Addressing Emotions within E-Learning Systems

ZURLONI, Valentino, et al.

Abstract

Emotions are attracting growing attention within the instructional design research community. However, clarification is still required as to how exactly to address emotions within the field of e-learning. The aim of this chapter is twofold. Firstly, we will focus on the reasons for including emotions within the instructional technology domain, and in particular, on the relevance of emotions to computer-based learning. The need for specific theory in this regard is heightened by the current drive to design instructional devices that interact with learners in a motivating, engaging, and helpful way. Secondly, within the framework affective computing paradigm, the different modalities for detecting emotions in instructional technology contexts will be systematically reviewed, and the strengths and limits of each will be discussed on the basis of the most up to date research outcomes. Finally, a tentative architecture for emotion recognition in computer-based learning will be proposed, focusing on the adoption of a multimodal approach to emotion recognition, in order to overcome the limitations and the difficulties associated [...]
Chapter LVII
Addressing Emotions within E-Learning Systems

Valentino Zurloni
CESCOM, University of Milan - Bicocca, Italy

Fabrizia Mantovani
CESCOM, University of Milan - Bicocca, Italy
ATN-P LAB, Istituto Auxologico Italiano, Italy

Marcello Mortillaro
CESCOM, University of Milan - Bicocca, Italy
CISA - University of Geneva, Switzerland

Antonietta Vescovo
CESCOM, University of Milan - Bicocca, Italy

Luigi Anolli
CESCOM, University of Milan - Bicocca, Italy

ABSTRACT

Emotions are attracting growing attention within the instructional design research community. However, clarification is still required as to how exactly to address emotions within the field of e-learning. The aim of this chapter is twofold. Firstly, we will focus on the reasons for including emotions within the instructional technology domain, and in particular, on the relevance of emotions to computer-based learning. The need for specific theory in this regard is heightened by the current drive to design instructional devices that interact with learners in a motivating, engaging, and helpful way. Secondly, within the framework affective computing paradigm, the different modalities for detecting emotions in instructional technology contexts will be systematically reviewed, and the strengths and limits of each will be discussed on the basis of the most up-to-date research outcomes. Finally, a tentative architecture for emotion recognition in computer-based learning will be proposed, focusing on the adoption of a multimodal approach to emotion recognition, in order to overcome the limitations and the difficulties associated with individual modalities.
INTRODUCTION

The Role of Emotions in Learning

There is growing recognition that emotions and affect play an important role in learning. The learning process is influenced by factors connected to the person, the task, and the context as well as the learner’s own on-going evaluation of the process itself. Situational characteristics and individual appraisals can trigger emotions (Efklides & Volet, 2005). In turn, as stated by Barrett and Salovey (2002), affect in learning facilitates the development of persistence and interest in a topic. Emotions can also influence learning through information processing activity and organization of recall (Pekrun, Goetz, Titz, & Perry, 2002). Furthermore, emotions can provide information about the learner’s own evaluation of the learning process, since they are linked to control- and value-related appraisals within a learning environment (Gläser-Zikuda & Mayring, 2003). For instance, positive emotions generally indicate that successful task control and interest have been experienced. Our learning, therefore, is heavily dependent on the emotional state we are in (LeDoux, 1998), and on the dynamic pattern of positive and negative emotions occurring in a given time period within a learning context (Sansone & Thoman, 2005).

The role of emotions can be relatively easily recognized and managed within face-to-face learning, where they have been shown to be significantly related to student motivation, learning strategies, cognitive resources, and achievement (Pekrun et al., 2002). What is worth considering is the role of emotions when students are remote from their teacher—even when computer-based education can be supported by a human tutor, the latter is likely to have a lesser awareness of the emotional state of students, and may thus more easily fail to provide a responsible teaching presence and appropriate leadership and direction (Wosnitza & Volet, 2005). Despite general awareness of the need to consider emotions in e-learning environments, it seems that, with the exception of computer anxiety, the emotions experienced during computer-based learning have not yet been analyzed in depth (Pekrun, 2005). Thus, there is a great need for e-learning projects to take the role of emotions in learning into account and to integrate this understanding into their pedagogical approach.

