STATEMENT:
The final version is published on the Proceedings of the European Simulation and Modelling Conference 2020
Available on : https://www.eurosis.org/cms/index.php?q=taxonomy/term/58
ISBN 978-9492859-12-9

TOWARDS THE MODELLING OF THE CONCENTRATED STATE OF LEARNERS. AN INTRA-SUBJECT MODELLING APPROACH

Ana Serrano-Mamolar
UNED
email: aserrano@dia.uned.es

Miguel Arevealillo-Herráez
Universidad de Valencia
email: miguel.arevalillo@uv.es

Jesús González-Boticario
UNED
email: jgb@dia.uned.es

KEYWORDS

Affective Computing, Learner modelling, Concentration modelling, Learner engagement

ABSTRACT

In this work, we present an experimental study that aims to explore the potential application of Hidden Markov Models (HMM) to predict the state of concentration of learners. The features used for the prediction are extracted from 4 commonly used physiological signals, namely heart rate, breath rate, skin conductance and skin temperature. Previous works that pointed how emotion translates into different physiological responses from one subject to another have motivated the study of intra-subject learner models. Thus in this work both inter-subjects and intra-subjects approaches have been evaluated. To this end, we have built a labelled dataset from an intensive data capture on 2 different subjects that developed different learning tasks across several sessions. Results are consistent across users and show that a high accuracy can be achieved when using HMM-based intra-subject models, but inter-subject models fail at the same task.

INTRODUCTION

Adaptive learning systems can be used to intelligently manage the affective dimension of the learner due to the interplay of affect and learning-related cognitive processes as reported in literature Blanchard et al. (2009). Different affective states can influence the learning process in a different way. For example, there are strong evidences coming from previous research that emotions have an important effect on the student's engagement and motivation, and consequently influence learning outcomes Pardos et al. (2014), Pekrun et al. (2010), Ainley (2006). Previous studies place engaged concentration as the most prevalent affect in a classroom context Pardos et al. (2014). Engaged concentration is a state of engagement with a task such that concentration is intense, attention is focused, and involvement is complete. Thus, an effective detection of the concentration state over time is a crucial task for adaptive learning systems that aim to take proper actions in order to improve the student's engagement.

A main issue confronted in this research area is how to improve the recognition performance. Related works show two different approaches to tackle the modelling of affective and cognitive states, inter-subject and intra-subject modelling approaches. Subject-independent (inter-subject) approaches focus on the affective state and build a global model by treating data from different subjects as if they all belonged to the same individual, as it occurs in Salmeron-Majadas et al. (2015), Purnamasari and Junika (2019), Nizetha Daniel et al. (2017), Reyes et al. (2016). However, the performance of physiological data-driven models are limited to the poor generalisation of the signals at reflecting emotional information across subjects. Indeed, studies reveal that the same stimulus can elicit different emotions and the same elicited emotion translates into different physiological responses from one subject to another, Hajlaoui et al. (2018). So, to deal with this problem other works produce an intra-subject model, which simultaneously considers the affective/mental state and the subject, and only use training data coming from the same targeted subject they try to predict on, Duan et al. (2013), Kim and André (2008). Despite there are some works that reflect relatively high prediction rates obtained with the first approach, inter-subject e.g. Ayesh et al. (2014) and Salmeron-Majadas et al. (2015), the benefits of the latter have recently been proved in Arevalillo-Herráez et al. (2019), Arevalillo-Herráez et al. (2019), Arnau-Gonzalez et al. (2020), showing that the subject-dependent component in Electroencephalography (EEG) signals is far stronger than the emotion-related component.

In this work, we study the performance of both inter-subject and intra-subject approaches at detecting concentration in an educational scenario, by using Hidden Markov Models (HMM) that are trained with physiological data. To this end, we run a data collection experience in which we captured a set of physiological signals that were already identified in Uria-rivas et al. (2019) as valuable information to identify changes in the subject's mental state, namely heart rate, breath rate, skin conductance and skin temperature. The data collected was labelled offline by a set of independent experts, who identified the concentration state by using both synchronised video and audio recordings and personal one-to-one interviews with the subjects to

verify and improve the labelling, as described in Saneiro et al. (2014). Finally, we used HMM to study whether the information contained in the physiological signals can effectively be used to predict concentration.
The results reveal that the concentration state can be accurately modelled by a multimodal HMM built from the physiological signals, but only when an intra-subject model is adopted. The results presented remark on the necessity to model each subject individually, unlike other related works which only focus on inter-subject user modelling. Intra-subject modelling implies a substantially more intense labelling effort from a practical perspective, due to the intensive data capturing process that is required. However, the large difference in detection accuracy between the two types of models suggests that more effort should be placed on the development of intra-subject models, in contrast to inter-subject models that fail to capture subject traits related to how physiological signals are affected by cognitive or affective states at an individual level.

