Affective Educational Technology: First steps on the introduction of the affective dimension in the core of online design

Eva de Lera
Universitat Oberta de Catalunya
Spain
edelera@uoc.edu

Carles Fernàndez
Universitat Oberta de Catalunya
Spain
cfernandezba@uoc.edu

Magí Almirall
Universitat Oberta de Catalunya
Spain
malmirall@uoc.edu

Abstract. The cognitive or functional focus has traditionally been the main key for instructional and technology designers, giving lower importance to the emotional dimension. This point collides with the idea supported by multiple research studies that cognitive and affective layers can not be separated and in order to enhance learning we have to promote appropriate affective states. In fact, designers tend to implicitly propose a particular emotional scenario when they design virtual learning environments, but these decisions are generally not based on affective learning theories or findings.

In this paper we aim to show the first steps that the Universitat Oberta de Catalunya is following in its effort to introduce the affective dimension in the core of the e-learning design process. We believe that our first stages could be useful for other institutions interested to include the affective dimension within the design of their online environment. We have concluded that three steps should be the starting point: a specific methodology to gather affective information, a set of general affective principles to guide designers' tasks and "testing & playing" with new tools and methodologies already available in the market.

Introduction

Most of the universities that are currently offering online courses usually have some kind of system for course design assessment. Traditionally these evaluation tools have been based on a set of recognized indicators provided by experts in graphic design, instructional designers, technologists… Up to now, the evaluation of affect in online learning usually means the use of questionnaires where students respond to items like 'do you think the course is motivating enough?' or 'has the course met your expectations?". Although questionnaires are a short and cost-effective way to gather data, we already know that users may answer just to offer an 'appropriate' opinion or are mediated by other variables. As Picard states (Picard, 2004) "self-report is colored by awareness of internal state, reflections on how such a report will be perceived, ability to articulate what one feels, etc.". Also, the use of questionnaires requires the learning experience to be interrupted (or in any case, they cannot be used at the same time learning is happening).

We aim to show the first steps that the Open University of Catalonia (UOC) is following in its effort to introduce the affective dimension in the core of the e-learning design process. We believe that our first stage could be useful for other institutions interested in including the affective dimension in the design of their online environment. We have concluded that three steps should be the starting point: a specific methodology to gather affective information, a set of general affective principles to guide designers' tasks and "testing & playing" with new tools and methodologies already available on the market. Our objective is to share our experiences in these three stages.

Nowadays, the subject of 'affective learning' is achieving major importance especially in the field of human-computer interaction where after many years of focusing on usability-functionality issues researchers are noticing that users' satisfaction cannot reach higher levels without taking emotions into account. Also, thanks to progress in brain research and even new insights in neuropsychology we are in a better position to carry out a comprehensive study based on emotions and specific affective states. Inasmuch as learning, and particularly online learning is an activity based on emotions, it is time to contribute more experiences and begin to build an appropriate theoretical framework for e-learning.

Since the birth of the Open University of Catalonia (UOC) in 1995, several teams within the staff of the university, made up of teachers, technologists or instructional designers, have made their contributions to the design of a rich but complex virtual learning environment.

Seven years ago we adopted a more systematic approach through the use of user-centered design methodologies and usability, which allowed us to gather more data on the interaction between our students and the virtual environment. Undoubtedly it was a big step on the way to reach real learners' behavior and allowed us to propose more accurate designs. However, although our questionnaires showed that the new designs reflected learners' needs, we found many students arguing that spaces were not engaging enough. In this sense, users were conscious of the difference between an environment that works and fits their needs and an environment that is also motivating. Our conclusion was that while classical usability studies were based on functional tasks, we needed a step forward to focus on fostering a more engaging overall experience. At this point of the process, we understood there was a deeper level, an emotional level acting as a powerful engine in the learning experience.

The process of adoption of an affective methodology prompted deep reflection among the teams involved in the design process. Inasmuch as affective e-learning is a very new, broad and controversial field, it showed us the need to focus our work on what we considered the 'first three essential tasks': to design a specific tool to measure affective data more objectively, to support designers in the creation of positive emotional environments and to experiment with existing and innovative tools that could provide affective data. With regard to the first point, we adapted a tool based on FACS (Facial Action Coding System) we called '*Ten Heuristics*', and we also started to work with eye trackers, pupilometry and students' self-assessment methods. As for support in affective design, we basically took advantage of the experiences in design we have gained from creating, evolving, and continuously evaluating a 15-year old learning environment, producing (with the help of an interdisciplinary team) a set of guidelines for designing engaging e-learning environments. Thirdly, we are just starting to analyze tools that provide neurophysiological data that may be related to specific affective states, based on sensors that gather electrophysiological data.

