

# Towards Affect Recognition through Interactions with Learning Materials

Esam Ghaleb, Mirela Popa, Enrique Hortal, Stylianos Asteriadis and Gerhard Weiss  
Department of Data Science and Knowledge Engineering

Maastricht University  
Maastricht, Netherlands

{esam.ghaleb, mirela.popa, enrique.hortal, stelios.asteriadis, gerhard.weiss}@maastrichtuniversity.nl

**Abstract**—Affective state recognition has recently attracted a notable amount of attention in the research community, as it can be directly linked to a student’s performance during learning. Consequently, being able to retrieve the affect of a student can lead to more personalized education, targeting higher degrees of engagement and, thus, optimizing the learning experience and its outcomes. In this paper, we apply Machine Learning (ML) and present a novel approach for affect recognition in Technology-Enhanced Learning (TEL) by understanding learners’ experience through tracking their interactions with a serious game as a learning platform. We utilize a variety of interaction parameters to examine their potential to be used as an indicator of the learner’s affective state. Driven by the Theory of Flow model, we investigate the correspondence between the prediction of users’ self-reported affective states and the interaction features. Cross-subject evaluation using Support Vector Machines (SVMs) on a dataset of 32 participants interacting with the platform demonstrated that the proposed framework could achieve a significant precision in affect recognition. The subject-based evaluation highlighted the benefits of an adaptive personalized learning experience, contributing to achieving optimized levels of engagement.

**Index Terms**—Affect Recognition, Learning Analytic, Interaction Tracking

## I. INTRODUCTION

Personalized learning is an important topic in education, and therefore a significant amount of research is now driven towards personalized interfaces that can motivate and, actually, support the student in interacting with learning material in a highly engaging manner [12], [33]. While currently there is a huge interest in Technology Enhanced Learning (TEL), such as Massive Open Online Courses (MOOCs), gamification of learning, and blogs, these systems, to a large extent, lack the capability of affect recognition which is a vital step towards achieving personalization. Conversely, in traditional education, teachers are capable of factoring the students’ affective states in their presentation and feedback strategies. Similarly, TEL must be able to understand the affective states of the users.

Goleman in [16], [17] stated that the brain operates to either boost or lower our performance in different domain of abilities such as learning and work. More specifically, he analyzed the relationship between one’s affective state and memory capacity. His research suggests that in educational contexts, stress, anxiety, and frustration experienced by a learner directly affect the learning outcome negatively. This



Fig. 1. Participants interacting with the learning game.

challenge can be tackled through various Machine Learning (ML) techniques [33]. ML’s role relies on providing a timely and an appropriate guidance of students’ cognitive states where high-level affective content can be predicted from low-level human-centered signals, using different sources such as depth and RGB cameras, inertial sensors, speech signals or log-data [1], [3], [4], [6], [21]. Along with these extensively utilized signals, log-data and contextual information of the learner while interacting with TEL can provide an essential channel for affect recognition due to its wide availability and relevance in digital learning materials [10], [24].

Within this paper, we address the problem of affective states recognition utilizing contextual information and interaction parameters as features, despite the fact that this input is usually viewed solely through the prism of performance. We demonstrate the feasibility of identifying student’s affective state such as boredom, engagement and frustration in a learning environment, inspired by the Theory of Flow [9], [30], which is a well-known and grounded paradigm of how affect is modeled during interacting with learning and gamified applications [5], [7], [9], [20], [22], [27], [32]. Towards this aim, the main goal of this work is to model and design a learning activity-based tracker, utilizing digital interaction parameters and learning analytics information related to performance, difficulty level, the complexity of the learning activity, number of attempts, and time needed to accomplish a task. Fig. 1 displays two students playing the learning game we use in our study, while the system is tracking their interaction features and is utilizing machine learning techniques to predict their affective state.

**Contributions.** The contribution of this work can be summarized as follows: We propose an affect recognition approach, by developing a generic framework that adopts interaction tracking for affect recognition to be used in a variety

of scenarios through utilizing a leading standard in learning analytics, the Experience API (xAPI) (Sect. III-A). Through this framework, we are able to describe a series of parameters according to user-platform interaction, related to performance, the time needed to accomplish a goal, and level of difficulty. The benefit of the proposed approach relies upon its suitability to be adopted in various setups and devices, where specific sources of information are not available, such as depth or RGB information. Secondly, we use the model of the Theory of Flow [30] (Sect. III-D) as a model for affect inference in the learning environment. This model has three different states (namely, boredom, engagement, and frustration), obtained through self-annotation in a real operating environment of students interacting with an e-learning platform (Sec. III). Finally, we present an extensive analysis of our case-study research (Sect. IV) and showcase the applicability of our model cross-subjects and per-subject for inferring affect during learning. Summarizing, we introduce in this paper an application of machine learning algorithms for predicting the user’s affective state based on interaction parameters contributing to the education domain and specifically supporting TEL.