RELATED WORK

Affect recognition has been successfully applied to marketing Krishna (2012), Garbas et al. (2013), health Tokuno et al. (2011), Moussa and Magnenat-Thalmann (2009) and more recently to learning systems Ghaleb et al. (2019), Farzaneh et al. (2019). Among the vast previous research on affect recognition in e-learning environments, some of them have focused on the detection of the level of concentration, as one of the most relevant cognitive states in relation to learning Baker et al. (2010). This is in support to other previous research works that have shown that concentration is significantly correlated with learning, engagement and student performance Corno and Mandinach (1983), Suryanti et al. (2019).
Different approaches are applied in detection of concentration, namely sensor-free approaches, based on visual information approaches and based on physiological data approaches.
With regard to sensor-free approaches, one method to detect the students' affect based on their interactions with an online learning system was proposed in Baker et al. (2012). Similar sensor-free interaction based approaches have been presented in other more recent works like Botelho et al. (2017), Ocumpaugh et al. (2014), Arevalillo-Herráez et al. (2017) and Cunha-Perez et al. (2018), but the accuracy achieved has not significantly improved the value reported in Baker et al. (2012).
As concerns visual sources of information, they have the advantage that some indicators are valid across subjects. Ekman's Ekman (1997) research is a universal coding system and reliably relates emotions with facial muscle activation. This enables the use of inter-subject models that are valid across the entire population, without significantly degrading performance. A variety of combinations of machine learning techniques and metrics from an eye-tracking system were attempted in Nizetha Daniel et al. (2017). The highest classification accuracy of the proposed system was achieved when a Multilayer Perceptron was used. However, they disregarded both the context and the subject traits, e.g. distractions caused by environmental elements not related to the course content. In Krithika and Lakshmi Priya (2016), concentration detection was based on a continuous monitoring of the learner's head rotation, eyelid opening, and visual focus on the content. A similar approach was also followed in Sharma et al. (2019), combining information about eyes and head movements with facial expressions to produce a concentration index in real-time while answering a quiz. Results revealed that the proposed system was unable to deal with the problem of face occlusion. As a consequence, such system assigns a low level of concentration to student's which, in view of their quiz results, were concentrated. Automatic real-time recognition of student engagement from students' facial expressions is also explored in Whitehill et al. (2014). In this work, four binary classifiers are trained for the automatic discrimination of engagement from static images of students' faces, based on the Action Units from The Facial Action Coding System, among other features. The results showed high inter-observer reliability when discriminating low versus high levels of engagement.
There are also some previous works that apply physiological sensors to detect concentration of learners. These methods are usually more intrusive than the ones in the visual category and require more expensive equipment. However, they are able to provide complementary data that can be used to enrich the information coming from visual sources and potentially increase the accuracy rate. In general, the existing literature suggests that methods based on physiological signals are generally less accurate. For example, in Hsu et al. (2012) a reading concentration monitoring system is proposed using sensor technologies. The sensors used in the study included a webcam, a heartbeat sensor, and a blood oxygen sensor. EEG signals are used to determine whether students were attentive or inattentive in Liu et al. (2013). In a similar way, EEG sensors were used in Kosmyna and Maes (2019) to directly calculate an engagement index through the value of the neural oscillations provided by the sensors. Nevertheless, we argue that the lower accuracy obtained with physiological signals is actually due to the higher influence of subjects traits on the way this type of signals are affected by cognitive or affective states. Most previous works on subject's concentration detection have generally followed an inter-subject approach, which is inappropriate because of the intrinsic nature of physiological signals. In this paper, we show that even

using some of the most simple and easily acquirable physiological signals we may reach an accuracy of the same order as visual methods, provided that subjects are modelled in an individual way.