Introducing affect in assessment: a reliable, feasible and cost-effective methodology

Our proposal consists of the use of three techniques: pupil size analysis, interpretation of facial and body expressions and students' self-assessment. All the data gathered through these techniques is compared to the real-time interaction

In order to validate the combination (henceforth triangulation) of these techniques as a specific methodology, we planned a test with seven students interacting with the new mainpage of the virtual learning environment. This new page is very similar to the recognized and awarded *Igoogle* personal area and students are allowed to import external modules (gmail, google calendar, news and many other applications) as well as information from the university (forums, groups, news…), and personalize the design and position of these modules.

The sample was selected following heterogenic criteria: age, gender, studies and ICT skills were representative according to our target students at the university.

We carried out this test with two objectives:

- To test whether the proposed methodology is reliable for affective assessment, feasible and cost-effective
- To test the new mainpage of UOC from an affective perspective

Brief state of the art of the techniques used

Pupil size has been widely studied by Picard and Hess (who coined the term in 1975), and accurately summarized by Partala and Surakka (2002).

Hess, in 1972, stated that pupil size depended on the kind of stimuli. While pleasant ones would increase pupil size, unpleasant ones would decrease it. In 1974, Janisse suggested that there is no pupil constriction with negative stimuli. The latest research on pupilometry has confirmed that the increase of pupil size has a linear relationship with affective arousal. Arousal is significantly increased except with neutral stimuli. As such, it is quite widely accepted that arousal can be measured through pupilometry.
Nevertheless, other factors such as cognitive load and light reflex may have an effect and must be appropriately corrected in any research study.

Pupilometry relies on arousal, FACS (Facial Action Code System) and other methods to analyze non-verbal information related to valence. To date, there is some consensus on the existence of six basic emotions that can be codified according to FACS (anger, disgust, fear, joy, sorrow and surprise). The technique created by Ekman codes expressions as a combination of 44 facial movements called Action Units (Michel and El Kaliouby, 2003).

Concerning self-assessment; this technique has been partially criticized due to the fact that adults' self-report is colored by awareness of internal state, reflections on how such a report will be perceived, ability to articulate what one feels, and other factors (Picard, 2004).

On the other hand, using self-assessment as a complementary method can bring interesting results. As a example of self-assessment interest, Sang-Hoon Jeong (2007) has developed a tool called VideoTAME, where subjects can view the recorded video of them performing tasks. This allows users to analyze the emotional changes that occurred. This data can be useful as complementary information for emotion inference or confirmation of affective states.

The methodology in the test design

Five tasks have to be carried out on the mainpage: read an email in their webmail, add a new module (widget), change the position of another module, access the forum of a subject and access the virtual library. All these tasks can be done directly from the new mainpage. More important than the tasks themselves are the *'key events'* that students need to perform in order to accomplish the task. For example, a key event in the task 'add a new module' is to find the right button that allows widgets to be added. Another key event for any of the tasks would be to know how to go back to the mainpage after the student uses webmail or the forums.

The key events are new stimuli within a 'steady/stable interaction', and represent opportunities for increase of arousal, especially when individuals cannot accomplish the event easily. These key events are the moments in the students' performance where we focus our analysis of affect through measure of pupil size, facial and gesture interpretation and self-assessment. Obviously, all this data is continuously analyzed along with real recorded interaction of the students.

We have identified the following key events:

- To find the key buttons such as: 'add a module', 'personalize your mainpage', 'virtual library', 'webmail' and 'forums' (a total of 5 events)
- To be able to go back to the main page after a task
- To understand the concept of module as a box that can be dragged and dropped

Apart from key events, we also identified a group of events that require special attention by the user such as typing the password, waiting for the pages to load, reading instructions, accessing wrong pages… We call them '*secondary events*'. As we show later in the paper, these events also showed higher levels in both pupilometry and heuristics, although key events were our main focus in the study.