Affect and Emotions in E-Learning Design

Very often, e-learning implies the presentation of information and material on a very rational basis, overlooking the role of emotions. Yet, computer-based learning can be affected by a range of emotions, including some which do not occur within face-to-face learning, such as emotions directed at technology. Nowadays almost all Web-based training platforms allow computer-mediated communication, where e-learning can take place within either solo or social situations. In solo learning, self-directed, task-directed, and technology-directed emotions have been identified. In social online learning, further emotions have been observed, such as emotions directed at another learner, at the group the learner belongs to, or at another group of learners that his/her group is interacting with (Wosnitza & Volet, 2005).

Moreover, as O’Regan (2003) has pointed out, there has been little exploration to date of the extent, nature, and significance of affect and emotions in e-learning design. If emotions are essential to human thinking and learning processes, virtual platforms and learning environments need to cater to the emotional factor in order to be successful. In particular, the computer graphical interface should not treat humans like information processing machines, but should take their
emotions into account. Therefore, it is critical that system designers consider the range of possible affective states that users may experience while interacting with the system.

The key problematic issue remains how to address emotions in human-computer interaction: This topic is currently the focus of a specific research field defined as affective computing (Picard, 1997). All studies with the aim of including emotions within information technology design can be ascribed to this domain, which deals specifically with three different levels of emotion integration: the detection of user emotions, the expression of emotions by computers, and ultimately, the possibility for a computer to “have” emotions. All three levels of emotion integration can greatly enhance the interaction of users with computer programs. Thus, software can nowadays be provided with a kind of emotional intelligence, conceived of as one of the most critical characteristics for successful human interaction (Salovey & Mayer, 1990). The current chapter will mainly focus on the issue of detection.

With specific reference to educational research, according to Picard et al. (2004), new technologies can play a double role: On the one hand, they can help provide new types of research data on the role of affect in learning, laying the bases for new approaches to education, and on the other, can provide enhanced computer-based learning environments to support the user more effectively through his/her learning process.

This chapter introduces the preliminary questions needing to be addressed in order to carry out emotion assessment within learning environments. Such questions highlight the need for wider use of recognition instruments. Subsequently, after reviewing a number of possible channels of emotion detection in e-learning applications, a multimodal approach is recommended in order to increase reliability of recognition outcomes. On the basis of this multimodal approach, a tentative architecture for automatic emotion recognition is outlined. Inclusion of a cognitive component is suggested, in order to integrate the outcomes from different channels, as well as the registration of meaningful events experienced by the user during the learning experience.

**DETECTING EMOTIONS WITHIN INSTRUCTIONAL TECHNOLOGY DOMAIN**

The first step towards the integration of emotional intelligence is the ability to detect users’ emotions. Two main emotion detection approaches can be identified within e-learning environments—that carried out by a human tutor, or teacher, participating in the learning environment, and that carried out by the e-learning system itself, termed automatic. The first is largely similar to the emotion recognition taking place in traditional learning environments and may also be used in computer supported learning, although the human tutor may have reduced access to emotional information. The second demands a more defined and specific basis and is more suitable for autonomous e-learning systems.

When addressing automatic emotion recognition, researchers must answer two main questions: First, when in the course of the learning interaction the system should be emotionally aware, and second, how it should recognize emotions, that is, based on which signals.

With regard to the timing issue, emotion recognition processes may be activated in three different periods vis-à-vis the learning path: (1) measurement immediately before and/or after the learning process, (2) measurement during the learning process, and (3) stimulated recall measurement of emotions after the learning process (Wosnitza & Volet, 2005).
Measuring emotions just before or after the learning process, on the one hand does not interrupt the learning experience but, on the other, requires specific pre-scripted moments to be integrated that may irritate the user and may lose efficacy due to the delay between the emotional experience and the emotional assessment. The delay factor also impairs the effectiveness of the “emotional coping” that the system should ideally carry out. Similar, but amplified, drawbacks apply to stimulated recall measurement.

