An EDM-based Multimodal Method for Assessing Learners' Affective States in Collaborative Crisis Management Serious Games

Ibtissem Daoudi
ENSI, COSMOS Laboratory
University of Manouba
ibtissem.daoudi@ensi-uma.tn

Erwan Tranvouez
LIS Laboratory
Aix-Marseille Université
erwan.tranvouez@lis-lab.fr

Raoudha Chebil
ENSI, COSMOS Laboratory
University of Manouba
raoudha.chebil@ensi-uma.tn

Bernard Espinasse
LIS Laboratory
Aix-Marseille Université
bernard.espinasse@lis-lab.fr

Wided Lejouad Chaari
ENSI, COSMOS Laboratory
University of Manouba
wided.chaari@ensi-uma.tn

ABSTRACT
Recently, Crisis Management Serious Games (CMSG) have proved their potential for teaching both technical and soft skills related to managing crisis in a safe environment while reducing training costs. In order to improve learning outcomes insured by CMSGs, many works focus on their evaluation. Despite its great interest, the learner emotional state is often neglected in the evaluation process. Indeed, negative emotions such as boredom or frustration degrade the learning quality since they frequently conduct to giving up the game. This research addresses this gap by combining gaming and affect aspects under an Educational Data Mining (EDM) approach to improve learning outcomes. Therefore, we propose an EDM-based multimodal method for assessing learners' affective states by classifying data communicated in text messaging and facial expressions. This method is applied to assess learners' engagement during a game-based collaborative evacuation scenario. The obtained assessment results will be useful for adapting the game to the different players' emotions.

Keywords
Serious game, crisis management, assessment, educational data mining, affective states, multimodal emotion detection.

1. INTRODUCTION
Recently, Serious Game (SG) development and usage have increased to improve learning benefits and to increase learners' motivation [1]. SGs applications reach out several domains such as crisis management, education, ecology, and health-care [2]. Indeed, collaborative Crisis Management Serious Games (CMSG) have proved their potential for teaching concepts related to managing different types of crisis situations such as natural disasters (earthquakes, floods), man-made disasters (terrorist attacks, pollution), and technological crises (industrial accidents, cyber attacks) in a fun way while reducing training cost and saving time [3].

Ibtissem Daoudi, Erwan Tranvouez, Raoudha Chebil, Bernard Espinasse and Wided Lejouad Chaari "An EDM-based Multimodal Method for Assessing Learners' Affective States in Collaborative Crisis Management Serious Games" In: *Proceedings of The 13th International Conference on Educational Data Mining (EDM 2020)*, Anna N. Rafferty, Jacob Whitehill, Violetta Cavalli-Sforza, and Cristobal Romero (eds.) 2020, pp. 596 - 600

Despite its obvious interest, the exploitation of the SG concept in learning processes is not always a guarantee of its effectiveness [4]. As any learning systems, SGs rely on the implicit alignment of the learning outcomes (knowledge or skills) and the game experience (engagement, motivation). In particular, the effectiveness of a collaborative CMSG depends on different learners' characteristics including cognitive, emotional and social aspects [4]. Consequently, there has been a lot of research focused on the evaluation of SGs and their effectiveness for Crisis Management (CM) training varying in terms of crisis situation, number of players, key indicators or characterization of learners [5,6,7,8,9,10]. However by studying the state of the art, we have noticed that there is a considerable lack of studies integrating the concept of affective computing, especially learners' affective states, in the evaluation process within collaborative CMSGs [4]. Besides, most of existing works use explicit techniques for analyzing learners' behaviors during playing like pre/post questionnaires, interviews and debriefing sessions. These techniques represent a subjective evaluation that relies on non-exhaustive players' opinions and disrupts the high level of engagement provided by the game; impacting thus negatively the accuracy of evaluation results [30]. So, improving players' engagement (and thus learning outcomes) requires detecting and assessing such emotional states in a non-intruding way [11].