As stated before, our methodology is based on a combination of pupilometry, gesture expression heuristics and student self-assessment. Each technique has its own objective. We used pupilometry to detect arousal increase, whether related to positive or negative affective states. Increased pupil size seems to indicate that 'something is happening internally' and may be a clear sign that impels us to also analyze the valence of the affective state through facial and gestural expressions. Moreover, triangulating pupil size with gestural heuristics and self-assessment will allow us to assess how precise pupil size is by comparing the average of increased pupil size, not related to affecting identification, with the rest of techniques. In this sense, it is very interesting to analyze the correlation between increased pupil size and a body expression that confirms this physical response in arousal. We consider that pupil size is increased when size is 10% bigger than the average pupil size in the test for a particular subject.

Secondly, our gesture coding system is used to determine the valence of emotions, since pupilometry itself cannot do anything with valence. Our own technique, called *ten heuristics*, is based on observation and does not require extra implementation effort since most interface evaluations are conducted observing and recording the user as he or she interacts with the interface. In such a scenario, facial and body expressions are often observed and recorded, but generally not measured in a structured manner. The smile heuristic included in this method represents the goal of the evaluation: to see a user with relaxed facials, therefore, without experiencing negative reactions or frustration.

The ten heuristic tool has been developed thanks to previous studies in which we tried to identify the most common expressions of students in virtual learning environments. In this sense, the current research is a study that allows us to confirm which of these expressions are the most used in online learning (De Lera & Garreta-Domingo, 2007).
This is a summary of the ten heuristics:

1. Frowning. Frowning can be a sign of a necessity to concentrate, displeasure or of perceived lack of clarity. Darwin (1872) wrote about how frowning is one of the signs of deep and "perplexed reflection". In their study, Partala and Surakka (2004) found that the frowning activity attenuated significantly after the positive interventions than the no intervention condition.
2. Brow Raising. Brow raising should also be considered a negative expressive reaction. To lift the arch of short hairs above the eye is a sign of uncertainty, disbelief, surprise and exasperation (Givens, 2005).
3. Gazing Away. The gazing away from the screen may be perceived as a sign of deception. For example, looking down tends to convey a defeated attitude but can also reflect guilt, shame or submissiveness (Givens 2005).
4. Smiling. A smile, or elevation of the cheeks, is a sign of satisfaction. The user may have encountered an element of joy during the evaluation process. Partala and Surakka (2002) found that smiling activity was significantly higher during the positive condition.
5. Compressing the Lip. Seeing the user compress his or her lips should be perceived as a sign of frustration and confusion. Lip and jaw tension clearly reflects anxious feelings, nervousness, and emotional concerns (Partala and Surakka, 2004).
6. Moving the Mouth. If the user is seen mouth gesturing or speaking to himself / herself, this is associated with a sign of being lost and of uncertainty.
7. Expressing Vocally. Vocal expressions such as sighs, gasps, coughs, as well as the volume of the expression, the tone or quality of the expression may be signs of frustration or deception.
8. Hand Touching the Face. Elevating the hand that is placed on the mouse to his / her face is a sign of confusion and uncertainty, generally a sign of the user being lost or tired.
9. Drawing Back on the Chair. The user may be experiencing negative or refusing emotions. By drawing back the chair, he / she may be showing a desire to get away from the present situation.
10. Forward Leaning the Trunk. Leaning forward and showing a sunken chest may be a sign of depression and frustration with the task at hand. Like with the previous heuristic, the user might be encountering difficulties but instead of showing refusal, leaning.

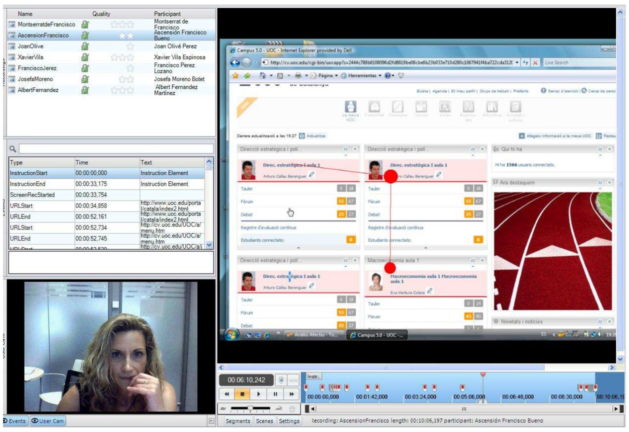

Figure 1: Recording data interface

Finally, as for the third technique, the objective of self-assessment in our methodology is to confirm that both pupilometry and gestural interpretation can really infer affective states.