EMPIRICAL RESULTS

Data collection

A data collection experience was conducted to evaluate the proposed approach to automatically detect concentration patterns from physiological signals. In order to verify that our methodology provides an added value in a practical setting, data collection took place in a real school, while students interacted with a series of learning tasks. The experience was repeated for 2 different users $U = \{u_1, u_2\}$, who previousy had signed a consent form.
The student interaction was recorded in video, within a framework that supports tracking and labelling from a single-subject experiment, including on site and offline data labelling Santos et al. (2016). At the same time, we used the Physiological Acquisition Shield (PhyAS) described in Uria-rivas et al. (2019) connected to an Arduino Uno to capture 4 simple but effective physiological signals at a rate of 10Hz: heart rate, breath rate, skin conductance and skin temperature. The video and physiological signals recordings were synchronized by using the system's clock.
In order to have enough information to build intra-subject models, each student participated in 4 sessions, each doing a different learning activity. There were two different types of sessions. The first three ones were focused on detecting the affective state of the user and the fourth included an additional task where users got feedback when they entered a state of excessive agitation that would hamper their performance. These followed the approach of related work Santos et al. (2016), but we used an improved shield that reduced the number of control units to a single one, in the same way as in Uria-rivas et al. (2019). The first two sessions consisted of a series of math exercises with an increasing level of difficulty. The third session consisted of a series of logic exercises that the student had to solve. These first three sessions involved tracking the students with the aforementioned sensors while they where performing some tasks with their keyboards and mouse. The fourth session was an oral test in a second language (English), which consists of a voice baseline and two oral activities with increasing level of difficulty.

Data labelling

The video recordings were used to manually label the data trough a flexible and detailed labelling approach Saneiro et al. (2014). Two different experts, one with a psychology background and the other with over 6 year experience in affective computing, independently labelled the dataset. They identified specific moments in the video where the user seemed to have reached one of a set of specific mental states, such as concentrated, and used the application described in Santos et al. (2016) to mark the time where they believed the peak was. After labelling, a validation meeting was held with each participant, correcting the labelling where appropriate.

Table 1: Number of samples per subject

	Concentrated samples	Non-concentrated samples
subject 1	8	78
subject 2	11	121

Data preparation

As it happens with most other mental states, concentration does not happen all of a sudden. We hence can reasonably assume concentration on a time frame of 5 seconds previous to the identified peaks. Based on this hypothesis, we took each time mark reported as a concentration peak by the experts and isolated the physiological measurements of the 5 seconds previous to it. This yielded a matrix of size 4×50 (4 physiological signals $\times$ 50 measurements per signal) for each concentration label reported, containing information about how the signals evolved until the concentration peak was reached. We call each of these matrices a concentrated sample.
In order to generate non-concentrated samples, we discarded the first second of the physiological signals and chose 5 second disjoint slots that were at least two minute apart from any identified concentration peak. We also guaranteed a minimum of a 5 second separation space between any two non-concentrated samples.
The number of positive and negative samples per user are shown in Table 1.

Description of experiments

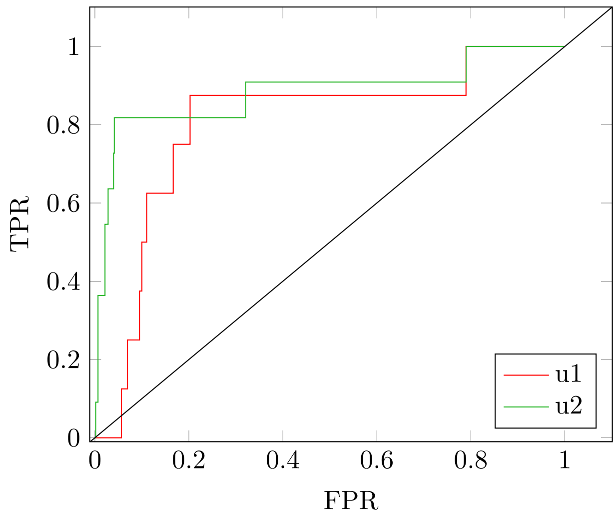
In order to analyze the validity of inter-subject and intra-subject models to predict concentration, we devised 3 experiments, which were repeated for the 2 users in U. In the first two experiments, we used two sets of samples S^+ and S^-, whose elements differed on whether they met a certain criterion C. We then applied a 5-fold cross-validation scheme based on the elements in S^+. For each fold, we used 80% of the samples in S^+ to train a HMM model. Once the model θ was built, we computed the probability $p(X|\theta)$ for each sample in the test set, which was composed of the remaining 20% of the samples in S^+ and all the samples in S^-. The scores produced for each sample in the test set were hence related to the probability that the sample met the specified criterion C. In order to evaluate the effectiveness of the model to discriminate samples according to whether they met the criterion C, the results from the 5-folds were used together to build a Receiver Operating Characteristic (ROC) curve and compute typical accuracy indicators. The accuracy indicators that were used are the Area Under the ROC Curve (AUC) and the Equal Error Rate (EER).