In this paper, we focus on addressing this gap. In doing so, we focus on the detection and analysis of learners' emotions expressed in textual and visual data to infer Flow game-play experience indicator (also called engagement) in collaborative CMSGs. To the best of our knowledge, players' engagement measure and impact on learning outcomes have not been investigated in such context. Hence, our contribution is to propose an emotion-based EDM method able to:

1) Assess the temporal dynamics of learners' affective states during a game-based session for CM training.
2) Evaluate their final states at the end of training process by classifying data communicated in text messaging and facial expressions.

3) Explore the final individuals' affective profiles to generate the group emotion at a global level.

The rest of this paper is organized as follows. Section 2 presents the proposed EDM-based multimodal method for learners' affective states assessment. Section 3 reports the application of our method on a collaborative CMSG used as a case study. Section 4 discusses our major findings. Section 5 summarizes the paper and presents our plans for future work.

2. AN EDM-BASED MULTIMODAL METHOD FOR ASSESSING LEARNERS' AFFECTIVE STATES

Our aim is to develop an automatic method for assessing learners' affective states (*engagement*, *frustration*, *confusion*, and *boredom*) using facial expressions and text analysis in collaborative CMSGs. To reach this objective, we need to perform five main steps corresponding to specific tasks namely *data collection*, *data annotation*, *data fusion*, *data analysis*, and *data visualization* as illustrated in Figure 1.

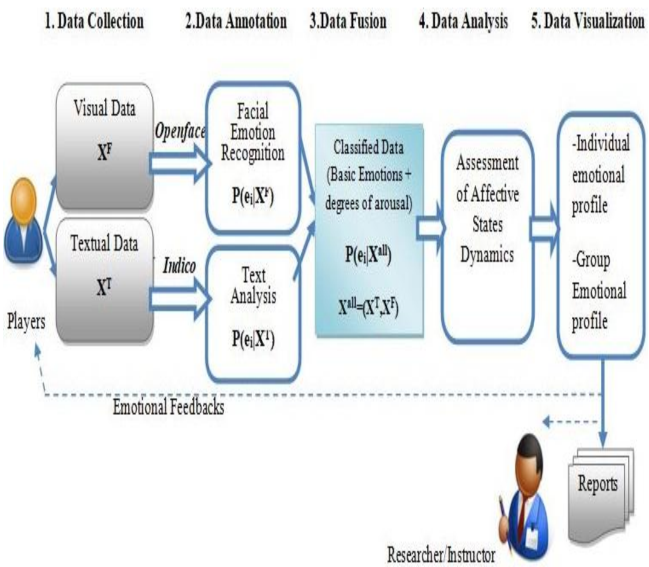

Figure 1 General Overview of the Proposed Method

2.1 Data Collection and Annotation Task

In order to collect data in a way that is more efficient and less intrusive compared to physiological measurements [12], we plan to extract the *messages* exchanged between players as well as the *video records* of learners' faces produced in real-time using a Webcam. These two kinds of data are annotated as follows:

- ***Text annotation***. The textual content of these messages represents a rich source to detect their emotions that are revealed by the annotator tool *indico.io.* It is a predictive analytics tool classified as one of the top AI APIs for emotion detection from raw text strings (shorter instances of text like conversations) using deep learning algorithms with 93.5% of accuracy [13]. The API gives as an output the probability that the text reflects the basic emotions as well as their intensities.
- ***Video annotation***. We have adopted *Openface 2.0*[14]: an automatic facial behavior analysis and understanding toolkit. Openface returns intensity and presence for each Facial Action Unit (FAU) estimated with several computer vision and machine learning algorithms. We exploit the output of FAU recognition system since it displays emotions according to [15]. Based on the EMotional Facial Action Coding System [16], mapping rules associate couples of FAU with basic emotions. For example, *joy* is associated with detection of *Cheek raiser* FAU and *Lip corner puller* FAU.