The triangulation between these three techniques has two objectives:

- To analyze the correlations between the data gathered with the techniques in order to assess the specific effectivity of each one
- To conclude whether the overall methodology is appropriate for affect measurement

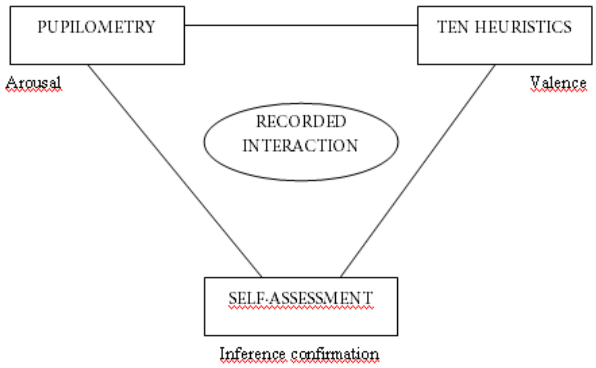

Figure 2: The methodology based on triangulation

Results and conclusions

Both pupilometry and 'ten heuristics' together were able to measure most of the key events and secondary events. We have analyzed the events where a key event could not be detected, the reason being that the student solved it without any problem, which means no arousal. The higher the level of success in a task, the less probable it is for arousal or non-verbal communication to appear.

These techniques have also showed that pupil size was increased or students reacted non-verbally to some events we had not considered as key or secondary. Pupil size, especially, helped us to detect these situations.
We have also compared whether pupil size or ten heuristics were more reactive to key and secondary events. Pupil size responded to 74% of these events while ten heuristics did so in 65%.

One of the most important questions for our proposed methodology is the correlation between positive responses in pupil size and positive responses in heuristics. For most of the students (five) only 40% of the events with increased pupil size also showed any gestural reaction. The two others showed over 80% of correlation. Our conclusion is that pupil size does not seem to be a very good predictor if we only want to use it to infer or to state that the student is 'feeling something'. The size of pupils does not seem to be enough to assess arousal in our particular environment or conditions. Also, almost 40% of the events with increased pupil size did not show any particular facial or body expression.

Very interestingly, the correlation between high pupil size and specific gesture (average of 47.8%) is higher than the correlation between normal pupil size-specific gesture (20.7%) and high pupil size-no gestures (31.1%). This data may show that affective states are more commonly accompanied by the presence of both high pupil size and non-verbal expression, though evidence is not great enough to focus on pupilometry alone.

The analysis of gestural expressions by itself is richer and allows a good determination of valence. Most of the key events showed a particular expression, and the emotions inferred through these were confirmed through students' self-assessment.

Through students' self-assessment we have particularly checked whether the expressive reactions in Ten Heuristics can be inferred as specific emotions. Some of the heuristics were frequently repeated in the pilot test, such as frowning, compressing the lips, expressing vocally or hands touching the face. Most of the time when pupil size was not increased students did not show any expression, smiled or their hands were touching their face (especially around the mouth area).

The amount of non-verbal expression depends on each individual, and is probably the most important factor for heuristics performance. There was a relationship between success in tasks and amount of expressions (more difficulties, more presence of body expressions), but there was an individual who had important problems in some of the tasks but did not show many expressions. In the self-assessment process she let us know that she is not very expressive.

Globally, students' self-assessment confirmed more than 96% of the events detected by both pupilometry and heuristics. Almost 100% of the heuristics were confirmed and more than 90% of the events detected through increased pupil size.

Through the qualitative analysis of each specific case study we have identified other variables that have a particular impact on affect and the techniques we have selected to measure:

1. Pupil size tends to be higher during approximately the first minute due to some kind of tension (all the subjects experienced this).
2. Sustained concentration makes pupil size increase but is not comparable to pupil size in presentation of new interfaces, errors produced or difficulties in performance.
3. Concerning patterns of timeline analysis, pupil size is reactive usually 500 ms. before the key event and it keeps increasing depending on how well the user finds a solution but generally maintains the maximum size during 2 seconds.
4. Pupil size is not especially high when users express something, probably because talking is a way to decrease arousal.
5. We have analyzed the highest moments of increased pupil size but could not find specific body expressions related to high arousal.
6. The two most successful users did not show high increases in pupil size although they also had to solve new problems. Overall confidence may be the explanation.

