Assessment of the Emotional State by Psycho-physiological and Implicit Measurements

Michael A. Bedek[1]
michael.bedek@tugraz.at

Ben Cowley[2]
ben.cowley@aalto.fi

Paul Seitlinger[1]
paul.seitlinger@tugraz.at

Martino Fantato[2]
martino.fantato@aalto.fi

Simone Kopeinik[1]
simone.kopeinik@tugraz.at

Dietrich Albert[1]
dietrich.albert@tugraz.at

Niklas Ravaja[2,3]
niklas.ravaja@aalto.fi

[1]Knowledge Management Institute
Graz University of Technology
Brückenkopfgasse 1/6, 8020 Graz
(+43) 3168739553

[2]Center for Knowledge & Innovation Research
Aalto Yliopisto, Helsinki
PO Box 21250, 00075 Aalto
(+358) 505631249

[3]Department of Social Research
University of Helsinki
PO Box 54, 000 14 University of Helsinki
(+358) 503605191

ABSTRACT

The ongoing assessment in a digital educational game is a prerequisite for providing in-game adaptations aiming to enhance the learner´s current emotional state. Our assessment procedure combines the results of two modalities: i) by interpreting the learner´s interactions with the virtual environment, and ii) by psychophysiological patterns covered by facial electromyography and electrodermal activity. The combination of signals and indicators is described separately for each modality and completed by details on the interpretation of the results, based on intrapersonal comparisons. Finally, we outline how to integrate the results of both modalities into a single value that indicates the current emotional state. We conclude with a short description of how we will validate the overall assessment procedure.

Categories and Subject Descriptors

K.8.0 [**PERSONAL COMPUTING**]: General – *games.*

General Terms

Measurement, Human Factors, Theory.

Keywords

Psychophysiological Measurements, Implicit Measurements, Emotion, State.

1. INTRODUCTION

The TARGET project aims to develop a new technology enhanced learning (TEL) platform to support competence development in the learning domains of project and innovation management and global sustainable manufacturing (http://www.reachyourtarget.org). The major component of the platform is a digital educational game (DEG). This DEG consists of game scenarios possessing critical incidents from the learning domains in order to support the knowledge transfer between the virtual world and the learner`s ordinary working life. Within this DEG, the learner is represented by an avatar and interacts with several non-playable characters (NPCs). It is necessary to communicate with different NPCs and to gather pieces of information in order to master the scenario. The learner may use several buttons on the keyboard which have the effect that the avatar expresses specific emotions by facial expressions or body movements. Additionally, several tools are available, for example a *chat tool* to communicate with the NPCs, a *teleport tool* to switch between different locations or a *face cam* which shows the avatar`s own face to enhance awareness of currently expressed emotions.

Even if a given DEG possesses the potential to provide a learning context which is intrinsically motivating and emotionally appealing to engage with, a DEG which *adapts* to the learner´s *current state* is expected to increase the probability that this engagement potential is realised. In TARGET, our aim is to provide in-game adaptations [7] tailored to optimize the learner´s current achievement motivation, cognitive workload, problem representation and emotional state for an efficient and sustainable learning process. The prerequisite for appropriate adaptations is a valid assessment of these constructs.

According to [2] emotion and related constructs encompass three aspects: i) a subjective cognitive experience (e.g. assessed by questionnaires), ii) behavioural expressions (i.e. actions and behavioural patterns assessed by implicit techniques) and iii) psychophysiological patterns. Since an explicit assessment by means of a short questionnaire appearing in regular time intervals would most likely disturb the learner`s flow experience [3], we aim to continuously assess the latter two aspects: behaviour and psychophysiology.

ICMI’11, November 14–18, 2011, Alicante, Spain

Such a multi-modal assessment depends on the two modalities being able to share a common model of a given construct, or else translate between different models that purport to describe the same construct. We apply the circumplex model of emotion proposed by [9], which consists of the two independent dimensions of *pleasantness* and *activation*. Pleasantness is a continuous bipolar dimension between pleasant and unpleasant. Activation, also called arousal, is a continuous unipolar dimension with the poles low and high activation (arousal).