In the first experiment, we aimed to test whether the information contained in the physiological signals of an individual can be used to train a HMM model that is able to detect when the subject is concentrated from an unseen sample of the signals. We repeated the experiment for each subject u_i, always focusing on samples from the same subject and defining the criterion C as whether the sample was labelled as concentrated. In this case, S^+ was composed of all concentrated samples for u_i and S^- contained all non-concentrated samples for the same individual (see definition of concentrated and non-concentrated samples in Section). In the second experiment, we studied whether subjects can be easily distinguished by how concentration affects their physiological signals. This time, we only considered concentrated samples and defined the criterion C as whether the sample belongs to a given user u_i. Thus, and again for each u_i, S^+ included all concentrated samples for u_i, and S^- all concentrated samples for the rest of the subjects $(U - \{u_i\})$.

Finally, the third experiment attempted to determine whether it is possible to build an inter-subject HMM model that is able to accurately predict concentration for an individual u_i, without using data from that subject. In this case, we followed a slightly different setting. For each user u_i we run a single validation experiment using all concentrated samples from subject u_i as training data. The test set was then built using all concentrated and non-concentrated samples from users in $U - \{u_i\}$. Results were assessed using the same measures as in the previous two experiments.

First Experiment

Fig. 1 represents the ROC curve for the first of the experiments, in which the model was fit by taking 80% of the concentrated samples from one subject and tested against a set that contained both concentrated and non-concentrated samples of the same subject. For each testing sample X, the probability $p(X|\theta)$ was used as a prediction of whether the user was concentrated. The AUC and EER for each user are also shown at the right-hand-side of the same figure. The lowest AUC corresponds to the second subject and it is still quite promising. An average AUC=0.84 suggests that concentration can be predicted in a relatively accurate way when the model has been trained with labelled data from the same subject.



(a) ROC curve

User	Training	Total test
1	32	398
2	44	616

(b) Experiment sample

User	AUC	EER
1	0.80	0.20
2	0.88	0.17
Average	0.84	0.185

(c) Accuracy results

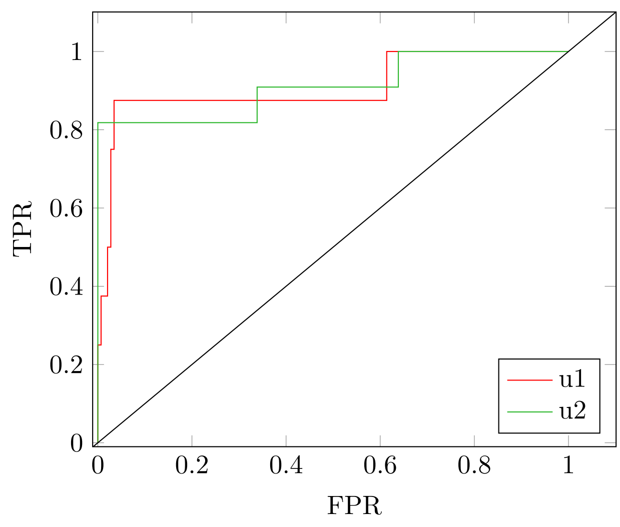
Figure 1: Results for each user u_i, using concentrated samples from u_i as training data and testing on concentrated and non-concentrated samples from the same user. The ROC curves indicates that the model is able to predict with a high accuracy whether a given sample belong to the concentrated or the non-concentrated category.

Second Experiment

Fig. 2 shows the results when the model was trained with 80% of the concentrated samples from a user u_i and tested against concentrated samples from all subjects. For each testing sample X, the probability $p(X|\theta)$ was used as a prediction score of whether the sample belonged to the user u_i. The high AUC and low error rates obtained in all cases clearly indicate that the way concentration reflects on the physiological signals is subject-dependent, to such an extend that from a concentrated sample we can accurately find out which user the sample belongs to.