## II. RELATED WORK

Affect recognition is a challenging problem due to its multimodal nature and the variability of affect expressivity across different individuals. These factors make sensing, modeling and recognizing affects a hard task [28]. This becomes even more evident in educational settings, where the research community has proposed various approaches for affect recognition, employing, most of the times, silo approaches that often make use of log-files [10], [24] and ambient sensing techniques [1]–[3], [21]. Unfortunately, these studies concluded with results that are not easily applicable on a satisfactorily wide range of educational settings.

The idea of correlating human behaviors with interaction context has been researched in a plethora of problems, aiming at interpreting the attention level of a user towards a certain goal using ML algorithms. For example, in educational settings, the authors in [24] studied affect prediction such as confusion, boredom and engagement through log-data of students in a math tutoring system. Their study shows a negative correlation between boredom and performance during problem-solving while engaged concentration and frustration are associated with positive learning. Similarly, authors in [10] used solely students’ interaction with a digital agent teaching algebra for affect recognition. Their work showed a performance which is better than chance at identifying engaged concentration, boredom, confusion and frustration. Moreover, many previous works have dealt with studying affect in gamification [31], [34], where the link between cognition and affect is the focus of attention.

In addition, facial expressions have been studied in TEL as an affect informative channel [1], [3]. However, many of these studies have focused on posed affect and did not take into account the nature of the classroom environment or the type of education provided, such as MOOCs, distance learning and

digital learning. For example, in [21], Kapoor et al. studied affect using multimodal sensory information from the face, postural shifts and learners’ activity on the computer. Their approach employed Mixture of Gaussian Processes for data fusion. Additionally, several studies have been using speech as an input for affect recognition. In [4], authors proposed a framework for online feedback based on the learner’s vocal affect expression. Asteriadis et al. in [2] analyzed the case of using head movements and gaze patterns towards learning material to study the mapping between these cues and learners’ affective states. This would eventually lead to adaptation of the learning process towards more personalized learning.

## III. DATA COLLECTION AND AFFECTIVE STATES MODEL

In TEL, there are different approaches proposed for data modeling and analysis with regards to affect recognition. One large family of approaches mainly deals with low-level data channels such as facial expressions, keystrokes, mouse events or gestures [2], [3], [21]. However, as mentioned in the related work discussion, approaches which focus on high-level features, for example, learner’s activities, digital interactions and so forth, have not been extensively studied in the field of affect analysis.

Therefore, the focus of this research is placed on the latter family of techniques, which targets the high-level interaction features. In learning analytics, currently, there are different methods of standardization, one of the most recent specifications applied in learning contexts is the Experience API (xAPI) which focuses on the high-level activities [11], [29]. We gathered data using the xAPI standard through an interaction with an e-learning platform, developed for the scope of this research. This section details the developed game, the way the xAPI tracker is used, the necessary steps and instructions for the users, the statistics of the collected dataset, and the affective states model used.

### A. Interaction Tracking

xAPI is a leading framework and an emerging standard for tracking and storing educational data [29]. The xAPI specification is two-fold, as it can both focus on the syntax of the data and, at the same time, define the characteristics of a Learning Record Storage (LRS), which serves as the end-point collecting and exchanging learning analytics traces [15], such as learning content, and learner’s activities and performance. When an activity is recorded, xAPI sends statements to the LRS. These experience statements are the core of xAPI, detailing the trajectories of learning activities for learning analytics, and providing the digital interactions for affect recognition in this study. The xAPI data format defines the experience statements with the attributes<sup>1</sup> as listed in Table I.

The rationale behind choosing this standard to track the interactions of the learner with learning materials is mainly due to the following reasons: (1) xAPI is event-centered, covering a larger range of actions, as the field verb in xAPI statements

<sup>1</sup><https://experienceapi.com/statement-design/>

TABLE I  
THE xAPI FRAMEWORK'S ATTRIBUTES.

Field's type	Attributes
Mandatory fields	Actor, Verb and Object
Optional fields	Result, Context, Timestamp (internal recording of timestamp), Authority, Version and Attachments

can define arbitrary actions; (2) it is designed for better interoperability in different educational systems; (3) xAPI has many adopters [8] and it is also supported by commercial Learning Record Store services such as Learning Locker [23]; and (4) the xAPI framework can be easily extended and integrated into new software components in a learning platform, allowing the adaptation of the system driven by the digitally tracked interactions in combination with ML techniques.

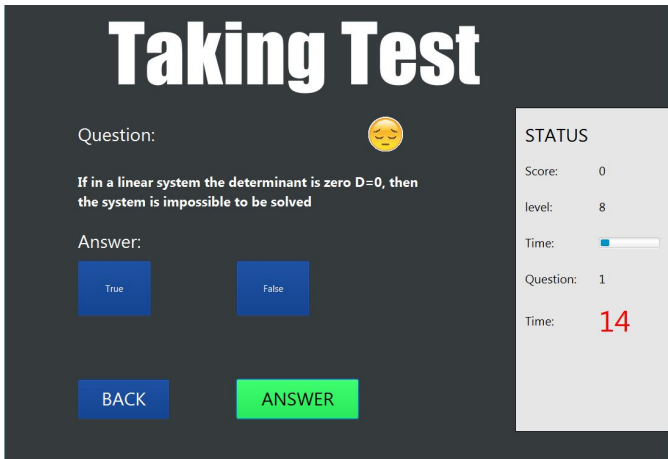


Fig. 2. A screen-shot of a question example from the learning game.