Emotion recognition during the learning process may be a better solution. This involves real-time emotion assessment, with the potential to be time effective when an emotionally relevant modification occurs. On the other hand, real-time emotion automatic recognition faces greater technical challenges than the other possible timing options, and furthermore needs to be as truly automatic as possible so as not to disrupt the learning process. In consequence, a balanced recommendation is for real-time emotion recognition to take place in specific and critical learning moments.

In regards on how to detect emotions, a number of channels are described in the next paragraph. Before outlining individual modalities it is worth mentioning that considering emotion to be a componential process (Scherer, 1984, 2001) enables simultaneous inclusion of many different measures (Pantic & Rothkrantz, 2003). This approach enhances the likelihood of recognizing an emotional state, and partially reduces the risk of distortion due to factors such as social desirability and deliberate control of emotional expression (Pekrun, 2005). Each channel presents its own limitations, but these may be partially overcome by the adoption of a multimodal approach to emotion recognition, which integrates information from multiple sources and may be regulated depending on the context of use.

A Multimodal Approach to Emotion Recognition

Providing e-learning systems with the ability to recognize user emotions is a prerequisite for the enhancement of computer-based learning by integrating emotional experience. Recognition systems face two core challenges: (1) detection of the emotion-related signals provided by the users and (2) inference of the emotional state of the user according to the signals detected.

With regard to modes of detection, in addition to self-report measures, a number of channels are considered emotion sensitive. For instance, studies have been carried out on the link between emotional states and modifications in each of the following channels: physiological correlates, facial expressions, gestures, verbal communication, vocal non-verbal communication.

Physiological correlates of emotions have been considered a reliable means of detection since the work of James (1884). Ekman, Levenson, and Friesen (1983) suggested that the identification of distinct emotions on the basis of autonomic nervous system (ANS) activity requires taking into account different indices simultaneously. Various instruments for the detection of physiological signals (e.g., blood volume pulse, respiration, skin conductance, electromyography, neurological response) are already in existence and are subject to constant enhancement, particularly in order to make them less intrusive by integrating them into more natural seeming devices, for example, incorporation of sensors in a mouse or wearable jackets. The main problem is recognized to be the extreme variability of these signals across people and situations and even within the same person. This is borne out by the fact that few studies have indicated the existence of specific patterns of autonomic activity for specific emotions (for a review see Cacioppo, Berntson, Larsen, Poehlmann, & Ito, 2000). The search for an invariant
Addressing Emotions within E-Learning Systems

relationship between emotions and physiological responses should be abandoned in favor of the analysis of, under what conditions, and for which emotions, differential physiological activity is observed (Cacioppo et al., 2000).

A number of attempts at automatic emotion recognition from physiological measures have been made. Some researchers have tried to directly link a set of values to a set of emotions via various statistical algorithms (Nasoz, Alvarez, Lisetti, & Finkelstein, 2004). Others, using a dimensional model of emotion (where emotions are represented along some main axes, typically arousal and valence) have tried to use each physiological measure as an index of a single dimension (Picard, Vyzas, & Healey, 2001; Prendinger, Mori, & Ishizuka, 2005). What emerged from most of these studies is that physiological measures alone cannot provide enough information for emotion recognition: other information is needed (Scheirer, Fernandez, Klein, & Picard, 2002). In addition, awareness of the context can have critical significance for the interpretation of physiological data (Stemmler, 2003; Ward & Marsden, 2003).

Facial expressions of emotions have been investigated in depth. According to Ekman (1994), each emotion goes with a specific pattern of facial actions, as predicted by the Facial Action Coding System (FACS) (Ekman & Friesen, 1978). Although this bi-univocal link has been severely criticized, most of the research on automated emotion recognition is based on the FACS model. An early work carried out by Kaiser and Wehrle (1992) described a procedure for automatic detection of facial behavior independently of individual physiognomic differences, through the use of plastic dots affixed to predefined facial regions. A pattern-recognition algorithm identifies dot patterns that are classified according to the FACS by an artificial neural network. More recently, Cowie et al. (2001) outlined a system for the recognition of facial expressions by identifying Facial Animation Parameter Units defined in MPEG-4 standard, but the system is still not fully automatic and requires human assistance. Conversely, Kapoor, Qi, and Picard (2003) proposed a fully automatic framework that requires no manual intervention to analyze and recognize upper facial actions, corresponding to the regions of eyes and eyebrows. Tian, Kanade, and Cohn (2000) developed an automated system to analyze subtle changes in facial expressions, based on both permanent (brows, eyes, mouth) and transient (deepening of facial furrows) facial features in a nearly frontal image sequence. The system is based on multi-state templates that require manual set up in the first frame of the sequence; thus the system is not fully automatic.

