and H. Deubel (eds), *The Mind's Eyes: Cognitive and Applied Aspects of Oculomotor Research* (pp. 473–491). Oxford: Elsevier Science.

(2004). Gazing and frowning as a new human–computer interaction technique. *ACM Transactions on Applied Perception*, ***1***, 40–56.

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

Vanhala, T. and Surakka, V. (2005). An agent framework for the development of psycho-physiologically interactive systems. In *Proceedings of HCI International 2005*. CD-ROM. Mahwah, NJ: Lawrence Erlbaum.

(2007a). Recognizing the effects of voluntary facial activations using heart rate patterns. In *Proceedings of the 11th WSEAS International Conference on Computers*, 628–632.

(2007b). Facial activation control effect (FACE). In A. Paiva, R. Prada, and R. W. Picard (eds), *Proceedings of ACII 2007, Lecture Notes in Computer Science*, vol. 4738 (pp. 278–289). Lisbon: Springer.

Vanhala, T., Surakka, V., Siirtola, H., Räihä, K.-J., Morel, B., and Ach, L. (2010). Virtual proximity and facial expressions of computer agents regulate human emotions and attention. *Computer Animation and Virtual Worlds*, ***21***, 215–224.

Vehkaoja, A. and Lekkala, J. (2004). Wearable wireless biopotential measurement device. In *Proceedings of the IEEE EMBS 2004* (pp. 2177–2179). San Francisco, CA: IEEE Press.

(2006). Wireless measurement band for EEG mismatch negativity registration in mobile activities. *IMEKO XVIII World Congress and IV Brazilian Congress of Metrology*.

Vehkaoja, A., Verho, J., Puurtinen, M., Nöjd, N., Lekkala, J., and Hyttinen, J. (2005). Wireless Head Cap for EOG and Facial EMG Measurements. *27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, ***6***, 5865–5868.

Weyers, P., Mühlberger, A., Hefele, C., and Pauli, P. (2006). Electromyographic responses to static and dynamic avatar emotional facial expressions. *Psychophysiology*, ***43***, 450–453.

Wolpaw, R. J., Birbaumer, N., Mcfarland, D. J., Pfurtscheller, G., and Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. *Clinical Neurophysiology*, ***113***, 767–791.

Zajonc, R. B. (1980). Feelings and thinking: preferences need no inferences. *American Psychologist*, ***35***(2), 151–175.

Appendix B

Publication II

Vanhala, T. and Surakka, V. (2008). Computer-assisted regulation of emotional and social processes. In Or, J. (Ed.) *Affective Computing*, pp. 405–420. Vienna, Austria: InTech Education and Publishing.

Available at:
`http://www.intechopen.com/articles/show/title/computer-assisted_regulation_of_emotional_and_social_processes`

22

Computer-Assisted Regulation of Emotional and Social Processes

Toni Vanhala and Veikko Surakka
Research Group for Emotions, Sociality, and Computing, TAUCHI, Department of Computer Sciences, University of Tampere Finland

1. Introduction

Imagine a person who has a fear of other people. Let us call her Anna. She is afraid of people watching her every move as she stands in a line or walks down the street. Meeting new people is almost impossible as she always feels stared at and judged by everyone. This fear, or maybe even a phobia, can make Anna's life very complicated. It is difficult for her to travel through public spaces in order to get to work, to deal with a bus or taxi driver, shop for groceries, etc. Anna's leisure time activities are also very limited. The situation is indeed a vicious cycle, as it is even difficult for her to seek treatment and go to a therapist.

In USA alone, there are approximately 15 million people like Anna who suffer from social anxiety disorder (Anxiety Disorders Association of America, 2008). A total of 40 million people suffer from different anxiety disorders. The associated yearly costs of mental health care exceed 42 billion U.S. dollars. Thus, emotional disorders are a significant public health issue. There is a need for demonstrably effective and efficient new methods for therapy.

Computer systems have recently been applied to the treatment of many emotional disorders, including different phobias (Krijn et al., 2004; Wiederhold & Bullinger, 2005). These systems provide controlled virtual exposure to the object of the disorder, for example, a computer simulation of a spider or a room filled with other people. In this form of behavioural therapy, patients are systematically desensitized by gradual exposure to a computer generated representation of the object of their fear (Weiten, 2007; Krijn et al., 2004). At first, the level of exposure is kept mild and constant, for example, by keeping the object of the fear visually distant and far away. Then, the level of exposure is increased little by little, for example, by moving a virtual spider closer or increasing the number of virtual people. The underlying theory is that such exposure replaces anxiety provoking memories and thoughts with more neutral ones that are acquired in a safe, controlled environment.

It has been shown that people react to computer generated stimuli in the same manner as to authentic, real-life stimuli. For example, socially anxious people are cautious about disturbing embodied artificial characters in virtual reality (Garau et al., 2005). People have also reported higher anxiety and shown increased somatic responses when speaking to negative as compared to neutral and positive audiences consisting of virtual agent characters (Pertaub et al., 2002). As these studies have shown that virtual characters are able to evoke emotions or anxiety, computer generated stimuli show clear potential as a new method for treating different social and emotional disorders by enabling controlled exposure to anxiety provoking stimuli.

Advantages of computer generated stimuli include accurate control of the grade of exposure, the relative easiness of creating diverse stimuli and varying their characteristics, and the cost-efficiency of therapy. For example, a person who has a phobia of flying can experience a whole virtual air plane trip from take-off to landing at a relatively low cost. Further, the experience can be replicated again and again with small variations to factors that would be very difficult to control in real situations. Virtual environments even allow the re-enactment of traumatic episodes, such as bombings and car accidents. In fact, there are various conditions that have been successfully treated using virtual exposure to artificial stimuli, including fear of flying, fear of driving, fear of confined spaces, fear of public speaking, social phobia, post-traumatic disorders, and panic disorders (Krijn et al., 2004; Wiederhold & Bullinger, 2005). However, Krijn et al. (2004) concluded in their review of virtual exposure methods that there is little conclusive evidence about the relative effectiveness of virtual reality and real, *in vivo* exposure. One particular concern was the lack of evidence for the effectiveness of virtual exposure therapy as a stand-alone treatment.

There is evidence that the effectiveness of exposure therapy can be further improved by applying physiological measurements (Wiederhold & Wiederhold, 2003). For example, physiological signals can be registered and displayed to the patient during exposure therapy (Wiederhold & Bullinger, 2005; Wiederhold & Wiederhold, 2003). This way, the patient can gain awareness of physiological processes and learn to voluntarily control them. Voluntary control of emotion-related physiological functions has been shown to influence emotional reactions associated with, for example, fear and facial expressions (Vanhala & Surakka, 2007a; Wiederhold & Wiederhold, 2003). In other types of setups, the clinician can monitor these signals, estimate the progress of therapy, and adjust its intensity accordingly.

Previous research has established a number of physiology-based measures of emotional states that can be used as a basis for adapting the therapy (Vanhala & Surakka, 2007a; Vanhala & Surakka, 2007b; Partala et al., 2006; Anttonen & Surakka, 2005; Wilhelm et al., 2006; Surakka & Vanhala, accepted). These measures include electrical brain and facial muscle activity, heart rate, respiration, and skin conductivity. However, it is not possible to use a single measure as an index of emotional states, as each individual measure is affected by a number of psychological and physiological factors (Ward & Marsden, 2003). Emotions themselves are often categorized according to a number of dimensions, such as arousal and emotional valence (Bradley & Lang, 1994). Further, emotional processes are tightly interconnected with other psychophysiological processes, including cognition and attention (Surakka et al., 1998). Thus, it is necessary to take other psychophysiological processes (e.g., attention) into account when recognizing and analyzing emotions (Ward & Marsden, 2003).

As multi-signal, online monitoring of human psychophysiology involves signals with several varying characteristics (e.g. sample rate and frequency content) and as each measure reflects several inter-linked physiological systems, the amount of information can easily overwhelm a human operator. One way to deal with this challenge is to build perceptual intelligence into computers themselves (Pentland, 2000). Signal analysis of measured psychophysiological signals and states could be performed automatically. Further, the role of human actors in this kind of a virtual therapy system could be changed. Currently, humans need to process all information that is used to control the parameters of a virtual therapy system. Proactive computing could be used to remove this bottleneck (Tennenhouse, 2000). A system that responds to the emotional and physiological state of a person could automatically adapt the computer system according to the rules of desensitization. This way, both the person being treated and the therapist could focus on training to regulate emotions instead of actively interpreting and estimating them.

The main goal of the present work is to present a new model for computer systems that proactively support emotion regulation. First, in the next section we present examples of single and compound measures of psychophysiological states that could be used to build perceptual intelligence for this kind of a system. Then, in the following section we discuss studies demonstrating the effectiveness of computers in regulating emotions. In the process we identify several computer-generated stimuli that could be used to influence emotional and social processes. In the fourth section we combine these findings into a model that supports both computer-assisted regulation and voluntary control of emotion related psychological and physiological processes. Finally, we discuss the advantages and challenges of this model and suggest pertinent research areas for future work.

2. Measuring emotions

As our aim is to support the regulation of emotions, we need to be able to evaluate the results of this regulation, that is, changes in emotional responding. Researchers generally view emotions as a concurrent change in experiential, behavioural, and physiological systems that organize human motivational behaviour (e.g., Frijda, 1986; Mauss et al., 2005). Thus, our first task is to identify measures that capture a wide view of emotional processes. There have been two research traditions of emotions. The first tradition views emotions as discrete states, such as, disgust, fear, joy, sadness, and surprise (Ekman, 1993). The second tradition views emotions as a three-dimensional space varying in emotional valence, arousal, and dominance (Bradley & Lang, 1994; 2000). These traditions have direct consequences especially for measuring the experiential component of emotions. For example, one common method is to ask people to rate their experiences using bipolar scales of emotional valence (i.e., from negative to positive), arousal (i.e., from calm to aroused), and dominance (i.e., from feeling of being in control to being controlled).

The measurement of the experiential component of emotion often requires that the person is interrupted and asked to report her or his experiences. For example, during exposure therapy patients are periodically asked to rate the intensity of their anxiety using a scale of subjective units of discomfort (SUD) ranging from 0 to either 10 or 100 (see, e.g., review by Krijn et al., 2004). The rating is used to evaluate when the level of anxiety has changed and requires the therapist to adapt the exposure. When the anxiety is very high, the exposure may be decreased, for example, with instructed relaxation. Low anxiety suggests that the patient is ready to proceed to a higher level of exposure, for example, to take one step closer to a spider. This way, the person is gradually exposed to the object of their fear and habituated to ever increasing amounts of exposure in the process.

The drawback of reporting subjective experiences is that it distracts the person's attention from any ongoing tasks that she or he may be performing. This may hinder a person's experience of being present in the virtual therapy environment. There is some evidence pro the view that this feeling is critical for the success of exposure therapy, as it is required for the experience of relevant emotions and learning to regulate them (Krijn et al., 2004). In this sense, behavioural and physiological components of emotion are somewhat more convenient to measure. It is feasible to acquire these measures continuously and in real-time without distracting the person (Öhman et al., 2000; Teller, 2004; Wilhelm et al., 2006; Mandryk & Atkins, 2007). This also creates potential for more accurate timing of emotional responses. For example, the exact time of a reaction to some surprising event is more readily identified from changes in facial behaviour as compared to a post study questionnaire.

Measures of facial behaviour have been frequently used for detecting emotional responses. For example, Ioannou and others (2005) reported results from using an adaptive system to classify the facial behaviour of one person. The system classified emotional facial expressions into three classes based on features extracted from video images. The classes represented three out of four quadrants of a two-dimensional emotional space (i.e., high arousal – negative, high arousal – positive, low arousal – negative). The classification accuracy of a general (i.e., person-independent) model was about 58%. After adapting this model to the particular person, the performance increased to approximately 78%.

In contrast to Ioannou et al. (2005), typically the classes in video-based classification of facial behaviour have been based on a view of discrete emotions (see, e.g., reviews in Donato et al., 1999; Cowie et al., 2001). The accuracies for these kinds of classifiers are impressive at their best. For example, Sohail & Bhattacharya (2007) reported an average accuracy of over 90% in classifying six emotional facial expressions. However, discrete classifiers usually do not address the intensity of emotional states which is used in adapting the amount of virtual exposure. As an exception, Bailenson (in press) recently developed a classifier for both the discrete facial expression and the intensity of the expression. In any case, most previously investigated discrete classifiers are limited in their applicability to virtual therapy.

Video-based measures can be used to detect facial activity in a non-invasive manner, for example, without restricting the movements of the person by electrode wires. However, video-based methods can only be used to detect clearly visible facial behaviour, while electrophysiological measures have the potential to register very small changes in muscle activity (Ekman et al., 2002). There is also evidence that physiological measures can reflect emotional responses that do not evoke observable behaviour (e.g., Gross & Levenson, 1997). Furthermore, video-based measures are very sensitive to lighting and head orientation as well as to inaccuracies in detecting facial landmarks (e.g., Cowie et al., 2005). For these reasons, physiological measures may be seen to reflect a more objective (e.g., context-independent) view of the emotional response.

A common method for measuring the physiological activity that underlies visible facial behaviour is electromyography (EMG). Facial EMG is performed by attaching electrodes that register the electrical activity of facial muscles over specific muscle sites (Tassinary & Cacioppo, 2000). Especially the EMG activity of the *corrugator supercilii* (activated when frowning) and the *zygomaticus major* (activated when smiling) muscles has been frequently found to co-vary with subjective experiences of emotional valence (e.g., Lang et al., 1993; Larsen et al., 2003). The *corrugator supercilii* muscle which knits and lowers the brow is located in the forehead. Its activity has been found to increase when a person experiences negative emotions and to decrease during positive experiences. The *zygomaticus major* is a relatively large muscle located in the cheek. When activated it pulls up the corner of the mouth. The intensity of its activity varies with emotional valence in the opposite manner to the *corrugator supercilii* muscle.

Although some physiological reactions are quite straight-forward to interpret, humans do not normally evaluate emotional expressions of other people from electrophysiological data. Even one electrophysiological signal can contain lots of information, which may overwhelm a human observer. For example, facial EMG activity may reflect both the intensity of facial muscle activations and the fatigue in muscles (Tassinary & Cacioppo, 2000).

Automatic analysis and interpretation of physiological signals can help in perceiving which changes in signals are related to emotional processes. There is evidence that even the subjective component of an emotion (i.e., emotional experiences) can be automatically estimated from electrical facial muscle activity. For example, Partala and others (2005; 2006) were able to build systems that automatically estimated and classified emotional experiences evoked by picture and video stimuli. The first system (Partala et al., 2006) was adapted to the individual responses of each person as follows. First, participants were shown a calibration block of 24 pictures selected from the standardized set of International Affective Picture System (IAPS). After each stimulus, the participant rated the emotional valence that she or he experienced using a 9-point bipolar scale. Then, the statistical models that estimated the emotional valence were adapted to the person based on the ratings and the EMG data from the calibration block. Finally, the system was tested using 28 pictures and six videos that showed a female or a male actor crying, smiling, and portraying a neutral facial expression. Subjective ratings of emotional valence were collected after each stimulus. These ratings and the system's estimate of emotional valence were compared in order to determine the accuracy of the system. The results showed that the best models were able to separate negative and positive emotional responses with accuracies of over 70 percent for pictures and over 80 percent in the case of video stimuli. Further, the largest correlation between the subjective ratings and the system's estimate of emotional valence on a 9-point scale was over 0.9. Thus, the results of the first system showed that subjective emotional experiences can be estimated based on measures of electrical facial activity with relatively simple models in real-time. Although there is still room for improvement, the accuracy achieved in this study is already sufficient for many applications.

The second system was person-independent and therefore did not require a separate calibration period (Partala et al., 2005). The valence of emotional experiences was estimated based on the direction of change in EMG activity from a baseline period of 0.5 seconds before stimulus onset. This system was able to distinguish between reactions rated as positive or negative at an accuracy of nearly 70 percent for pictures and over 80 percent for videos. In summary, facial activity shows clear promise as a reliable measure for automatic, real-time classification of emotional valence, as both person-adapted and person-independent systems were demonstrated to perform at a reasonable accuracy.

In addition to measures of electrical facial muscle activity, there is a wide variety of other physiological measures that have been shown to vary between emotional reactions, such as the mean heart rate and its frequency components (Anttonen & Surakka, 2005; Levenson & Ekman, 2002; Bradley, 2000; Malliani et al., 1991). For example, Rainville and others (2006) investigated classification of emotional responses using a large set of heart activity and respiration related features. Participants recalled and experientially relived one or two autobiographical episodes associated with the experience of fear, anger, sadness, or happiness. The system was able to detect which of the four emotions the participant was experiencing (i.e., according to subjective ratings) at an accuracy of about 65%.

One challenge that has rarely been investigated in previous classification studies is the recognition and accurate timing of emotional responses. In other words, participants themselves have typically reported the onset and offset of emotional responses and data has been segmented by hand. It is clear that in order to react to the events in real-time, a system should be able to segment the collected data without human intervention. Vanhala & Surakka (2007b) recently reported a study of this kind of an online system. The system automatically detected the onset and offset of emotion related events (i.e., voluntary smiling and frowning) based on less than half a second of heart rate data. The onset of activity was detected with a statistically significant accuracy of 66.4% and the offsets were detected with an accuracy of 70.2%. However, the rate of false recognitions was 59.7% which is quite high. Thus, the results showed that the heart rate can be used to support recognition and classification of emotional responses, but it should be used as one of several corroborative measures in practical applications.

In fact, previous studies have usually employed more than one measure in classifying emotional states (e.g., Kim et al., 2004; D'Mello et al., 2007; Mandryk & Atkins, 2007). Otherwise, recognizing mental states and responses can be challenging, as physiological responses are person-dependent and they reflect several overlapping reactions and mental processes. For example, Bailenson and others (in press) compared classifiers that used facial activity as such or combined it with several physiological measures of heart activity, skin conductance, and finger temperature. The use of physiological measures significantly improved the precision of classification (i.e., the proportion of correctly classified samples in each classified group) as compared to classifiers that used only hand-coded facial features. The best improvements were over 15% for classifying sadness and about 9% for classifying amusement. Similarly, Zeng and others (2004) were able to improve the accuracy of their emotion classification system to 90 percent when both facial expressions and prosodic cues of speech were used. When only one of these modalities was used, the accuracies dropped to 56 and 45 percent, respectively. Busso and others (2004) achieved similar results with a system that recognized emotions from speech and facial expressions. In an earlier work, Picard and others (2001) identified specific physiological responses from four physiological signals (i.e., facial electromyogram, blood volume pressure, skin conductance, and respiration) and used these response patterns in classifying emotional experiences to eight classes. They achieved a classification accuracy of 81 percent.

The measurement of bioelectric signals can be criticized based on the complex arrangements (e.g., electrodes, amplifiers, and skin cleaning) that are required for measuring them. Recently, several wireless and non-invasive technologies have been developed for measuring physiological signals, including facial EMG (e.g., Anttonen & Surakka, 2005; Teller, 2004; Wilhelm et al., 2006). For example, the electrical activity of forehead muscles (e.g., *corrugator supercilii*) can be measured with an easy-to-wear wireless headband that contains embroidered silver thread electrodes (Vehkaoja & Lekkala, 2004; Nöjd et al., 2005). As another example of non-invasive and easily applied measurement technology, Anttonen and Surakka (2005; 2007) were able to reliably measure emotion related heart rate changes with a regular looking office chair. The chair contained embedded electromechanical sensors in the seat, arm rests, and back rest. The sensors can be used to detect pressure changes due to heart activity, body movement, or changes in posture. Based on these recent advances in non-invasive technologies, physiological measures are quickly catching up on the current benefits of video-based methods for tracking changes in emotion related behaviour.

In summary, there are several well-tried methods for measuring the different aspects of emotion. Our present review suggested that especially physiological measures show potential as objective and sensitive measures of emotion related processes. Thus, there is no need to rely on any single measure of emotional processes, such as SUD in adjusting the exposure in virtual therapy. In fact, typically several measures have been fused together in order to derive more accurate compound measures. This also helps in interpreting the data, as it can be pre-processed into a form that is more accessible to a human observer. Further, physiological measures are less prone to distract the person as they can be continuously acquired without intervention. However, monitoring emotion related processes can still require considerable human effort after integration and interpretation by the computer. The model that we present in the current paper is aimed to facilitate this work.

3. Regulating emotions with computers

Social and emotional cues from computers have been found to evoke significant responses in their human observers. For example, synthesized speech with emotional content has been found both to evoke positive emotions and to enhance problem solving activity (Partala & Surakka, 2004; Aula & Surakka, 2002). Aula & Surakka (2002) used synthesized speech to provide neutral, negatively, or positively worded auditory feedback that seemed to reflect participant's performance in solving arithmetic problems. In reality, the content of feedback was random and independent of the participant's performance. Nonetheless, positive feedback significantly facilitated the speed of solving problems. In a later study, Partala & Surakka (2004) investigated emotionally worded interventions after a pre-programmed mouse delay during computerized puzzle solving tasks. Similar to the previous study, problem solving performance was significantly better after positively worded interventions. In terms of facial EMG measurements, participants also smiled more and frowned less after positive interventions as compared to facial activity after neutral and negative interventions. These kinds of studies have shown that explicit feedback and interventions from computers can affect human cognitive and emotional processes. There is also evidence that even more subtle social and emotional cues are effective in human-computer interaction. For example, in one of the first studies of virtual proximity, Partala and others (2004) investigated reactions to the simulated distance of a virtual head. When the head appeared to be closer, participants rated that they felt dominated by it. Vice versa, when the head was further away, participants felt that they were controlling it. Vanhala and others (submitted) recently found similar subjective dominance reactions to the simulated proximity of an embodied computer agent. Some researchers have even described computers as social actors, meaning that people have a strong tendency to behave socially when interacting with computers (Nass et al., 1994; Reeves & Nass, 1996).

The effectiveness of virtual stimuli in evoking social and emotional reactions is the basis for virtual exposure therapy. The idea is that new neutral memory structures are formed during virtual exposure. These memory structures should replace the previous anxiety related structures when responding to real-life situations (Krijn et al., 2004). In other words, people should react to provoking virtual stimuli in the same manner as to authentic, real-life stimuli. There are some studies that support his view. For example, socially anxious people get highly distressed when they talk to or need to disturb embodied artificial characters in virtual reality (Pertaub et al., 2002; Garau et al., 2005). Further, the effects of virtual exposure to spiders have been found to generalize to real-life behaviour as measured by the Behavioural Avoidance Test (Garcia-Palacios et al., 2002). That is, people were able to approach a real spider more easily after exposure to a virtual one.