2.2 Data Fusion Task

In our study, we perform a multimodal fusion at the decision-level which refers to the process of combining data collected from many modalities after being pre-classified independently to obtain the final classification. In fact, each classified modality, using the previous annotators, provides one hypothesis on labeled emotion categories; and this integration method gives a global estimate based on partial results [17].

$X^{all} = (X^T, X^F)$ represents the global feature vector consisting of the text feature vector, X^T, and the face feature vector, X^F.

In decision-level fusion, two separate classifiers provide the posterior probabilities $P(e_i|X^T)$ and $P(e_i|X^F)$ for text and face, respectively, having to be combined into a single posterior probability $P(e_i|X^{all})$; where e_i represents one of six possible classes of basic emotions (e_1=*joy*, e_2=*sadness*, e_3=*surprise*, e_4=*anger*, e_5=*fear and* e_6=*disgust*).

The face modality is assumed to be the main modality in our multimodal approach (but the text modality is not neglected). Hence, we assign weights as follows: $\mu_T=0.3$ for the text modality and $\mu_F=0.7$ for the face modality. We adopt this weighting proposed and validated by works referenced by [18] and [19]. Then, we apply the averaging formula using these weights in order to compute the average probability of the two modalities defined as follows [20]:

$$P(e_i|X^{all} = X^T and\ X^F) = \left(\mu_T * P(e_i|X^T) + \mu_F * P(e_i|X^F)\right) / 2$$

2.3 Data Analysis Task

In this task, we perform a fine-grained analysis of the dynamics of learners' affective states based on facial features during playing by studying the impact of stress on affective transitions, and we produce a summative evaluation of their emotional states at the end of training process:

- ***Stress detection.*** The stress is one of the most frequently occurring emotions inherent to CM since it affects the actors' way to manage crisis situations [4]. Given stress is related to emotions; also facial expressions have been used to detect stress by linking some of basic emotions as features [21]. In fact, many works have proved that, in different contexts like driving and working environments, stress is detected if either anger, fear, or a combination of these two negative emotions is detected constantly within a fixed time interval [22,23]. In particular, they focus on some specific FAU and their activation level extracted in each video frame, described as an indicator for fear and/or anger.
- ***Mapping between affective states and basic emotions.*** Affective states are particular combinations of basic emotions as demonstrated by [24] using association rules mining. In our study, we adopt the existing mapping as described in [24, 19]; and we propose some novel interpretations of basic emotions combinations

allowing us to deduce affective states based on existing theories of emotions [15,26].

Flow/*engagement* is defined by a high level of *surprise* and a low *sadness* level [19]. Since joy and sadness are opposite emotions as validated by [15] and flow is characterized by a full involvement and enjoyment in the activity [26], we can affirm that *flow* can be defined also by a high level of *surprise* and a high *joy* level.

Frustration is detected at the presence of a high degree of *anger* and a low degree of *joy* [19]. Likewise, frustration can be defined by a high level of *anger* and a high *sadness* level. Moreover, basing on the definition of frustration state [26], it can be mapped to a high level of fear as well as a high level of sadness. In the same way, frustration can be defined by a high level of *fear* and a low level of *joy*.

Boredom can be mapped to a high level of *disgust* as well as a low level of *joy*. In the same manner, *boredom* can be defined by a high level of *disgust* and a high *sadness* level [19].

The state in which *all the levels of six basic emotions* are low will represent the *confusion* affective state [19].

2.4 Data Visualization Task

This final task concerns visualization of our analysis results at two levels: individual and global. On the one hand, we visualize the summative individual emotional profiles which contain relevant information about affective states expressed by each player at the end of training process by selecting the dominant and the most pronounced emotion. On the other hand, we visualize the aggregation of all individual emotional profiles based on a decision tree algorithm to decide on the *polarity* of global emotion (*positive* or *negative*) and then to constitute the group emotion [27]. So, we apply *J48 decision tree classifier*, an implementation of *C4.5 algorithm in Weka*, to generate a decision on the group emotion based on individual affective states with a default confidence value=0.25. The principle is to decide the class label of group emotion (*positive* or *negative*) by learning decision rules inferred from training data (rates of individuals affective states). According to our experimental results, this tree-based method reaches an accuracy of *81%* using *5-fold cross-validation*. Figure 2 shows the decision tree model of group emotion.