Third Experiment

The relatively higher accuracy values obtained in the second experiment as compared to those obtained in the first one suggest that the subject's influence in the physiological signals is higher than that caused by the mental state itself. This finding motivated this third experiment, to test results when a model is created by using data coming from subjects other than



(a) ROC curve

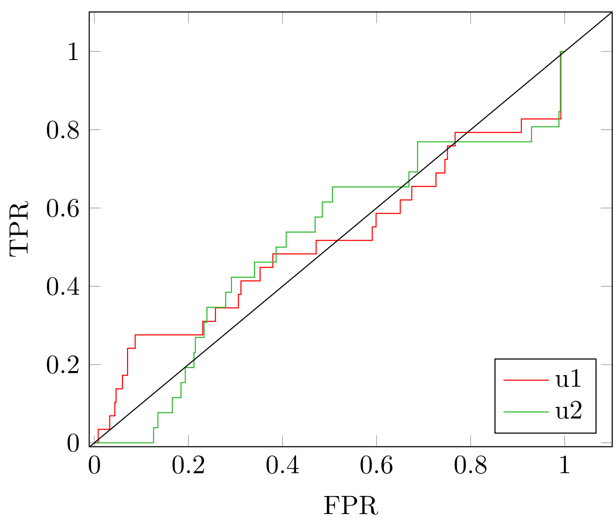
User	Training	Total test
1	32	153
2	44	141

(b) Experiment sample

User	AUC	EER
1	0.90	0.13
2	0.97	0.18
Average	0.93	0.15

(c) Accuracy results

Figure 2: Results for each user u_i, using concentrated samples from a user u_i and testing on concentrated samples from all subjects in U. The ROC curves indicates that, given a concentrated sample, it is possible to accurately identify which user it belongs to.



(a) ROC curve

User	Training	Total test
1	8	398
2	11	352

(b) Experiment sample

User	AUC	EER
1	0.51	0.49
2	0.50	0.47
Average	0.50	0.48

(c) Accuracy results

Figure 3: Results for each user u_i, using concentrated samples from a user u_i and testing on concentrated samples from all subjects in U. The ROC curves indicates that, given a concentrated sample, it is possible to accurately identify which user it belongs to.

the target. Fig. 3 shows the ROC curve for each user u_i, when the model is trained with concentrated samples from all other users, and tested against positive and negative samples from user u_i. AUC and EER values are also shown at the right-hand-side of the plot. Results reported are extremely closed to a random classifier and reinforce the hypothesis that the subject-related component of the physiological signals is stronger than the concentration-related one, and reveal the inadequacy of inter-subject models in this particular context.

Discussion of Results

HMM intra-subject models have shown to be extremely powerful at detecting concentration. Results reported in the first experiment are encouraging, and endorse the use of HMM models in this context. Positive results were consistently obtained for the 2 users, with an average EER of 0.185, despite that the labeling methodology is not exempt from potential mistakes. For example, and because data was relatively scarce, all 5 second signal frames were labelled as concentrated or non-concentrated, instead of discarding frames where experts expressed doubts. We believe that improving the labelling methodology could cause a further reduction of the EER and yield even better results. Furthermore, a true ROC curve is smooth. However the lack of enough positive samples in the cross-validation experiments make the ROC curves very jagged. Therefore, the improvement in the labelling will also make the curves smother. In addition, and thinking of practical settings, there is still room for improvement by combining different approaches, e.g. considering the scores in consecutive signal frames and setting a threshold based of the proportion of positive judgements on the last frames.

On the negative side, there are factors that limit the applicability of the intra-subject approach. Apart from the intrusiveness of the devices (something we have been always trying to reduce to the minimum by using smart bracelets and belts that measure physiological signals such as those considered in Uria-rivas et al. (2019)), this presented method suffers from the intrinsic difficulty of training a model with large amounts of data from the same subject. First, this restriction makes the approach not suitable to predict concentration on previously unseen subjects. Second, intra-subject data collection is generally time and effort-consuming. In our case, we required a planned data collection experience for each subject, which involved an average time of 42 minutes of user interaction, plus the data labelling effort and validation meetings held with experts. In all each user devoted to the experience one hour over four consecutive day. This was meant to help the user feel rested and avoid fatigue after every experience.

