Scientific Research Publishing

Automatic Detection of Learner Engagement Using Machine Learning and Wearable Sensors

Meredith Carroll[1], Mitchell Ruble[2], Mark Dranias[2], Summer Rebensky[1], Maria Chaparro[1], Joanna Chiang[2], Brent Winslow[2]*

[1]College of Aeronautics, Florida Institute of Technology, Melbourne, FL, USA
[2]Design Interactive, Inc., Orlando, FL, USA
Email: *brent.winslow@designinteractive.net

Abstract

Training can now be delivered on a large scale through mobile and web-based platforms in which the learner is often distanced from the instructor and their peers. In order to optimize learner engagement and maximize learning in these contexts, instructional content and strategies must be engaging. Key to the development and study of such content and strategies, and adaptation of instructional techniques when learners become disengaged, is the ability to objectively assess engagement in real-time. Previous self-reported metrics, or expensive EEG-based engagement measures are not appropriate for large-scale platforms due to their complexity and cost. Here we describe the development and testing of a measurement and classification technique that utilizes non-invasive physiological and behavioral monitoring technology to directly assess engagement in classroom, simulation, and live training environments. An experimental study was conducted with 45 students and first responders in a unmanned aircraft systems (UAS) training program to assess the ability to accurately assess learner engagement and discriminate between levels of learner engagement within classroom, simulation and live environments via physiological and behavioral inputs. A series of engagement classifiers were developed using cardiovascular, respiratory, electrodermal, movement, and eye-tracking features that were able to successfully classify engagement levels at an accuracy level of 85% with eye-tracking features included or 81% without eye-tracking features. This approach is capable of monitoring, assessing, and tracking learner engagement across learning situations and contexts, and providing real-time and after action feedback to support instructors in modulating learner engagement.

Keywords

Learner Engagement, Physiological Monitoring, Eye Movements,

Electrodermal Activity, Cardiovascular Physiology

1. Introduction

Digital technologies are increasingly being exploited in self-directed or guided educational settings to provide individualized opportunities for learning. Emerging technology can now deliver training on a larger scale through platforms such as massive open online courses (MOOCs), instructional use of games, and virtual classrooms [1]. These emerging training environments rely heavily on learner-initiated involvement and motivation before, during, and after training [2]. Self-motivation guides personal effort and the allocation of resources toward work [3], and leads to significantly greater declarative knowledge and skill acquisition, higher post-training self-efficacy, and improved performance [4]. Personal motivation represents the driving force behind learning activities and can lead to learner engagement, a person's active involvement in a learning activity [5]. High levels of engagement have been associated with a state of flow, in which an individual becomes completely engaged in a task marked by high levels of interactivity, challenge, and feedback; a state which has been found to lead to improved task performance and learning outcomes [6]. Learner engagement is influenced by a range of factors related to the individual learner, the learning tasks, and the learning environment [7]. The ability to measure and optimize learner engagement during training holds the potential to increase the transfer of training to practice, leading to enhanced digital training effectiveness [6].

A number of approaches to measure learner engagement have been reported previously. Self-report questionnaires and interviews provide non-invasive and inexpensive methods to assess learner engagement, but do not provide results in real-time [8], and are associated with bias which affects the reliability and validity of such approaches [9]. Analysis of facial expressions using computer vision can determine affective components of learner engagement and can be applied to groups [10], but may be considered obtrusive and suffer from privacy concerns [11]. Cognitive engagement classifiers have been reported using laboratory grade physiological sensors, but require high cost sensor technology, and are complex from a usability and data analysis perspective. For example, a previous electroencephalography (EEG)-based engagement classifier returns a 3-level score between relaxed wakefulness and high engagement based on individual performance in a protracted psychomotor vigilance task [12]. However, the expense associated with EEG equipment, complex data interpretation, and sensor montaging have limited the use of EEG in learning settings. Alternatively, there are a range of physiological measures that provide an opportunity for validly and non-invasively measuring engagement. Measures such as electrodermal activity (EDA) [13], heart rate and heart rate variability (HRV) [14], and gross body movement [15] have been used to monitor learner engagement and learning performance. Previous research has also indicated that analysis of eye tracking data can offer insight in-

to engagement during gaming [16], online search [17], and online conversations [18]. While there is theoretical support for the use of these measures to capture engagement, to date there is limited empirical research identifying valid measures of engagement that are noninvasive, easy to use, and scale to large populations.

There is a need to evaluate emerging sensor technology to measure learner engagement objectively in real-time, via noninvasive, wireless sensors. Here we describe the development and testing of a measurement and classification technique that utilizes non-invasive physiological monitoring technology with subjective self-report measures to directly assess engagement in classroom, simulation, and live training environments to support instructor interventions to increase engagement. Using unmanned aircraft systems (UAS) instructional materials, we delivered instructional content via video-based, simulated flight, and live flight training tasks that were designed to vary in cognitive engagement based on varying levels of interactivity, challenge, feedback, and immersion. We hypothesized that participant engagement, and associated participant physiology and behavior, would differ between tasks, with increasing levels from video to simulation to live training. This was also hypothesized to provide features necessary to develop a high accuracy classifier of engagement state.