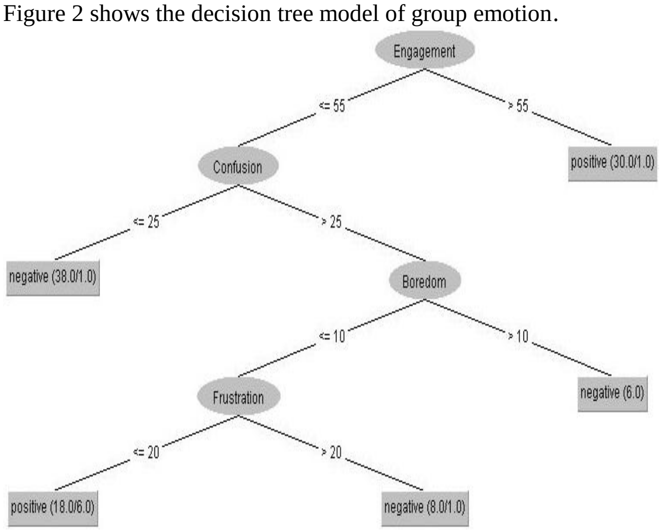

Figure 2 Decision Tree Model of Group Emotion

3. EXPERIMENT AND RESULTS

3.1 Game Description

We have developed a collaborative scenario for building evacuation training in case of a fire emergency situation. This scenario is implemented on the iScen software platform [29], specifically intended for crisis simulation, management and training. The scenario aims to train people (staff or students) of a Tunisian university building on evacuating all the present persons during a fire emergency triggered in the coffee shop as shown in Figure 3. The evacuation exercise involves a group of 30 participants (including player and virtual characters) having different roles namely *coordinator*, *security responsible*, *firefighter*, *warden* and *deputy* who must collaborate and coordinate their actions in order to manage an emergency evacuation procedure. This scenario allows learners to reach two main pedagogical objectives consisting of: (1) acquiring personal fire safety skills both in general and specifically in a university context, and (2) teaching best evacuation practices required to manage any fire emergency in an efficient and rapid manner.

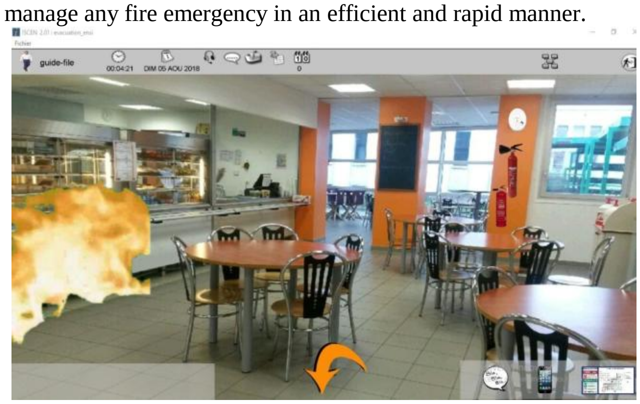

Figure 3 Screen Capture of Crisis Situation

3.2 Dataset

We analyze (n=30) students' behavior trace interaction data obtained after the game session that lasts approximately 25 minutes. These data are video recordings of the participants' faces while playing using webcams as well as exchanged messages using the text chatting system. We analyze affective dynamics experienced by participants by tracking emotions at a fine-grained level using facial features**.** We make judgments on what affective states were present in each 20-second interval basing on the mapping described above. In addition, pre- and post-test questionnaires were completed individually by all students before starting the game session (pre-test) and immediately after finishing it (post-test). Pre-test questions address personal information concerning prior game experience and CM knowledge. Learners are then categorized as *novice*, *intermediate* and *expert* to be confronted afterward to the experimental results. Post-test questions aim to measure the level of engagement based on the Game Engagement Questionnaire. Both pre-test and post-test are on a 5-point Likert scale ranging from 1 (not at all) to 5 (extremely).