In addition to these computer generated stimuli that regulate emotional responses, emotions can also be actively self-regulated. Gross & Thompson (2007) have described the development of emotion self-regulation as a continuum. In the first stages emotions are consciously regulated. Later, emotion regulation becomes more automatic and effortless. Thus, the process of learning to regulate emotions resembles the process of skill acquisition in general (Anderson, 2000). In this view, less skilled emotion regulators use cognitive processes extensively to support the regulation. For example, they may deliberately rely on instructions and examples of successful regulation. After practise, the regulation of emotions becomes autonomous and efficient, demanding much less cognitive processing. For example, a skilled self-regulator does not need to explicitly apply instructions (e.g., from a therapist) in order to regulate emotions.

Instructions that support emotion regulation may be relatively simple. For example, Vanhala & Surakka (2007a) investigated whether computer-guided voluntary facial activations have an effect on autonomous nervous system activity. Participants were instructed to activate either the *corrugator supercilii* muscle or the *zygomaticus major* muscle at one of three intensity levels (i.e., low, medium, or high). Instructions for each task and real-time feedback about the intensity of facial muscle activations were provided to the participant on the computer screen. Subjective ratings of emotional valence were collected after the activation. It was found that different muscle activations produced both task-specific emotional experiences and significant changes in heart rate and heart rate variability. Thus, the results showed that relatively simple computer-instructions allow people to actively influence their involuntary physiological reactions and subjective experiences that are associated with emotions.

Physiological feedback as such can also help in learning to regulate emotions. Usually, either skin conductivity or breathing patterns are registered and displayed to the patient or the therapist during computer-assisted therapy sessions (Wiederhold & Bullinger, 2005). This way, a person can become aware of unconscious physiological responses and processes, which can enable voluntary control over them. As an impressive example in favour of the effectiveness of virtual exposure therapy, Wiederhold & Wiederhold (2003) followed the behaviour of a group of 10 patients inflicted with fear of flying who were treated using virtual exposure and physiological feedback. As the terrorist attacks on September 11th, 2001 were quite directly related to flying, they could have caused relapses in terms of intensifying the fear of flying in these patients. However, everyone of this group was able to fly without medication or further treatment just four months after the attacks.

Physiological data can also be collected for later reflection. For example, Lindström and others (2006) presented an "affective diary" that provided a multimodal (i.e., auditory and visual) representation of sensor data. A measure of arousal was extracted from the physiological measures and the estimated level of arousal affected the posture of a virtual character displayed on the screen. Users of the diary could later reflect their experiences and manipulate the character in order to match it to their recollection of those feelings. This application illustrates the interplay of involuntary emotion related physiological reactions (i.e., visually coded sensor data) and voluntary regulation of emotions (i.e., later reflection and adjustment in "affective diary"). However, a crucial component for supporting the training of emotion regulation is the online adjustment of emotional stimulation, for example, the amount of exposure to virtual stimuli. This requires a real-time system for the evaluation and reflection of psychological and physiological processes.

In summary, computer systems show potential for regulating human emotions. First, studies have shown that people react socially and emotionally to computers and virtual environments. Second, the effects of virtual stimuli (e.g., habituation of anxiety responses) have been further facilitated when feedback of emotion related physiological activity has been provided. Third, we found that voluntary regulation of emotion related processes seems to be a potential key factor both in learning the regulation as such but also in modulating the functioning of involuntary mechanisms activated during emotion related processes. Fourth, we reviewed a large number of physiological measures that show potential as sensitive and objective measures of emotional responses. The relevant information from these measures could be extracted using automatic classification of emotional responses. This way, it would be possible to avoid overwhelming a human observer, while still using all of the available information in order to maximally support emotion regulation. In the next section we present a model that supports this goal by integrating perceptual intelligence to the system.

4. Adaptive support for emotion regulation

Figure 1 shows a model of how virtual exposure therapy is currently performed. The model contains a set of different actors that take part in an interaction loop. First, the relevant emotional state is observed using different emotion related measures. Then, a human facilitator monitoring these measures decides how the virtual stimuli should be adapted. For example, if the patient reports a relatively low subjective experience of discomfort, the facilitator may proceed to increase the amount of exposure, for example, by moving a virtual spider closer. Note that the facilitator may in fact be the person who is being measured and treated, that is, the person may choose to control the amount of stimulation her or himself. Finally, the interaction loop in Figure 1 is closed after the adaptation by the newly modified stimulation. For example, if the virtual spider was moved closer, it may now provoke stronger anxiety reactions. These anxiety reactions are then reflected in the measures of emotion related processes, which leads to another cycle of interaction. The underlying idea of these continuous cycles is that the person learns to regulate emotional responses to increasing levels of stimulation.

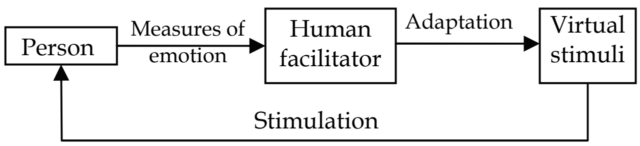


Fig. 1. A diagram of the current model used in virtual exposure therapy. Different actors are presented as boxes and labelled arrows represent the flow of information.

Although the model is quite compact and straight-forward, there are three major challenges involved when it is applied. First, although previous work has shown that virtual stimulation is effective in evoking similar emotional and social responses as real-life stimuli, the effects of virtual stimulation and its online adaptation have not been extensively investigated. It has also been found that computer-generated stimuli may significantly facilitate cognitive processing and effectively support regulation of anxiety responses. However, more information is still needed about how adapting the different parameters of stimulation in real-time affects emotion related processes. This challenge should be resolved by controlled experimental studies of each virtual stimulus in the future.

The second challenge is that there are several emotion related measures that provide complementary, non-overlapping information. There is a large amount of information contained in each of these measures. A human observer is often forced to choose between a broad view of the emotional state and an in depth analysis of it. Perceptual intelligence can be used to solve this challenge. Different methods for building computers that perceive emotion from physiological and behavioural measures were reviewed in the second section. These methods form the basis for the perceptual intelligence in our new model.

Figure 2 presents a model where perceptual intelligence has been included into the system. The model is similar to the previous one with the exception that the interpretation of emotion related measures is performed automatically. Thus, the facilitator has access to a higher level representation or a summary of information that is relevant to therapy. Simply stated, the computer acts as a kind of a translator that deciphers the information in the measured signals into a summary that is more accessible to the human observer. This way, the facilitator is less likely to be overwhelmed with the load of information available from different experiential, behavioural, and physiological measures. However, the actual adaptation is still controlled by a human facilitator acting on the basis of the summarized information.

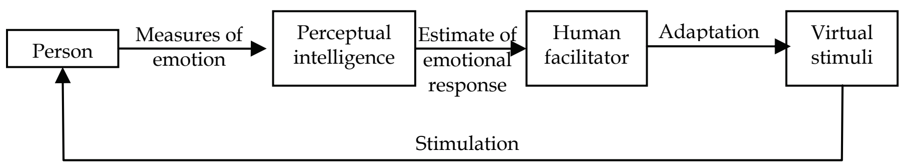


Fig. 2. A diagram of a perceptually intelligent model to be used in virtual exposure therapy. Different actors are presented as boxes and labelled arrows represent the flow of information.

The third and final challenge in using the conventional model is that it places the human facilitator as a part of the real-time system (see Figs. 1&2). This requires that a person must continuously attend to the measurements and decide when and how to react to any changes or even to a lack of changes. Figure 3 shows a final model designed to more efficiently support emotion regulation in virtual exposure therapy. The continuous monitoring of emotion related information is now built into the computer system itself. In contrast to conventional computer systems that place humans as a part of the processing loop, this model can support emotion regulation without distracting the person or requiring constant attention. The system provides this support by taking the initiative and adapting the stimulation when it is appropriate, that is, by being proactive (Tennenhouse, 2000).

In this kind of a system, the role of the human facilitator is to supervise the process of therapy. In order to perform this task, the supervisor needs information about the therapy process and the functioning of the system. Further, this information should be concise if we are to retain the main advantage of automatic signal analysis and reasoning which was to allow people to focus on the task at hand. One potentially efficient way to do this is to provide an explanation of the system's reasoning to the supervisor. This type of a model fits the definition of an expert system which solves problems in a narrow domain of expertise (i.e., virtual exposure therapy) and is able to explain its own reasoning (Bratko, 2001). For example, if the system moved the object of the phobia closer to the person, it could be asked why it did this. A brief explanation could be that, for instance, the physiological signals showed that the current level of anxiety was very low. Then, the person could further query the specifics of these signals, if she or he so desires. This way, the users confidence of the system's functioning could be increased by making it transparent to the user. Of course, this would also allow the system's reasoning to be monitored and tuned when appropriate.

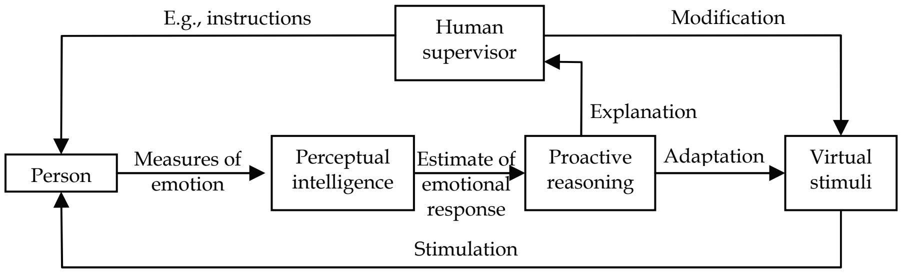


Fig. 3. A diagram of a proactive model for virtual exposure therapy. Different actors are presented as boxes and labelled arrows represent the flow of information.

During the operation of the system its supervisor may exert control over the system either by adapting the stimulation directly or through interaction with the person who is being trained. For example, a therapist may instruct the person to relax by performing controlled breathing exercises. This voluntary control affects the physiological state which is further reflected in the collected emotion related measures. As a result, the changes in these measures lead to corresponding adaptation of the stimulation. As another example, the person may directly influence those measures that affect the intensity of stimulation. This is feasible as the same expressive channels that reflect spontaneous emotional reactions can also be voluntarily controlled. For instance, the person may voluntarily frown in order to signal a high level of discomfort and move the spider further away. This also highlights another advantage of proactive adaptation. As responses to measured changes (i.e., voluntary activity) are explicit pre-programmed reactions, they can be guaranteed to be consistent. This might not be the case if the responses were selected by a human operator.
In summary, a perceptually intelligent and proactive system enables a wide variety of information to be used in regulating emotions. First, perceptual intelligence enables more efficient processing of measured emotion related signals. This enables the monitoring of a larger set of emotional measures, which then results in a more comprehensive and reliable view of the emotional state, for example, attending to multiple physiological and behavioural changes. Second, proactive reasoning may be used to adapt the stimulation in an appropriate and consistent manner. The adaptation can be based on findings that show how virtual stimulation affects human emotions and cognitive processing. As a whole, a system that uses this model can function independently without constant human supervision, helping people to regulate emotions without distracting them.

5. Discussion and future work

The current work presented a model for a computer system that supports the regulation of emotion related processes during exposure to provoking stimuli. We identified two main challenges for these kinds of systems. First, emotions as such are complex, multi-component processes that are measured with several complementary methods. The amount of information can overwhelm a human operator. Second, the adaptation of stimulation requires real-time reasoning about the current emotional state and the effects of adaptation. This reasoning may distract a human facilitator from tasks related to emotion regulation. Further, a human operator may respond inconsistently to changes in emotional processes, which effectively removes the control of the system from the person who is being trained.

In the present work we addressed the first challenge of identifying emotional reactions by including perceptual intelligence to our model. Several measures for automatic analysis of emotional state have been investigated in previous studies. Especially physiological measures were found to show potential as objective and reliable measures of emotional processes. For example, there are several new wireless and wearable measurement technologies that enable continuous and non-invasive measurement of emotion related physiology. Thus, automatic analysis of emotion related physiological activity can help to identify significant emotional responses during virtual exposure therapy. For a human observer, this pre-processed data is easier to interpret and apply to emotion regulation.

The second challenge of adapting virtual stimulation was addressed with proactive reasoning. First, we reviewed studies of human responses to virtual stimulation. These studies showed that human cognitive functioning and emotional responses may be significantly regulated using different computer-generated social and emotional cues, for example, virtual proximity. Second, we suggested a model where the computer automatically adapts the virtual stimulation according to the emotional state that it has perceived. This way, perceptual intelligence and artificial reasoning result in a proactive system that does not require humans to process data in real-time. In other words, when our model is applied to virtual exposure therapy, a person can focus on the training itself instead of monitoring and responding to measured physiological signals.

In spite of the promising findings from previous studies, there are still open questions in the computer perception of emotional responses to provoking stimuli. For example, some findings suggest that physiological reactions of phobics and healthy people may be significantly different (Wilhelm & Roth, 1998). Although the responses may be similar in terms of direction of change from a baseline (e.g., heart rate accelerated in both phobics and healthy subjects exposed to provoking stimulation), the differences in the magnitude of change may affect the results of automatic recognition. This raises the question whether automatic classification methods for emotional responses in healthy people provide information that is applicable to treating emotional disorders (i.e., abnormal emotional responses). Thus, there is a need to study systems where automatic perception has been included in a virtual therapy system.

The previous research on automatic classification of emotional states has used both person-independent methods and methods that are calibrated to each individual person. These two types of methods are suited for different kinds of applications (Bailenson et al., in press). Systems based on a universal model of emotional responses are suited when lots of people use the same interface, for example, a public computer at a library. An idiosyncratic model that adapts to each person is more suitable when the same person repeatedly uses the interface. The latter case is typical in virtual therapy applications, as the person is treated over multiple similar sessions (Krijn et al., 2004; Wiederhold & Wiederhold, 2005). However, a person-independent model could be used as a starting point for the adaptation, similar to the video-based system by Ioannou and others (2005). This would enable the system to provide estimates of emotional experiences even before a set of person-specific physiological data is collected and calibration is performed. Then, later adaptation of the model could be performed to improve its accuracy.

Another challenge in perceptual intelligence that has received little attention is the automatic segmentation of collected measures. Most previous studies of automatic classification of emotional responses have used hand-segmented data. Typically, a participant reports the onset and the offset of an emotional state and the data is segmented off-line. In contrast to these methods, virtual exposure therapy requires a system that analyses the data online and adapts to the emotional state of a person in real-time. If perceptual intelligence is to be included in this kind of a system, there is a need to investigate online, automatic segmentation of physiological data. Our preliminary results of heart rate responses have shown that such automatic segmentation is feasible (Vanhala & Surakka, 2007b). However, there is a need to investigate systems that use multiple complementary signals in order to improve the reliability and accuracy of methods.

On a general level, our review suggested that people appreciate computer systems that respond to their emotions, for example, display empathy (Klein et al., 2002; Brave et al., 2005). Although it seems a small step to assume that people would appreciate computers that administer virtual exposure therapy by responding to anxiety, there can be a fundamental difference. Emotion regulation aims not only to respond but also to change the emotional reactions to virtual and real-life emotional stimulation. There is a need to study how people experience this kind of a system and whether it helps in regulating emotions.

In summary, the present work showed how automatic perception of emotions and proactive adaptation of a computer system could help in facilitating virtual exposure therapy. The skill of regulating emotions is gradually acquired by adapting virtual stimuli according to the emotional state of the person. This principle is applicable to other emotionally intelligent applications as well. For example, we might be less likely to loose our temper if the desktop computer could display empathy when an important document gets accidentally deleted. Thus, research on perceptual intelligence and proactive reasoning in virtual exposure therapy systems has the potential to improve the quality of human-technology interaction in general. The current work identified the state-of-the-art and the future research that will help in reaching this goal.

6. Acknowledgements

This research was financially supported by the Graduate School in User-Centered Information Technology and the Academy of Finland (project number 1115997).

7. References

Anderson, J. R. (1999). *Learning and memory: An integrated approach*, Wiley, New York.

Anttonen, J. & Surakka, V. (2005). Emotions and heart rate while sitting on a chair, *Proceedings of CHI 2005*, 491-499, ACM.

Anttonen, J. & Surakka, V. (2007). Music, heart rate, and emotions in the context of stimulating technologies, *Affective Computing and Intelligent Interaction, LNCS,* 4738, 290-301.

Anxiety Disorders Association of America. (2008). *Statistics and facts about anxiety disorders*, http://www.adaa.org/AboutADAA/PressRoom/Stats&Facts.asp, Jan 30th, 2008.

Aula, A. & Surakka, V. (2002). Auditory Emotional Feedback Facilitates Human-Computer Interaction, *Proceedings of HCI 2002*, 337-349, Springer.

Bailenson, J. N.; Pontikakis, E. D.; Mauss, I. B.; Gross, J. J.; Jabon, M. E.; Hutcherson, C. A.; Nass, C. & John, O. (in press). Real-time classification of evoked emotions using facial feature tracking and physiological responses. *International Journal of Human-Computer Studies*.

Bradley, M. M. (2000). Emotion and Motivation, In: *Handbook of Psychophysiology*, Cacioppo, J. T.; Tassinary, L. G. & Berntson, G. G. (Ed.), 602-642, Cambridge University Press.

Bradley, M. M. & Lang, P. J. (2000). Affective reactions to acoustic stimuli. *Psychophysiology,* 37, 204-215.

Bradley, M. M. & Lang, P. J. (1994). Measuring emotions: the self-assessment manikin and the semantic differential. *Journal of Behavioral Therapy and Experimental Psychiatry*, 25, 1, 49-59.

Bratko, I. (2001). *Prolog Programming for Artificial Intelligence*, 3rd ed., Pearson.

Brave, S.; Nass, C. & Hutchinson, K. (2005). Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent. *International Journal of Human-Computer Studies*, 62, 161-178.

Busso, C.; Deng, Z.; Yildirim, S.; Bulut, M.; Lee, C. M.; Kazemzadeh, A.; Lee, S.; Neumann, U. & Narayanan, S. (2004). Analysis of emotion recognition using facial expressions, speech and multimodal information, *Proceedings of ICMI*, 205-211.

Cowie, R.; Douglas-Cowie, E.; Taylor, J.; Ioannou, S.; Wallace, M. & Kollias, S. (2005). An intelligent system for facial emotion recognition, *Proceedings of IEEE ICME*, 4 pp.

Cowie, R.; Douglas-Cowie, E.; Tsapatsoulis, N.; Votsis, G.; Kollias, S.; Fellenz, W. & Taylor, J. (2001). Emotion Recognition in Human-Computer Interaction. *IEEE Signal Processing Magazine*, 18, 1, 32-80.

D'Mello, S.; Graesser, A. & Picard, R. W. (2007). Toward an affect-sensitive AutoTutor. *IEEE Intelligent Systems*, 22, 4, 53-61.

Donato, G.; Bartlett, M. S.; Hager, J. C.; Ekman, P. & Sejnowski, T. J. (1999). Classifying facial actions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21, 974-989.

Ekman, P. (1993). An argument for basic emotions. *Cognition and Emotion*, 6, 3, 169-200.

Ekman, P.; Friesen, W. V. & Hager, J. C. (2002). *Facial Action Coding System: Investigator's Guide*, A Human Face, Salt Lake City, USA.

Frijda, N. H. (1986). *The Emotions*, Cambridge University Press.

Garau, M.; Slater, M.; Pertaub, D. & Razzaque, S. (2005). The Responses of People to Virtual Humans in an Immersive Virtual Environment. *Presence: Teleoperators & Virtual Environments*, 14, 1, 104-116.

Garcia-Palacios, A.; Hoffman, H.; Carlin, A.; Furness, T. A. & Botella, C. (2002). Virtual reality in the treatment of spider phobia: a controlled study. *Behaviour Research and Therapy*, 40, 9, 983-993.

Gross, J. J. & Levenson, R. W. (1997). Hiding feelings: The acute effects of inhibiting negative and positive emotion. *Journal of Abnormal Psychology*, 106, 1, 95-103.

Gross, J. J. & Thompson, R. A. (2007). Emotion regulation: Conceptual foundations, In: *Handbook of Emotion Regulation*, Gross, J. J. (Ed.), 3-24, Guilford Press.

Ioannou, S. V.; Raouzaiou, A. T.; Tzouvaras, V. A.; Mailis, T. P.; Karpouzis, K. C. & Kollias, S. D. (2005). Emotion recognition through facial expression analysis based on a neurofuzzy network. *Neural Networks*, 18, 4, 423-435.

Kim, K. H.; Bang, S. W. & Kim, S. R. (2004). Emotion recognition system using short-term monitoring of physiological signals. *Medical and Biomedical Engineering and Computing*, 42, 3, 419-427.

Klein, J.; Moon, Y. & Picard, R. W. (2002). This computer responds to user frustration: Theory, design, and results. *Interacting with Computers*, 14, 119-140.

Krijn, M.; Emmelkamp, P.; Olafsson, R. & Biemond, R. (2004). Virtual reality exposure therapy of anxiety disorders: A review. *Clinical Psychology Review*, 24, 259-281.

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.

Larsen, J. T.; Norris, C. J. & Cacioppo, J. T. (2003). Effects of positive and negative affect on electromyographic activity over the zygomaticus major and corrugator supercilii. *Psychophysiology*, 40, 776-785.

Levenson, R. W. & Ekman, P. (2002). Difficulty does not account for emotion-specific heart rate changes in the directed facial action task. *Psychophysiology*, 39, 397-405.

Lindström, M.; St, A.; Höök, K.; Sundström, P.; Laaksolahti, J.; Combetto, M.; Taylor, A. & Bresin, R. (2006). Affective diary: designing for bodily expressiveness and self-reflection, *CHI '06 extended abstracts*, 1037-1042, ACM, New York.

Malliani, A.; Pagani, M.; Lombardi, F. & Cerutti, S. (1991). Cardiovascular neural regulation explored in the frequency domain. *Circulation*, 84, 482-492.

Mandryk, R. L. & Atkins, M. S. (2007). A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. *International Journal of Human-Computer Studies*, 66, 329-347.

Mauss, I. B.; Levenson, R. W.; McCarter, L.; Wilhelm, F. H. & Gross, J. J. (2005). The tie that binds? Coherence among emotion experience, behavior, and physiology. *Emotion*, 5, 2, 175-190.

Nass, S.; Steuer, J. & Tauber, E. R. (1994). Computers are Social Actors, *Proceedings of CHI'94*, 72-78.

Nöjd, N.; Puurtinen, M.; Niemenlehto, P.; Vehkaoja, A.; Verho, J.; Vanhala, T.; Hyttinen, J.; Juhola, M.; Lekkala, J. & Surakka, V. (2005). Wireless wearable EMG and EOG measurement system for psychophysiological applications, *Proceedings of the 13th Nordic Baltic Conference on Biomedical Engineering and Medical Physics*, 144-145

Öhman, A.; Hamm, A. & Hugdahl, K. (2000), Cognition and the Autonomic Nervous System: Orienting, Anticipation, and Conditioning, In: Handbook of Psychophysiology, Cacioppo, J. T.; Tassinary, L. G. & Berntson, G. G. (Ed.), 533-575.