The standard is used in our framework as follows: “Actor” defines who is the subject that performs an activity, either a learner, a tutor or the device being used. Field “Verb” describes an action taken by the actor (e.g. passed a question). “Object” represents the subject of the actor’s action (e.g. question). In this work, “Timestamp” provides essential information that allows the system to get historical information about all the actions performed and the exact moment when they occurred while it also enables the calculation of meaningful information such as the time required to perform specific activities. The “Result” element gives the outcome of the experience (such as parameters regarding success, completion or score during the activity).

### B. Game Description

A serious game has been developed for the scope of this research, consisting of a learning platform which, according to its complexity and levels of difficulty of its materials, is expected to generate different affect states to the student, who will then report them at the end of each session. Fig. 2 shows a screen-shot of this game, where the students can perform a question-based learning interaction with the platform. The game contains three types of questions namely: “choose an

answer”, “true or false” and “fill in the blank”, where students are required to enter the whole answer. For collecting as much data as possible, a database that includes a total of 800 questions has been developed, covering four major topics: Mathematics, History, Sports and Geography. Moreover, the questions used have a varying difficulty level from 1 to 9 (in ascending order). Further details about the game flow and how the students can use it are listed below:

- A student logs in after creating an account, which includes her/his demographics information, competence (skills) on each topic presented in the game and her/his education level.
- Then, the student is directed towards an interface where she/he can choose a specific topic. Subsequently, a new learning session starts.
- In each learning session, the student is asked consecutively seven questions related to the chosen topic (belonging to one of the aforementioned three types of questions).
- After answering each question, the student is able to see whether her/his answer was correct or not.
- At the end of each session, the student is asked to provide self-annotations with regards to the experienced affective state, as boredom, engagement or frustration.
- Next, the platform provides the student with her/his score in the topic so far, and a comparison with the results of other students.
- The same procedure can be followed up to four different trials for each topic, while the level of difficulty increases after each session. Finally, the student has the choice of either continuing playing the game or logging out.

Regarding the modification of the level of difficulty, initially, the first level of a topic was selected randomly among the lowest three ones, while the subsequent levels were increased according to the student performance. It is important to note that users were given oral instructions about the game process they had to follow, the number of sessions they should complete and regarding self-annotation of the affective states. Their self-annotation has been used as the ground truth label of their interaction in a learning session.

### C. xAPI Statement Description

Each of the steps mentioned above is associated with an xAPI statement for recording students’ activities and gathering their interaction parameters. In order to achieve this goal, a set of statements has been defined. The initial set which has been applied in this study is listed in Table II.

Using these pre-defined statements, it is possible to track, for example, when a student starts a new session (triggering the “initialized” verb), when a question is presented (“Game asked a question”) or when a learner responds correctly/wrongly (“Learner passed/failed a question”). To track when a learner does not provide an answer or skips a question, the “completion” field of the “Result” element is sent as *False* to indicate this event. This set of statements could be easily expanded to monitor a broader range of interactions.

TABLE II  
THE xAPI STATEMENTS USED TO TRACK FEATURES WHILE INTERACTING WITH THE SERIOUS GAME.

Actor	Verb	Object	Related activities and features
Tutor/game	initialized	an interaction	Timer is triggered and student's profile information is collected such as skill level, study background, and age.
Learner	passed/failed	a question	Performance and question-related data are streamed (for example, time needed to answer, score attained, difficulty level and topic of the question, and number of attempts).
Game	asked	a question	A learner is asked a question related to the session topic.
Learner	terminated	an interaction	Self-annotation of affective state is asked, and the summary of interaction details is stored.

#### D. Affective States Modelling

Previous studies used various models in affect recognition, such as the discrete model, proposed by Ekman [13] or dimensional ones that model affects on the valence and arousal dimensions [14]. In [13], Ekman described a theory that addresses the basic emotions (namely: anger, contempt, disgust, fear, interest, happiness, sadness and surprise). This model is widely used in affect recognition. However, in [26], authors suggested the use of a simplified model that students mostly use to express their state, interest, distress (frustration) or pleasure (enjoyment) rather than the general models of basic affects. Another model based on Yerkes-Dodson Law, the so-called "sweet spot for achievement [17]", describes affect through three major states, namely disengagement, frazzle, and flow. These states have an impact on a person's performance whether at work or during learning activities.