3.3 Obtained Results

Comparing to several predictions proposed by the Cognitive Disequilibrium Model [24], it appeared that some of these predictions have been validated while others not addressed by the model are identified by our method. This model addresses transitions between affective states of learners while solving complex activities in relatively short learning sessions [24].

The supported predictions include the transitions from the state of engagement into confusion, confusion into frustration, and frustration into boredom which naturally occurred. In fact, analyzing transitions between affective states are so important because they provide insight into how learners enter into an affective state since engagement and confusion is correlated with higher performance, while frustration and boredom are correlated with poorer performance.

The two predictions that have been identified, but were unexpected in the model, include the transitions from frustration to confusion and boredom to frustration. First, even though the transition from frustration into confusion occurred rarely, we believe that some frustrated participants, could view the situation as a challenge and become more energized; and ultimately enter the confusion state while trying to resolve the current misunderstanding. Second, the transition from boredom into frustration occurred significantly when we detect a high activation level of some FAU characterizing the stress emotion. To resume, our findings suggest that some aspects of the cognitive disequilibrium model might need refinement and some transitions can occur due to a specific characteristic of the context of CM training namely the stress.

When aggregated across the all participants at the end of training process, our results indicated that 25% of learners felt engagement, 50% expressed boredom, 25% felt frustration, and 0% experienced confusion. Figure 4 displays a global view of the all individual affective states. This global view allowed us to decide the polarity of group emotion by applying our decision tree model. Hence, we can deduce that the global emotion is negative *(25% engagement + 0% confusion + 50% boredom + 25% frustration=> negative class).*

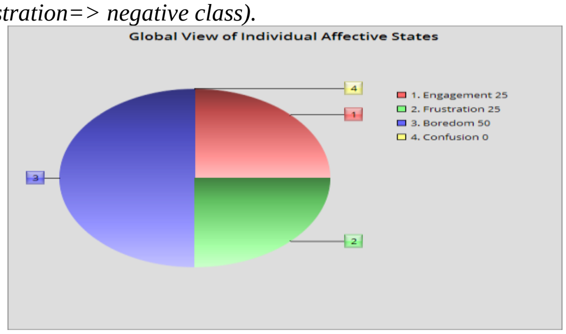

Figure 4 Global View of Individual Affective States of Players

3.4 Results Validation

For validation purpose, summative evaluation results are compared to the answers to the Game Engagement Questionnaire (GEQ) analyzing self-reported subjective descriptions and collected after the game session. This questionnaire is one of the most commonly used self-report questionnaires in the player experience field for measuring engagement specifically elicited while playing games [28]. The core module of GEQ is a 33-item scale which is designed to measure game players' experience across seven dimensions namely *Immersion, Flow, Competence, Positive Affect, Negative Affect, Tension,* and *Challenge*. Dimension scores are computed as the average value of their items. The descriptive statistics obtained from learners responses are reported in Table 1.

Table 1 Descriptive statistics for dimensions of the GEQ

Dimension	Mean	Standard deviation	Max	Min
Immersion	2.13	0.60	4.00	1.00
Flow	2.34	0.62	3.55	1.13
Positive affect	2.00	0.50	3.00	1.00
Negative affect	4.28	0.85	5.00	2.00
Tension	3.96	0.77	5.00	1.65
Challenge	3.43	0.68	5.00	2.00
Competence	2.43	0.56	3.00	1.20