Partala, T. & Surakka, V. (2004). The effects of affective interventions in human-computer interaction. *Interacting with Computers*, 16, 2, 295-309.

Partala, T.; Surakka, V. & Vanhala, T. (2006). Real-time estimation of emotional experiences from facial expressions. *Interacting with Computers*, 18, 208-226.

Partala, T.; Surakka, V. & Vanhala, T. (2005). Person-independent estimation of emotional experiences from facial expressions, *Proceedings of IUI 2005*, 246-248, ACM.

Pentland, A. (2000). Perceptual intelligence. *Communications of the ACM*, 43, 3, 35-44.

Pertaub, D.; Slater, M. & Barker, C. (2002). An Experiment on Public Speaking Anxiety in Response to Three Different Types of Virtual Audience. *Presence: Teleoperators & Virtual Environments*, 11, 1, 68-78.

Picard, R. W.; Vyzas, E. & Healey, J. (2001). Toward machine emotional intelligence: Analysis of affective physiological state. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23, 1175-1191.

Rainville, P.; Bechara, A.; Naqvi, N. & Damasio, A. R. (2006). Basic emotions are associated with distinct patterns of cardiorespiratory activity. *International Journal of Psychophysiology*, 61, 5-18.

Reeves, B. & Nass, C. (1996). *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*, Cambridge University Press, New York.

Sohail, A. S. M. & Bhattacharya, P. (2007). Classification of Facial Expressions Using K-Nearest Neighbor Classifier, *Proceedings of MIRAGE 2007*, 555-566, Springer.

Surakka, V. & Vanhala, T. (accepted). Emotions in human-computer interaction, In: *Multi-channel communication on the Internet*, Kappas, A. (ed.), Cambridge University Press.

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.

Tassinary, L. G. & Cacioppo, J. T. (2000). The skeletomotor system: Surface electromyography, In: The Handbook of Psychophysiology, Cacioppo, J. T.; Tassinary, L. G. & Berntson, G. G. (Ed.), 163-199, Cambridge University Press.

Teller, A. (2004). A platform for wearable physiological computing. *Interacting with Computers*, 16, 917-937.

Tennenhouse, D. (2000). Proactive computing. *Communications of the ACM*, 43, 5, 43-50.

Vanhala, T. & Surakka, V. (submitted). Virtual proximity and facial expressions regulate human emotions and attention. *International Journal of Human-Computer Studies*.

Vanhala, T. & Surakka, V. (2007a). Facial activation control effect (FACE), *Affective Computing and Intelligent Interaction*, *Lecture Notes in Computer Science*, 4738, 278-289.

Vanhala, T. & Surakka, V. (2007b). Recognizing the Effects of Voluntary Facial Activations Using Heart Rate Patterns, *Proc. of the 11th WSEAS Int. Conf. on Computers*, 628-632.

Vehkaoja, A. & Lekkala, J. (2004). Wearable wireless biopotential measurement device, *Proceedings of the IEEE EMBS 2004*, 2177-2179.

Ward, R. D. & Marsden, P. H. (2003). Physiological responses to different WEB page designs. *International Journal of Human-Computer Studies*, 59, 199-212.

Weiten, W. (2007). *Psychology: Themes and variations*, Thomson.

Wiederhold, B. K. & Bullinger, A. H. (2005). Virtual reality exposure for phobias, panic disorder, and posttraumatic disorder: A brief sampling of the literature, *Proceedings of HCII 2005*, CD-ROM, Lawrence Erlbaum Associates.

Wiederhold, B. K. & Wiederhold, M. D. (2003). Three-Year Follow-Up for Virtual Reality Exposure for Fear of Flying. *CyberPsychology and Behavior*, 6, 441-445.

Wilhelm, F. H.; Pfaltzl, M. C. & Grossman, P. (2006). Continuous electronic data capture of physiology, behavior and experience in real life: towards ecological assessment of emotion. *Interacting with Computers*, 18, 171-186.

Wilhelm, F. H. & Roth, W. T. (1998). Taking the laboratory to the skies: Ambulatory assessment of self-report, autonomic, and respiratory responses in flying phobia. *Psychophysiology*, 35, 5, 596-606.

Zeng, Z.; Tu, J.; Zhang, T.; Rizzolo, N.; Zhang, Z.; Huang, T. S.; Roth, D. & Levinson, S. (2004). Bimodal HCI-related affect recognition, *Proceedings of ICMI*, 205-211.

Appendix C

Publication III

Vanhala, T., Surakka, V., and Anttonen, J. (2008). Measuring bodily responses to virtual faces with a pressure sensitive chair. In *Proceedings of the 5th Nordic Conference on Human-Computer Interaction (NordiCHI '08)*, pp. 555–559. ACM.

Available at:
`http://dx.doi.org/10.1145/1463160.1463243`

Measuring Bodily Responses to Virtual Faces with a Pressure Sensitive Chair

Toni Vanhala, Veikko Surakka, Jenni Anttonen
Research Group for Emotions, Sociality, and Computing
Tampere Unit for Computer-Human Interaction
Department of Computer Sciences
University of Tampere, FI-33014 Tampere, Finland
Corresponding author: Toni.Vanhala@cs.uta.fi, +358 3 3551 4036

ABSTRACT
The present aim was to study emotion related body movement responses using an unobtrusive measurement chair that is embedded with electromechanical film (EMFi) sensors. 30 participants viewed images of a male and a female computer agent while the magnitude and direction of body movements were measured. The facial expressions (i.e., frowning, neutral, smiling) and the size of the agents were varied. The results showed that participants leaned statistically significantly longer towards the agent when it displayed a frowning or a smiling expression as compared to a neutral expression. Also, their body movements were reduced while viewing the agents. The results suggest that the EMFi chair is a promising tool for detecting human activity related to social and emotional behaviour. In particular, the EMFi chair may support unobtrusive measurement of bodily responses in less strictly controlled contexts of human-computer interaction.

Categories and Subject Descriptors
H.5.2. User Interfaces: Input devices and strategies.

General Terms
Algorithms, Measurement, Experimentation, Human Factors.

Keywords
Affective computing, non-invasive sensors, body movements, virtual embodied agents, approach and withdrawal motivation.

1. INTRODUCTION
The posture and dynamics of the body provide cues of our internal motivation towards other people. For example, we tend to approach (e.g. lean towards) things and people we like and care about, and withdraw (e.g. lean away) from people and objects that threaten us [6,16,18]. People can also recognize the emotions and intentions of others based on their bodily behaviour [1,17]. In one study observers identified some posed postural expressions of sadness, anger, and happiness with over 90% accuracy [11].

NordiCHI 2008: Using Bridges, 18-22 October, Lund, Sweden.
Copyright 2008 ACM ISBN 978-1-59593-704-9. $5.00

Postural cues are also significant in social and emotional human-computer interaction (HCI). For example, the posture of and the distance to an embodied agent have been found to modulate subjective emotional experiences and physiological responses of human observers [11,20,24]. Further, there is evidence that the regulation of a person's posture can facilitate cognitive and emotional processes [8,15]. Ahn and others [2] found that a robotic computer that manipulated the posture of its user could increase the user's persistence in solving problems.

In order to intelligently react to the posture of its user, the computer should be able to recognize socially and emotionally significant postural cues. In laboratory studies the participant typically stands on a force platform that is used to measure bodily behaviour (e.g., [14,18]). The measurement is obvious, which may influence the emotional and social responses that are investigated. More unobtrusive measurements could be performed by integrating measurement sensors to everyday objects, such as chairs [3,4,23]. Using such a chair, episodes of boredom, confusion, delight, flow, and frustration have been successfully recognized during a learning task [13].

Truly unobtrusive (e.g., unnoticeable) measurements of body movements have not yet been demonstrated. Furthermore, previous research on automatic recognition of postural cues has not been firmly based on established theories of emotion. One prominent line of research maps emotions into a three dimensional bipolar space varying in emotional valence (i.e. from unpleasant to pleasant), arousal (i.e. from calm to aroused), and dominance (i.e. from feeling of being controlled to feeling of being in control) [7]. It has been argued that especially the emotional valence of a stimulus and behaviour are connected at a very basic motivational level. In this view, unpleasant stimuli activate defensive motivation which manifests itself as withdrawal or freezing behaviour [6,16]. Pleasant stimuli promote appetitive motivation which evokes approach behaviour. Further, these basic motivational tendencies and behavioural changes should be reflected in postural (e.g. leaning towards an appealing stimulus) and other body movements. In line with this reasoning, Facchinetti and others [14] found that negatively arousing pictures reduced sideways movement, whereas positively arousing pictures reduced sway mostly in the forward-backward direction.

There is also evidence that body movements are related to other dimensions of emotion. For example, when people or computers have rated images of body postures using scales of valence, arousal, and dominance, the highest agreement has been reached for the ratings of arousal [10,13,17]. Further, objects that appear to be closer (e.g. larger or approaching objects) have been rated more dominating (i.e. more controlling) than objects that are further away [12,20]. In sum, previous results provide support for grounding a study of body movements in the dimensional view of emotions. We have recently been developing a device called the EMFi chair for this purpose. The EMFi chair is a regular looking office chair containing embedded electromechanical film (EMFi) sensors in the seat, the back rest, and the arm rests [Figure 1].

Figure 1. The EMFi chair. The back rest contained three strips of electromechanical film (EMFi). The back cushion is folded and the sensors are revealed for demonstration.

Movements of a person sitting on the chair generate small changes in pressure that are exerted to the EMFi sensors in the chair. When external force or pressure is exerted to the film's surface, the thickness of the film is changed and a charge proportional to the force or pressure acting on the film is generated [19]. The measurement of these charges enables the use of the EMFi film as a sensor for detecting dynamic forces. Earlier work has shown that emotion related heart rate activity can be reliably measured by using the EMFi chair [3,4]. The chair detects heart activity non-invasively by extracting ballistocardiography (BCG) from the EMFi signal. The signal should also reflect changes in bodily activity in general (e.g., freezing or increased level of activity).

The present work investigated the feasibility of the EMFi chair in detecting bodily responses during emotionally arousing and socially engaging stimulation. 30 participants viewed a male and a female agent character in three different sizes and with a frowning, a neutral, or a smiling facial expression. The three sizes were used to simulate different distances for interacting with the agent, as different social and emotional cues have been found to be prominent on different distances [20]. Pressure changes were measured with the EMFi chair during stimulation. Emotional experiences were measured using the three dimensional affective space (i.e., valence, arousal, dominance) by Bradley and Lang [7].

2. METHODS

2.1 Participants

15 female and 15 male subjects participated in the experiment. The participants' mean age was 25.14 ranging from 19 to 45 years. All participants had normal or corrected to normal vision.

2.2 Equipment

The EMFi chair was used to measure pressure changes resulting from body movements. The back rest of the chair covered three 6.0 × 18.5 cm strips of EMFi [Figure 1]. The three strips were connected to a charge amplifier and acted as a single sensor. This sensor registered a negative charge when pressure was exerted to any of the film strips, that is, when the person leaned back in the chair. When the person leaned forward, pressure was released and a positive charge was generated. This way, body movements (i.e., backward or forward) could be detected.

The EMFi signal was sampled at 500 Hz using a Quatech DAQP-16 A/D-card and a PC running Windows XP. Stimuli were presented on a 19-inch CRT monitor with a resolution of 1024×768 pixels. Stimulus presentation was controlled using E-Prime experiment generator running on a PC with Windows XP [21]. Stimulus on- and offset markers were sent through a serial cable from the E-Prime computer to the Quatech A/D-card.

2.3 Stimuli

Two (i.e., female and male) realistic humanlike characters created by Cantoche [9] were used as stimuli in the present study. The agents were displayed in three different heights of 9 cm, 18 cm, and 27 cm on the 19-inch monitor. The participants were seated at approximately 50 cm from the monitor. The agents had realistic body proportions and their head and body were fully displayed. The character displayed a negative (i.e., frowning), a neutral, or a positive (i.e., smiling) facial expression. There were a total of 18 stimuli (2 genders × 3 sizes × 3 facial expressions).

The stimuli were counter-balanced so that 10 people viewed a smiling, a frowning, or a neutral character as the first stimulus. The rest of the stimuli were presented in a random order.

2.4 Procedure

The experiment was performed in a sound-attenuated and electro-magnetically shielded laboratory. First, the laboratory and the procedure were introduced to the participant. The participant was told a cover story that the purpose of the study was to investigate changes in body temperature when viewing different agent characters. Then, the participant was seated on the EMFi chair. An ear clip was attached to support the cover story.

Participants were instructed to relax and breathe calmly while viewing the stimuli. Stimuli were displayed at the centre of the screen with a white background for 30 seconds. 10-second long inter-stimulus intervals were used, except for a 30 second interval before the first stimulus. After the stimulus presentation, the ear sensor was detached.

Next, the participant was asked to rate the stimuli using the three bipolar 9-point scales of emotional valence, arousal, and dominance. The midpoint (i.e., the value 5) of each scale represented a neutral value. The ratings were administered on paper. Each stimulus remained on the screen until the participant proceeded to the next one by pressing the left mouse button. The stimuli were presented in a random order during the ratings.

Finally, the participant was debriefed and interviewed. None of the participants reported having noticed anything special about the chair they were seated on (i.e. the EMFi chair). The experiment lasted approximately one hour for each participant.

2.5 Data analysis

Body posture and body movement data were extracted from the EMFi signal. First, the EMFi signal was pre-processed with a DC notch filter in order to remove any possible bias (Eq. 1).

$$y(n) = x(n) - x(n-1) + 0.999y(n-1) \qquad (1)$$

Then, posture was estimated by integrating the values over time, starting from the stimulus onset. Data was divided into six intervals by extracting the posture value every five seconds.

The magnitude of body movements was computed using the root mean square (RMS) of the EMFi signal [22]. Similar to previous studies, the level of activity was compared to a baseline value [5,14]. The RMS was computed for a five second pre-stimulus baseline and a mean of the six interval values during stimulation.

A 3 × 3 (size × expression) repeated measures multivariate analysis of variance (MANOVA) was performed for posture data. Post hoc analyses of significant effects were performed with univariate ANOVAs for each interval. A 2 × 3 × 3 (stimulation × size × expression) repeated measures ANOVA was performed for body movement data. Greenhouse-Geisser correction was used. Post hoc pairwise comparisons were performed using Bonferroni corrected paired samples t-tests.

3. RESULTS

The subjective ratings (Mean ± Standard Error of the Mean; S.E.M.) showed that smiling expressions were rated as more pleasant (5.86 ± .11) and frowning expressions as more unpleasant (3.51 ± .11) than neutral facial expressions (5.11 ± .10). Similarly, smiling expressions were rated as more calming (3.39 ± .13) and less dominating (6.43 ± .15) than neutral expressions (arousal: 3.66 ± .13, dominance: 6.29 ± .15), while frowning expressions were rated as the least calming (4.86 ± .14) and close to neutral dominance (5.32 ± .16).

Regarding the mean body postures, Figure 2 shows that in general participants leaned forward while viewing the agents. After an initial movement towards the stimulation that was a common response to all types of agents, the postural responses began to differ during the third interval (i.e. 10 – 15 seconds after stimulus onset). A 3 × 3 (size × expression) MANOVA for posture data at the six intervals showed a statistically significant main effect of expression, $F(12,108) = 1.869$, $p < .05$. There were no other statistically significant main or interaction effects.

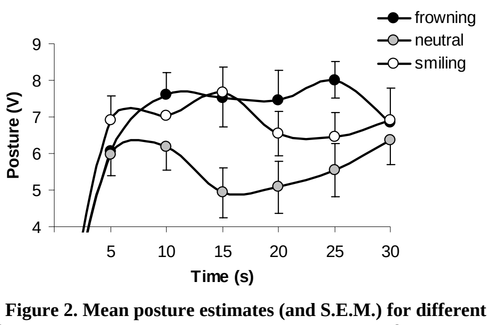


Figure 2. Mean posture estimates (and S.E.M.) for different facial expressions in volts. Increasing values signify movement towards and decreasing values away from stimulation.

Post hoc ANOVA tests showed statistically significant main effects of expression within the third $F(2,58) = 5.869$, $p < .01$, the fourth $F(2,58) = 3.926$, $p < .05$, and the fifth interval $F(2,58) = 6.597$, $p < .01$. Post hoc pairwise comparisons showed statistically significant differences between the negative and the neutral MD = 2.57, $p < .05$, and the positive and the neutral MD = 2.72, $p < .05$ expressions within the third interval. Post hoc pairwise comparisons also showed statistically significant differences between the negative and the neutral expressions within the fourth MD = 2.38, $p < .05$ and the fifth MD = 2.47, $p < .05$ interval.

The mean magnitude of body movements (i.e. the level of movement activity) decreased from 20.88 ± 2.37 mV during the pre-stimulus baseline and remained at a lower level with an average of 15.41 ± 1.87 mV during the whole stimulation [Fig. 3]. A 2 × 3 × 3 (stimulation × size × expression) ANOVA for body movement data showed a statistically significant main effect of stimulation $F(1,29) = 8.893$, $p < .01$. There were no other statistically significant main or interaction effects. A post hoc pairwise comparison confirmed that movement was significantly reduced after stimulus onset MD = 5.47, $p < .01$.

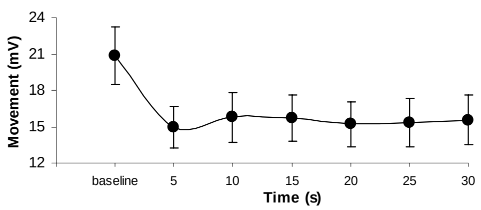


Figure 3. Mean magnitude of body movements (and S.E.M.) in millivolts for the baseline and six intervals.

4. DISCUSSION

Our results showed that the EMFi chair successfully detected bodily responses to emotionally significant stimulation. First, the EMFi signal showed a positive change suggesting an initial movement towards the agent regardless of its facial expression. This change can be seen as an initial orientation response that has been suggested to be common to all motivationally significant and emotionally engaging stimuli [6].

The facial expression modulated how long the initial postural response was sustained. No statistically significant difference between expressions was found within the first 10 seconds of stimulation. During the third interval, the responses began to differ. The posture was statistically significantly more forward leaning when viewing negatively rated or positively rated agents as compared to viewing agents rated as more neutral. The posture was sustained for the longest time during the frowning expression, as the difference between frowning and neutral expressions remained statistically significant during the fourth and the fifth intervals. The difference between smiling and neutral expressions, on the other hand, diminished earlier as participants leaned back on the fourth interval during positive stimulation. Thus, the participants leaned more towards a virtual embodied agent when it displayed a negative or a positive facial expression as compared to a neutral expression. This may reflect a sustained engagement of social attention resources to the more expressive stimulations.

The present results also showed reduced body movement activity while viewing the agents as compared to the pre-stimulus baseline. Previous studies have found similar bodily responses to both negative (e.g., mutilation) and positive pictures (e.g., babies) in terms of reduced movement or "freezing" responses [5,14].

The EMFi chair provides several benefits for investigating behavioural responses in HCI. For example, reduced awareness of the measurement may help to avoid biased results (e.g., due to voluntary regulation of body movements). Further, the technology supports the extension of body movement studies from force platforms in laboratories to less strictly controlled settings with embedded sensors.

In summary, the EMFi chair was a promising tool for measuring behavioural responses to social and emotional stimulation in HCI. Statistically significant body movement responses to virtual emotionally expressive agents simulating proximity cues could be detected using the EMFi chair. Future work includes studying body movements in an interactive scenario and developing real-time classification of bodily responses. This way, the EMFi chair could facilitate HCI in a regular office or home setting, for example, by enabling behavioural regulation using automated postural analysis [2,8,13,15]. Thus, the EMFi chair paves the way for integrating social and emotional cues for real-time HCI.

5. ACKNOWLEDGMENTS

The EU 6th Framework Programme project AtGentive (IST-4-027529-STP), the Academy of Finland (project number 1115997), and the Graduate School in User-Centered Information Technology (UCIT) have financially supported this research. We thank Benoît Morel from Cantoche for providing the agent characters for this study.

6. REFERENCES

[1] Adams Jr., R. B., Ambady, N., Macrae, C. N., and Kleck, R. E. Emotional expressions forecast approach-avoidance behaviour. *Motivation and Emotion*, 30 (2006), 179-188.

[2] Ahn, H., Teeters, A., Wang, A., Breazeal, C., and Picard, R. Stoop to conquer: posture and affect interact to influence computer users' persistence. In *Proceedings of ACII 2007, Springer LNCS*, 4738 (2007), 582-593.

[3] Anttonen, J., and Surakka, V. Music, heart rate, and emotions in the context of stimulating technologies. In *Proc. of ACII 2007, LNCS, 4738* (2007), 290-301.

[4] Anttonen, J., and Surakka, V. Emotions and heart rate while sitting on a chair, In *Proceedings of CHI 2005*, 491-499, 2005.

[5] Azevedo, T. M., Volchan, E., Imbiriba, L. A., Rodrigues, E. C., Oliveira, J. M., Oliveira, L. F., Lutterbach, L. G., and Vargas, C. D. A freezing-like posture to pictures of mutilation, *Psychophysiology*, 42 (2005), 255-260.

[6] Bradley, M. M., Codispoti, M., Cuthbert, B. N., and Lang, P. J. Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion*, 1 (2001), 276-298.

[7] Bradley, M. M., and Lang, P. J. Measuring emotions: the self-assessment manikin and the semantic differential. *Journal of Behavioral Therapy and Experimental Psychiatry*, 25 (1994), 49-59.

[8] Cacioppo, J. T., Priester, J. R., and Berntson, G. G. Rudimentary determinants of attitudes: II. Arm flexion and extension have differential effects on attitudes. *Journal of Personality and Social Psychology*, 65 (1993), 5-17.

[9] Cantoche. The Living Actor^{TM} Avatars humanize interfaces. http://www.cantoche.com. (April 19th, 2008)

[10] Castellano, G., Villabla, S. D., and Camurri, A. Recognising Human Emotions from Body Movement and Gesture Dynamics. In *ACII 2007, LNCS, 4738* (2007), 71-82.

[11] Coulson, M. Attributing emotion to static body postures: recognition accuracy, confusions, and viewpoint dependence. *Journal of Nonverbal Behavior*, 28 (2004), 117-139.

[12] Detenber, B. H., and Reeves, B. A bio-informational theory of emotion: Motion and image size effects on viewers. *Journal of Communication*, 46 (1996), 66-84.