In conclusion, the methodology of triangulation we have proposed for data collection and affect inference is useful, but we cannot rely on one technique alone, especially on pupilometry, since it showed that some key events and some difficulties experienced by users did not increase pupil size (less than half of events). Nevertheless, it is convenient to collect pupil size data since it can make us concentrate on an event where the subject did not express anything.

Concerning other considerations of implementation such as feasibility, usability or time or cost-effectiveness, we can argue that these techniques do not require a big investment and the way they gather and export data is very useful for research. Pupilometry only requires an eyetracker with infrared rays, facial expression monitoring is carried out with a simple webcam, and for students' self-assessment we only need learners. These technologies are affordable although there is, in fact, specific software, such as Tobi, which allows monitoring of eye movement, pupil size, facial gestures and real-time interaction through use of the same software, facilitating data analysis and interpretation.

With the appropriate guidelines for integrating this methodology in course assessment, staff should not notice a significant increase in work, or they should at least feel a balance between time investment and higher quality in course assessment and design.

Introducing the affect in design: the Enjoy guidelines

As a further step for a real learner-centered design we concluded that we needed a set of guidelines or main principles that should promote elements for engagement. Teachers, instructional designers, technologists and even other staff involved in design issues participated in a process in order to identify these principles. We finally summarized all these conclusions in twelve principles as a methodology to help design engaging and motivating online learning environments for e-learners (De Lera & Almirall, 2008). The basic idea for any of the principles is to keep high levels of the intrinsic motivation that learners bring to the process while adding other elements of extrinsic motivation that improve the whole learning experience.

The principles are still in process of evaluation. Up to now we have only identified them and we are incorporating them in a few existing courses that we will test in terms of performance and impact on affect.
The following are the 12 easy-to-follow guidelines for designers, developers, learning technologists and other participants in the e-learning design process:

1. Personalization – the environment must make the student feel like a person and not like a user. Use of communication strategies that are more personal, common language and options for this person to participate in this environment. According to the current trends in e-learning, the learner must have the possibility to design a learning space that makes sense for himself and incorporates areas or tools that do not necessarily provide our virtual campus but exist in the Internet. The final objective of personalization is the generation of a real 'Personal learning environment'.
2. Identity – utilizing real and consistent images to help the student identify him/herself with the values and the community in a quicker and more efficient way. The idea of generating a sense of belonging.
3. Brand – ensuring that the brand and the brand values are reflected throughout the virtual environment to reinforce the relationship between the student and the institution. Justifying the value proposition, why this University and not another.
4. Community – offering options to communicate, relate and participate. Offering options to make friends, to meet in small groups and large groups (i.e. organizing cultural activities that involve learners) and to promote associations. Making community options very visible and easily accessible.
5. Surprise – introducing positive surprise elements or special events in the initial entry pages or in strategic locations to make the students feel that they are part of a creative and dynamic community, or to act as reminders of why this is the place, a place that provides enriching stimulus. Creating the idea that 'something is changing, something is moving' all the time, even promoting some kind of arousal (Shen et al., 2007).
6. Innovation – integrating innovative elements in the virtual environment, those that they may begin hearing or reading about in the media and other trend environments.
7. Zen – ensuring that there is not an overload of text in the screen, that white spaces are used, as well as photographic or graphic elements. Need to avoid unnecessary noise and obstacles, there is a need to design to safe time and be more efficient.

8. Search – providing shortcuts to students that have little time, ensuring that they can find the information they need by doing a simple search.
9. Clarity – utilizing lively and bright colors to facilitate interaction, reading and information visualization, providing guidance through the design.
10. Situation – ensuring that the student quickly recognizes the structure or map of the environment in a glimpse, without needing to scroll or to read a lot.
11. Aesthetics – ensuring a consistent aesthetic throughout, to help guide the student through his or her tasks and objectives.
12. Recognition – utilizing standard icons and symbols that can be easily and quickly understood without requiring the alternative text or an extra click to understand it.

. In our study we also wanted to focus on other aspects of the learning environment, which we consider informal learning spaces that should not be underestimated: the homepage, the community tools, the structure, design, functionalities and other elements of the virtual campus that could, if designed properly, contribute to motivating and engaging the student.

ENJOY guidelines have been generated from the information gathered from several studies, which include user analysis and information from stakeholders, including the institution. The key aspects identified during the data gathering have been translated into design guidelines, in such a way that those participating in the design of any tools, service or application for a virtual campus would understand what the key elements are that should be taken into account and thus, avoid omitting them.