One of the models that has been proposed for modelling affect is the Theory of Flow (ToF), developed by Csikszentmihályi in [9], [30]. Flow is the mental state experienced when a person conducts an activity in a state of deep involvement, in a feeling of energized focus, maximum engagement and enjoyment in the process [30]. ToF has been shown to be a very effective tool for measuring affect in the learning procedure, as it encapsulates the way learners perceive the learning content as a function of their skills and the challenge imposed by the learning materials [22], [27], [32]. In the research reported in [32], the authors studied students' engagement using longitudinal data of 526 schools in U.S.. This work investigated the conditions of adolescents under which they reported their engagement. The research stated the increase of engagement when there is a balance between the students' skill and the challenge of the given task. Moreover, the work in [22] examined students' affective responses in distance learning by studying the cause and effect of flow experience and the impact of interaction on it. In addition, ToF has been widely applied in serious games [18], [19] and this theory has been extensively used and investigated in the domain of computational intelligence in serious games and the video game industry [5], [7], [20]. Therefore, ToF is presented as a convenient affect model in serious games, since it includes important requirements such as: clear goals, clear challenge-skill balance, immediate feedback, concentration, timelessness, and the ability of experiencing temporary loss of the feeling of self-consciousness.

Motivated by ToF and its use in both education and serious games, and due to the gamified character of our learning application, we adopted the basic model of ToF as defined

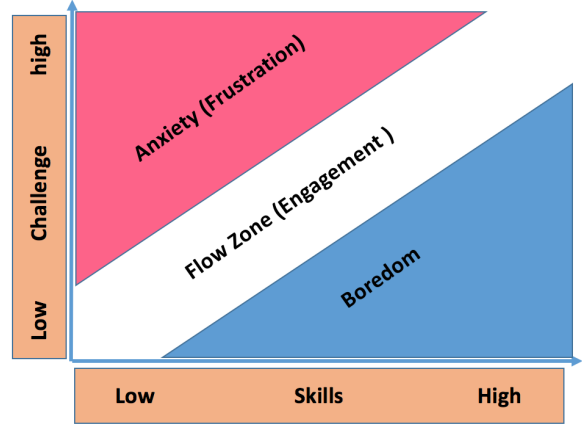


Fig. 3. The model of the Theory of Flow consisting of the three affective states relevant to learning experience as defined in [9].

in early work of Csikszentmihályi in [9] and described in [7]. Fig. 3 presents the details of this model, which consists of three different zones, namely, boredom, flow (engagement) and anxiety (frustration). These states are the most relevant to the users' skills and the amount of difficulty imposed by the learning application. In this model, for a certain user, keeping the balance between skill level and challenge dimensions leads to a positive effect, while a wrong selection can provoke either learners' frustration or boredom.

The ground truth utilized in the collected data is based on the self-annotation of the experienced affective state by the participants. This annotation could be considered reliable since it was performed by the users themselves (bachelor and master students) following the guidelines outlined by the researchers and based on the ToF model.

#### E. Dataset Statistics

During data collection, diversity in participants' population has been carefully considered in order to collect data from interactions of users with varying profiles in terms of age, education level and gender. Specifically, a dataset of 32 subjects has been collected where 18 females and 14 males voluntarily participated in the process. The users were either bachelor or masters students (12 master and 20 bachelor) with an average age of  $22.40 \pm 3.13$ . The participants can be grouped in two different knowledge profiles: 22 students from the Engineering department, and 10 are from the Faculty of Psychology at the University of Maastricht. The total duration of each experiment per participant lasted an average of  $26 \pm 5$  minutes. Each student was asked to perform four learning sessions on the available four topics. However, some students



performed less than four sessions leading to a total number of 459 recorded sessions, instead of 512. Data gathering was conducted in adjusted environments like classrooms.

#### IV. DATA EVALUATION

This section presents statistical data related to the gathered sessions and the interaction features extracted and modeled according to the xAPI standard stream. The approaches applied to evaluate the data (cross-subjects and subjects-based) and the implications of the results are discussed. Various classification methods had been applied, including Adaboost, Random Forest classifier, while the best results, which are reported next were obtained using radial basis Support Vector Machines (SVM).

##### A. Sessions Statistics

The collected 459 sessions were labeled as follows: 57 sessions as boredom, 272 sessions as engagement and 130 sessions as frustration using the provided self-annotations by the users. As we intentionally collected data from two different faculties (Engineering and Psychology), Fig. 4 presents the distribution of each affective state for the collected sessions, in each topic of the game among the students of both faculties independently. For example, the second column shows the ratio of the reported self-annotation in mathematics among engineering students. The table suggests a similar distribution of the states among both profiles with relatively higher frustration for psychology students in Mathematics, while engineering students show more frustration in the History topic.

##### B. Feature Vectors

Table III details the constructed feature vector for each session according to each student's performance in it, the session information and the profile information provided by the students. Since a session consists of seven questions, there are 14 features related to the obtained scores and the time needed to answer each question. In addition, each session contains questions related to the same difficulty level and a trial value which represents the student's stage in the game for a given topic. We also include the session duration, and a summary of the results in it through computing the ratio of correct, wrong and skipped answers. In the third column, four features regarding the student's profile are included, where skill level is based on the students' report of their level of knowledge in each topic presented in the game (graded in a scale ranging between 0 and 10), while the level of education can be either bachelor or master. Therefore, the resulted feature vector's length is 24. These features provide overall descriptors, while our aim is to associate them with the learner's affective state. Feature vectors were normalized by subtracting the mean and dividing them by the standard deviation.