This paper focuses on the description of two modalities to assess the learner's emotional state. The first modality is based on modelling and interpreting the learner´s interactions with the virtual environment via keyboard and mouse. The second modality uses psychophysiological signals which have been empirically linked to emotional states. We will describe the signals and indicators for both modalities, their theoretical basis and how they are combined and interpreted. This is a work-in-progress and as evaluation is on-going the paper will remain at a schematic level with no results section. However, we conclude by outlining how we will validate the assessment of both modalities.

2. MEASUREMENTS

In the following, we describe the different indicators and signals separated for both sources, starting with the implicit assessment technique.

2.1 Behavioural Indicators

The implicit behavioral indicators are based on the interpretation of a learner´s actions and interactions within the virtual environment [1]. A set of behavioural indicators for the assessment of emotion as state is provided in table 1. Some of them are primarily related to *activation*, such as click rate (#1) or inactivity (#9). Others are primarily related to pleasantness, such as using defined keys to express positive or negative emotions (#7 and #8). In order to continuously compute each indicator, the game play is divided into consecutive and equally long time slices, lasting for e.g. 30 seconds. All indicators are operationalized and described in detail in [1]. The indicators 10 to 14 are directly derived from the theory of information foraging [12], which describes the strategies that people apply to search and gather information.

Table 1. Set of Behavioural Indicators

#	Behavioural Indicator
1	Click rate
2	Length of mouse movements
3	Relative exploitation of available tools
4	Frequency of tool-usage (of each available tool)
5	Frequency of communication tool-usage
6	Frequency of interactions with NPCs
7	Frequency of expressing positive emotions via keys
8	Frequency of expressing negative emotions via keys
9	Inactivity [sec.]
10	Within-patch processing [sec.]
11	Between-patch Processing [sec.]
12	Extent of NPC-interactions weighted by amount of Within-Patch processing
13	Information gained
14	Rate of information gain

In the theory of information foraging, human search behaviour is regarded as adaptive to the environment to gain information from external sources as effectively and efficiently as possible. External sources are called patches (e.g. online documents). An ideal information forager maximizes the rate of gaining valuable information by seeking for a balanced ratio between explorative and exploitative search behaviour. Available time needs to be divided into the search for new sources bearing valuable information (*Between-Patch processing*) as well as into elaborated processing of these patches to extract relevant information (*Within-Patch processing*). By concentrating solely on one single patch (e.g. a single paper) valuable information from external resources does not become available. Contrariwise, solely explorative search will lead to ignorance of important details.

Even if the theory of information foraging has been initially developed in the context of navigation on the web, we apply the principles and adapt some of the indicators to the area of games because of two reasons: i) the learner has to search for and to communicate with several NPCs in order to collect all information necessary to master the game scenarios and ii) it is assumed that a successful information forager experiences a positive emotional state more often than an unsuccessful one.

2.2 Psychophysiological Signals

For a thorough review of psychophysiological methods for game-based experiments see [8]. Such methods are useful for objectively examining game experiences, because the physiological processes measured are mostly unwilled. Thus these measures have several advantages over traditional self-report: (i) measurements can be performed continuously with high temporal resolution; (ii) processes of interest can be covertly assessed; and (iii) these measures provide information on emotional and motivational processes that are not available to conscious awareness [13].

Facial electromyography (EMG) provides a direct measure of the electrical activity associated with facial muscle contractions that are an important form of emotional expression [15]. A number of studies have shown that the processing of unpleasant emotions is associated with greater activity over the corrugator supercilii (brow) muscle region and that processing pleasant emotions prompts greater activity over the zygomaticus major (cheek) muscle region [10][14][18]. In addition, increased activity at the orbicularis oculi (periocular) muscle area is involved in the expression of enjoyment smile and genuine pleasure [5]. Several studies have shown that orbicularis oculi activity is particularly high during pleasurable high-arousal emotions [14][18].