In any case, the FACS model is not free of problems. Although prototypic expressions of some emotions, for example happiness, are natural, they occur infrequently in everyday life, since people tend to communicate more through subtle facial actions. In addition, emotions like confusion, boredom, and frustration do not have corresponding prototypic expressions. Furthermore, this method lacks temporal and detailed spatial information (Russell & Fernández-Dols, 1997).

Gestures are not generally linked to specific emotions. Although some studies have identified expression of emotional meaning through body movements (Wallbott, 1998), psychologists generally maintain that the meaning of gestures is mainly determined by the specific interaction context. Consequently, it is difficult to assume a direct link between gestures and specific emotional states.

The situation with regard to posture is even more complicated, since in the field of nonverbal behavior research there is no established and generalized criterion about how to classify postures or about the association between posture and emotional state. However, some attempts have been made to automatically detect emotion from...
Addressing Emotions within E-Learning Systems

Mota and Picard (2003) presented a system for recognizing naturally occurring postures and associated affective states, relating to the interest level of children while performing a learning task on a computer. The system is capable of two kinds of recognition: (1) recognition of a static posture position and (2) recognition of a sequence of postural behaviors. The system is not fully automatic because it requires preliminary selection and coding of posture action units by human observers. Similarly, affective states linked to postures have to be selected and labeled manually by observers.

Table 1. Assessment of different detection channels

<table>
<thead>
<tr>
<th>CHANNEL EVALUATION</th>
<th>INTRUSIVENESS</th>
<th>TECHNICAL COMPLEXITY</th>
<th>PRICE / COSTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physiological measures</td>
<td>Advantages: Rapid and synchronous modifications. Biological foundations. Uncontrollable by the participant. Limits: Easily influenced by non-emotional events and environment. Wide variations across people, situations, and even within the same individual.</td>
<td>Very high. Future development of wearable devices may reduce it.</td>
<td>Hardware difficulties: Integration of specific devices. Software difficulties: (1) real-time data processing; (2) recognition algorithms. Medium-high for physiological detection device.</td>
</tr>
<tr>
<td>Facial expressions, gestures, and posture</td>
<td>Advantages: Relatively spontaneous; easy to observe; objective. Limits: Influenced by non-emotional factors and by environmental events (context). May be culturally influenced. Gestures are linked to conversation.</td>
<td>Medium. People can be annoyed by cameras. Some systems require cameras positioned on participants' head.</td>
<td>Hardware difficulties: Integration of cameras. Software difficulties: (1) AU extraction; (2) posture tracking and extraction; 3) real-time processing. High for cameras and AU recognition software.</td>
</tr>
<tr>
<td>Speech (verbal)</td>
<td>Advantages: Non intrusive; easy to collect. Limits: For emotion lexicon, difficult to know whether the reported emotion is a conceptualisation or the effective state; whereas linguistic indexes require long conversational sequences to be detected.</td>
<td>Low</td>
<td>Software difficulties: Integration of word recognition software and emotional semantic map. Low for microphone. Medium for word recognition software.</td>
</tr>
<tr>
<td>Vocal non-verbal measures</td>
<td>Advantages: Link to physiological changes; wide research tradition on emotional vocal patterns; low intrusivity. Limits: Technical difficulties; problems in noisy environments; possible problems with phonetic variability; few results for naturally collected data; recognition rate varies according to the database (number of emotions, mode of data collection, number of speakers); systematic confusion for some emotions.</td>
<td>Low</td>
<td>Hardware difficulties: Integration of a spectrograph. Software difficulties: (1) real-time data processing.; (2) recognition algorithms. Low for the microphone and for spectrograph.</td>
</tr>
<tr>
<td>Self-report measures</td>
<td>Advantages: Simple to administer to the user. Limits: distortions due to social desirability, errors in the self-monitoring process, cultural differences, potential disruption of learning path.</td>
<td>Medium/low. Requires definition of timing, may disrupt the learning process.</td>
<td>None Low</td>
</tr>
</tbody>
</table>
Addressing Emotions within E-Learning Systems