##### C. Classification and Evaluation Schemes

We trained radial basis SVM<sup>2</sup> classifiers making use of a one-vs-all strategy, in order to fit a classifier for each class against all other classes. This gives us the advantage of inspecting each class and its corresponding classifier. For instance, a separate classifier has been trained to infer frustration against both boredom and engagement classes. A similar approach was followed for boredom and engagement. During the testing phase, the three classifiers were fit on the test samples. Consequently, for a test sample, a classifier with maximal value was selected as a predictor of the affective state corresponding to a given session. We introduce next two evaluation strategies, *cross-subject* for testing the generality of our approach across students with different profiles, personality and knowledge base and *subject-based* for enabling an adaptive learning according to user's personality, capable of detecting changes in the affective state of the user and proposing a way to improve his/her productivity and learning experience.

##### D. Cross-subject Evaluation

The cross-subject evaluation assesses the system performance in predicting the self-reported affective states, using cross-validation. As we aim at testing the system applicability to a new subject, we employ the leave-one-out cross-validation scheme (LOOCV), by dividing the data according to the users where, in each fold, we exclude one users' data for testing while the rest of the users' data is used for training. Furthermore, we also evaluate the system using 10-fold cross-validation, by dividing each class samples into ten partitions.

The performance on the collected data is reported in terms of precision, recall, and f1 score. The second and the third column in Table IV display the results of the two evaluation schemes on the collected data. In LOOCV, the classifier achieves good results obtaining a precision of 66%. Although the expectation is that experienced affective states and learning experience are person-specific notions, the achieved performance indicates the potential of the affect recognition system in TEL based on interaction parameters. Likewise, the results of the 10-folds evaluation in the third column are promising. The performance, in this case, is similar to the LOOCV which supports the generality of the framework for the prediction of the self-reported affective states in the digital learning environment.

Table V details the confusion matrix of the three states obtained from the LOOCV. As expected, the accuracy of engagement and frustration is higher compared to boredom. Indeed, most of the participants in the data collection reported engagement or frustration while only 50% of them reported the boredom state in any of their self-annotations. Additionally, this result can be due to the fact that the boredom state was

<sup>2</sup>SVM implementation provided by scikit-learn [25] is used. Scikit-learn is a machine learning library for Python programming language. Soft margin cost function parameter C and the regulation parameter gamma are set to 100 and 0.1, respectively.

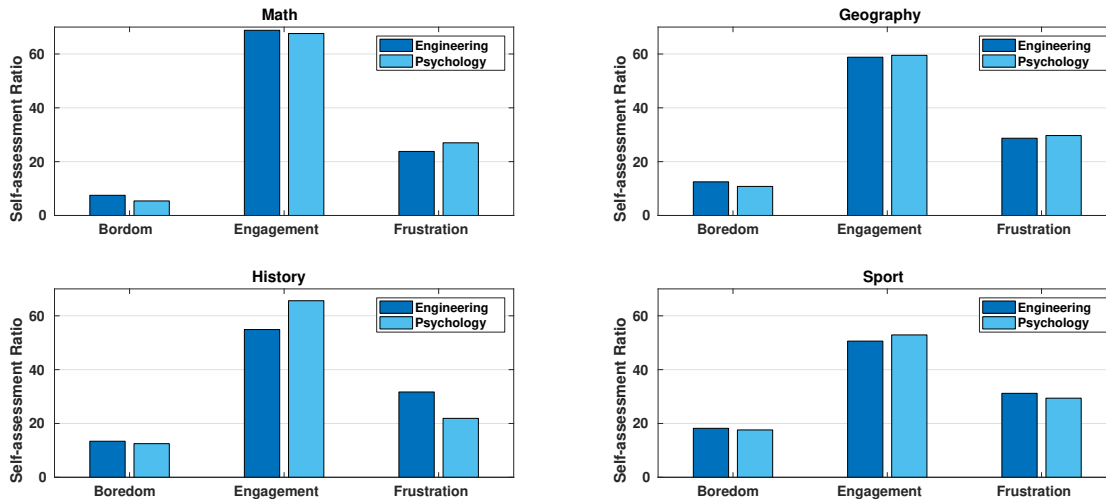


Fig. 4. Percentage of the self-assessment (its distribution) among the students in two faculties for each game's topics.

TABLE III  
FEATURE VECTOR PER SESSION

Performance features	Sessions' summary features	Profile information
<ul style="list-style-type: none"> <li>Score attained and time needed to answer each question</li> </ul>	<ul style="list-style-type: none"> <li>Ratio of correct, wrong and skipped answers.</li> <li>Difficulty level of the session</li> <li>Duration of the session</li> <li>Number of trial</li> </ul>	<ul style="list-style-type: none"> <li>Skill level in the session's topic,</li> <li>Level of education,</li> <li>Age</li> <li>Gender</li> </ul>

TABLE IV  
PRECISION, RECALL AND F1-SCORE RESULTS OBTAINED BY SVM USING THE TWO VALIDATION SCHEMES.