[13] D'Mello, S., Graesser, A., and Picard, R. W. Toward an affect-sensitive AutoTutor. *IEEE Intelligent Systems*, 22 (2007), 53-61.

[14] Facchinetti, L. D., Imbiriba, L. A., Azevedo, T. M., Vargas, C. D., and Volchan, E. Postural modulation induced by pictures depicting prosocial or dangerous contexts, *Neuroscience Letters*, 410, 1 (2006), 52-56.

[15] Flack Jr., W. F. Peripheral feedback effects of facial expressions, bodily postures, and vocal expressions on emotional feelings. *Cognition and Emotion*, 20 (2006), 177-195.

[16] Hillman, C. H., Rosengren, K. S., and Smith, D. P. Emotion and motivated behavior: postural adjustments to affective picture viewing. *Biological Psychology*, 66 (2004), 51-62.

[17] Kleinsmith, A., and Bianchi-Berthouse, N., Recognizing affective dimensions from body posture. In *Proceedings of ACII 2007, Springer LNCS*, 4738 (2007), 48-58.

[18] Marsh, A. A., Ambady, N., and Kleck, R. E. The Effects of Fear and Anger Facial Expressions on Approach- and Avoidance-Related Behaviors, *Emotion*, 5 (2005), 119-124.

[19] Paajanen, M., Lekkala, J., and Kirjavainen, K. ElectroMechanical Film (EMFi) - a new multipurpose electret material. *Sensors and Actuators A: Physical*, 84, 1-2 (2000), 95-102.

[20] Partala, T., Surakka, V., and Lahti, J. Affective effects of agent proximity in conversational systems. In *Proceedings of NordiCHI 2004*, 353-356, ACM, 2004.

[21] Schneider, W., Eschman, A., and Zuccolotto, A. *E-Prime User's Guide*. Psychology Software Tools, Inc., Pittsburgh, PA, USA, 2002.

[22] Smith, S. W. *The scientist and engineers guide to digital signal processing*. California Technical Publishing, San Diego, CA, USA, 1997.

[23] Tan, H. Z., Slivovsky, L. A., and Pentland, A. A Sensing Chair Using Pressure Distribution Sensors. *IEEE/ASME Transactions on Mechanotronics*, 6 (2001), 261-268.

[24] Vinayagamoorthy, V., Brogni, A., Steed, A., and Slater, M. The role of posture in the communication of affect in an immersive virtual environment. In *Proceedings of VRCIA 06*. ACM, New York, NY, USA, 229-236, 20

Appendix D

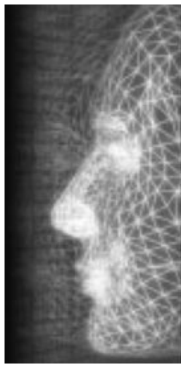
Publication IV

Vanhala, T., Surakka, V., Siirtola, H., Räihä, K.-J., Morel, B., and Ach, L. (2010). Virtual proximity and facial expressions of computer agents regulate human emotions and attention. *Computer Animation and Virtual Worlds*, 21 (3–4), 215–224.

Available at:
`http://dx.doi.org/10.1002/cav.366`

COMPUTER ANIMATION AND VIRTUAL WORLDS
Comp. Anim. Virtual Worlds 2010; **21**: 215–224
Published online 27 May 2010 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/cav.366

Virtual proximity and facial expressions of computer agents regulate human emotions and attention



By Toni Vanhala, Veikko Surakka, Harri Siirtola, Kari-Jouko Räihä, Benoît Morel and Laurent Ach*

Emotion- and attention-related subjective and physiological responses to virtual proximity and facial expressions of embodied computer agents (ECA) were studied. Thirty participants viewed female and male characters with a neutral, unpleasant, or pleasant facial expression. Agents' size was used to simulate three levels of proximity. Participants' electrical facial muscle and heart activity were registered, and subjective ratings of emotional and attentional experiences collected. Unpleasant and large (i.e., closer) agents were more alerting (i.e., unpleasant, arousing, and dominating) and attracted more stimulus-driven attention than neutral, pleasant, and smaller (i.e., further away) agents. Pleasant agents attracted more voluntary attention than neutral and unpleasant agents. Heart rate (HR) responded to agent proximity, while the valence of the agent affected electrical facial muscle activity. Thus, the imitation of human social emotional cues in embodied computer agents (ECAs) could be used to regulate human–computer interaction.

KEY WORDS: embodied agents; proximity; facial expressions; heart rate; emotion; attention

Introduction

Virtual computer agents have significant potential for evoking social and emotional responses in human–computer interaction. Even non-embodied agents can affect a person's emotional and cognitive processes.[1] However, results from several studies suggest that the mere presence of an embodied agent may enhance, for example, learning experiences.[2] Further, the current highly proficient work on embodied agents provides broad and vivid simulation of human behavior, enabling the use of social and emotional cues that are similar to human–human interaction (see Reference [3] for a recent review). These cues, such as facial expressions and proximity, are very much operational in human–computer interaction, due to our strong and automatic tendency for social behavior.[4]

Emotional and social responses to human facial expressions have been extensively studied. For example, it is well-known that facially expressed emotions evoke congruent experiences and physiological activity in the perceiver such as pleasant experiences and smiling in response to positive human facial cues.[5] There is a growing interest for replicating and extending these findings using artificial facial cues.[6]

Reproduction of human social cues like facial expressions can be relatively challenging (see Reference [7] for a comprehensive approach). On the other hand, some cues can be easily simulated. For example, the retinal size of a visual stimulus (e.g., agent) is directly related to its perceived distance (i.e., proximity).[8] Different proximities are preferred for different forms of interaction.[9] Especially very close proximity to other people can cause feelings of discomfort.[10] In general, larger images have been found to accentuate ratings of arousal and physiological responses (e.g., facial activity).[11] In agents, the simulation of a closer proximity to a virtual humanlike head decreased subjective dominance, that is, the feeling of being in control of the virtual stimulus.[12]

Emotional experiences have been frequently studied with self-report scales that form a three-dimensional bipolar space of emotional valence, arousal, and dominance.[13] The valence dimension varies from

*Correspondence to: T. Vanhala, Tampere Unit for Computer–Human Interaction (TAUCHI), Department of Computer Sciences, University of Tampere, Tampere 33014, Finland. E-mail: toni.vanhala@cs.uta.fi

unpleasant to pleasant emotional experience. The arousal dimension ranges from calm to aroused and the dominance dimension ranges from the feeling of being controlled to the feeling of being in control of the situation or stimulus. The midpoint of each scale represents a neutral (e.g., neither unpleasant nor pleasant) experience.

Emotions in general are viewed as concurrent changes in experiential, physiological, and behavioral systems that organize human motivational behavior.[14] They are known to influence several cognitive processes including memory, problem-solving, and decision-making.[1,15] Thus, emotional responses to artificial communication (e.g., virtual facial expressions and proximity) could modulate both the experiences and the performance of a person.

Visual emotional information in general has been found to significantly modulate early (i.e., within 150–300 ms time from stimulus onset) change detection and attention-related auditory processing.[16] Generally, emotions and attention have both been associated with the same basic neurocognitive systems: the Behavioral Approach System and the Behavioral Inhibition System.[17] Thus, emotion and attention can be viewed as inherently tied responses.

In contrast to research on emotions, there are no well-tried and widely accepted scales for subjective ratings of attention-related factors. Attention is the means to select a limited amount of information for processing.[18] These means include both stimulus-driven (i.e., exogenous) and voluntary processes. Stimulus-driven processes are more directly influenced by the properties of the stimulus. They are often involuntary and automatic. On the other hand, voluntary processes may be seen as mostly goal-directed and controlled.[19] These two types of processes are constantly interacting. For example, stimuli with strong exogenous properties are often used as distractors while subjects perform voluntary tasks.[19]

Responses to embodied agents in computer interfaces are often voluntary. For example, we may choose whether to interact with an agent based on how competent the agent seems for the task at hand. However, there is evidence that we also respond to subtle social and emotional cues from the agents.[3,4,12] Although a major part of these responses are involuntary and automatic, we may also become aware of them by voluntary reflection of our reactions.[15]

There has been some interest for developing and using scales for ratings of attention, especially in virtual reality and gaming research (e.g., Reference [20]). Apparently, there is a clear need for developing measures for subjective experiences of attention. Thus, a first tentative set of attentional scales was designed in the present work.

We aimed to cover both stimulus-driven and voluntary attentional responses using three scales: conspicuousness, interestingness, and concentration. Merriam-Webster[21] defines conspicuous as "attracting attention," while interesting is defined as "holding the attention." Concentration is defined as the "direction of attention to a single object." Distraction can be seen as its opposite, as Merriam-Webster[21] defines it as "to draw or direct (as one's attention) to a different object or in different directions at the same time." These concepts have similar meanings in the Finnish language.[22]

Based on these definitions, conspicuous objects are more likely to affect stimulus-driven (i.e., exogenous) processes, while voluntary concentration should reflect voluntary (i.e., endogenous) processes. Strong conspicuous stimuli should also act as distractors and negatively affect voluntary concentration. Finally, ratings of interestingness should reflect how intensely the stimulus holds the attracted attention (i.e., the intensity of both stimulus-driven and voluntary processes).

We settled on the three dimensions for measuring attention-related experiences. The nine-point bipolar self-report scales in Finnish ranged from inconspicuous (Fin. *huomiota herättämätön*) to conspicuous (*huomiota herättävä*), from uninteresting (*mielenkiinnoton*) to interesting (*mielenkiintoinen*), and from distracting (*häiritsee keskittymistä*) to assisting (*auttaa keskittymään*) concentration. The midpoint of each scale represented a neutral experience (e.g., neither inconspicuous nor conspicuous).

In addition to subjective experiences, emotional and attentional stimulation is known to affect several physiological processes, including heart activity, facial muscle activity, and skin conductance.[11,23] A well-tried measure of emotion-related facial activity is electromyography (EMG) which reflects the electrical activity of muscles. Specifically, the activity of *zygomaticus major* muscle (activated when smiling) has been shown to decrease during negative and to increase during positive emotional experiences.[24] The activity of *corrugator supercilii* muscle in the forehead (activated when frowning) varies with emotional valence in the opposite manner. However, other social and cognitive processes (e.g., affiliation with the expresser) also affect facial activity.[25] Thus, measures of facial activity (e.g., EMG) are not fully specific to emotional phenomena.

Similarly, heart rate (HR) measures are also sensitive to many types of emotional and cognitive responses.

A series of studies has shown that the mean HR first decelerates during most emotional stimulations, but the following pattern of accelerations and decelerations varies between different types of positive and negative stimuli.[11,23] Positive visual stimuli have evoked larger peak accelerations, while negative stimulation has evoked greater initial decelerations of the HR. Generally, HR deceleration can be seen as an orientation response that reflects a heightened level of attention, especially during negative stimulation.[11,23] Acceleration of the HR, on the other hand, has been linked with preparation for action. Thus, mean HR may provide a relatively unintrusive measure that suggests the level of attention versus action.

In the present work we studied the effects that virtual proximity and facial expressions of embodied agents can have on cognitive and emotional processes. First, participants viewed a female and a male character in three different sizes and with either an unpleasant, a neutral, or a pleasant facial expression. HR and facial EMG activity were measured during this session. Then, participants viewed the stimuli again and reported their experiences using six bipolar scales after each stimulus. The first three scales measured the three-dimensional emotion space of valence, arousal, and dominance. The other three scales of conspicuousness, concentration, and interestingness aimed to measure attentional effects of the stimuli (i.e., a subjective attentional space). Principal component analysis (PCA) was used to explore the structure of the ratings in relation to the theorized concepts of stimulus-driven and voluntary attention.

Methods

Participants

Fifteen female and 15 male subjects participated in the experiment. The participants' mean age was 25.1 years ranging from 19 to 45 years. All participants had normal or corrected to normal vision and normal heart functioning by their own reports. The acquisition of the HR data failed for two participants (one female and one male) who were excluded from all analyses.

Equipment

HR was measured with a Tunturi photoplethysmographic (PPG) sensor attached to the earlobe. Signal was sampled at 500 Hz using a Quatech DAQP-16 A/D-card and Windows XP.

Facial EMG was registered from the left side of the face above the *corrugator supercilii* and *zygomaticus major* muscle sites using Ag–AgCl sintered electrodes. A pair of electrodes was attached below and above the right eye to detect eye blinks. The ground electrode was over the mastoid bone and an active reference on the forehead close to the center of the hairline. Guidelines of Fridlund and Cacioppo were followed.[26] EMG and blink data were acquired using Neuroscan SynAmp2™ amplifiers and a server computer. Data were stored by the computer registering PPG data and connected to the Neuroscan server.[27] Passband for EMG was 0.1–1 kHz, amplification 2010 times, and sample rate 5 kHz. Passband for blinks was 0.05–30 Hz and off-line highpass from 10 Hz. All electrode impedances were below 10 kΩ.

Stimuli were presented with E-Prime© software on a 19-inch CRT monitor with a resolution of 1024×768 pixels.[28] E-Prime© sent stimulus onset and offset markers through a serial cable to the Quatech A/D-card.

Stimuli

Small facial cues and full-body agents were designed in order to imitate a real person in every-day human–human interaction (i.e., fully visible body and non-exaggerated facial expressions). Two (i.e., female and male) realistic humanlike characters with identical body postures (Figure 1 a–c) created by Cantoche[29] were used as stimuli. The agents were displayed as static images in three different heights of 9, 18, and 27 cm at a distance of 50 cm. The character displayed either an unpleasant, a neutral, or a pleasant facial expression. Thus, there was a total of 18 stimuli (2 agents $\times$ 3 sizes $\times$ 3 expressions). Ten people each first viewed either an unpleasant, a pleasant, or a neutral character. The rest of the stimuli were in random order.

Procedure

First, the sound-attenuated and electro-magnetically shielded laboratory was introduced to the participant. A cover story told that the study aimed to investigate changes in skin temperature. Then, the participant was seated at a distance of about 50 cm from the monitor and the electrodes were attached, including a dummy electrode on the back of the left hand in order to support the cover story. Participant was also told that the PPG sensor was used to isolate measurement artifacts. EMG and PPG sensors were attached to head-boxes in the same room.

DOI: 10.1002/cav

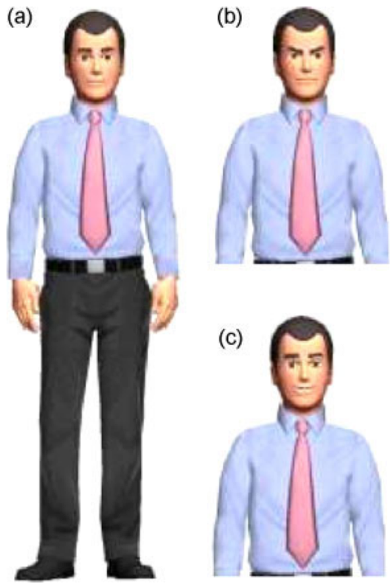


Figure 1. The male character in small size displaying (a) neutral, (b) unpleasant, and (c) pleasant expressions. Lower bodies for (b) and (c) were identical to (a). Printed by permission of Cantoche.[29]

Then, the participant viewed the stimuli. Participants were instructed to relax and breathe calmly. Each stimulus was centered on a white background. A 30-second long stimulation was used to allow low frequency changes to the HR, although the mean HR responds to visual stimulation in a few seconds.[23] There were 10-second long intervals between consecutive stimuli. The first stimulus was preceded with an interval of 30 seconds. After the last stimulus was viewed, sensors were detached.

Next, the participant rated the stimuli using six bipolar nine-point scales of experienced emotional valence, arousal, dominance, conspicuousness, interestingness, and concentration. The scales were administered on paper and participants were instructed to rate their own subjective experiences. Ratings were self-paced to allow carefully considered responses. The participant could move on to the next stimulus by pressing the left button of a computer mouse. The stimuli were in random order. Finally, the participant was debriefed after rating all the stimuli. The session lasted about one hour.

Data Analysis

HR data were extracted from PPG data as follows. First, HR data were non-uniformly sampled as inter-beat intervals. Then, the data were re-sampled at 5 Hz using a tachiogram.[30] EMG data were analyzed by averaging rectified EMG sample values. Potential artifacts were removed from the data by discarding samples taken when blink activity exceeded 50 μV. Mean HR for the whole stimulus and mean EMG for the first 5 seconds from the stimulus onset was computed using a one-second prestimulus baseline correction.

The intended categories (i.e., unpleasant, neutral, and pleasant) of expressions were confirmed with a manipulation check using a $2 \times 2 \times 3 \times 3$ (participant's gender $\times$ agent $\times$ expression $\times$ size) mixed-model repeated measures analysis of variance (ANOVA) on valence ratings.

Next, subjective ratings and mean physiological responses were within-subject standardized. Standardization is often used to correct different ranges of responses between participants and measurement sessions.[31] However, standardization (i.e., z-scores) of physiological measures may misrepresent results, if the relative order of averaged responses is not preserved. Thus, raw scores for physiological measures are reported as recommended by Fridlund and Cacioppo.[26]

Then, PCAs were applied to the rating scales. Components with eigenvalues statistically significantly over 1.0 were identified with bootstrapping ($n = 504$, rep. $= 2000$, $p < 0.05$) and factor scores extracted.[32]

Finally, 3×3 (facial expression $\times$ size) within-subject repeated measures ANOVAs were performed on HR, *corrugator supercilii*, and *zygomaticus major* EMG, and PCA scores. One-way ANOVAs were used for *post hoc* simple effect analyses. *Post hoc* pairwise comparisons were performed using Fisher's least significant difference (LSD) for paired samples t-tests [33, p. 368]. ANOVAs were Huynh–Feldt corrected.

Results

Manipulation Check

The mean ratings of valence were congruent with the facial expression of the agent (Figure 2). Mean valence ratings were lowest for unpleasant, close to neutral (i.e., value of 5) for neutral, and highest for pleasant facial expressions. Further, facial expressions were rated consistently regardless of the size of the agent.

Comp. Anim. Virtual Worlds 2010; **21**: 215–224
DOI: 10.1002/cav

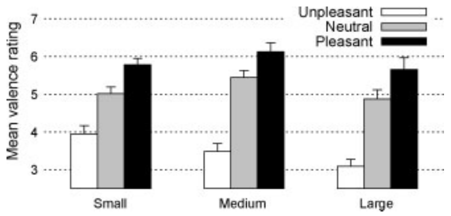


Figure 2. Mean valence ratings and Standard Error of the Mean (SEM) for different facial expressions and sizes of the agents.

The $2 \times 2 \times 3 \times 3$ mixed-model ANOVA for ratings of valence showed statistical significance for the main effects of facial expression $F(2, 52) = 48.46$, $p < 0.001$, size $F(2, 52) = 4.69$, $p < 0.05$, and the interaction effect of facial expression and size $F(4, 104) = 2.54$, $p < .05$. Main and interaction effects of participant's gender and agent were not statistically significant. The interaction effect was likely due to the difference in the effect of size within each expression (Figure 2). Agents with unpleasant expressions were rated the more unpleasant the larger they were, whereas within neutral and pleasant expressions the effect was curvilinear.

Post hoc comparisons showed that valence ratings were lower for unpleasant than neutral $t(27) = 9.18$, $p < 0.001$, and pleasant expressions $t(27) = 7.23$, $p < 0.001$. Valence ratings were higher for pleasant than neutral expressions $t(27) = 3.48$, $p < 0.01$. Valence ratings were higher for medium sized than large agents $t(27) = 2.99$, $p < 0.01$. Other pairwise comparisons were not statistically significant.

	Component		
	1	2	3
Eigenvalue	1.933[a]	0.598	0.469
Valence	0.787[a]	−0.536[a]	0.306[a]
Arousal	−0.837[a]	0.018	0.546[a]
Dominance	0.783[a]	0.557[a]	0.277[a]

Table 1. Component matrix of emotional scales.
[a] $p < 0.05$

	Component		
	1	2	3
Eigenvalue	1.477[a]	1.105[a]	0.418
Conspicuousness	0.900[a]	−0.050	0.434[a]
Concentration	−0.471[a]	0.825[a]	0.313[a]
Interestingness	0.668[a]	0.649[a]	−0.364[a]

Table 2. Component matrix of attentional scales.
[a] $p < 0.05$

Principal Component Analysis

PCA of the emotion space (i.e., valence, arousal, and dominance) produced one component with eigenvalue statistically significantly over 1.0. The component was directly related to emotional valence and dominance, and inversely related to arousal (Table 1). We labeled the component as *comfort* that ranges from an alerting (i.e., unpleasant, aroused, and dominating) to a comfortable (i.e., pleasant, calm, and controlled) experience.

PCA of the attention-related ratings (i.e., conspicuousness, concentration, and interestingness) produced two components with eigenvalues statistically significantly over 1.0. The first component was directly related to conspicuousness and interestingness, and inversely related to concentration (Table 2). Stimuli with high scores of this component strongly attracted attention (i.e., were conspicuous), held attention (i.e., were interesting), and distracted concentration. Stimuli with low scores were rated as inconspicuous and uninteresting. Thus, the component was labeled as *stimulus-driven attention*.

The second component was positively related to both concentration and interestingness. Thus, stimuli with high scores of this component were voluntarily concentrated on and they held attention more intensely than stimuli with lower scores. We labeled the component as *voluntary attention*.

Emotion and Attention Scores

Table 3 shows the results of the 3×3 (facial expression $\times$ size) ANOVAs for comfort, stimulus-driven attention, and voluntary attention scores.

The main effects of expression and size on comfort scores were statistically significant. Figure 3 shows that the scores conformed to the agent's expression. Unpleasant expressions received a negative comfort score, neutral expressions received a higher score, and pleasant expressions received the highest positive score. Figure 3 also shows that comfort scores were lower when the agent was larger.

Comp. Anim. Virtual Worlds 2010; **21**: 215–224
DOI: 10.1002/cav

	Expression *F*(2,54)	Size *F*(2,54)	Expression × Size *F*(4,108)
Comfort	96.39[c]	16.00[c]	1.15
Stimulus-driven	36.09[c]	64.63[c]	6.81[b]
Voluntary	25.39[c]	6.14[b]	3.71[a]

Table 3. ANOVA results for comfort, stimulus-driven attention, and voluntary attention scores.
[a] $p < 0.01$
[b] $p < 0.001$

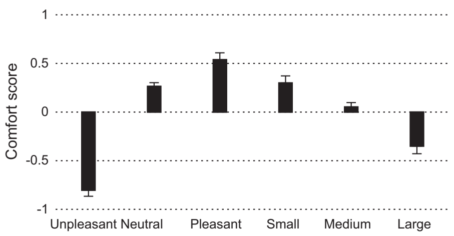


Figure 3. Mean comfort scores (and SEM) for different facial expressions and sizes of the agents.