Electrodermal activity (EDA), commonly known as skin conductance, is an important physiological index of arousal and is innervated entirely by the sympathetic nervous system [4]. Several studies using the picture-viewing paradigm have shown that EDA is highly correlated with self-reported emotional arousal [10]. That is, arousing pictures of either valence result in increased EDA as compared to low-arousal pictures.

In the laboratory and in deployment, these signals are measured by the Varioport-ARM mobile psychophysiological data acquisition system (Becker Meditec, Karlsruhe, Germany). This device is lightweight and battery operated, and thus, ideal for deployment on-site with clients.

3. INTEGRATION & INTERPRETATION

In the following, we describe how to combine the different signals and indicators for each modality, how to interpret the results in terms of high and low values and finally, how to combine the results of both modalities.

3.1 Integrating Measurements per Modality

The reason for integrating the measurements for both modalities separately in a first step is that not every learner may have access to the Varioport device. In consequence, the validity of each assessment technique has to be evaluated independently.

For the implicit assessment technique we will apply a multiple regression model as suggested by [11]. A linear model would be the simplest case; however, we will evaluate whether it delivers better results (i.e. if it explains more variance) than other models (such as a Generalized Linear Model which can *incorporate* non-linear covariates in the coefficients). For each time slice and both dimensions of the emotional model, i.e. activation and pleasantness, a regression equation in the following form has to be calculated:

$$x_i = d_i + w_{1i} * BI_{1i} + ... + w_{ji} * BI_{ji} + ... + w_{14i} * BI_{14i} \qquad (1),$$

with x_i as the initial value for the dimension i in the given time slice, d_i as the constant intercept, BI_{ji} as the "raw-values" for the 14 predictors (i.e. the behavioural indicators) and finally, w_{ji} as the predictors' weights. The intercepts and weights for both regression equations are to be conducted by means of a validation study which is briefly outlined in section 4.

The integration of psychophysiological signals involves finding suitable functions to characterise all signals on the same scale so that they are comparable. Both EDA and EMG transforms begin by pre-processing. A low pass Butterworth filter is applied where computational power allows (i.e. when the requisite frequency domain transform can be performed by a hardware routine – this is available on the Varioport device, but may not be on others).

3.2 Interpretation

The interpretation of the current emotional state as indicated by the initial values x_i in terms of an optimal or suboptimal state depends on their comparison with a baseline. The baseline can be conducted either by the learner´s values in previous time-slices of the same game scenario (intrapersonal comparison) or by the values of other learners. The latter approach is feasible when an extensive database for interpersonal comparison is available. However, we prefer an intrapersonal comparison which takes the learner´s *gaming history* into account since individual learner`s baselines may differ to a great extent.

For the implicit assessment technique we standardize both initial values by a z-transformation:

$$z_i = (x_i – M_i) / SD_i \qquad (2),$$

with z_i and x_i as standardized and initial values of the dimension i, respectively. The average values of the dimension i represented by M_i is computed by averaging the initial values x_i of all *previous* time-slices. Thus, M_i does not incorporate the current time-slice. The standard deviation of initial values of all previous time slices is indicated by SD_i. Since the reliability of the standard deviation depends on the amount of incorporated data, the deviation of the current time-slice from the averaged previous time-slices in terms of standard deviations is not taken into account until the fourth time-slice has passed. Hence, the computation of the standardized values begins 120 seconds after the learner starts to play the scenario.

Finally, in order to gather manifestations of a continuous variable, whose values range from 0 and 1, the standardized value z_i is inserted into the following logistic function:

$$p(z_i) = 1 / (1 + e^{-z_i}) \qquad (3),$$

with $p(z_i)$ indicating the probabilistic value of the dimension i. The logistic function is positively accelerated and differentiates primarily in a range between -3 and +3.