Measure	Subject-based validation	10-fold validation
Precision	66 %	67%
Recall	59 %	59%
F1 score	61 %	61%

TABLE V  
CONFUSION MATRIX OBTAINED WITH THE SUBJECT-BASED LOOCV.

States	Boredom	Engagement	Frustration
Boredom	29.8%	31.6%	38.6%
Engagement	16.2%	62.1%	21.7%
Frustration	23.8%	11.5%	64.6%

harder to experience due to the dynamic nature of the game and the possibility to leave the session anytime.

**Engagement vs Non-engagement Evaluation:** A further analysis was applied on the collected data for investigating the performance of the classification when considering the scenario of engagement versus non-engagement states. In this strategy, boredom and frustration sessions were assumed to

TABLE VI  
PERFORMANCE OF ENGAGEMENT VERSUS NON-ENGAGEMENT EVALUATION.

Measure	Engagement vs Non-engagement
Precision	75%
Recall	74%
F1 score	74%

represent the non-flow zone in the adopted affect model, as shown in Fig. 3. This simplification can assist in cases when the focus is on detecting the non-engagement states to adapt the learning experience according to learner's competence and by prompting the right level of difficulty. Table VI illustrates the results of this scenario using LOOCV. Recall and precision have high values, fact which proves the ability of the proposed approach to distinguish between the states of engagement and non-engagement.

Fig. 5 gives a closer look at the results of this scenario among the two different faculties and the topics used in the game. When observing the obtained results of each faculty independently, we notice a higher detection accuracy of engagement than non-engagement for engineering students with an exception in sport, whereas for psychology students, in all topics, the detection of non-engagement is relatively higher than engagement. Indeed, the overall performance of engineering students was better than psychology students which might explain their engagement in playing the game, while for psychology students the non-engagement can be due to the perceived challenge and their lower interest towards the game topics. These observations and results are consistent with the Theory of Flow model. The obtained system accuracy shows how the game's topic, challenge level, and students' educational background can contribute to affective learning technologies.

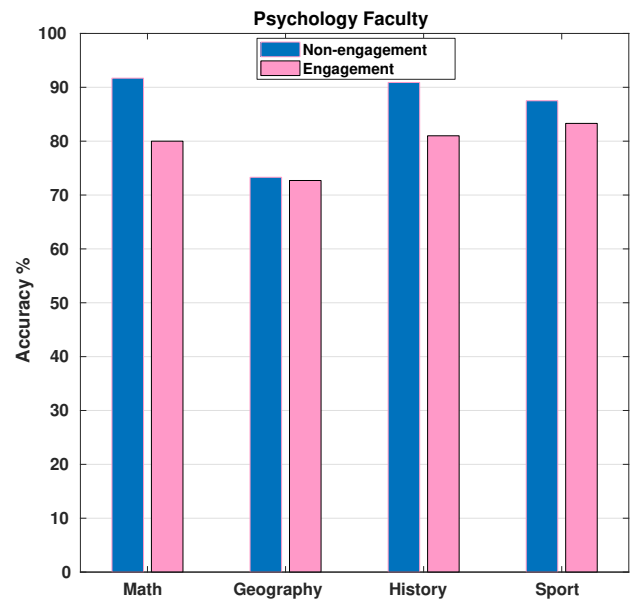
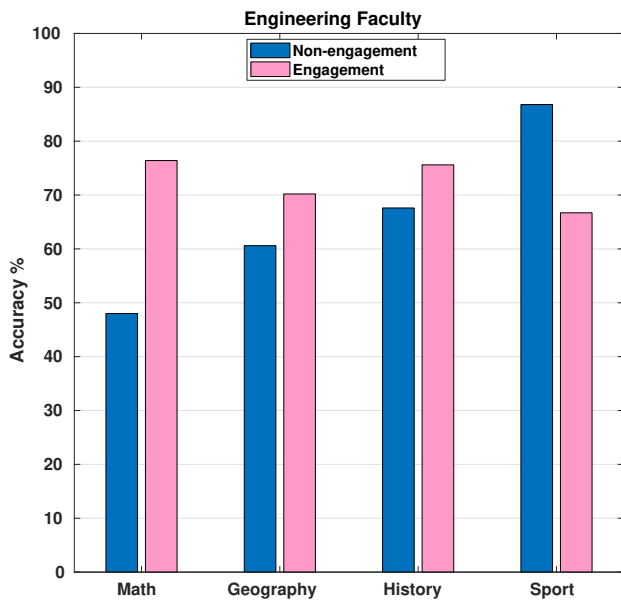


Fig. 5. Accuracy of the engagement and non-engagement detection among the students in two faculties and different topics.