4. DISCUSSION

As shown in Table 1, positive feelings are much less severe and less frequently experienced compared to negative feelings (lower than the mid-value of the scale). In fact, participants reported the level of *positive affect* to be low (2.00). More specifically, results analysis shows that *immersion* (reflecting how players felt strongly connected with the game) and *flow* (indicating whether players lost track of their own effort and/or the passage of time during the game) receive respectively average degrees (2.13 and 2.34). The dimension *negative affect* receives the highest value of all (4.28). This result indicates that playing the game engendered some negative emotional experiences in particular boredom. In addition, participants experience a certain high degree of *tension* (3.96) in the form of specific negative emotions like frustration. Moreover, in terms of *challenge*, participants report that the game environment is difficult and challenging (3.43) according to their level of *competence* (2.43). All these results confirm the negative group emotion detected after the application of our method on the same CM scenario. Basing on this result, we can conclude that the team performance is also negative. In fact, this interpretation can be explained by the fact that all participants are situated, for the first time, in an emergency evacuation procedure based on a virtual training environment. It can also be a consequence of limited learners' guidance and assistance carried out by the instructor during the training process in order to better achieve the game objectives.

To summarize, the final affect annotations obtained via our method correlate well with subjective responses to the GEQ. In comparison to the GEQ, our method represents an objective and rapid manner to analyze learners' emotions and to infer their affective states without distracting them from game-play using EDM techniques. Hence, our contribution is intended to support learning, maintain motivation, and increase learners' engagement in the virtual word of the game.

5. CONCLUSION AND FUTURE WORK

This paper investigates a multimodal learner analytics approach to assess emotional states in collaborative CMSGs. Specifically, decision tree models were trained to predict learners' affective states utilizing bimodal data including textual messages and facial expressions. Affective states predicted by the model are evaluated with learners' self-reported engagement scores reported after the game session. In future work, we want to extend this study to a larger sample of participants within another multi-players CMSG which is currently under development using Unity 3D game engine. Furthermore, we plan to analyze the quality of social interactions during a collaborative game session in order to more deeply understand the dynamics of affective states over time.

6. ACKNOWLEDGMENTS
The authors would like to thank EVERSIM (Simulation & Serious Games) society for providing them the iScen platform and assisting them in its use.

7. REFERENCES
[1] Aghababyan, A. 2014. E3: Emotions, Engagement and Educational Games. *International Educational Data Mining Society.*

[2] Daoudi I., Chebil R. and Lejouad Chaari W. 2018. A Novel Tool to Predict the Impact of Adopting a Serious Game on a Learning Process. *In Proceedings of the 20th International Conference on Enterprise Information Systems. 1:585-592.*

[3] Walker, W. E., Giddings, J., and Armstrong, S. 2011. Training and learning for crisis management using a virtual simulation/gaming environment. *Cognition, Technology & Work, 13(3), 163-173.*

[4] Daoudi I., Chebil R., Tranvouez E., Lejouad Chaari W., and Espinasse B. 2017. Towards a Grid for Characterizing and Evaluating Crisis Management Serious Games: A Survey of the Current State of Art. *International Journal of Information Systems for Crisis Response and Management (IJISCRAM), 9, Issue 3, 76-95.*

[5] A. Haferkamp N., and Kraemer N. C. Linehan, C., Schembri, M. 2011. Training disaster communication by means of serious games in virtual environments. *Entertainment Computing, 2(2), 81-88.*

[6] Mendez, G., Avramides, K., de Freitas, S., and Memarzia, K. 2009. Societal impact of a Serious Game on raising public awareness: the case of FloodSim. *In Proceedings of the ACM SIGGRAPH Symposium on Video Games,15-22.*

[7] Oulhaci M.A., Tranvouez E., Fournier S., and Espinasse B. 2015. Improving Players' Assessment in Crisis Management Serious Games: The SIMFOR Project. *In Information Systems for Crisis Response and Management in Mediterranean Countries, 85-99.*

[8] Taillandier F., and Adam C. 2018. Games Ready to Use: A Serious Game for Teaching Natural Risk Management. *Simulation & Gaming. 49, 441-470.*

[9] Silva, V., Dargains, A., Felício, S., and Carvalho, P. and al. 2014. Stop disasters: serious games with elementary school students in Rio de Janeiro. *In 8th International Technology, Education and Development Conference,1648-1659.*