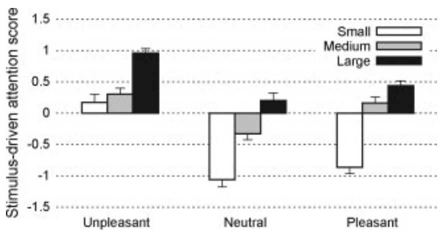


Figure 4. Mean stimulus-driven attention scores (and SEM) for different facial expressions and sizes of the agents.

Post hoc pairwise comparisons between expressions showed that comfort scores were significantly lower for unpleasant than neutral, $t(27) = 13.89$, $p < 0.001$, and pleasant expressions, $t(27) = 10.45$, $p < 0.001$, and significantly lower for neutral than pleasant expressions, $t(27) = 2.90$, $p < 0.05$. *Post hoc* pairwise comparisons between sizes showed that comfort scores were significantly higher for small than medium size agents, $t(27) = 2.64$, $p < 0.05$, and large agents, $t(27) = 4.52$, $p < 0.001$, and significantly higher for medium size than large agents, $t(27) = 3.85$, $p < 0.001$.

For stimulus-driven attention scores, Table 3 shows statistically significant main effects of expression and size. The interaction of expression and size was also statistically significant. Figure 4 shows that the mean stimulus-driven attention scores were higher for larger agents. The effect was similar within each expression. However, large unpleasant agents had a markedly high score, whereas small neutral and pleasant agents had a markedly low score. Figure 4 also shows that the valence of the agent's facial expression had a curvilinear effect within each size. Unpleasant (i.e., negative valence) agents received the highest mean score, neutral agents received the lowest score, and pleasant (i.e., positive valence) agents received a slightly higher mean score than neutral agents.

The interaction effect most likely reflected the markedly low and high mean scores of some agent categories (e.g., large unpleasant agents). However, we performed within-size and within-expression simple effect analyses in order to confirm the main effects. Simple effect of expression was significant within small, $F(2, 54) = 29.24$, $p < 0.001$, medium size, $F(2, 54) = 12.10$, $p < 0.001$, and large agents, $F(2, 54) = 17.99$, $p < 0.001$. Simple effect of size was significant within unpleasant, $F(2, 54) = 19.074$, $p < 0.001$, neutral, $F(2, 54) = 31.35$, $p < 0.001$, and pleasant expressions, $F(2, 54) = 56.04$, $p < 0.001$. As all simple effects were statistically significant, we proceeded to *post hoc* pairwise comparisons.

Post hoc comparisons between expressions showed that stimulus-driven attention scores were significantly higher for unpleasant than neutral, $t(27) = 7.54$, $p < 0.001$, and pleasant agents, $t(27) = 5.48$, $p < 0.001$, and significantly lower for neutral than pleasant agents, $t(27) = 3.32$, $p < .05$. *Post hoc* comparisons between sizes showed that stimulus-driven attention was significantly lower for small than medium size agents, $t(27) = 5.83$, $p < 0.05$, and large agents, $t(27) = 10.98$, $p < 0.001$, and significantly lower for medium than large sized agents, $t(27) = 5.77$, $p < 0.001$.

For voluntary attention scores, Table 3 shows statistically significant main effects of expression and size. The

DOI: 10.1002/cav

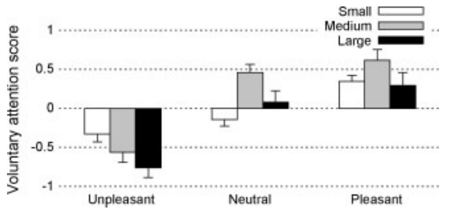


Figure 5. Mean voluntary attention scores (and SEM) for different facial expressions and sizes of the agents.

interaction of expression and size was also statistically significant. Figure 5 shows that the effect of agent size on voluntary attention scores was different between expressions. Unpleasant agents received lower scores the larger they were, whereas neutral and pleasant agents received highest scores when they were medium sized. On the other hand, the effect of agent's facial expression was consistent within each size. Unpleasant agents received the lowest score, neutral agents received a higher score, and pleasant agents received the highest mean score within each size.

Figure 5 suggests that the interaction effect reflected differences in response to size between different facial expressions. However, we performed both within-expression and within-size simple effect analyses in order to test both main effects. The simple effect of size was statistically significant within the unpleasant expression, $F(2, 54) = 4.63$, $p < 0.05$. The simple effect of size was not statistically significant within neutral and pleasant expressions. *Post hoc* pairwise comparisons showed that voluntary attention scores were significantly lower for large than small unpleasant agents, $t(27) = 2.92$, $p < 0.01$. Other pairwise comparisons were not statistically significant.

The simple effect of expression was significant within small, $F(2, 54) = 15.25$, $p < 0.001$, medium size, $F(2, 54) = 19.66$, $p < 0.001$, and large agents, $F(2, 54) = 11.60$, $p < 0.001$. As within-size simple effects of expression were significant, we proceeded to *post hoc* pairwise comparisons. Pairwise comparisons showed that voluntary attention score was lower for unpleasant than neutral, $t(27) = 5.53$, $p < 0.001$, and pleasant expressions, $t(27) = 5.85$, $p < 0.001$. Voluntary attention score was higher for pleasant than neutral agents, $t(27) = 2.27$, $p < 0.05$.

	Standardized	Raw Value (μV or BPM)
Corrugator supercilii EMG		
Unpleasant	-0.17 ± 0.08	-0.07 ± 0.05
Neutral	0.09 ± 0.07	0.05 ± 0.06
Pleasant	0.09 ± 0.06	0.06 ± 0.05
HR		
Small	0.02 ± 0.05	-0.39 ± 0.94
Medium	-0.12 ± 0.05	-1.70 ± 0.92
Large	0.10 ± 0.05	0.26 ± 0.37

Table 5. Mean ($\pm$ SEM) changes from baseline in *corrugator supercilii* EMG for different expressions and HR for agent sizes.

Physiological Responses

Table 4 shows the results of the 3×3 (facial expression $\times$ size) ANOVAs for standardized *corrugator supercilii* EMG, *zygomaticus major* EMG, and HR.

The main effect of expression on *corrugator supercilii* EMG was statistically significant. Table 5 shows that unpleasant agents suppressed, while neutral and pleasant agents increased *corrugator supercilii* EMG activity.

Post hoc pairwise comparisons showed that *corrugator supercilii* EMG was significantly lower during unpleasant than pleasant expressions, $t(27) = 2.14$, $p < 0.05$. Other pairwise comparisons were not statistically significant.

	Expression F(2,54)	Size F(2,54)	Expression $\times$ Size F(4,108)
Corrugator	3.25*	0.88	0.17
Zygomaticus	1.11	2.36	0.17
HR	0.34	3.24*	1.15

Table 4. ANOVA results for *corrugator supercilii* EMG, *zygomaticus major* EMG, and HR.
$^*p < 0.05$

Comp. Anim. Virtual Worlds 2010; **21**: 215–224
DOI: 10.1002/cav

For HR, Table 4 shows statistically significant main effect of size. Table 5 shows that HR clearly decelerated in response to medium sized agents and accelerated in response to large agents. Raw HR responses to small agents in beats per minute (BPM) show some variability, but standardized scores more clearly suggest that HR did not markedly respond to small agents.

Post hoc pairwise comparisons showed that HR was significantly lower in response to medium size than large agents, $t(27) = 2.45$, $p < 0.05$. Other pairwise comparisons were not statistically significant.

Discussion

Our results showed that both the expressions and the virtual proximity of embodied agents had statistically significant effects on both subjective ratings and physiological activity. According to subjective ratings, unpleasant agents were more alerting (i.e., unpleasant, arousing, and dominating) and attracted more stimulus-driven attention than neutral and pleasant agents. On the other hand, pleasant agents were rated as the most comforting (i.e., pleasant, calming, and controlled) and attracted more voluntary attention than unpleasant and neutral agents.

The size of the agent (i.e., apparent distance) affected so that small agents were rated as comforting, while large agents were rated as alerting. Agents also attracted more stimulus-driven attention the larger they were. A closer proximity intensified the effect of facial expressions, as ratings of voluntary attention decreased even lower when unpleasant agents increased in size.

HR was averaged over each stimulus presentation. This analysis revealed HR deceleration in response to medium size agents and acceleration in response to large agents. Small agents did not markedly change the HR. Large agents also received the highest mean score of stimulus-driven attention and they were rated as the most alerting. These subjective responses and the observed HR acceleration may both be linked with the need to prepare for action.[23]

These results are in line with previous work showing that larger or closer negative stimulation, in general, evokes accentuated subjective and physiological responses.[11] These responses are tied to the basic neurocognitive systems that direct us to approach positive and withdraw from negative stimuli.[17] However, although these basic tendencies are thought to be universal, there is evidence that evaluations of virtual facial expressions are culture-dependent to some extent.[34] Thus, several factors (e.g., culture) should still be considered when designing artificial facial cues in order to convey the intended emotional tone (e.g., pleasant).

The present *corrugator supercilii* EMG responses were opposite to the emotional tone of the agent, as unpleasant agents suppressed and pleasant agents increased *corrugator supercilii* EMG activity. Similar counter or neutral responses to facial cues have been previously reported, for example, when the observer does not affiliate with the expresser (due to, e.g., opposite gender).[25] However, at least our manipulation check did not show any significant gender effects. Thus, the exact nature of the present facial responses is not fully clear.

In sum, our results showed that the simulated emotional tone and bodily distance to a virtual agent activated both subjective and physiological responses. Thus, virtual computer agents have the means for effectively regulating the human psychophysiological emotion and attention response systems. It is widely accepted that these systems are central in human motivational behavior.[14,23] In conclusion, fine grained and vivid emotional and social cues may become a major tool for future interfaces to engage and motivate their users.

ACKNOWLEDGMENTS

The EU 6th framework project AtGentive (IST-4-027529-STP), the Academy of Finland (project number 1115997), and the Graduate School in User-Centered Information Technology (UCIT) financially supported this research.

References

1. Partala T, Surakka V. The effects of affective interventions in human–computer interaction. *Interacting with Computers* 2004; **16**(2): 295–309.
2. Moundridou M, Virvou M. Evaluating the persona effect of an interface agent in a tutoring system. *Journal of Computer Assisted Learning* 2002; **18**: 253–261.
3. Beale R, Creed C. Affective interaction: how emotional agents affect users. *International Journal of Human–Computer Studies* 2009; **67**: 755–776.
4. Nass S, Steuer J, Tauber ER. Computers are social actors. *Proceedings of CHI'94*, 1994; 72–78.
5. Hess U, Blairy S. Facial mimicry and emotional contagion to dynamic emotional facial expressions and their influence on decoding accuracy. *International Journal of Psychophysiology* 2001; **40**: 129–141.
6. Weyers P, Mühlberger A, Hefele C, Pauli P. Electromyographic responses to static and dynamic avatar emotional facial expressions. *Psychophysiology* 2006; **43**: 450–453.

7. Martin J-C, Niewiadomski R, Devillers L, Buisine S, Pelachaud C. Multimodal complex emotions: gesture expressivity and blended facial expressions. *International Journal of Humanoid Robotics* 2006; **3**(3): 269–291.
8. Loftus GR, Harley EM. Why is it easier to identify someone close than far away? *Psychonomic Bulletin & Review* 2005; **12**: 43–65.
9. Hall ET. *The Hidden Dimension*. Doubleday Press: Garden City, NY, USA, 1996.
10. Hayduk LA. Personal space: where we now stand. *Psychological Bulletin* 1983; **94**(2): 293–335.
11. Codispoti M, De Cesarei A. Arousal and attention: picture size and emotional reactions. *Psychophysiology* 2007; **44**: 680–686.
12. Partala T, Surakka V, Lahti J. Affective effects of agent proximity in conversational systems. *Proceedings of NordiCHI*, 2004; 353–356.
13. Bradley MM, Lang PJ. Measuring emotions: the self-assessment manikin and the semantic differential. *Journal of Behavioral Therapy and Experimental Psychiatry* 1994; **25**(1): 49–59.
14. Mauss IB, Levenson RW, McCarter L, Wilhelm FH, Gross JJ. The tie that binds? Coherence among emotion experience, behavior, and physiology. *Emotion* 2005; **5**(2): 175–190.
15. Sternberg RJ. *Cognitive Psychology*. Wadsworth: Belmont, CA, USA, 2006.
16. Surakka V, Tenhunen-Eskelinen M, Hietanen JK, Sams M. Modulation of human auditory information processing by visual emotional stimuli. *Cognitive Brain Research* 1998; **7**: 159–163.
17. Amodio DM, Master SL, Yee CM, Taylor SE. Neurocognitive components of the behavioral inhibition and activation systems: implications for theories of self-regulation. *Psychophysiology* 2008; **45**: 11–19.
18. Posner MI. Cumulative development of attentional theory. *American Psychologist* 1982; **37**(2): 168–179.
19. Arvidson PS. A lexicon of attention: from cognitive science to phenomenology. *Phenomenology and the Cognitive Sciences* 2003; **2**(2): 99–132.
20. Schuemie MJ, van der Straaten P, Krijn M, van der Mast CAPG. Research on presence in virtual reality: a survey. *CyberPsychology & Behavior* 2001; **4**(2): 183–201.
21. Merriam-Webster. Retrieved on: 17 August 2009. Available at: http://www.merriam-webster.com/dictionary
22. Lehtinen M, Grönros E-R, Eronen R, *et al*. (eds). *MOT Kielitoimiston sanakirja 1.0*. Kotimaisten kielten tutkimuskeskuksen julkaisuja 132. Kotimaisten kielten tutkimuskeskus ja Kielikone Oy: Helsinki, Finland, 2004.
23. Bradley MM. Emotion and Motivation. In *Handbook of Psychophysiology* (2nd edn). Cambridge University Press: Cambridge, UK, 2000; 602–642.
24. Larsen JT, Norris CJ, Cacioppo JT. Effects of positive and negative affect on electromyographic activity over the zygomaticus major and corrugator supercilii. *Psychophysiology* 2003; **40**: 776–785.
25. Bourgeois P, Hess U. The impact of social context on mimicry. *Biological Psychology* 2008; **77**(3): 343–352.
26. Fridlund AJ, Cacioppo JT. Guidelines for human electromyographic research. *Psychophysiology* 1986; **23**: 567–589.
27. Neuroscan. *Acquire 4.3: online acquisition of neurophysiological data*. Compumedics Neuroscan, 2003. Document number 2202, Revision E.
28. Schneider W, Eschman A, Zuccolotto A. *E-Prime User's Guide*. Psychology Software Tools, Inc., Pittsburgh, PA, USA, 2002.
29. Cantoche. http://www.cantoche.com/
30. PhysioNet. Physiotoolkit. http://www.physionet.org/
31. Gratton G. Biosignal Processing. In *The Handbook of Psychophysiology*. Cambridge University Press: Cambridge, UK, 2000; 900–923.
32. Mehlman DW, Shepherd UL, Kelt DA. Bootstrapping principal components analysis: a comment. *Ecology* 1995; **76**(2): 640–643.
33. Howell DC. *Statistical Methods for Psychology* (6th edn). Thomson Wadsworth: Belmont, CA, USA, 2007.
34. Koda T, Ishida T, Rehm M, André E. Avatar culture: cross-cultural evaluations of avatar facial expressions. *AI & Society* 2009; **24**(3): 237–250.

Authors' biographies:

Toni Vanhala is a postgraduate student at the University of Tampere, Finland. His thesis aims to create a physiology-based adaptive computer system for the treatment of anxiety disorders. He works as a researcher in the Research Group for Emotions, Sociality, and Computing at TAUCHI (Unit on Computer–Human Interaction).

Veikko Surakka is a professor in interactive technology. He has MA, Psych. Lic., and doctoral degrees in psychology. He is the head of the Research Group for Emotions, Sociality, and Computing (http://www.cs.uta.fi/hci/ESC/index.html). The group focuses especially in research on emotions in human–technology interaction.

Comp. Anim. Virtual Worlds 2010; **21**: 215–224
DOI: 10.1002/cav

Harri Siirtola works as a postdoctoral researcher in the Visual Interaction Research Group (VIRG) of Tampere Unit for Human–Computer Interaction (TAUCHI). His research centers on information visualization and graphical user interfaces, focusing on interactive techniques for information visualization.

Kari-Jouko Räihä received his education in Computer Science at the University of Helsinki, where he obtained his PhD in 1982. Since 1985 he has been a full professor of Computer Science at the University of Tampere, Finland. He founded TAUCHI (Unit on Computer–Human Interaction (http://tauchi.cs.uta.fi/)) in mid-1990s and has led it since. Räihä's research interests span HCI broadly, with special interest in new interaction techniques, especially those based on the use of eye gaze.

Benoît Morel is the cofounder and CEO of the French company, Cantoche that is the inventor of the Living Actor(tm) technology. After experiences as Sound Engineer, Radio Producer, Video Game artist, his main interest is the use of avatars for companies and enduser applications.

Laurent Ach is the CTO of the French company, Cantoche that is the inventor of the Living Actor(tm) technology. After he graduated as an engineer from Ecole Centrale of Lyon in 1990, his main interests were the fields of 3D visualization technologies, Virtual Reality, and Artificial Intelligence. He worked in Thales Group, at Sagem, and in startup companies.

DOI: 10.1002/cav

Appendix E

Publication V

Vanhala, T. and Surakka, V. (2007). Facial activation control effect (FACE). In *Affective Computing and Intelligent Interaction (ACII '07), Lecture Notes In Computer Science*, 4738, 278–289. Springer.

Available at:
`http://dx.doi.org/10.1007/978-3-540-74889-2_25`

Appendix F

Publication VI

Vanhala, T., Surakka, V., Courgeon, M., and Martin, J.-C. (in press). Voluntary facial activations regulate physiological arousal and subjective experiences during virtual social stimulation. To appear in *Transactions on Applied Perception*, 9 (1). ACM.

Voluntary Facial Activations Regulate Physiological Arousal and Subjective Experiences During Virtual Social Stimulation

TONI VANHALA and VEIKKO SURAKKA, Research Group for Emotions, Sociality, and Computing, TAUCHI, University of Tampere
MATTHIEU COURGEON and JEAN-CLAUDE MARTIN, LIMSI-CNRS, Paris South XI University

Exposure to distressing computer-generated stimuli and feedback of physiological changes during exposure have been effective in the treatment of anxiety disorders (e.g., social phobia). Here we studied voluntary facial activations as a method for regulating more spontaneous physiological changes during virtual social stimulation. 24 participants with low or high level of social anxiety activated either the *corrugator supercilii* (used in frowning) or the *zygomaticus major* (used in smiling) facial muscle to keep a female or a male computer character walking towards them. The more socially anxious participants had higher level of skin conductance throughout the trials as compared to less anxious participants. Within both groups, short-term skin conductance responses were enhanced both during and after facial activations, and *corrugator supercilii* activations facilitated longer term electrodermal relaxation. *Zygomaticus major* activations had opposite effects on subjective emotional ratings of the less and the more socially anxious. In sum, voluntary facial activations were effective in regulating emotional arousal during virtual social exposure. Especially *corrugator supercilii* activation was found to be a promising method for facilitating autonomic relaxation.

Categories and Subject Descriptors: H.1.2 [**Models and Principles**]: User/Machine Systems—*human factors*; J.4 [**Social and Behavioral Sciences**] psychology; H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*input devices and strategies*

General Terms: Experimentation, Human Factors

Additional Key Words and Phrases: Affective computing, virtual characters, social anxiety, facial muscles, electrodermal activity

ACM Reference Format:

Vanhala, T., Surakka, V., Courgeon, M., and Martin, J.-C. 2012. Voluntary facial activations regulate physiological arousal and subjective experiences during virtual social stimulation. ACM Trans. Appl. Percept. 9, 1, Article 0 (2012), 20 pages.
DOI = 10.1145/0000000.0000000 http://doi.acm.org/10.1145/0000000.0000000

1. INTRODUCTION

Exposure to virtual stimulation has proven to be an effective method for overcoming excessive anxiety and fears of, for example, spiders, flying, and small confined places (i.e., claustrophobia) [North et al. 1998;

This study was financially supported by the Graduate School in User-Centered Information Technology and the Academy of Finland (project number 1115997). Author's addresses: T. Vanhala, Current address: VTT Technical Research Centre of Finland, P.O. Box 1300, 33101 Tampere, Finland; V. Surakka, Research Group for Emotions, Sociality, and Computing, Tampere Unit for Computer-Human Interaction (TAUCHI), School of Information Sciences, 33014 University of Tampere, Finland; M. Courgeon, J.-C. Martin, Computer Sciences Laboratory for Mechanics and Engineering Sciences (LIMSI), National Center for Scientific Research (CNRS), Paris South XI University, BP 133, 91403 Paris, France.

DOI 10.1145/0000000.0000000 http://doi.acm.org/10.1145/0000000.0000000

Rothbaum et al. 2000; Gerardi et al. 2010]. In one traditional treatment of anxiety disorders (e.g., phobias), a person is gradually exposed and desensitized to increasing levels of distressing stimulation, for example, by reducing the distance to a real spider [Foa and Kozak 1986; Öst 1989]. Computer-generated virtual stimulation has several advantages as compared to traditional face-to-face therapy. For example, the level of exposure (e.g., detailed movements of a virtual spider) and even the environment (e.g., simulated weather during a flight) can be controlled at a very fine-grained level [Wiederhold and Rizzo 2005; Vanhala and Surakka 2008; Courtney et al. 2010; Gerardi et al. 2010]. A recent meta-analysis of controlled studies on virtual reality exposure therapy showed that virtual reality exposure has been slightly more efficient in affecting the outcomes of treatment (i.e., reducing a variety of phobic symptoms) as compared to *in vivo* exposure [Powers and Emmelkamp 2008]. The relatively high control over stimulus exposure could in part explain why virtual stimulation is so effective. Further, some potent therapeutic virtual setups would be very hard or practically impossible to arrange in traditional therapy. For example, virtual reality has enabled vivid and immersive re-enactments of traumatizing events, which has helped people to overcome their post-traumatic stress reactions originating from such events (e.g., car accidents and terrorist bombings) [Beck et al. 2007; Josman et al. 2008].

One particularly challenging form of anxiety disorders is social phobia, that is, excessive fears related to social interaction. Social phobia is characterized by and is in part diagnosed based on avoidance of social situations and inhibition of social behavior [Mick and Telch 1998; Connor et al. 2000; Gerardi et al. 2010; Rinck et al. 2010]. This means that suitable settings for treating social anxiety are difficult to arrange in regular face-to-face therapy, as the person may avoid social contact in the first place and she or he may be unwilling to participate in group-therapy [Olfson et al. 2000; Anderson et al. 2003]. For these reasons, it is not surprising that social phobia remains largely untreated in the general population [Kessler 2003]. However, there is evidence that virtual reality exposure may be more acceptable than real-world exposure [Garcia-Palacios et al. 2007]. Thus, virtual reality environments could support a person to seek treatment.