Joint observation of results from the first and second experiment suggest that physiological signals are modulated by both the subject and the mental state. They also reveal that when we fix one of the variables, we can accurately predict the other. However, results from the third experiment support the idea that subject traits have a larger contribution to the physiological signal than the component related to the concentration mental state. This is aligned with recent findings for Electroencephalography (EEG) signals Arevalillo-Herráez et al. (2019), and discourages the use of inter-subject approaches that aim to generate a single model that is valid across the entire population.

CONCLUSSIONS AND FUTURE WORK

One current challenge in user-centered adaptive systems is an accurate detection of relevant mental states that contribute to improve adaptation capabilities. One such a state that directly influences engagement and the success of the learner is concentration. In this work, we have proved the success of intra-subject HMM models at detecting concentration from a set of physiological signals, and also shown the inability of inter-subject models in this particular context.

Despite the positive results reported in this paper, there are still a number of ways in which this work can and will be extended. First, there is a need to improve labelling methodologies, to be able to work with mistake-free data that allows for a better estimation of accuracy measures. Second, the already available labels in the same data set can be used to build predicting tools for other mental/affective states. Third, despite the negative results at using inter-subject models, we consider that they have to be further explored using other alternatives such as the subject-based normalization proposed in Arevalillo-Herráez et al. (2019) for EEG signals, and also a larger dataset with more subjects involved. The development of inter-subject models that can be used on previously unseen subjects is a key issue from a practical perspective, and would open the door to a seamless integration of this type of technology on today's learning applications.

In order to further study these aspects and advance the methodological approach and developments that open the door towards the development of affect-aware user-centered adaptive systems in realistic educational scenarios, we have started two new projects financed by the Spanish Ministry of Education. These are ITS-MathPS and INT2AFF. The first of these projects attempts to improve the learning of word problem solving by using, in between other data, the student's personal cognitive and affective characteristics as a solver. The second aims to advance the methodological and practical developments required to address the intertwine relationship between the learner's affective and cognitive states, as the centre and the target of a multisensorial affect-aware user-centred adaptive learning system, which considers the given context in order to provide the most appropriate response to a particular learner in a given situation.

ACKNOWLEDGMENTS

This research was partly supported by the Spanish Ministry of Economy and Competitiveness through projects TIN2014-59641-C2-1-P, PGC2018-096463-B-I00, TIN2014-59641-C2-2-P and the Spanish Ministry of Science, Innovation and Universities through the project PGC2018-102279-B-I00.

REFERENCES

Ainley M., 2006. *Connecting with learning: Motivation, affect and cognition in interest processes. Educational Psychology Review*, 18, no. 4, 391–405.

Arevalillo-Herráez M.; Chicote-Huete G.; Ferri F.J.; Ayesh A.; Boticario J.G.; Katsigiannis S.; Ramzan N.; and González P.A., 2019. *On using EEG signals for emotion modeling and biometry.* In *33rd European Simulation and Modelling Conference.* European Multidisciplinary Society for Modelling and Simulation Technology, 229–233.

Arevalillo-Herráez M.; Cobos M.; Roger S.; and García-Pineda M., 2019. *Combining inter-subject modeling with a subject-based data transformation to improve affect recognition from EEG signals. Sensors (Switzerland)*, 19, no. 13.

Arevalillo-Herráez M.; Marco-Giménez L.; Arnau D.; and González-Calero J.A., 2017. *Adding sensor-free intention-based affective support to an Intelligent Tutoring System. Knowl-Based Syst*, 132, 85–93.

Arnau-Gonzalez P.; Arevalillo-Herraez M.; Katsigiannis S.; and Ramzan N., 2020. *On the influence of affect in EEG-based subject identification. IEEE Transactions on Affective Computing*, Early Acess.

Ayesh A.; Arevalillo-Herraez M.; and Ferri F.J., 2014. *Cognitive reasoning and inferences through psychologically based personalised modelling of emotions using associative classifiers.* In *Proceedings of 2014 IEEE 13th International Conference on Cognitive Informatics and Cognitive Computing, ICCI*CC 2014.* Institute of Electrical and Electronics Engineers Inc., 67–72.

Baker R.S.; D'Mello S.K.; Rodrigo M.M.T.; and Graesser A.C., 2010. *Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive-affective states during interactions with three different computer-based learning environments. International Journal of Human Computer Studies*, 68, no. 4, 223–241.