TABLE VII  
OVERALL PERFORMANCE OF SUBJECT-BASED EVALUATION.

Measure	Subject based evaluation
Precision	74%
Recall	73%
F1 score	73%

TABLE VIII  
CONFUSION MATRIX OBTAINED WITH THE SUBJECT-BASED EVALUATION.

States	Boredom	Engagement	Frustration
Boredom	61.4%	26.3%	12.3%
Engagement	8.5%	79.4%	12.1%
Frustration	7.7%	26.9%	65.4%

### E. Subject-based Evaluation

Both affect and personality traits contribute to the learning process. In addition, the difficulty level of the presented learning material can be perceived individually according to user's competence, which can also influence her/his affective state. For this reason, a subject-based analysis is performed to study the ability of the system to predict the self-reported affective states, based on interaction features. This evaluation was applied for each user independently, by employing each session once for testing and the rest of the user sessions for training. The obtained results are included in Table VII, having a high precision of 74%, which is supported by the confusion matrix presented in Table VIII where a good accuracy detection is achieved for each state in the adopted affect model. It is important to mention, that the system was able to detect sudden changes of affective states (e.g. frustration after several engagement sessions or vice versa). The achieved results highlight the benefit of employing interaction features in customized and personalized learning activities according to the user's affective state.

## V. CONCLUSION

In this study, we presented a framework for affect recognition in TEL, based on contextual information and the tracking of interaction features during learning activities. For this purpose, we utilized a standard tool, the xAPI framework, for learning analytics and high-level activities tracking. We have developed a serious game, as a learning platform for data collection and evaluation of the proposed framework. We adopted the Theory of Flow as an affect model, used by the students to self-report their experienced affective state during the interaction with the platform. The results of affect recognition within the proposed framework indicate the potential usage of interaction parameters with learning materials as a useful channel for measuring students' affective states. Specifically, we employed two evaluation strategies in combination with machine learning algorithms. The cross-subjects analysis highlighted the generalization ability of our system across students with different profiles (namely engineering and psychology), obtaining a good precision of 67% for three affective states and 75% when considering engagement vs. non-engagement states. Furthermore, the subject-based evaluation is especially useful in boosting the adaptive nature of the learning process to enhance the outcome of the learning experience, maximizing the learners' knowledge acquisition and enabling personalization. The obtained results are promising in this direction, with a precision of 74% for the recognition of the three affective states.

## ACKNOWLEDGMENT

This work was supported by the Horizon 2020 funded project MaTHiSiS (Managing Affective-learning THrough Intelligent atoms and Smart InteractionS) nr. 687772 (<http://www.mathisis-project.eu/>).