[10] Theo van, R., Igor, M., and Mark de, B. 2015. Multidisciplinary coordination of on-scene command teams in virtual emergency exercises. *International Journal of Critical Infrastructure Protection, 9, 13-23.*

[11] Shute, V. J. 2011. Stealth assessment in computer-based games to support learning. *Computer games and instruction, 55(2), 503-524.*

[12] Guthier B., Dörner R., and Martinez H.P. 2016. Affective Computing in Games. *Entertainment Computing and Serious Games, 9970, 402-441.*

[13] https://indico.io/blog/docs/indico-api/text-analysis/

[14] Tadas B., Amir Z., Yao C.L., and Louis-Philippe M. 2018. OpenFace 2.0: Facial Behavior Analysis Toolkit. *13th IEEE International Conference on Automatic Face & Gesture Recognition.*

[15] Ekman, P. 1999. Basic emotions. *In T. Dalgleish & M. J. Power (Eds.),* Handbook of cognition and emotion, *45-60. New York, NY, US: John Wiley & Sons Ltd.*

[16] Friesen, W., and Ekman, P. 1983. EMFACS-7: Emotional Facial Action Coding System. *Unpublished manual, University of California. https://www.paulekman.com/*

[17] Soujanya P., Erik C., Rajiv B., and Amir H. 2017. A review of affective computing: From unimodal analysis to multimodal fusion. *Information Fusion, 37, 98-125.*

[18] Tan C. T., Rosser D., Bakkes S., and Pisan Y. 2012. A Feasibility Study in Using Facial Expressions Analysis to Evaluate Player Experiences. *In Proceedings of The 8th Australasian Conference on Interactive Entertainment: Playing the System, New York, NY, USA.*

[19] Ramin T., Ashish A., Troy M., and Sethuraman P. 2018. Real-time stealth intervention for motor learning using player flow-state. *IEEE 6th International Conference on Serious Games and Applications for Health (SeGAH).*

[20] Kuncheva L. I. 2002. A theoretical study on six classifier fusion strategies. *IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (2), 281 –286.*

[21] Aigrain, J., Spodenkiewicz, M., Dubuisson, S., Detyniecki, M., Cohen, D., and Chetouani, M. 2016. Multimodal stress detection from multiple assessments. *IEEE Transactions on Affective Computing.*

[22] Alberdi, A., Aztiria, A., and Basarab, A. 2016. Towards an automatic early stress recognition system for office environments based on multimodal measurements: A review. *Journal of biomedical informatics, 59, 49-75.*

[23] Gao, H., Yüce, A., and Thiran, J. P. 2014. Detecting emotional stress from facial expressions for driving safety. *In IEEE International Conference on Image Processing (ICIP), 5961-5965.*

[24] D'Mello, S., and Graesser, A. 2012. Dynamics of affective states during complex learning. *Learning and Instruction, 22(2), 145-157.*

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

[26] Csikszentmihalyi M. 1990. Flow: The psychology of optimal experience. *Harper & Row.*

[27] Michalis, F. 2016. A Review of Emotion-Aware Systems for e-Learning in Virtual Environments. *Chapter11, Formative Assessment, Learning Data Analytics and Gamification. In ICT Education Intelligent Data-Centric Systems, 217-242.*

[28] Jeanne H.B., Christine M.F., Kathleen A.C., Evan M. , Kimberly M.B., and Jacquelyn N.P. 2009. The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology, 45, Issue 4, 624-634.*

[29] http://www.i-scen.com/home.php?langue=en

[30] Daoudi I., Tranvouez E., Chebil R., Espinasse B., and Lejouad Chaari W. 2017. Learners' Assessment and Evaluation in Serious Games: Approaches and Techniques Review. In: Dokas I., Bellamine-Ben Saoud N., Dugdale J., Díaz P. (eds) Information Systems for Crisis Response and Management in Mediterranean Countries. ISCRAM-med 2017. Lecture Notes in Business Information Processing, vol 301. Springer, Cham.