The verbal communication of emotions concerns terms and words produced in concurrence with an emotion, and it takes into account the cultural differences reflected in the terms adopted and in their meaning (the so-called emotional lexicon). The language-emotion relationship has been studied from various perspectives: lexicographic research on words that refer to emotions, highly related to the universality or culture-relativity of human emotions (Wierzbicka, 1995); syntactic research focusing on emotion verbs like to fear; and so forth; investigation of certain types of language use, such as hyperbole, repetition, the use of strong metaphors; conversational analysis, which studies the rules regulating the occurrence of emotionally expressive behavior in interaction.

Attempts at automatic emotion recognition based on verbal communication can be grouped into a number of categories: (1) keyword spotting, where text is classified into affect categories based on the presence of fairly unambiguous affect words like distressed, enraged, and happy, for example, affective reasoner (Elliott, 1993) and Ortony’s affective lexicon (Ortony, Clore, & Collins, 1988); (2) lexical affinity, detecting more than just obvious affect words, the approach assigns arbitrary words a probabilistic “affinity” with a particular emotion; (3) statistical natural language processing, that is, by feeding a machine a learning algorithm with a large training corpus of affectively annotated texts, it is possible for the system to learn the affective valence of affective keywords as well as that of other arbitrary keywords (as in the lexical affinity approach), punctuation, and word co-occurrence frequencies.

The vocal-nonverbal communication of emotions deals with the differences that can be identified in some phonatory variables (e.g., intensity, tone, rhythm) when an individual is, for example, expressing happiness or sadness: Some patterns of vocal supra-segmental traits have been described (for a review see Juslin & Laukka, 2003; Scherer, 2003). Several studies have shown that each emotion is associated with a distinctive vocal profile in a systematic way.

A growing research corpus about automatic recognition of emotions from voice can be found in scientific literature. It should be pointed out that the motivations behind these studies range from theoretical and scientific to application based. The first category focuses on acted speech and on a broader range of emotions, while the second includes, for instance, work on speech from automatic call centers (Batliner et al., 2000) and research on detection of stress during driving (Fernandez & Picard, 2003). As pointed out by Oudeyer (2003), research on automated vocal production and recognition of emotion has only been carried out for a short number of years. As well as systems aimed at automatically detecting some basic emotions independently of context, systems have been developed to focus on one/two target emotions, of particular relevance to specific application fields. For example, Batliner et al. (2000) proposed a system for the recognition of anger to enhance the effectiveness and usability of automatic dialog processing.

Self-report measures have also been proposed to detect emotions. Some researchers have developed questionnaires to be administered to participants in order to obtain information about their emotional experience (Russell & Mehrabian, 1977; Scherer, 1988). Besides questionnaires, other self-report measures based on nonverbal and pictorial methods (e.g., Self Assessment Manikin Scale; Bradley & Lang, 1984) were put forward; here users choose the images that best represent their emotional states, partially overcoming the limits inherent to linguistic and verbal expression.

This means of detection, on the one hand is non-intrusive, fast and cheap, but, on the other, can be affected by distortion due to social desirability factors and to the potential inability of the user to correctly express his/her own emotional experience.
What emerges from this review is that each channel has its own strengths and weaknesses that should be taken into account when trying to implement an automatic recognition model. Table 1 provides a brief critical assessment of the different channels reviewed, both from psychological and implementational viewpoints.

Guidelines for Designing an Inferential System of Emotion Recognition

Once emotion-related signals have been detected, the recognition system needs an inferential system in order to attribute an emotional state to the users.

The inferential system can be defined as a cognitive architecture that is able to integrate the various input signals with other elements, so as to infer the user’s emotional state.

Specifically, we propose that a cognitive architecture of emotions should be designed in order to output a possible emotional state, by working from four main inputs: (1) individual characteristics, (2) initial signal profile, (3) real-time signals, and (4) context modeling.