## REFERENCES

- [1] Mohamed Ben Ammar, Mahmoud Neji, Adel M Alimi, and Guy Gouardères. The affective tutoring system. *Expert Systems with Applications*, 37(4):3013–3023, 2010.
- [2] Stylianos Asteriadis, Paraskevi Tzouveli, Kostas Karpouzis, and Stefanos Kollias. Estimation of behavioral user state based on eye gaze and head pose application in an e-learning environment. *Multimedia Tools and Applications*, 41(3):469–493, 2009.
- [3] Kiavash Bahreini, Rob Nadolski, and Wim Westera. Towards multimodal emotion recognition in e-learning environments. *Interactive Learning Environments*, 2016.
- [4] Kiavash Bahreini, Rob Nadolski, and Wim Westera. Towards real-time speech emotion recognition for affective e-learning. *Education and Information Technologies*, 21(5):1367–1386, 2016.
- [5] Riccardo Berta, Francesco Bellotti, Alessandro De Gloria, Danu Pranantha, and Carlotta Schatten. Electroencephalogram and physiological signal analysis for assessing flow in games. *IEEE Transactions on Computational Intelligence and AI in Games*, 5(2):164–175, 2013.
- [6] Nigel Bosch, Sidney D’Mello, Ryan Baker, Jaclyn Ocuppaugh, and Valerie Shute. Temporal generalizability of face-based affect detection in noisy classroom environments. In *International Conference on Artificial Intelligence in Education*, pages 44–53. Springer, 2015.
- [7] Jenova Chen. Flow in games (and everything else). *Communications of the ACM*, 50(4):31–34, 2007.
- [8] Alberto Corbi and Daniel Burgos Solans. Review of current student-monitoring techniques used in elearning-focused recommender systems and learning analytics: The experience api & lime model case study. *IJIMAI*, 2014.
- [9] M Csikszentmihaly. Beyond boredom and anxiety: Experiencing flow in work and play, 1975.
- [10] Ryan SJ d Baker, Sujith M Gowda, Michael Wixon, Jessica Kalka, Angela Z Wagner, Aatish Salvi, Vincent Alevan, Gail W Kusbit, Jaclyn Ocuppaugh, and Lisa Rossi. Towards sensor-free affect detection in cognitive tutor algebra. *International Educational Data Mining Society*, 2012.
- [11] Ángel Del Blanco, Ángel Serrano, Manuel Freire, Iván Martínez-Ortiz, and Baltasar Fernández-Manjón. E-learning standards and learning analytics. can data collection be improved by using standard data models? In *Global Engineering Education Conference (EDUCON), 2013 IEEE*, pages 1255–1261. IEEE, 2013.
- [12] Sidney D’Mello, Nathan Blanchard, Ryan Baker, Jaclyn Ocuppaugh, and Keith Brawner. Affect-sensitive instructional strategies. *Design Recommendations for Intelligent Tutoring Systems: Volume 2-Instructional Management*, 2:35, 2014.
- [13] Paul Ekman. An argument for basic emotions. *Cognition & emotion*, 6(3-4):169–200, 1992.
- [14] Lisa A Feldman. Valence focus and arousal focus: Individual differences in the structure of affective experience. *Journal of personality and social psychology*, 1995.
- [15] Christian Glahn. Using the adl experience api for mobile learning, sensing, informing, encouraging, orchestrating. In *Next Generation Mobile Apps, Services and Technologies (NGMAST), 2013 Seventh International Conference on*, pages 268–273. IEEE, 2013.
- [16] Daniel Goleman. *Emotional intelligence*. Bantam, 2006.
- [17] Daniel Goleman. The brain and emotional intelligence: New insights. *Regional Business*, 94, 2011.
- [18] Juho Hamari and Jonna Koivisto. Measuring flow in gamification: Dispositional flow scale-2. *Computers in Human Behavior*, 40:133–143, 2014.
- [19] Juho Hamari, Jonna Koivisto, and Harri Sarsa. Does gamification work?—a literature review of empirical studies on gamification. In *System Sciences (HICSS), 2014 47th Hawaii International Conference on*, pages 3025–3034. IEEE, 2014.
- [20] Daniel Johnson and Janet Wiles. Effective affective user interface design in games. *Ergonomics*, 46(13-14):1332–1345, 2003.
- [21] Ashish Kapoor and Rosalind W Picard. Multimodal affect recognition in learning environments. In *Proceedings of the 13th annual ACM international conference on Multimedia*, pages 677–682. ACM, 2005.
- [22] Li-Fen Liao. A flow theory perspective on learner motivation and behavior in distance education. *Distance Education*, 27(1):45–62, 2006.
- [23] Maurizio Megliola, Gianluigi De Vito, Roberto Sanguini, Fridolin Wild, and Paul Lefrere. Creating awareness of kinaesthetic learning using the experience api: current practices, emerging challenges, possible solutions. In *CEUR Workshop Proceedings*, volume 1238, pages 11–22, 2014.
- [24] Zachary A Pardos, Ryan SJD Baker, Maria OCZ San Pedro, Sujith M Gowda, and Supreeth M Gowda. Affective states and state tests: Investigating how affect throughout the school year predicts end of year learning outcomes. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, pages 117–124. ACM, 2013.
- [25] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. Scikit-learn: Machine learning in python. *Journal of machine learning research*, 12(Oct):2825–2830, 2011.
- [26] Rosalind W Picard. Building an affective learning companion. In *Intelligent Tutoring Systems*, pages 811–811, 2006.
- [27] Maria Ofelia Z San Pedro, Ryan SJ d Baker, Sujith M Gowda, and Neil T Heffernan. Towards an understanding of affect and knowledge from student interaction with an intelligent tutoring system. In *International Conference on Artificial Intelligence in Education*, pages 41–50. Springer, 2013.
- [28] Enver Sanginetto, Gloria Zen, Elisa Ricci, and Nicu Sebe. We are not all equal: Personalizing models for facial expression analysis with transductive parameter transfer. In *Proceedings of the 22nd ACM international conference on Multimedia*, pages 357–366. ACM, 2014.
- [29] Jose Luis Santos, Katrien Verbert, Joris Klerkx, Erik Duval, Sven Charleer, and Stefaan Ternier. Tracking data in open learning environments. *Journal of Universal Computer Science*, 21(7):976–996, 2015.
- [30] Martin EP Seligman and Mihaly Csikszentmihalyi. *Positive psychology: An introduction*. Springer, 2014.
- [31] Noor Shaker, Stylianos Asteriadis, Georgios Yannakakis, and Kostas Karpouzis. A game-based corpus for analysing the interplay between game context and player experience. *Affective Computing and Intelligent Interaction*, 2011.
- [32] David J Shernoff, Mihaly Csikszentmihalyi, Barbara Schneider, and Elisa Steele Shernoff. Student engagement in high school classrooms from the perspective of flow theory. In *Applications of Flow in Human Development and Education*, pages 475–494. Springer, 2014.
- [33] Beverly Park Woolf. Ai and education: Celebrating 30 years of marriage. In *Workshop on Les Contes du Mariage: Should AI stay married to Ed?*, page 38, 2015.
- [34] Georgios N Yannakakis and John Hallam. Real-time game adaptation for optimizing player satisfaction. *IEEE Transactions on Computational Intelligence and AI in Games*, 1(2):121–133, 2009.