Exposure treatment of social phobia is a relatively recent approach, but the results so far have been promising [Anderson et al. 2003]. For example, public speaking anxiety has been successfully treated using exposure to a group of peers [Heimberg et al. 1990]. There is also evidence that virtual computer-generated stimulation can be similarly effective in inducing anxiety, for example, when speaking to a virtual crowd of computer-generated characters [James et al. 2003; Anderson et al. 2003; Gerardi et al. 2010]. Further, continued exposure to speaking in front of a virtual audience has been found both to reduce self-reported public speaking anxiety and to decrease the mean heart rate during a speech [Harris et al. 2002].

More generally, several studies have found that human responses to computer-generated social cues (e.g., the facial expressions of computer-generated human characters) are similar to those in human-human interaction [Nass et al. 1994; Pertaub et al. 2002; Zanbaka et al. 2007; Beale and Creed 2009]. For example, simulated distance to a virtual human-like character has been found to affect how dominant the character appears to the participants [Partala et al. 2004; Vanhala et al. 2010]. In another recent study, virtual characters approached participants who were instructed to stand still in the virtual environment [Llobera et al. 2010]. It was found that their physiological arousal after the approach was higher when the virtual characters came closer. Similarly, in a study using an opposite setup, it was found that more socially anxious participants remained at a farther distance from virtual human characters that they could freely approach in a virtual reality environment [Rinck et al. 2010]. Based on the above, virtual computer systems seem to offer controlled stimulation that can affect human social processes, even to the level of inducing significant social anxiety, and could thus also be used as a part of exposure therapy.

A relatively simple computer-assisted setup for controlling the level of exposure to distressing virtual stimuli could be based on the variation of simulated distance to a stimulus. One of the strongest cues of distance is the retinal size of the visual object, that is, larger images appear to be closer [Loftus and Harley 2005]. In general, large images or images that appear to approach the person have been found to accentuate both subjective and physiological arousal [Reeves et al. 1999; Codispoti and De Cesarei 2007; Vanhala et al. 2010]. In particular, closer distance to a feared object (e.g., a snake) accentuates the subjective (e.g., self-reported level of fear) and physiological (e.g., heart rate) responses to the stimulus [Teghtsoonian and Frost 1982]. Impressive results have been achieved in the treatment of phobias with a specific object (e.g., spiders) by instructing participants to cope with an increasingly closer distance to a distressing stimulus [Öst 1989; Hellström and Öst 1995]. In many cases a single session of about 2 hours has been sufficient to suppress the excessive fear and allow the person to function normally in her or his daily life.

There is evidence that monitoring and feedback of physiological reactions can further improve the effectiveness of exposure therapy [Gatchel and Proctor 1976; Wiederhold and Wiederhold 2003; Wiederhold and Rizzo 2005]. The person may be taught methods for voluntarily regulating her or his physiological processes (e.g., heart rate and blood pressure) prior to the therapy, which may then enable her or him to gain awareness of and control over excessive reactions during therapy [Gatchel and Proctor 1976; Hellström and Öst 1995; Gerardi et al. 2010]. Another option is to provide continuous feedback about physiological processes during the exposure, which could be especially suitable in an already computerized setup (i.e., virtual exposure). There is some evidence that such real-time feedback about the physiological processes could improve the effects of virtual exposure therapy. For example, in one follow-up study participants who received visual feedback about their physiology during virtual exposure therapy for the fear of flying were all able to fly without medication 3 years after the treatment [Wiederhold and Wiederhold 2003]. Physiological measurements can also provide additional information about the progress of therapy to the therapist. For example, heart rate during exposure therapy can be expected to decrease from session to session as treatment progresses [Grayson et al. 1982].

In traditional therapy without technological aids, the progress of therapy is typically assessed by verbal reports. Such assessments are necessary for deciding when the person is ready for higher levels of exposure (e.g., take a step closer to the spider) and when support should be given, for example, by guided relaxation. A common measure used in guiding a therapeutic session is the Subjective Units of Discomfort (SUD) scale which is a person's verbal assessment of her or his experienced distress, for example, on a scale from 0 (i.e., none) to 100 (i.e., maximum) [Wiederhold and Wiederhold 2003; Krijn et al. 2004]. A more detailed approach to measuring emotional responses divides them into three components and corresponding bipolar rating scales: emotional valence (i.e., from unpleasant to pleasant), arousal (i.e., from relaxed to aroused), and dominance (i.e., from being controlled to being in control of the stimulus or situation) [Bradley and Lang 1994]. However, it is clear that more detailed subjective measures of emotion (e.g., the use of several rating scales) can interfere the process of systematic desensitization, for example, by distracting attention from the therapy itself.

In comparison to subjective ratings of emotional experiences, physiological measures have been suggested as relatively non-invasive methods for monitoring emotional processes during exposure therapy and controlled stimulation in general [Ward and Marsden 2003; Wiederhold and Wiederhold 2003; Vanhala and Surakka 2008]. For example, anxiety related emotional arousal has long been associated with changes in electrodermal activity (EDA), that is, changes in the function of sweat glands as measured by the conductivity of electricity on skin [Schlosberg 1954; Dawson et al. 2000]. Due to the high specificity of EDA measures, some researchers have even suggested using them as direct measures of anxiety [Fowles 1988; Dawson et al. 2000]. EDA is an elecrophysiological measure that specifically indicates the level of sympathetic activation of the autonomic nervous system (ANS). Higher level of sympathetic activation is typically associated with higher overall skin conductance level (SCL), enhanced increases in skin conductance after the onset of a stimulus (i.e., event-related skin conductance response, ER-SCR), increased frequency of skin conductance changes in the absence of a specific stimulus (i.e., non-specific skin conductance responses, NS-SCR), and increased magnitude of NS-SCR [Dawson et al. 2000]. ER-SCR can also be seen as a part of the orientation response, that is, greater magnitude of response is typically associated with enhanced attention to the eliciting stimulus [Frith and

Allen 1983; Dawson et al. 2000]. More commonly EDA measures have been tightly connected with arousal induced by emotionally relevant (i.e., pleasant or unpleasant) visual and auditory stimuli [Witvliet and Vrana 1995; Bradley 2000; Bradley et al. 2001; Mauss and Robinson 2009].

Perhaps the most well-established physiological measure associated with emotional valence is facial electromyography (EMG) which reflects the electrical activity of facial muscles [Larsen et al. 2003]. The activity of *zygomaticus major* muscle in the cheek (activated when smiling) is known to typically decrease during unpleasant and increase during pleasant emotional experiences. The activity of *corrugator supercilii* muscle in the forehead (activated when frowning) varies with emotional valence in the opposite manner. More generally, it has been suggested that these muscles are common to many facial expressions of positive (e.g., *zygomaticus major* activations in pleasure and contentment) and negative (e.g., activations of forehead muscles in distress and fear) emotion [Ekman 1979; 1993; 2004]. In the context of computer systems, automatic analysis of the changes in EMG activity can provide real-time estimates of the experienced emotional valence. For example, our recent system was able to match its EMG-based estimations of emotional valence (i.e., negative or positive) to the participants' ratings of subjective emotional experience with accuracies of over 70% for picture stimulation and over 80% for video stimulation, although the accuracy of the system did vary quite much between participants [Partala et al. 2006]. Average accuracies ranged subject-wise from 47.3% to 87.1% for picture stimuli and from 41.6% to 85.0% for video stimuli. Thus, although in general there is a systematic pattern in facial responses to negatively and positively valenced stimuli, individual differences in the association of emotion and facial activity could be studied in further detail (e.g., which characteristics between participants could account for such differences).

Similar to more spontaneous activity, voluntary activations of facial muscles are also associated with specific effects on emotional experience and physiological measures [Levenson et al. 1990; Coan et al. 2001]. In one series of experiments, Levenson and his colleagues [Levenson et al. 1990; Levenson et al. 1992; Levenson and Ekman 2002] instructed participants muscle-by-muscle to voluntarily produce facial configurations associated with certain emotions (e.g., happiness or anger). For example, the facial configuration for anger required the participant to activate the *corrugator supercilii* muscle and five other facial muscles. The results showed that the voluntary activations induced significant numbers of self-reported emotional experiences that matched the facial configuration and specific physiological differences between emotions (e.g., fear increased skin conductance more than happiness). In our own studies, we have found that computer-assisted activations (i.e., on-screen instructions and visual biofeedback) of either the *zygomaticus major* or the *corrugator supercilii* muscle can significantly affect concurrent emotion related ANS activity in terms of heart rate and heart rate variability [Vanhala and Surakka 2007].

There is some debate whether such physiological responses are distinct patterns of physiological activity associated with a set of discrete emotions (e.g., fear and happiness) [Boiten 1996; Rainville et al. 2006; Mauss and Robinson 2009]. However, evidence for the correlation of autonomic activity and emotional dimensions (e.g., positively and negatively valenced emotions) has been quite consistently found, also when reviewing the effects of voluntary activations [Cacioppo et al. 2000; Levenson and Ekman 2002; Christie and Friedman 2004; Mauss and Robinson 2009]. In any case, taken at face value the above results indicate that voluntary facial activations have significant physiological effects. In general, researchers see physiological changes as one of the essential parts of emotion [Mauss et al. 2005]. Further, recent research suggests that both cognitive and emotional knowledge may be fundamentally embodied, that is, physiological changes may be a core part of all human information processing [Niedenthal 2007]. Thus, voluntary activations of facial muscles, which are mainly controlled by the central nervous system, could provide means for significantly affecting more spontaneous physiological processes, which are more tightly controlled by the ANS. This could also facilitate control over heightened physiological activity which is a core part of excessive anxiety and fear [Connor et al. 2000; Wiederhold and Wiederhold 2003].

In summary, previous research suggests that computer-assisted systems can provide effective means for

Table I. Gender distributions, mean ± standard error of the mean (S.E.M.), minimum, and maximum values for the age and SPIN-FIN scores of all students and the experiment groups

	gender		age	SPIN-FIN test			SPIN-FIN retest		
group	male	female	mean	mean	min	max	mean	min	max
All	59	27	24.2±.5	15.9±1.1	0	47			
LSA	10	2	24.1±1.0	5.5±.8	1	8	4.1±.9	0	12
HSA	5	7	22.4±.8	32.2±2.8	22	47	36.3±2.2	12	51

regulating physiological arousal and subjective anxiety. Similar to *in vivo* exposure, virtual exposure has been effective in evoking and reducing subjective and physiological responses to distressing stimulation. Physiological monitoring and feedback has been found to further improve the outcome of such virtual exposure. However, monitoring and voluntary regulation of physiological responses during exposure may require effort and distract attention from the therapy itself. Perhaps voluntary facial activations could offer a more seamless method for regulating physiology during therapy by using facial activity to control the exposure itself, that is, without explicit biofeedback.

The present aim was to study whether voluntary facial activation could be used to regulate emotion related subjective experiences and physiological arousal during artificial social stimulation. 24 participants were recruited based on their level of social anxiety assessed with the Finnish translation of the Social Phobia Inventory (SPIN-FIN) [Ranta et al. 2007]. During the experiment, participants viewed realistic human-like female and male characters that were controlled by voluntary *corrugator supercilii* or *zygomaticus major* activations. The characters appeared to walk closer towards the participant, stop at a close distance, and deliver an arithmetic equation using speech synthesis. The participant was to attend to the equation, verbally repeat it, and report whether the equation was correct or false. Skin conductance was measured during the tasks to assess the level of arousal in relation to the level of social anxiety and the effects of facial activations. Subjective ratings of emotional valence, arousal, and dominance were also collected before and after the facial activations in order to investigate changes in subjective experiences.

2. METHODS

2.1 Participants

Voluntary participants were invited to the study based on their SPIN-FIN score. The SPIN-FIN questionnaire was completed by 88 students at the Department of Computer Sciences, University of Tampere, Finland. 2 questionnaires were excluded due to incomplete data. The present study was one of several that were offered for the students for partial course credit. Table I shows two groups of 12 participants each that were recruited to the present experiment based on their SPIN-FIN score. The Low Social Anxiety (LSA) group was recruited from the lower quartile of SPIN-FIN scores, while the High Social Anxiety (HSA) group was recruited from the upper quartile of SPIN-FIN scores. Cut-off scores were at 8 points for the LSA group, and at 22 points for the HSA group. This exceeded the suggested cut-off score of 19 points for screening sub-clinical social phobia and social phobia using the SPIN and SPIN-FIN questionnaires [Connor et al. 2000; Ranta et al. 2007]. A SPIN-FIN retest was administered at the beginning of the experiment session which was held 8–22 days after the first test. All participants had normal or corrected to normal vision. An informed written consent was obtained from each participant before the experiment.

2.2 Materials

Female and male realistic humanlike characters in Figures 1 and 2 were used in the study. The characters were presented using the Multimodal Affective and Reactive Character (M.A.R.C.) software [Courgeon et al. 2008]. M.A.R.C. is a platform which simulates a virtual human in real-time using high-quality graphical rendering

Fig. 1. The female character standing at the initial pose, in mid gait, and at the closest distance.

of face and body using high resolution textures and 3D models (about 80,000 polygons on each character). M.A.R.C. is designed to provide a controlled platform for studies of human perception [Courgeon et al. 2009]. In the present study, M.A.R.C. was used to simulate a person that walks towards the participant, stops at a close distance, and speaks to the participant. M.A.R.C. produced eye blinking animations, automatic lip-synchronization, and motion-capture based walking animation for the characters. The animations were generated in real time so that, for example, if walking was paused at mid gait, the character returned to an upright stance with feet on the ground.

The character was displayed on a white background, that is, no 3D environment was displayed in order to encourage the participant to focus on the approaching behavior of the character. The character was set to show a neutral facial expression throughout the experiment. The virtual camera was positioned so that each character's head would be approximately 23 cm high at the closest distance, while keeping the time it took to walk from the initial distance to the closest distance equal between the characters. Thus, different camera positions were used for the two characters. Initial heights of the characters at the onset of stimulation were 54 mm for the female and 49 mm for the male character. Starting at the initial distance and walking without pauses, it took 8010 ± 70 ms for the character to reach the closest distance.

Festival speech synthesis (Finnish female and male voices) was used in creating verbal arithmetic tasks to be answered by the participants. The aim was to direct attention to the artificial stimulation by requiring participants to concentrate on the speech of the character. Further, the aim was to choose a task that would be unlikely to confound the group-wise effects of the facial activation with the demands of the task. Previous studies suggest that mental arithmetic is suitable for this purpose, as social anxiety does not affect performance in such a task [Larkin et al. 1998; Gramer and Saria 2007]. The tasks were modeled after previous studies using subtraction tasks that require moderate effort to answer correctly (e.g., Geary, French, & Wiley, 1993). 18 pre-generated tasks were created by randomly choosing a two-digit number and a one-digit number. The tasks were in the form of an equation where the one-digit number was subtracted from the two-digit number. The right-side of the equation (i.e., difference) was either correct (e.g., $95 - 9 = 86$; 10 equations), or incorrect so that the result was either smaller by one (4 equations) or greater by one (e.g.,

Fig. 2. The male character standing at the initial pose, in mid gait, and at the closest distance.

74 - 3 = 72; 4 equations). The verbal expression of the character consisted of a sentence with an invariant first part followed by the equation. For example, the character said "In my opinion, seventy-four minus three is equal to seventy-two". The duration of speech samples ranged from 5170 to 7270 ms. Speech samples for the female and the male voice had equal duration (± 3.0 ms).

100 ms long sound tones for indicating the onset (1000 Hz tone) and offset (1100 Hz tone) of facial activation tasks were created using sinusoidal wave and the Hanning window. All sound materials (i.e., tones and speech samples) were in 16-bit 22.05 kHz mono audio.

2.3 Equipment

The experiment was performed at an electro-magnetically shielded and sound-attenuated laboratory. Participants were seated at a distance of 50 cm from a 19-inch Dell 1908FPb flat-panel computer monitor that was rotated to a vertical orientation (see Figure 3). Resolution of 1024 × 768 pixels was used. EMG was measured using bipolar pre-gelled Ag-AgCl electrodes (Spes Medica 15 × 20 mm) placed above the *corrugator supercilii* and the *zygomaticus major* muscle sites on the left side of the face. A reference electrode was placed on the mastoid process behind the left ear. Guidelines of Fridlund and Cacioppo (1986) were followed in EMG preparation.

EDA was measured from the non-dominant hand using Ag-AgCl sintered electrodes attached to the intermediate phalanges (i.e., second section) of index and middle fingers. Participants were instructed to wash and carefully dry their hands before electrodes were attached. Grass Technologies EC 33 electrode paste was used.

EMG and EDA were measured with a NeXus-10 physiological monitoring device (Mind Media B.V.) that was connected to a laptop computer using a wireless Bluetooth communications link. The sample rates were 2048 Hz for EMG and 256 Hz for EDA. EMG was analyzed according to common procedure as follows [Tassinary and Cacioppo 2000]. Analog high-pass filter of .5 Hz was used and EMG was further digitally pass-band filtered (7-th order Butterworth) from 20 to 500 Hz. Finally, a 500 ms moving average filter was applied to rectified EMG samples in order to derive a sample-by-sample estimate of muscle tension.

Fig. 3. The first author seated at a distance of 50 cm from the monitor while wearing the EMG and EDA electrodes.

An AKG CK 80 microphone placed at the base of the computer monitor was used in recording the participant's speech. Mono audio was captured at 22.05 kHz with 16-bit resolution and stored without compression. Sound materials were presented to the participant using Creative Cambridge Soundworks® SW310 speakers. The experiment was conducted using a single Dell E6410 laptop (Microsoft Windows XP SP3, Intel i5 2.53 GHz CPU, 3 GB RAM, NVidia Quadro NVS 3100M) which ran the M.A.R.C. software and the experimental software that communicated with the M.A.R.C. software using UDP. The experimental software registered and processed the EMG and EDA measurements and audio recordings in real-time. The software was implemented in C++ using Microsoft Foundation Classes and DirectX® technologies.

2.4 System Calibration

The system was adapted for each participant by measuring the minimum and maximum levels of EMG activity during a separate calibration session prior to the experiment. The participant was given a mirror to help in monitoring her or his own facial behavior and was instructed to practice the activation of both facial muscles. A computer screen with two vertical black rectangles was shown to the participant during the calibration. The height of the rectangles varied according to the measured values of EMG activity. The left rectangle showed the level of *corrugator supercilii* activity in relation to the minimum and maximum values registered during the calibration, whereas the right rectangle showed the level of *zygomaticus major* activity. The registered range of EMG activity was used to define a threshold for facial activation tasks in the experiment by adding 20 % of the registered range of activity to the minimum value. A horizontal red line was shown on both rectangles to indicate the current threshold.

The participant was guided to produce at least two maximal activations, first with the *corrugator supercilii* muscle and then with the *zygomaticus major* muscle, while relaxing between activations. After maximal activations, the participant was informed about the required threshold and asked to produce longer activations that were maintained above the threshold level. The calibration session ended after several controlled activations had been produced and the participant felt confident in her or his ability to voluntarily activate each muscle at the required level.

2.5 Trials

Participants were to perform trials by activating one of the two muscles (i.e., *corrugator supecilii* or *zygomaticus major*) or without activating either muscle. Each trial consisted of a relaxation period, a facial activation task, and the response to a verbal arithmetic task. A single muscle, a single character, and a randomly chosen arithmetic task was used in a trial.

First, participants were to relax their facial muscles for 6 consecutive seconds according to text instructions displayed on the computer screen. When both the *corrugator supercilii* and the *zygomaticus major* muscles were relaxed for more than 1 second, a counter starting from 5 and counting to 1 was displayed instead of the instructions. If the level of EMG activity of either muscle exceeded the defined threshold, the counter was reset and the instructions shown again.

After relaxation, the display faded in from a black screen to showing either the female or the male agent standing at the initial distance. After one second, the 1000 Hz tone indicated the start of the facial activation task. When the EMG activity of the muscle that was to be used in the task exceeded the defined threshold, the character appeared to walk towards the participant. If the intensity level of EMG activity dropped and remained for 250 ms below the threshold, the character stopped walking. The activity had to remain for 250 ms above the threshold to restart the movement. The third task condition did not require voluntary facial activation, that is, the character started walking after the tone and did not stop until reaching the closest distance.

When the character reached the closest distance, the 1100 Hz tone was sounded to indicate the end of the facial task. After 3000 ms, the character started to speak with a synthesized voice. The character delivered one of the verbal arithmetic tasks. The participant was to repeat the equation, and to say whether the equation was correct or false. The experimenter was seated in another room during the trials and wore headphones to listen and detect when the response was completed. Then, he used a keyboard to manually proceed to the next trial. The screen faded to black and the next trial started from the relaxation period.

In addition to the experimental trials, the participant performed two training trials when the task condition changed, that is, when another muscle was to be used in the following trial. Tasks without facial activation were not trained. A training trial was similar to an experimental trial so that it consisted of a relaxation period, followed by a facial activation task. A white ball displayed on a black background was controlled with the facial activation, instead of the computer character. Similar to actual trials, the participant was to bring the ball closer until the second tone was sounded. Arithmetic tasks were not further trained.