Firstly, the system should profile the user in terms of different psychological variables in order to establish his/her individual personality characteristics. In other words, the user should be emotionally and motivationally profiled (e.g., Matsubara & Nagamachi, 1996), and these characteristics, subsequently, should be taken into account to weight the possible outcomes of the context modeling.

Secondly, an initial stage of interaction should be provided for, in order to detect the baseline values of physiological and vocal nonverbal signals for each user (physiological and vocal profiles). Due to the interpersonal variation for these signals, the inferential system needs a period to register the signal levels of the user at rest: these values will be considered the baseline for modifications detected during the interaction (Berntson, Uchino, & Cacioppo, 1994).

Thirdly, the inferential system should work on the outputs of different modality-specific modules (speech: vocal analysis module; visual analysis module; physiological signal module; self-report measures module; speech: verbal module). Each of these has its own limitations, and none can be fully relied upon to detect emotions in isolation from the others, whereas multimodal analysis may at least partially reduce problems (Pantic & Rothkrantz, 2003). All the modules provide information about behavior and physiological variations displayed by the user at a certain stage of the interaction: Variations should be processed by the inferential system in order to assess emotional state using an advanced statistical approach, including different learning and classification algorithms. To this end, it is critically necessary to define a training database, as a fundamental prerequisite for developing reliable classification algorithms (Anolli et al., 2005).

Fourthly, context modeling is one of the main features to be included, and it may be one of the more promising ways to enhance inferential reliability. If only physiological and communication measures are considered, distortions and mistaken emotional attributions may occur, since a key component of emotional experience is still lacking. Emotions result from both an arousal component and from a cognitive component, the latter encompassing the appraisal of arousal and its attribution to an emotion eliciting event. Appraisal theories suggest that emotions are elicited and shaped by an individual’s subjective evaluation of situations or events that affect the individual’s needs or goals (Frijda, 1986; Scherer, 1984; Weiner, 1986). Environmental or proprioceptive stimuli are evaluated by the individual vis-à-vis a number of criteria and dimensions linked to individual meaning. As a result, the same stimulus...
Addressing Emotions within E-Learning Systems

Figure 1. Ideal representation of a possible emotion recognition system

...can provoke very different emotional reactions among people (Lazarus, 1966), depending on the outcome of the subjective appraisal.

Appraisal theorists suggest that individuals use a fixed number of dimensions or criteria for evaluating the significance of the events. These criteria can be categorized into four major classes: (1) intrinsic characteristics, such as novelty; (2) significance for the individual’s needs or goals; (3) ability to influence or cope with the consequences of an event; and (4) compatibility with social or personal standards, norms, or values (Scherer, 1984).

The model of emotions proposed by Ortony et al. (1988) is one of the most used within affective computing, since it can be implemented in computer software. The authors do not represent affect as a set of basic emotions nor as emotions defined within a dimensional space, but group emotions according to the eliciting cognitive conditions.

The inclusion of context modeling to address the cognitive component of emotion may require the definition of an additional module (registry module). This module would take into account relevant information about events happening to the user, pertinent both to the learning context and the learning process, for example, a failure while using a certain software or the loss of some data due to a system error.

While it is not possible to register what is happening outside of the concrete interaction, it may be useful to consider the events of the learning interaction itself, for example, failure in a test phase could be presumed to be negative valenced and to influence emotional state according to corresponding appraisal criteria. In other words, the system could be provided with some appraisal...
criteria that would partially define the emotional significance of events occurring to users.

The cognitive architecture should be able to process variations within each modality, integrate them, and attribute a potential meaning to them depending on the initial state and profile of the specific user and output of the registry module.

**CONCLUSION**

A topic of key interest to educational researchers is the search for strategies capable of inspiring the interest and active participation of learners (Bransford, Brown, & Cocking, 1999). In traditional face-to-face learning this hinges on the ability of the teacher. Teasing out and replicating all the components of this ability remains a huge challenge for e-learning designers. One possible way forward, according to Picard et al. (2004), is to approach these learning goals by incorporating affective information into the e-learning path. This line of enquiry appears all the more feasible, as a growing number of relevant technologies are coming on stream, especially in the field of emotional detection.