With regards to the psychophysiological modality, after applying the low pass Butterworth filter, EDA continues by splitting the signal into tonic and phasic components. Tonic values are obtained at each time t by estimating the mean of the signal over the previous 8 seconds (8 is a standard value reflecting the speed of EDA reactivity). Phasic values are then obtained by subtracting the tonic from the overall at time t. The EMG transformation proceeds by rectification.

Thereafter, for each signal, our relative approach is based on comparing a current value at time t to the average from a time-slice of n previous values. The size of a change which constitutes an interesting variation (as opposed to minor fluctuations in the signal), and thus reflects actual change in the player's underlying psychological state, is Δ. The method of indexing psychophysiological signals involves determining a general value for n and for Δ, for each signal. Currently, the learning part of our approach is offline so that a learned model will be fixed for every use of the system. Offline learning involves statistically estimating the overall fit of each combination of n and Δ values using General Estimating Equations on a dataset of live recordings from presentations of emotionally evocative stimuli to subjects.

With continuous psychophysiological data for both emotional dimensions (i.e., pleasantness and arousal), it is possible to combine these data to 'locate' the subject more specifically in emotional state space. That is, a stressed emotional state is characterized by a combination of displeasure (e.g., high corrugator EMG activity) and high arousal (e.g., high EDA); a positively excited state is characterized by a combination of pleasure (e.g., high zygomatic EMG activity) and high arousal (e.g., high EDA); depression/boredom is characterized by combined displeasure (e.g., high corrugator EMG activity) and low arousal (e.g., low EDA); and positive relaxation is a combination of pleasure (e.g., high zygomatic EMG activity) and low arousal (e.g., low EDA).

This is to be interpreted as a probable tendency toward the stated emotion, rather than an exact emotional state per se. Probability can be measured by the distance of the combined signals from their theoretical maximum.

3.3 Integrating Results of both Modalities

The procedure to integrate the results of both modalities is similar to the multiple regression equation (1) above. The difference is that there would not be a constant intercept and only two pairs of predictors and appropriate weights (one pair for each modality). The weights will be fixed by the amount of variance of an external criterion the predictor explains relative to the amount of shared variance R^2 of both predictors and the external criterion. Put simply, the higher the predictive validity of a modality (predictor), the higher its weight and contribution to the final assessment result.

4. OUTLOOK: EVALUATION

For the evaluation of the indicators' validity we adopt approaches suggested by [6] and [16] in order to have a non-invasive measurement procedure and to elicit (as a dependent measure) the third remaining part of the emotional trinity, i.e. the subjective cognitive experience. This aspect of an emotional state can be measured by self-report and will be used as an external criterion to be compared with values of the physiological and behavioral indicators.

The self-report will be mediated by a pop-up screen intermittently occurring during game-play and displaying small sets of (at most four) items about the current emotional state, which can be taken for example from the PANAS Scale [17]. The learner can respond through slider scales: by moving the position of a slider between two poles of a graphical intensity-dimension, the learner indicates the extent to which she or he respectively agrees or disagrees on a particular item. The different questions or items will cover different aspects of the two emotional dimensions and will be selected randomly for presentation.

Finally, a regression analysis will be conducted to determine the nature and significance of the relationship between the indicators and the self-report. Standardized Beta-Coefficients will support the identification of valid indicators as well as their weights for equation (1). Additionally, the correlation of each modality with the external criteria delivers the basis for equation (2) which accounts for a substantial amount of emotional variance. The regression equation can be cross-validated by comparing the predicted with the empirical results derived from an additional sample.

5. ACKNOWLEDGMENTS

This paper is part of the EC-Project TARGET funded by the Seventh Framework Programme of the European Commission (ICT-231717). The authors are solely responsible for the content of this paper.