2.6 Procedure

First, the sound-attenuated and electro-magnetically shielded laboratory was introduced to the participant and a written informed consent obtained. A cover story told that there were three computer characters in the study. The purpose of the cover story was to provide a similar context for all subjective ratings by leading the participant to believe that the experiment would continue after the final ratings were collected. The SPIN-FIN retest was administered after the consent. Then, the participant was instructed to wash her or his hands and seated in front of the computer monitor. EMG and EDA electrodes were attached. Next, system calibration was performed. After calibration, the experimenter instructed the participant in how to

Table II. The experimental procedure for each participant consisted of system calibration followed by subjective ratings, training trials, and experimental trials. Different parts are shown in chronological order from left to right

group	id																		
				female character								male character							
LSA	1	C	R	T_C	E_C	R	T_Z	E_Z	R	E_N	R	T_C	E_C	R	T_Z	E_Z	R	E_N	R
LSA	2	C	R	T_Z	E_Z	R	T_C	E_C	R	E_N	R	T_Z	E_Z	R	T_C	E_C	R	E_N	R
									...										
LSA	6	C	R	E_N	R	T_Z	E_Z	R	T_C	E_C	R	E_N	R	T_Z	E_Z	R	T_C	E_C	R
HSA	7	C	R	T_C	E_C	R	T_Z	E_Z	R	E_N	R	T_C	E_C	R	T_Z	E_Z	R	E_N	R
									...										
HSA	12	C	R	E_N	R	T_Z	E_Z	R	T_C	E_C	R	E_N	R	T_Z	E_Z	R	T_C	E_C	R
				male character								female character							
LSA	13	C	R	T_C	E_C	R	T_Z	E_Z	R	E_N	R	T_C	E_C	R	T_Z	E_Z	R	E_N	R
									...										
LSA	18	C	R	E_N	R	T_Z	E_Z	R	T_C	E_C	R	E_N	R	T_Z	E_Z	R	T_C	E_C	R
HSA	19	C	R	T_C	E_C	R	T_Z	E_Z	R	E_N	R	T_C	E_C	R	T_Z	E_Z	R	E_N	R
									...										
HSA	24	C	R	E_N	R	T_Z	E_Z	R	T_C	E_C	R	E_N	R	T_Z	E_Z	R	T_C	E_C	R

id = participant's identifier, C = calibration, R = subjective ratings, T = two training trials, E = three experiment trials; Sub-indices: C = *corrugator supercilii* activation, Z = *zygomaticus major* activation, N = no facial activation

perform the arithmetic tasks. Two examples of a task and an answer were presented. Then, the participant was asked to answer two more tasks presented by the experimenter. The experimenter left the room before the experimental trials started.

The participants were to complete a total of six blocks each consisting of two training trials and three experimental trials using the same facial muscle. Blocks were counter-balanced within the LSA and HSA group similarly as illustrated in Table II. Half (6) of the participants in each group performed the tasks first with the female and then with the male character, while the other half performed tasks first with the male followed by the female character. A different permutation of the blocks (i.e., *corrugator supercilii*, *zygomaticus major*, and the block without facial activation) was assigned to each of the 6 participants within each group. After the completion of six blocks, the participant was debriefed.

Subjective ratings of emotional valence, arousal, and dominance were collected before each block and after the last block. Bipolar nine-point scales ranging from -4 (e.g., unpleasant) to +4 (e.g., pleasant) were used. The middle value of 0 represented a neutral rating (e.g., neither unpleasant nor pleasant) on each scale. Scales were administered on paper and participants were instructed to rate their own subjective experience. Extra rating scales were provided to the participant in order to support the cover story.

2.7 Data Reduction and Analysis

Participants whose SPIN-FIN re-test score did not meet cut-off criteria (11 and 12 in the LSA group, 12 and 21 in the HSA group) were excluded from further analysis. 1 female and 1 male participant was excluded from each group. Due to difficulties with EMG measurement, one participant in each group was not able to complete the *corrugator supercilii* block with the female agent.

Four different electrodermal measures of SCL, NS-SCR frequency, NS-SCR magnitude, and ER-SCR magnitude were extracted as follows. The base level of skin conductance (i.e., SCL) was calculated as the mean value of the EDA signal during the last 3 seconds of each relaxation period. The change in SCL from a relaxation period to the next was analyzed, that is, changes in SCL occurring during the facial activation were not analyzed. In order to detect skin conductance responses (i.e., SCR), a 62.5 ms moving average filter was applied to the EDA signal. Responses were scored automatically based on subsequent minima and maxima in the filtered signal. An amplitude change of over .02 μS was considered a response. This type of analysis is in effect similar to baseline correction of physiological data, as it excludes the base level of activity (i.e., SCL) and quantifies the magnitude of an individual response [Dawson et al. 2000]. The frequency of NS-SCRs was computed as the mean number of responses occurring during the facial activation task. The magnitude of NS-SCRs was also computed and analyzed. The ER-SCR elicited by the verbal arithmetic task was calculated as the mean magnitude of a detected response occurring within 1–5 s of the onset of speech synthesis (i.e., during speech synthesis).

Changes in subjective ratings were analyzed by computing the mean ratings on each scale and subtracting the pre-block rating from the post-block rating (i.e., using the pre-block rating as a baseline). Performance in the arithmetic tasks was scored based on the audio recordings. The answer was judged as incorrect, if either the repeated equation or the answer (i.e., correct or false) was incorrect. Performance in the facial task was calculated as the mean number of errors, that is, times when the EMG activity fell below the threshold, and the mean time to complete the task (i.e., time it took for the character to reach the closest distance). The two complementary performance measures (i.e., errors and time) were used in order to gain a more complete view of the objective functionality of facial activations as a method of voluntary control. It seems likely that difficulty in reaching the threshold in the first place would mainly extend the time required to complete the task, while difficulty in maintaining facial activations (i.e., unstable activations) should in addition lead to more errors in the task as EMG activity would repeatedly exceed and fall under the threshold.

2×3 (Group × Task) mixed-model Analyses of Variance (ANOVA) were performed for subjective ratings, electrodermal measures, and the mean number of correct answers to the arithmetic task. Task completion times did not vary in tasks without facial activation, as they were not dependent on participant's performance (i.e., no errors could be made in the task). These tasks were excluded from analysis of performance times and the number of errors. Thus, a 2×2 (Group × Task) ANOVA was performed for performance times and errors. For significant interactions, simple effects of Group were analyzed using between-subjects ANOVA. Pairwise comparisons were performed using Bonferroni corrected t-tests. Greenhouse-Geisser correction was applied to within-subject factors and the corresponding epsilon (ϵ) is reported. Reported effect size measures are the partial eta squared (η_p^2) for ANOVA and Cohen's d for t-tests.

3. RESULTS

3.1 Subjective Ratings

Figure 4 shows that in general the mean ratings of valence, arousal, and dominance were on the positive side of the scales. Thus, participants rated the situation as slightly pleasant, arousing, and being in their own control during the whole experiment session. However, the patterns of ratings were different between the LSA and HSA groups, especially for the *zygomaticus major* activation condition. After *zygomaticus major* trials, participants in the LSA group rated their subjective experience as more pleasantly valenced, less aroused, and more in control of the situation, while mean ratings of the HSA group show the opposite: decreased valence, increased arousal, and decreased subjective dominance.

The 2×3 ANOVA for the change in valence ratings did not show statistically significant effects of Group, $F(1{,}18) < 1$, or Task, $F(2{,}36) < 1$. The interaction effect of Task and Group was statistically significant, $F(2{,}36) = 4.47$, $p = .021$, $\eta_p^2 = .20$, $\epsilon = .94$. The simple effect of Group was statistically significant within the *zygomaticus major* task, $F(1{,}18) = 5.93$, $p = .025$, $\eta_p^2 = .25$. The post hoc comparison between the LSA and HSA groups within the *zygomaticus major* task was statistically significant, $t(18) = 2.44$, $p = .025$, $d = .54$. The simple effect of Group was not statistically significant within the *corrugator supercilii* task, $F(1{,}18) = 1.04$, $p = .322$, or the task without facial activation, $F(1{,}18) = 2.90$, p = .110.

The 2×3 ANOVA for the change in arousal ratings did not show statistically significant effects of Group, $F(1{,}18) = 1.13$, p = .303, or Task, $F(2{,}36) = 1.29$, p = .288. The interaction effect of Task and Group was statistically significant, $F(2{,}36) = 4.25$, $p = .023$, $\eta_p^2 = .19$, $\epsilon = .98$. The simple effect of Group was statistically significant within the *zygomaticus major* task, $F(1{,}18) = 9.60$, $p = .006$, $\eta_p^2 = .35$. The post hoc comparison be-

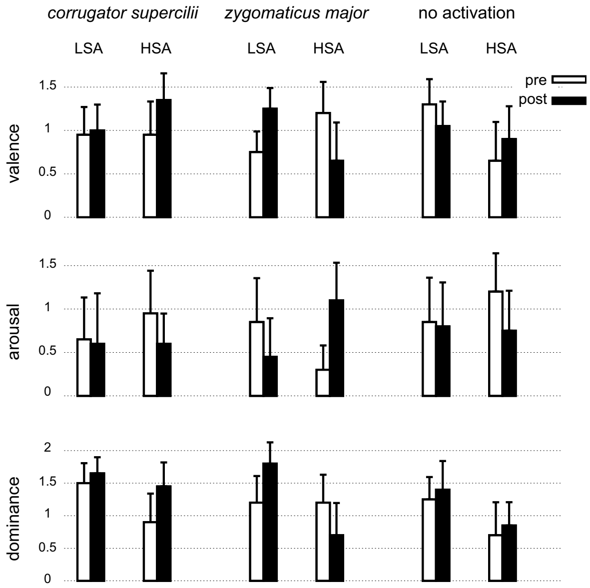


Fig. 4. Mean pre- and post-block ratings of valence, arousal, and dominance for different facial activation tasks and the LSA and HSA groups.

tween the LSA and HSA groups within the *zygomaticus major* task was statistically significant, $t(18) = 3.10$, $p = .006$, $d = .69$. The simple effect of Group was not statistically significant within the *corrugator supercilii* task, $F(1,18) < 1$, or the task without facial activation, $F(1,18) = 1.315$, p = .27.

The 2×3 ANOVA for the change in dominance ratings did not show statistically significant effects of Group, $F(1,18) = 2.10$, p = .164, or Task, $F(2,36) < 1$. The interaction effect of Task and Group was statistically significant, $F(2,36) = 3.72$, $p = .045$, $\eta_p^2 = .17$, $\epsilon = .80$. The simple effect of Group was statistically significant within the *zygomaticus major* task, $F(1,18) = 7.56, p = .013$, $\eta_p^2 = .30$. The post hoc comparison between the LSA and HSA groups within the *zygomaticus major* task was statistically significant, $t(18) = 2.75$, $p = .013$, $d = .61$. The simple effect of Group was not statistically significant within the *corrugator supercilii* task, $F(1,18) = 2.30$, $p = .146$, or the task without facial activation, $F(1,18) < 1$.

3.2 Performance Measures

Table III shows that there was little difference in the performance of the LSA and HSA groups in facial activation and arithmetic tasks. Both groups made more errors and the tasks took longer to complete when using the *zygomaticus major* as compared to the *corrugator supercilii* muscle. The verbal arithmetic task that followed the facial activations was performed with good accuracy overall, as the mean percentage of correct answers was above 90% in all conditions. Especially performance of the HSA group was very accurate (98.3% correct) following *corrugator supercilii* activations.

The 2 × 2 ANOVA for the number of errors in the facial activation task showed a statistically significant main effect of Task, $F(1,18) = 6.31$, p = .022, $\eta_p^2 = .26$. The main effect of Group, $F(1,22) < 1$, and the interaction effect of Group and Task, $F(1,18) < 1$, were not statistically significant. The post hoc comparison

Table III. Mean values ($\pm$ S.E.M.) of number of errors made during the facial activation tasks, the time to complete each task in seconds, and the percentage of correct answers to the verbal arithmetic task for different facial activation tasks and the LSA and HSA groups

	corrugator supercilii		*zygomaticus major*		no activation	
	LSA	HSA	LSA	HSA	LSA	HSA
errors	.13±.09	.42±.14	1.37±.65	1.33±.66		
task time	8.1±.1	8.3±.1	9.4±.7	9.0±.4	8.0±.0	8.0±.0
correct (%)	93.3±2.7	98.3±1.7	95.0±2.5	95.0±3.6	93.3±3.7	93.3±3.7

Table IV. Mean values ($\pm$ S.E.M.) of the change in SCL during a trial, the number of NS-SCRs during the task, the magnitude of NS-SCR during the task, and the magnitude of the ER-SCR to the verbal arithmetic task in μS for different facial activation tasks and the LSA and HSA groups

	corrugator supercilii		*zygomaticus major*		no activation	
	LSA	HSA	LSA	HSA	LSA	HSA
change in SCL	-7.9±1.5	-2.4±1.3	-4.8±1.2	-1.9±.9	-3.9±1.2	-1.8±.5
NS-SCR frequency	.9±.1	.9±.2	1.3±.3	1.1±.1	.9±.1	.9±.1
NS-SCR magnitude	6.9±1.0	4.7±1.0	7.5±1.2	5.0±1.2	5.2±.9	3.7±.7
ER-SCR magnitude	2.4±.4	2.2±.5	2.5±.4	2.2±.4	1.9±.3	1.2±.2

between the *corrugator supercilii* and the *zygomaticus major* tasks was statistically significant, $t(19) = 2.51$, $p = .022$, $d = .56$.

The 2 × 2 ANOVA for the task completion times showed a statistically significant main effect of Task, $F(1,18) = 6.74, p = .018, \eta_p^2 = .27$. The main effect of Group, $F(1,18) < 1$, and the interaction effect of Group and Task, $F(1,18) < 1$, were not statistically significant. The post hoc comparison between the *corrugator supercilii* and the *zygomaticus major* tasks was statistically significant, $t(19) = 2.60$, $p = .018$, $d = .58$.

The 2 × 3 ANOVA for the number of correct answers in the arithmetic task did not show statistically significant effects of Group, $F(1,18) < 1$, Task, $F(2,36) < 1$, or the interaction of Group and Task, $F(2,36) < 1$.

3.3 Electrodermal Measures

Table IV shows that the level of skin conductance decreased during all trials. The decrease was markedly greater within the LSA as compared to the HSA group during all tasks. Within both groups the *corrugator supercilii* task induced a greater decrease in SCL as compared to the other two tasks. The frequency of non-specific skin conductance responses was similar in both groups and for each task, although responses were slightly more frequent during *zygomaticus major* activations than during the other two tasks. Tasks involving facial activation induced greater skin conductance responses both during the activation (i.e., NS-SCR) and after the activation (i.e., ER-SCR to the verbal arithmetic task), as compared to the task without facial activation.

The 2×3 ANOVA for the SCL change showed statistically significant main effects of Group, $F(1,18) = 6.92$, $p = .017$, $\eta_p^2 = .28$, and Task, $F(2,36) = 4.37$, $p = .036$, $\eta_p^2 = .20$, $\epsilon = .69$. The interaction effect of Group and Task was not statistically significant, $F(2,36) = 2.41$, $p = .125$. A post hoc comparison confirmed that SCL decreased more within the LSA group than within the HSA group, $t(18) = 2.63$, $p = .017$, $d = .59$. Post hoc comparisons also showed that SCL decreased more during the *corrugator supercilii* trials as compared to trials without facial activation, $t(19) = 2.64$, $p = .050$, $d = .59$. Post hoc comparisons did not show significant differences between the *zygomaticus major* task as compared to the task without facial activation, $t(19) = .94, p = 1$, and as compared to the *corrugator supercilii* task, $t(19) = 1.81, p = .260$.

The 2 × 3 ANOVA for the frequency of NS-SCRs did not show statistically significant effects of Group, $F(1,18) < 1$, Task, $F(2,36) = 3.72$, $p = .051$, or the interaction of Group and Task, $F(2,36) < 1$.

The 2 × 3 ANOVA for the magnitude of NS-SCRs showed a statistically significant main effect of Task, $F(2{,}36) = 9.06, \mathrm{p} < .001, \eta_p^2 = .34,\ \epsilon = .96$. The main effect of Group, $F(2{,}36) = 2.34, \mathrm{p} = .144$, and the interaction effect of Group and Task, $F(2{,}36) < 1$, were not statistically significant. Post hoc comparisons showed that as compared to the task without facial activation, the magnitude of NS-SCRs was greater during the *corrugator supercilii* task, $t(19) = 2.73, p = .041, d = .61$, and during the *zygomaticus major* task, $t(19) = 4.15, p = .002, d = .93$. The magnitude of NS-SCR was not statistically significantly different between the *corrugator supercilii* and the *zygomaticus major* tasks, $t(19) = 1.26, p = .671$.

The 2 × 3 ANOVA for the magnitude of ER-SCRs showed a statistically significant main effect of Task, $F(2{,}36) = 7.99, \mathrm{p} = .003, \eta_p^2 = .31,\ \epsilon = .78$. The main effect of Group, $F(1{,}18) < 1$, and the interaction effect of Group and Task, $F(2{,}36) < 1$, were not statistically significant. Post hoc comparisons showed that as compared to the task without facial activation, the magnitude of ER-SCRs was greater after *corrugator supercilii* activation, $t(19) = 2.68, p = .046, d = .60$, and after the *zygomaticus major* activation, $t(19) = 4.82, p < .001, d = .1.08$. The magnitude of ER-SCR was not statistically significantly different between the *corrugator supercilii* and the *zygomaticus major* tasks, $t(19) = .38, p = 1$.

4. DISCUSSION

The present results showed that voluntary activations of facial muscles had significant effects on both physiological arousal and subjective ratings of emotional experiences. The mean level of skin conductance decreased after every task within both groups, which suggests that both the less and the more socially anxious participants were able to relax more as a trial progressed. However, the decrease in skin conductance was smaller within the socially anxious participants, which suggests that social anxiety restrained them from becoming similarly accustomed to the arousing situation as compared to the less socially anxious participants.

Within both groups, *corrugator supercilii* activations (i.e., frowning) were associated with a significantly higher decrease in the level of skin conductance (i.e., SCL) as compared to the task without facial activations. Although the mean level of SCL was decreased after every task, the decrease was both greater and happened faster (i.e., the task time was shorter) during *corrugator supercilii* activations as compared to *zygomaticus major* activations, suggesting a higher rate of decrease. These decreases in long-term SCL suggest that the general level of sympathetic arousal was reduced following tasks involving *corrugator supercilii* activations [Dawson et al. 2000]. However, as compared to the task without facial activation, both *corrugator supercilii* and *zygomaticus major* activations enhanced the magnitude of skin conductance responses that occurred either during the task (i.e., NS-SCR) or in response to the subsequent speech synthesis (i.e., ER-SCR). This suggests that, in the short term, muscle activations increased the level of autonomic arousal.

Subjective ratings of emotional experiences revealed interesting differences between the less and the more socially anxious participants. The ratings of the low social anxiety group increased in pleasantness after *zygomaticus major* activations (i.e., smiling), while participants in the high social anxiety group rated their experience as less pleasant. Socially anxious participants also rated their subjective experience as more aroused after *zygomaticus major* activations (i.e., smiling), while the mean ratings of less anxious participants showed a decrease in arousal. Further, the ratings of subjective dominance showed that *zygomaticus major* activations enhanced the feeling of subjective control for the less socially anxious participants, while the ratings of the more socially anxious participants showed a decrease in the feeling of subjective control. These results suggests that socially anxious participants did not feel as comfortable as the less socially anxious participants in smiling to the approaching virtual character.

In terms of the objective functionality of voluntary facial activations as a method for controlling the stimulation, all participants were able to complete the present tasks without notable difficulties and within reasonable time. However, the results suggest that voluntary *corrugator supercilii* activations were slightly better controlled than *zygomaticus major* activations in the present context. Both groups of participants (i.e., LSA and HSA) were able to maintain the required level of activation in the majority of *corrugator supercilii* tasks, whereas about one error per task was made on average during *zygomaticus major* activations. Thus, at least with relatively simple tasks and without a longer training for producing the controlled voluntary facial activations, both the less and the more socially anxious participants could control the virtual social stimulation slightly more accurately using *corrugator supercilii* activations.

The present ratings of *zygomaticus major* activity were in line with the behavioral tendencies associated with social anxiety. Social anxiety commonly leads to the avoidance of social interaction and the inhibition of pro-social behavior [Mick and Telch 1998; Connor et al. 2000]. Smiling, on the other hand, is a pro-social cue that may indicate that a person can be approached and affiliated with, that is, she or he is willing to participate in social interaction [Hess et al. 2004]. Further, a person with a high level of social anxiety is more likely to associate negative perceptions to social interactions, while *zygomaticus major* activity is typically associated with positive emotion [Levenson et al. 1990; Rapee and Heimberg 1997; Larsen et al. 2003; Partala et al. 2006]. Thus, it is quite natural that the more anxious participants were not as comfortable as the less anxious participants in smiling to the virtual character.

On the other hand, the mean changes in subjective ratings after *corrugator supercilii* activations suggest that especially the more socially anxious participants were more comfortable in frowning to the character. This pattern of increased pleasure, decreased arousal, and increased dominance may at first seem surprising from the point of view that emotions are fundamentally embodied phenomena and that bodily changes are tightly coupled with emotional experience [Coan et al. 2001; Levenson and Ekman 2002; Niedenthal 2007]. It could be argued that as spontaneous *corrugator supercilii* activity is associated with negative emotion, also voluntary activity of the same muscle should induce a similar physiological state and thus lead to increasingly negative ratings. However, like other channels of expression, facial activity serves also social functions (e.g., conversational signaling) and is neither exclusively emotional nor a direct reflection of emotional state as such [Ekman 1979; Mauss and Robinson 2009]. In particular, *corrugator supercilii* activity has also been associated with both the effort of performing a task and dominating a social interaction [Van Boxtel and Jessurun 1993; Waterink and Van Boxtel 1994; Carroll and Russell 1997; Hietanen et al. 1998]. In the present setup, the activation of *corrugator supercilii* muscle may have matched particularly well with the participants' task of taking voluntary control over (i.e., more or less dominating) a virtual character, which may have led to a more comfortable subjective experience.

It could be desirable that a socially anxious person would eventually learn to use and perceive also *zygomaticus major* activations (i.e., smiling) as a comfortable and natural part of social situations, that is, rate their own experience after these activations similarly as the less socially anxious participants: more pleasantly valenced, less aroused, and being more in their own control. Voluntary facial activations could have a significant role in achieving this as a part of exposure treatment, where a person is gradually exposed and desensitized to increasing levels of distressing stimulation, for example, by reducing the distance to a real or a virtual person [Foa and Kozak 1986; Öst 1989; Gerardi et al. 2010]. First, during continued exposure to distressing stimulation (e.g., an approaching virtual human), both voluntary *corrugator supercilii* and *zygomaticus major* activations could provide accurate control over the virtual stimulation during exposure treatment. Then, through continuous training, the responses to voluntary *zygomaticus major* activations could be modified from negative (i.e., decreased pleasantness) to positive (i.e., increased pleasantness). This could finally lead to the modification of subjective and behavioral tendencies that upkeep excessive anxiety.

Physiological changes associated with voluntary facial activations could further facilitate these changes in anxiety responses. The results showed that voluntary *corrugator supercilii* and *zygomaticus major* activations were both effective in inducing physiological arousal (i.e., sympathetic activation) during exposure to the virtual social stimulation. In general, physiological activation is a central factor in desensitization and habituation of fear [Foa and Kozak 1986]. It is hypothesized that increased physiological activity correlates with the activation of anxiety-relevant cognitive-emotional fear structures that can then be modified. For example, there is evidence that phobic persons who react more strongly in terms of physiological changes are also those who benefit most from exposure to their objects of fear [Kozak et al. 1988]. Thus, in terms of applying the present results, voluntarily induced (e.g., by facial activations) increases in the level of physiological activity could provide support for the activation and habituation of fear during exposure.