Introducing the affect in experimentation

This third stage in our initial phase of affective design of virtual learning environments is probably the most challenging one. It covers tools that have already been used in different disciplines such as psychology, neurology, marketing, artificial intelligence or human-computer interaction. They gather physiological or neurological data, and the level of efficacy of these tools and devices has already been tested in their respective fields (Zeng, Z. et al., 2007, Allanson and Wilson. 2002). For educational or other psychological issues, the main challenge remains: how to translate all this neurophysiological data into emotions. The difficulty is considerable since the taxonomies regarding emotions are not always shared (such as Kort or Krathwohl's proposals) and individual or cultural elements can vary the spectrum and conceptualization of emotions.

Although we cannot report data, we are presently experimenting with two tools that have already been used for affective computing: a sensitive mouse as a galvanic response device, and an electrophysiological sensor system. The pressure sensitive mouse is allowing us gather the level of arousal and muscle tension of the users while interacting with an application. Together with our MORAE system, we are able to relate facial and body gestures to levels of excitement and help us understand what users are experiencing. Galvanic Skin Response (GSR) is a device to help measure arousal. This is also used with the pressure sensitive mouse, and independently, during other methodologies such as interviews and focus groups.

We have selected a tool that can provide direct information about brain activity. Specifically, we are interested in three variables of this activity: EEG (Electroencephalogram), ECG (Electrocardiogram) and EOG (Electrooculogram). We are using these values to research how brain attention works (and especially distraction). There are several works that indicate the relationship that exists between spectral characteristics of EEG and attention (Lutsyuk, 2006), whereas this study shows how students with better attention show a better proportion on beta rhythm than on alpha rhythm and how the most successful students had higher proportions in alpha rhythms in some parts of the brain. Also, we are interested to study how gamma rhythm is related to the level of students' distractibility during the learning process. At any rate, it is still not very clear how emotions are reflected in EEG patterns, and we are still in the initial phase of utilization of this tool.

References

Allanson, J. & Wilson, G. (2002). Physiological computing. *CHI Workshop 2002: Changing the world, changing ourselves*. Minnesota, 912-913.

Darwin, C. (1872). *The expression of the emotions in man and animals*. Oxford University Press, New York, NY, 1998.

De Lera, E. & Almirall, M. (2008). ENJOY: guidelines for designing engaging eLearning environments. *In:proceedings iLearning Forum.* París. 1-4.

De Lera, E. & Garreta-Domingo, M. (2007). Ten Emotion Heuristics: Guidelines for assessing the user's affective dimension easily and cost-effectively. In: *Proceedings of the 21st BCS HCI Group Conference HCI 2007.* Lancaster University. 1-3.

Givens, D. B. (2005). The nonverbal dictionary of gestures, signs and body language cues. http://members.aol.com/nonverbal2/diction1.htm#The%20 NONVERBAL%20DICTIONARY. PDF edition

Lutsyuk, N.V., Éismont, E. V. & Pavlenko, V. B. (2006). Correlation of the characteristics of EEG potentials with the indices of attention in 12-to 13-year-old children. *Neurophysiology*, 38

Michel, P. & El Kaliouby, R. (2003). Real time facial expression recognition in video using support vector machines. In: *proceedings ICMI International Conference on Multimodal Interfaces 2003*, Vancouver, ACM Press, 258-263.

Partala, T. & Surakka, V. (2002). Pupil size variation as an indicator of affective processing. *International Journal of Human-Computer Studies*, 59 (1-2), 185-198

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

Picard RW., Papert. S., Bender, W., Blumberg, B., Breazeal, C., Cavallo, D., Machover,T., Resnick, M., Roy, D & Strohecker, C. (2004). Affective learning – a manifesto. *BT Technology Journal* 22 (4), 255.

Sang-Hoon, J., (2007). Suggestion of Methods for Understanding User's Emotional Changes While Using a Product. In M.J. Smith, G. Salvendy (Eds.) *Human Interface* (pp. 59-67), Berlin: Springer-Verlag

Shen, L., Leon, E., Callaghan, V. & Shen, R. (2007). Exploratory research on an Affective e-Learning Model. *Workshop on Blended Learning.* Edinburgh, 276

Zeng, Z., Pantic, M., Roisman, G. & Huang, T. (2007). A survey of affect and recognition methods: Audio, visual, and spontaneous expressions. *ICMI' 08*, Nagoya, 126-133