Recently, a number of projects have addressed the inclusion of affect in learning and education environments (see Kapoor, Mota, & Picard, 2001; Zhou & Conati, 2002). For instance, as cited previously, a preliminary system for the automatic detection of children’s interest levels during learning situations, based on a combination of posture and facial expressions, was designed at the Massachusetts Institute of Technology Media Lab (Picard et al., 2004).

In this paper we have described a possible way to enhance an automatic emotion recognition system. The key defining characteristics of our proposal are multimodality, that is, the inclusion of multiple detection channels and contextual awareness of events occurring within the learning path. Both of these characteristics can contribute to creating a more effective system.

Nonetheless, some key limitations of the proposal should be discussed. Currently, it is not possible to identify the optimum combination of real-time measures and other sources of information required for emotion recognition. A “golden truth” is still lacking, both for the individual channels reviewed and how best to combine them. As already stressed, each channel presents specific reliability problems, and some have additional drawbacks that partially hinder implementation (e.g., intrusiveness in detecting physiological signals). In order to minimize technical obstacles, more suitable and comfortable devices are needed, which may be wearable or integrated in normally used instruments (mouse, keyboards, etc.).

In addition to technical and scientific limitations, researchers should be aware of ethical issues raised by the implementation of devices connected with emotions. Ethical considerations should be continuously taken into account at all stages of research and development, from the preliminary phase of emotional data collection (in particular if data is acquired through induction procedures) to the final concrete implementation of recognition platforms within everyday applications, because of their potential to detect learners’ emotions beyond their conscious wish to disclose them.

In conclusion, it should be pointed out that a recognition system is not an end in itself, but a means to design instructional technology environments to be more effective in supporting user learning experience. Once a system has detected and recognized an emotional state, it can take this information into account in selecting those subsequent actions held most appropriate to improve learner performance, according to the pedagogical strategy in use. Although some scattered attempts have been made to incorporate emotions into models of learning (e.g., Kort, Reilly, & Picard, 2001; Mandler, 1984), no widely shared theoreti-
Addressing Emotions within E-Learning Systems

cal model can be identified. The main issue to be investigated along with the definition of emotion recognition systems remains the development of a generally accepted theoretical model of how to use affective information in learning.

ACKNOWLEDGMENT

The present work was partially supported by the European Commission FP6 research project MYSELF (SME-2003-1-508258). Website: www.myself-proj.it

REFERENCES


Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students’ self-
Addressing Emotions within E-Learning Systems

regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist, 37*(2), 91-105.


Addressing Emotions within E-Learning Systems


**KEY TERMS**

**Affective Computing:** Affective computing is a research field that aims at including emotions within information technology design. It deals specifically with three different levels of emotion integration: (1) the detection of user emotions, (2) the expression of emotions by computers, and ultimately, (3) the possibility for a computer to “have” emotions.

**Appraisal:** Appraisal refers to the cognitive evaluation antecedent to an emotional episode. Appraisal theoretical models are characterized and differentiated by the appraisal dimensions included.

**Autonomic Nervous System (ANS):** ANS is the part of the nervous system that regulates individual organ function and homeostasis, and for the most part, is not subject to voluntary control. It is usually divided into sympathetic and parasympathetic.

**Computer Anxiety:** Computer anxiety is the individual fear or apprehension of using a computer directly or the anticipation of having to use it.

**Emotional Intelligence:** Emotional intelligence is the underlying general competence that comprises a variety of emotional skills. Such skills vary according to the different models of emotional intelligence proposed.

**Facial Action Coding System (FACS):** FACS is a system originally developed by Paul Ekman and Wallace Friesen in 1978 to taxonomize human facial expression. It is the most used method to measure and describe facial behaviors, coded through action units (AU).

**Modality (of emotional detection):** Modality is each of the different channels that are considered to be emotion-sensitive (i.e., physiological measures, vocal non-verbal measures, self-report measures, facial expressions, posture and gestures, and verbal content).

**Multimodal Emotion Recognition:** Multimodal emotion recognition is where one is inferring an emotional state and integrating inputs coming from multiple emotion-sensitive sources.