Further, the skin conductance response to the arithmetic task was also enhanced after voluntary facial activations as compared to tasks without facial activation, which suggests that both types of facial activations induced a stronger orientation to the task [Frith and Allen 1983]. This is also in line with both types of facial activations leading to slightly higher mean percentages of correct answers to the arithmetic task as compared to tasks without facial activation. This suggests that participants focused more on the synthetic speech stimulation after voluntary facial activations. Especially participants in the HSA group gave very accurate answers following *corrugator supercilii* activations with over 98 % of correct answers. In terms of exposure treatment, higher level of attention to anxiety-relevant stimulation has been found to facilitate the habituation of fear responses [Foa and Kozak 1986]. Thus, the present results suggest that voluntary facial activations could be used to facilitate behavioral and physiological changes that support the outcome of exposure treatments.

The present study can offer only limited insight into the mechanisms behind the physiological and subjective effects of voluntary facial activations. However, the present results do suggest that the social aspects of facial expression were a significant factor influencing subjective responses to voluntary *zygomaticus major* activations, as the ratings of the less and the more social anxious participants showed completely opposite patterns. On the other hand, we did not observe any significant differences between the two muscles or the LSA and HSA groups in respect to the frequency or magnitude skin conductance responses during or shortly after (i.e., during synthesized speech) the voluntary activations. Thus, the present work did not provide evidence for physiological patterning (i.e., "embodied emotion") underlying the emotional response to voluntary activations, that is, no direct coupling of physiology and experience. However, it is possible that facial muscle specific physiological response patterns could have been observed with other complementary physiological measures. For example, while the presently used EDA measures are strongly associated with arousal, heart rate is known to correlate well with the subjective experience of emotional valence (e.g., negative or positive) [Bradley 2000; Bradley et al. 2001; Mauss and Robinson 2009]. A more detailed analysis of multi-component physiological responses to voluntary activations of individual facial muscles could be a promising direction for future work.

On the other hand, the present work can already provide a basis for more straight-forward interpretation and application of the results. First, the present setup showed that controlling a virtual character with voluntary facial expressions is feasible and functional with relatively little preparation, including very brief training and instructions for the person. Thus, similar setup would be practical to use, for example, in a clinical setting. Second, by using voluntary facial activations to control the stimulation, participants were able to significantly affect their more spontaneous physiological activity (i.e., sympathetic ANS arousal) without explicit biofeedback. As the present setup couples feedback about facial activity with changes in the stimulation itself, this kind of system would encourage a person to focus on the anxiety relevant stimulation, supporting the goals of exposure treatment. Third, the present setup provides a platform that could be used as such to implement an exposure session which could consist of similar tasks as used in the present study. It could be expected that continuing exposure to the arousing tasks would eventually lead to the habituation of responses and arousal. This is in line with the current finding that the level of skin conductance decreased, indicating reduction of autonomic arousal, as a trial progressed (i.e., more tasks were performed).

In general, exposing a socially anxious person to a distressing (i.e., negatively arousing) setup where they voluntarily smile (i.e., activate the *zygomaticus major* muscle) to bring a virtual character closer, would be in line with the method of traditional exposure treatment of anxiety disorders. However, although the present setup was effective in producing significant subjective and physiological effects, there are several factors that may have contributed to the results. For example, the role of (the participant's perception of) the virtual characters and synthetic speech could be significant and should be investigated in order to optimize the effectiveness of stimulation (e.g., design virtual characters that specifically induce social anxiety). Further, the functionality of such a setup in inducing and reducing clinical levels of social anxiety is to be confirmed by further study, in particular due to the small sub-clinical (i.e., not clinically diagnosed) sample used in the present study as well as its short-term nature. For example, in order to generalize the effects of virtual stimulation to more complex real scenarios, it could be beneficial to create more wide variety of tasks and stimulation that are administered over several sessions.

The present setup was designed to be quite strictly controlled and simple in terms of the social communication, that is, the virtual character could only approach when a certain facial muscle was activated. However, real-life interactions involve other people whose reactions cannot be controlled and always predicted. Thus, a socially anxious person who progresses to the level of learning to deal with everyday human-to-human interaction could benefit from a richer virtual environment. Fuller social scenarios could be constructed by using one or several virtual characters that react more expressively to a person's behavior, for example, by having virtual audience members that lean forward smiling or withdraw frowning when the person speaks [Pertaub et al. 2002].

There are also several other parameters that could be modified in real time due to the flexibility of computer-generated stimulation. For example, a more complete 3D environment could be gradually introduced to increase the feeling of presence, that is, the experience of sharing a real physical space with the virtual characters [Schuemie et al. 2001]. Such an environment could also facilitate the perception of distance to the virtual characters by enabling more sophisticated cues of depth (e.g., stereoscopy) to be used. High-end interaction techniques could also be used to facilitate the realism of the setup. For example, the point-of-view was currently static and somewhat unrealistic, as the character was standing but appeared to approach the seated participant at the eye level. The use of a head-mounted display with movement tracking could have enabled each participant to look around using a realistic point-of-view matching their own head movement. In general, these kinds of richer technologies would offer more versatility for virtual exposure treatment and we intend to incorporate them to our future work.

5. CONCLUSIONS

The present results showed that voluntary facial activations could provide a method for regulating physiological and subjective arousal during exposure to artificial social stimuli. Especially *corrugator supercilii* activations were found to facilitate the long-term relaxation of physiological arousal (i.e., SCL) and they were also relatively well controlled in the present context. On the other hand, the subjective ratings of emotional experiences following *zygomaticus major* suggested that the more socially anxious participants were less comfortable in smiling to the virtual character. Detecting such informative response patterns and monitoring their change could be useful for assessing social anxiety, for example, in order to follow the progress of exposure therapy. Further, both types of facial muscle activations produced short-term physiological changes (i.e., enhanced magnitude of NS-SCR and ER-SCR) that were compatible with and potentially beneficial for the aims of exposure treatment. In summary, the present results provide several promising directions for research and form a solid basis for continuing our work in studying voluntary facial activations as a method for computer-assisted regulation of emotions.

ACKNOWLEDGMENTS

We thank Dr. Klaus Ranta for the permission to use the SPIN-FIN questionnaire.

REFERENCES

Anderson, P., Rothbaum, B. O., and Hodges, L. F. 2003. Virtual reality exposure in the treatment of social anxiety. *Cognitive and Behavioral Practice 10*, 3, 240 – 247.

BEALE, R. AND CREED, C. 2009. Affective interaction: How emotional agents affect users. *International Journal of Human-Computer Studies 67*, 755 – 776.

BECK, J. G., PALYO, S. A., WINER, E. H., SCHWAGLER, B. E., AND ANG, E. J. 2007. Virtual reality exposure therapy for PTSD symptoms after a road accident: An uncontrolled case series. *Behavior Therapy 38,* 1, 39–48.

BOITEN, F. 1996. Autonomic response patterns during voluntary facial action. *Psychophysiology 33*, 123–131.

BRADLEY, M. M. 2000. *Handbook of Psychophysiology* 2nd Ed. Cambridge University Press, Cambridge, UK, Chapter Emotion and Motivation, 602–642.

BRADLEY, M. M., CODISPOTI, M., CUTHBERT, B. N., AND LANG, P. J. 2001. Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion 1,* 3, 276–298.

BRADLEY, M. M. AND LANG, P. J. 1994. Measuring emotions: the self-assessment manikin and the semantic differential. *Journal of Behavioral Therapy and Experimental Psychiatry 25,* 1, 49–59.

CACIOPPO, J. T., BERNTSON, G. G., LARSEN, J. T., POEHLMANN, K. M., AND ITO, T. A. 2000. *The Handbook of Emotion.* Guildford Press, New York, Chapter The psychophysiology of emotion, 173–191.

CARROLL, J. M. AND RUSSELL, J. A. 1997. Facial expressions in Hollywood's portrayal of emotion. *Journal of Personality and Social Psychology 72,* 164–176.

CHRISTIE, I. C. AND FRIEDMAN, B. H. 2004. Autonomic specificity of discrete emotion and dimensions of affective space: a multivariate approach. *International Journal of Psychophysiology 51,* 2, 143–153.

COAN, J. A., ALLEN, J. J. B., AND HARMON-JONES, E. 2001. Voluntary facial expression and hemispheric asymmetry over the frontal cortex. *Psychophysiology 38*, 912–925.

CODISPOTI, M. AND DE CESAREI, A. 2007. Arousal and attention: Picture size and emotional reactions. *Psychophysiology 44*, 680–686.

CONNOR, K. M., DAVIDSON, J. R. T., CHURCHILL, L. E., SHERWOOD, A., FOA, E., AND WEISLER, R. H. 2000. Psychometric properties of the social phobia inventory (SPIN). *British Journal of Psychiatry 176*, 379–386.

COURGEON, M., BUISINE, S., AND MARTIN, J.-C. 2009. Impact of expressive wrinkles on perception of a virtual characters facial expressions of emotions. In *Intelligent Virtual Agents*, Lecture Notes in Computer Science, vol. 5773. Springer, 201–214.

COURGEON, M., MARTIN, J.-C., AND JACQUEMIN, C. 2008. MARC: A multimodal affective and reactive character. In *Proceedings of the ICMI-MLMI 2008 Workshop on Affective Interaction in Natural Environments.*

COURTNEY, C. G., DAWSON, M. E., SCHELL, A. M., IYER, A., AND PARSONS, T. D. 2010. Better than the real thing: Eliciting fear with moving and static computer-generated stimuli. *International Journal of Psychophysiology 78,* 2, 107 – 114.

DAWSON, M. E., SCHELL, A. M., AND FILION, D. L. 2000. *Handbook of Psychophysiology.* Cambridge University Press, Cambridge, UK, Chapter The electrodermal system, 200–223.

EKMAN, P. 1979. *Human Ethology.* Cambridge University Press, Cambridge, UK, Chapter About brows: emotional and conversational signals.

EKMAN, P. 1993. An argument for basic emotions. *Cognition and Emotion 6,* 3, 169–200.

EKMAN, P. 2004. *Language, Knowledge, and Representation.* Kluwer, Netherlands, Chapter Emotional and conversational nonverbal signals, 39–50.

FOA, E. B. AND KOZAK, M. J. 1986. Emotional processing of fear: Exposure to corrective information. *Psychological Bulletin 99,* 1, 20–35.

FOWLES, D. C. 1988. Psychophysiology and psychopathology: A motivational approach. *Psychophysiology 25*, 373–391.

FRIDLUND, A. J. AND CACIOPPO, J. T. 1986. Guidelines for human electromyographic research. *Psychophysiology 23*, 567–589.

FRITH, C. D. AND ALLEN, H. A. 1983. The skin conductance orienting response as an index of attention. *Biological Psychology 17,* 1, 27 – 39.

GARCIA-PALACIOS, A., BOTELLA, C., HOFFMAN., H. G., AND FABREGAT, S. 2007. Comparing acceptance and refusal rates of virtual reality exposure vs. in vivo exposure by patients with specific phobias *CyberPsychology & Behavior 10,* 5, 722–724.

GATCHEL, R. J. AND PROCTOR, J. D. 1976. Effectiveness of voluntary heart rate control in reducing speech anxiety. *Journal of Consulting and Clinical Psychology 44,* 3, 381–389.

GEARY, D. C., FRENSCH, P. A., AND WILEY, J. G. 1993. Simple and complex mental subtraction: Strategy choice and speed-of-processing differences in younger and older adults. *Psychology and Aging 8*, 2, 242 – 256.

GERARDI, M., CUKOR, J., DIFEDE, J., RIZZO, A., AND ROTHBAUM, B. O. 2010. Virtual reality exposure therapy for post-traumatic stress disorder and other anxiety disorders. *Current Psychiatry Reports 12,* 4, 298–305.

GRAMER, M. AND SARIA, K. 2007. Effects of social anxiety and evaluative threat on cardiovascular responses to active performance situations. *Biological Psychology 74,* 1, 67–74.

Grayson, J., Foa, E., and Steketee, G. 1982. Habituation during exposure treatment: Distraction vs attention-focusing. *Behaviour Research and Therapy 20,* 4, 323–328.

Harris, S. R., Kemmerling, R. L., and North, M. M. 2002. Brief virtual reality therapy for public speaking anxiety. *CyberPsychology & Behavior 5,* 6, 543–550.

Heimberg, R. G., Dodge, C. S., Hope, D. A., Kennedy, C. R., Zollo, L. J., and Becker, R. E. 1990. Cognitive behavioral treatment of social phobia: comparison to a credible placebo control. *Cognitive Therapy and Research 14*, 1–23.

Hellström, K. and Öst, L.-G. 1995. One-session therapist directed exposure vs two forms of manual directed self-exposure in the treatment of spider phobia. *Behaviour Research and Therapy 33,* 8, 959 – 965.

Hess, U., Adams Jr., R. B., and Kleck, R. E. 2004. Facial appearance, gender, and emotion expression. *Emotion 4,* 4, 378–388.

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

James, L., Lin, C., Steed, A., Swapp, D., and Slater, M. 2003. Social anxiety in virtual environments: Results of a pilot study. *CyberPsychology & Behavior 6,* 3, 237–243.

Josman, N., Reisberg, A., Weiss, P. L., Garcia-Palacios, A., and Hoffman., H. G. 2008. Busworld: An analog pilot test of a virtual environment designed to treat posttraumatic stress disorder originating from a terrorist suicide bomb attack. *CyberPsychology & Behavior 11,* 6, 775–777.

Kessler, R. C. 2003. The impairments caused by social phobia in the general population: implications for intervention. *Acta Psychiatrica Scandinavica 108*, 19–27.

Kozak, M. J., Foa, E. B., and Steketee, G. 1988. Process and outcome of exposure treatment with obsessive-compulsives: Psychophysiological indicators of emotional processing. *Behavior Therapy 19,* 2, 157 – 169.

Krijn, M., Emmelkamp, P., Olafsson, R., and Biemond, R. 2004. Virtual reality exposure therapy of anxiety disorders: A review. *Clinical Psychology Review 24*, 259–281.

Larkin, K. T., Ciano-Federoff, L. M., and Hammel, D. 1998. Effects of gender of observer and fear of negative evaluation on cardiovascular reactivity to mental stress in college men. *International Journal of Psychophysiology 29*, 311–318.

Larsen, J. T., Norris, C. J., and Cacioppo, J. T. 2003. Effects of positive and negative affect on electromyographic activity over the zygomaticus major and corrugator supercilii. *Psychophysiology 40*, 776–785.

Levenson, R. W. and Ekman, P. 2002. Difficulty does not account for emotion-specific heart rate changes in the directed facial action task. *Psychophysiology 39*, 397–405.

Levenson, R. W., Ekman, P., Heider, K., and Friesen, W. V. 1992. Emotion and autonomic nervous system activity in the Minangkabau of West Sumatra. *Journal of Personality and Social Psychology 62*, 972988.

Levenson, R. W., Ekman, P., and W. V. Friesen, W. V. 1990. Voluntary facial action generates emotion-specific autonomic nervous system activity. *Psychophysiology 27*, 363384.

Llobera, J., Spanlang, B., Ruffini, G., and Slater, M. 2010. Proxemics with multiple dynamic characters in an immersive virtual environment. *ACM Transactions on Applied Perception 8*, 1–12.

Loftus, G. R. and Harley, E. M. 2005. Why is it easier to identify someone close than far away? *Psychonomic Bulletin & Review 12*, 43–65.

Mauss, I. B., Levenson, R. W., McCarter, L., Wilhelm, F. H., and Gross, J. J. 2005. The tie that binds? coherence among emotion experience, behavior, and physiology. *Emotion 5,* 2, 175–190.

Mauss, I. B. and Robinson, M. D. 2009. Measures of emotion: A review. *Cognition and Emotion 23,* 2, 209–237.

Mick, M. A. and Telch, M. J. 1998. Social anxiety and history of behavioral inhibition in young adults. *Journal of Anxiety Disorders 12,* 1, 1 – 20.

Nass, S., Steuer, J., and Tauber, E. R. 1994. Computers are social actors. In *Proceedings of CHI'94*. 72–78.

Niedenthal, P. M. 2007. Embodying emotion. *Science 316*, 1002–1005.

North, M., North, S., and Coble, J. 1998. Virtual reality therapy: an effective treatment for phobias. *Studies in Health Technology and Informatics 58*, 112–119.

Olfson, M., Guardino, M., Struening, E., Schneier, F. R., Hellman, F., and Klein, D. F. 2000. Barriers to the treatment of social anxiety. *American Journal of Psychiatry 157*, 521527.

Öst, L.-G. 1989. One-session treatment for specific phobias. *Behaviour Research and Therapy 27,* 1, 1 – 7.

Partala, T., Surakka, V., and Lahti, J. 2004. Affective effects of agent proximity in conversational systems. In *Proceedings of NordiCHI 2004*. 353–356.

Partala, T., Surakka, V., and Vanhala, T. 2006. Real-time estimation of emotional experiences from facial expressions. *Interacting with Computers 18*, 208–226.

Pertaub, D.-P., Slater, M., and Barker, C. 2002. An experiment on public speaking anxiety in response to three different types of virtual audience. *Presence: Teleoperators & Virtual Environments 11,* 1, 68–78.

Powers, M. B. and Emmelkamp, P. M. G. 2008. Virtual reality exposure therapy for anxiety disorders: A meta-analysis. *Journal of Anxiety Disorders 22,* 3, 561–569.

Rainville, P., Bechara, A., Naqvi, N., and Damasio, A. R. 2006. Basic emotions are associated with distinct patterns of cardiorespiratory activity. *International Journal of Psychophysiology 61*, 5–18.

Ranta, K., Kaltiala-Heino, R., Rantanen, P., Tuomisto, M. T., and Marttunen, M. 2007. Screening social phobia in adolescents from general population: The validity of the social phobia inventory (spin) against a clinical interview. *European Psychiatry 22,* 4, 244 – 251.

Rapee, R. M. and Heimberg, R. G. 1997. A cognitive-behavioral model of anxiety in social phobia. *Behaviour Research and Therapy 35,* 8, 741 – 756.

Reeves, B., Lang, A., Kim, E. Y., and Tatar, D. 1999. The effects of screen size and message content on attention and arousal. *Media Psychology 1*, 49–67.

Rinck, M., Rörtgen, T., Lange, W.-G., Dotsch, R., Wigboldus, D. H. J., and Becker, E. S. 2010. Social anxiety predicts avoidance behaviour in virtual encounters. *Cognition and Emotion 24,* 7, 1269–1276.

Rothbaum, B. O., Hodges, L., Smith, S., Lee, J. H., and Price, L. 2000. A controlled study of virtual reality exposure therapy for the fear of flying. *Journal of Consulting and Clinical Psychology 68,* 6, 1020–1026.

Schlosberg, H. 1954. Three dimensions of emotion. *Psychological Review 61,* 2, 81–88.

Schuemie, M. J., van der Straaten, P., Krijn, M., and van der Mast, C. A. P. G. 2001. Research on presence in virtual reality: a survey. *CyberPsychology & Behavior 4,* 2, 183–201.

Tassinary, L. G. and Cacioppo, J. T. 2000. *Handbook of Psychophysiology.* Cambridge University Press, Cambridge, UK, Chapter The skeletomotor system: Surface electromyography, 163–199.

Teghtsoonian, R. and Frost, R. O. 1982. The effects of viewing distance on fear of snakes. *Journal of Behavior Therapy and Experimental Psychiatry 13,* 3, 181–190.

Van Boxtel, A. and Jessurun, M. 1993. Amplitude and bilateral coherency of facial and jaw-elevator EMG activity as an index of effort during a two-choice serial reaction task. *Psychophysiology 30,* 6, 589–604.

Vanhala, T. and Surakka, V. 2007. Facial activation control effect (FACE). In *Affective Computing and Intelligent Interaction*, A. Paiva, R. Prada, and R. W. Picard, Eds. Springer-Verlag, 278–289.

Vanhala, T. and Surakka, V. 2008. *Affective Computing.* InTech Education and Publishing, Vienna, Austria, Chapter Computer-Assisted Regulation of Emotional and Social Processes, 405–420.

Vanhala, T., Surakka, V., Siirtola, H., Räihä, K.-J., Morel, B., and Ach, L. 2010. Virtual proximity and facial expressions of computer agents regulate human emotions and attention. *Computer Animation and Virtual Worlds 21,* 3-4, 215–224.

Ward, R. D. and Marsden, P. H. 2003. Physiological responses to different web page designs. *International Journal of Human-Computer Studies 59*, 199–212.

Waterink, W. and Van Boxtel, A. 1994. Facial and jaw-elevator EMG activity in relation to changes in performance level during a sustained information processing task. *Biological Psychology 37*, 183–198.

Wiederhold, B. K. and Rizzo, A. 2005. Virtual reality and applied psychophysiology. *Applied Psychophysiology and Biofeedback 30,* 3, 183–185.

Wiederhold, B. K. and Wiederhold, M. D. 2003. Three-year follow-up for virtual reality exposure for fear of flying. *CyberPsychology and Behavior 6*, 441–445.

Witvliet, C. V. and Vrana, S. R. 1995. Psychophysiological responses as indices of affective dimensions. *Psychophysiology 5*, 436–443.

Zanbaka, C. A., Ulinski, A. C., Goolkasian, P., and Hodges, L. F. 2007. Social responses to virtual humans: implications for future interface design. In *Proceedings of CHI'07*, 1561–1570.

Received November 2010; revised May 2011; accepted May 2011

Publications in the **Dissertations in Interactive Technology** series

1. **Timo Partala:** Affective Information in Human-Computer Interaction
2. **Mika Käki:** Enhancing Web Search Result Access with Automatic Categorization
3. **Anne Aula:** Studying User Strategies and Characteristics for Developing Web Search Interfaces
4. **Aulikki Hyrskykari:** Eyes in Attentive Interfaces: Experiences from Creating iDict, a Gaze-Aware Reading Aid
5. **Johanna Höysniemi:** Design and Evaluation of Physically Interactive Games
6. **Jaakko Hakulinen:** Software Tutoring in Speech User Interfaces
7. **Harri Siirtola:** Interactive Visualization of Multidimensional Data
8. **Erno Mäkinen:** Face Analysis Techniques for Human-Computer Interaction
9. **Oleg Špakov:** iComponent — Device-Independent Platform for Analyzing Eye Movement Data and Developing Eye-Based Applications
10. **Yulia Gizatdinova:** Automatic Detection of Face and Facial Features from Images of Neutral and Expressive Faces
11. **Päivi Majaranta:** Text Entry by Eye Gaze
12. **Ying Liu:** Chinese Text Entry with Mobile Devices
13. **Toni Vanhala:** Towards Computer-Assisted Regulation of Emotions