6. REFERENCES

[1] Bedek, M. A., Seitlinger, P., Kopeinik, S. and Albert, D. Multivariate Assessment of Motivation and Emotion in Digital Educational Games. *5th European Conference on Games Based Learning*, (Athens, Greece, 2011), in press.

[2] Bradley, M. M. and Lang, P. Measuring emotion: behavior, feeling and physiology. in Lane, R. and Nadel, L. eds. *The Cognitive Neuroscience of Emotion*, Oxford University Press, New York, 2000, 242-276.

[3] Csikszentmihalyi, M. *Flow: The psychology of optimal experience*. Harper & Collins, New York, 1990.

[4] Dawson, M. E., Schell, A. M. and Filion, D. L. *The electrodermal system*. Cambridge University Press, Cambridge, 2000.

[5] Ekman, P., Davidson, R. J. and Friesen, W. V. 1990. The Duchenne smile: emotional expression and brain physiology. II. *J Pers Soc Psychol*, *58* (2), 342-353.

[6] Insko, B. E. Measuring Presence: Subjective, behavioral and physiological methods. in Riva, G., Davide, F. and Ijsselsteijn, W. A. eds. *Being There: Concepts, effects and measurement of user presence in synthetic environments*, IOS Press, Amsterdam, 2003, 110-118.

[7] Kickmeier-Rust, M. D. and Albert, D. 2010. Micro Adaptivity: Protecting Immersion in Didactically Adaptive Digital Educational Games. *J Comp Assist*, *26* (2), 95-105.

[8] Kivikangas, J. M., Chanel, G., Cowley, B., Ekman, I., Salminen, M., Järvelä, S. and Ravaja, N. 2011, in press. A review on the use of psychophysiological methods in game research. *Journal of Gaming and Virtual Worlds*.

[9] Larsen, R. J. and Diener, E. Promises and problems with the circumplex model of emotion. in Clark, M. S. ed. *Review of personality and social psychology: Emotion*, Sage, Newbury Park, 1992, 25-59.

[10] 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* (3), 261-273.

[11] Margolis, M. J. and Clauser, B. E. A regression-based procedure for automated scoring of a complex medical performance assessment. in Williamson, D.W. Mislevy, R. J. and Bejar, I. I. eds. *Automated scoring of complex tasks in computer-based testing*, Lawrence Erlbaum Assoc, Mahawah, 2006, 123-168.

[12] Pirolli, P. and Card, S. K. 1999. Information Foraging. *Psychol Rev*, *106* (4). 643-675.

[13] Ravaja, N. 2004. Contributions of Psychophysiology to Media Research: Review and Recommendations. *Media Psychol*, *6*(2), 193-235.

[14] Ravaja, N., Saari, T., Kallinen, K. and Laarni, J. 2006. The Role of Mood in the Processing of Media Messages From a Small Screen: Effects on Subjective and Physiological Responses. *Media Psychol*, *8*(3), 239-265.

[15] Tassinary, L. G., & Cacioppo, J. T. The skeletomotor system: Surface electromyography. in J. T. Cacioppo, L. G. Tassinary & G. G. Berntson eds. *Handbook of psychophysiology*, Cambridge University Press, Cambridge, 2000, 163–199.

[16] Van Reekum, C. M., Johnstone, T., Banse, R., Etter, A., Wehrle, T. and Scherer, K. R. 2004. Psychophysiological responses to appraisal dimensions in a computer game. *Cognition and Emotion*, *18*(5), 663-688.

[17] Watson, D., Clark, L. A. & Tellegen, A. 1988. Development and validation of brief measures of positive and negative affect: The PANAS Scale. *J Pers Soc Psychol*, *54*, 1063-1070.

[18] Witvliet, C. V. and Vrana, S. R. 1995. Psychophysiological responses as indices of affective dimensions. *Psychophysiology*, *32*(5), 436-443.