Baker R.S.J.D.; Kalka J.; Aleven V.; Rossi L.; Gowda S.M.; Wagner A.Z.; Kusbit G.W.; Wixon M.; Salvi A.; and Ocumpaugh J., 2012. *Towards Sensor-Free Affect Detection in Cognitive Tutor Algebra.* Tech. rep.

Blanchard E.G.; Volfson B.; Hong Y.J.; and Lajoie S.P., 2009. *Affective artificial intelligence in education: From detection to adaptation.* In *Frontiers in Artificial Intelligence and Applications*. IOS Press, vol. 200, 81–88.

Botelho A.F.; Baker R.S.; and Heffernan N.T., 2017. *Improving Sensor-Free Affect Detection Using Deep Learning.* Tech. rep. URL http://tiny.cc/affectdata.

Corno L. and Mandinach E.B., 1983. *The Role Of Cognitive Engagement in Classroom Learning and Motivation. Educational Psychologist*, 18, no. 2, 88–108.

Cunha-Perez C.; Arevalillo-Herráez M.; Marco-Giménez L.; and Arnau D., 2018. *On Incorporating Affective Support to an Intelligent Tutoring System: an Empirical Study. IEEE-RITA*, 13, no. 2, 63–69.

Duan R.N.; Zhu J.Y.; and Lu B.L., 2013. *Differential entropy feature for EEG-based emotion classification.* In *International IEEE/EMBS Conference on Neural Engineering, NER*. 81–84.

Ekman R., 1997. *What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS).* Oxford University Press, USA.

Farzaneh A.H.; Kim Y.; Zhou M.; and Qi X., 2019. *Developing a deep learning-based affect recognition system for young children.* In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Springer Verlag, vol. 11626 LNAI, 73–78.

Garbas J.U.; Ruf T.; Unfried M.; and Dieckmann A., 2013. *Towards robust real-time valence recognition from facial expressions for market research applications.* In *Proceedings - 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction, ACII 2013*. 570–575.

Ghaleb E.; Popa M.; Hortal E.; Asteriadis S.; and Weiss G., 2019. *Towards Affect Recognition through Interactions with Learning Materials.* In *Proceedings - 17th IEEE International Conference on Machine Learning and Applications, ICMLA 2018*. Institute of Electrical and Electronics Engineers Inc., 372–379.

Hajlaoui A.; Chetouani M.; and Essid S., 2018. *EEG-based Inter-Subject Correlation Schemes in a Stimuli-Shared Framework: Interplay with Valence and Arousal.* Tech. rep.

Hsu C.C.; Chen H.C.; Su Y.N.; Huang K.K.; and Huang Y.M., 2012. *Developing a Reading Concentration Monitoring System by Applying an Artificial Bee Colony Algorithm to E-Books in an Intelligent Classroom. Sensors*, 12, no. 10, 14158–14178. URL http://www.mdpi.com/1424-8220/12/10/14158.

Kim J. and André E., 2008. *Emotion recognition based on physiological changes in music listening. IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30, no. 12, 2067–2083.

Kosmyna N. and Maes P., 2019. *AttentivU: An EEG-Based Closed-Loop Biofeedback System for Real-Time Monitoring and Improvement of Engagement for Personalized Learning. Sensors*, 19, no. 23, 5200. URL https://www.mdpi.com/1424-8220/19/23/5200.

Krishna A., 2012. *An integrative review of sensory marketing: Engaging the senses to affect perception, judgment and behavior. Journal of Consumer Psychology*, 22, no. 3, 332–351.

Krithika L.B. and Lakshmi Priya G.G., 2016. *Student Emotion Recognition System (SERS) for e-learning Improvement Based on Learner Concentration Metric.* In *Procedia Computer Science*. Elsevier B.V., vol. 85, 767–776.

Liu N.H.; Chiang C.Y.; and Chu H.C., 2013. *Recognizing the Degree of Human Attention Using EEG Signals from Mobile Sensors. Sensors*, 13, 10273–10286. URL https://www.mdpi.com/1424-8220/13/8/10273.

Moussa M.B. and Magnenat-Thalmann N., 2009. *Applying Affect Recognition in Serious Games: The PlayMancer Project.* Tech. rep.

Nizetha Daniel K.; Kamioka E.; Daniel K.N.; and Kamioka E., 2017. *Detection of Learner's Concentration in Distance Learning System with Multiple Biological Information. Journal of Computer and Communications*, 5, 1–15. URL http://www.scirp.org/journal/jcchttp://creativecommons.org/licenses/by/4.0/.

Ocumpaugh J.; Baker R.; Gowda S.; Heffernan N.; and Heffernan C., 2014. *Population validity for* *educational data mining models: A case study in affect detection. British Journal of Educational Technology*, 45, no. 3, 487–501.

Pardos Z.A.; Baker R.S.; San Pedro M.; Gowda S.M.; and Gowda S.M., 2014. *Affective States and State Tests: Investigating How Affect and Engagement during the School Year Predict End-of-Year Learning Outcomes. Journal of Learning Analytics*, 1, no. 1, 107–128. URL `https://epress.lib.uts.edu.au/journals/index.php/JLA/article/view/3536`.

Pekrun R.; Goetz T.; Daniels L.M.; Stupnisky R.H.; and Perry R.P., 2010. *Boredom in Achievement Settings: Exploring Control-Value Antecedents and Performance Outcomes of a Neglected Emotion. Journal of Educational Psychology*, 102, no. 3, 531–549.

Purnamasari P.D. and Junika T.W., 2019. *Frequency-based EEG human concentration detection system methods with SVM classification.* In *Proceedings: CYBERNETICSCOM 2019 - 2019 IEEE International Conference on Cybernetics and Computational Intelligence: Towards a Smart and Human-Centered Cyber World*. Institute of Electrical and Electronics Engineers Inc., 29–34.

Reyes F.M.; Bolivar C.B.; Olivas V.C.A.; and Serna J.G.G., 2016. *KAPEAN: A supportive tool for observing performance and concentration of children with learning difficulties.* In *Proceedings of 2015 International Conference on Interactive Collaborative and Blended Learning, ICBL 2015*. Institute of Electrical and Electronics Engineers Inc., 52–56.

Salmeron-Majadas S.; Arevalillo-Herráez M.; Santos O.C.; Saneiro M.; Cabestrero R.; Quirós P.; Arnau D.; and Boticario J.G., 2015. *Filtering of spontaneous and low intensity emotions in educational contexts.* In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Springer Verlag, vol. 9112, 429–438.

Saneiro M.; Santos O.C.; Salmeron-Majadas S.; and Boticario J.G., 2014. *Towards emotion detection in educational scenarios from facial expressions and body movements through multimodal approaches. Scientific World Journal*, 2014.

Santos O.C.; Uria-Rivas R.; Rodriguez-Sanchez M.C.; and Boticario J.G., 2016. *An open sensing and acting platform for context-aware affective support in ambient intelligent educational settings. IEEE Sensors Journal*, 16, no. 10, 3865–3874. URL `http://ieeexplore.ieee.org/document/7425146/`.

Sharma P.; Joshi S.; Gautam S.; Filipe V.; Reis M.; and Reis M.C., 2019. *IET Computer Vision Student Engagement Detection Using Emotion Analysis, Eye Tracking and Head Movement with Machine Learning. Student Engagement Detection Using Emotion Analysis, Eye Tracking and Head Movement with Machine Learning.* Tech. rep.

Suryanti S.; Arifani Y.; Zawawi I.; and Fauziyah N., 2019. *Student's engagement behaviour and their success in abstract algebra: Structural equation modelling approach.* In *Journal of Physics: Conference Series*. Institute of Physics Publishing, vol. 1188.

Tokuno S.; Tsumatori G.; Shono S.; Takei E.; Yamamoto T.; Suzuki G.; Mituyoshi S.; and Shimura M., 2011. *Usage of emotion recognition in military health care.* In *2011 Defense Science Research Conference and Expo, DSR 2011*.

Uria-rivas R.; Rodriguez-sanchez M.C.; Santos O.C.; Vaquero J.; and Boticario J.G., 2019. *Impact of physiological signals acquisition in the emotional support provided in learning scenarios. Sensors (Switzerland)*, 19, no. 20, 4520. URL `https://www.mdpi.com/1424-8220/19/20/4520`.

Whitehill J.; Serpell Z.; Lin Y.C.; Foster A.; and Movellan J.R., 2014. *The faces of engagement: Automatic recognition of student engagement from facial expressions. IEEE Transactions on Affective Computing*, 5, no. 1, 86–98.