2. Methods

2.1. Participants

All methods involving participants were approved by an Institutional Review Board (IRB; Florida Institute of Technology [FIT]). A priori power analysis with $\alpha = 0.05$, $\beta = 0.8$, and d = 0.4 indicated a sample size of 41 participants needed, with an additional 20% included to address participant attrition. Forty-nine adult participants from the Melbourne, FL, USA area were recruited from an FIT UAS course, FIT drone club, and UAS operators from local first responder agencies. Participation was voluntary and all student participants were given the option to perform an alternate task if they were not interested in participating in the study. Student participants were given extra credit in the UAS course for their participation.

2.2. Materials

A demographic survey was used to determine participant gender, age, and UAS experience. The Flow Short Scale (FSS) was used to assess self-reported engagement [19], as a state of flow is proposed to represent a high level of learner engagement [7]. A Dell Precision M2800 laptop PC, with 8 GB RAM and a quad core CPU was used to provide video and simulation training. Simulation training used Real Flight 7.5 software, along with a U818A-1 quadcopter controller (UDIRC Technology, Guangdong, China). For live UAS training, the U818A-1 quadcopter was flown through a series of physical obstacles in a high bay learning environment (Figure 1).

Figure 1. Experimental tasks. Top panel: participants viewed instructional flight videos while holding UAS controllers and mimicking movements in the video task. Center panel: participants controlled a virtual UAS in a number of flight tasks within Real Flight 7.5 software as part of the simulation task. Bottom panel: participants controlled a UAS through a series of physical obstacles as part of the live flight task. Depicted participants provided written informed consent for publication of this figure.

The Equivital EQ02 system (Hidalgo; Cambridge, UK) was used to collect electrocardiography (ECG; 256 Hz), electrodermal activity (EDA; 2 Hz), respiratory rate (0.0667 Hz), and accelerometry (25.6 Hz). The VT3 mini eye tracker (EyeTech; Mesa, AZ, USA) was used to quantify participants' gaze location (45 Hz) during computer-based tasks (video training and simulation, described below).

2.3. Experimental Procedure

Upon arrival participants provided written, informed consent and were briefed on the experimental learning tasks. Participants then responded to a demographic questionnaire and were fitted with the Equivital device. Participants were instructed to stand comfortably and read a magazine for 5 minutes while a physi-

ological baseline was collected. After the baseline, participants were directed to the adjacent room for a series of computer-based UAS tasks with varying levels of engagement. Participants stood in front of a workstation, with height adjusted to ensure adequate alignment with the eye tracker, and the eye tracker was calibrated. Participants then watched a series of videos on how to use the UAS controllers while holding the UAS controller and mimicking the movements displayed in the videos for a duration of six minutes. The video training task was hypothesized to have low levels of engagement, due to low levels of interactivity, challenge, goal clarity, feedback, and immersion. The next task consisted of performing a series of UAS obstacles in the Real Flight 7.5 simulator, and was hypothesized to have moderate to high levels of engagement due to high levels of interactivity, challenge, goal clarity and feedback and moderate levels of immersion. Finally, participants moved to a live, indoor, high-bay, learning environment where they completed the live training task, which included flying a UAS quadcopter through a series of physical obstacle courses. The live task was hypothesized to induce high levels of engagement due to high levels of interactivity, challenge, goal clarity, feedback and immersion of the learning environment. In both simulation and live training tasks, participants completed as many obstacle course challenge levels as they could, within the six minute session, up to the maximum of 10 levels. Challenge levels increased in difficulty, requiring increasing levels of skill such as hovering, flying in various orientations, and landing. Participants in simulation and live training also received clear goals related to their objective prior to the challenge and received performance feedback after each failed attempt at a level. Two performance measures were captured during the simulation and live tasks. Outcome performance was assessed based on the number of levels completed. Errors were assessed based on the number of times the UAS bumped into an object or crashed. Participant performance was not assessed during the video task. Participants responded to the FSS survey immediately following each task. All participants completed the tasks in the same order. Eye-tracking was not captured during the live training task.

2.4. Data Modeling

Physiological data analysis and classifier development consisted of synchronization between data capture systems, feature extraction, and data modeling. Data analysis was implemented in Python 3.7 with scikit-learn [20]. Data from each sensor was graphically visualized and synchronized with events of interest including: baseline; video training; simulation training; and live training. Any epochs of signal loss, such as when participants were not looking at the screen for eye tracking, were also identified.

In order to preserve temporal relationships among variables, timestamps for data arising from the Equivital system were synchronized. Respiration rate and heart rate were up sampled with linear interpolation to 0.133 Hz. The Equivital and EyeTech systems were synchronized with an NIST clock and the eye-tracker

was adjusted to synchronize with the Equivital clock. Classification was performed on 30 second blocks of data and synchronization errors were constrained below 5 seconds.

Features were extracted from the raw physiological data, baseline normalized, and updated every 30 seconds to provide input to an engagement classifier. Features which were not predictive of the user state were removed, and dimensionality reduction was performed to combine features as needed. A random forest algorithm was utilized to rank feature significances collected from the Equivital and eye tracking systems. Features from the Equivital system included mean EDA, respiration rate, accelerometer signal magnitude area (SMA), mean heart rate (HR), QRS duration, QRS sum [21], and heart rate variability (HRV) measured by the ratio of low frequency (LF) HRV, defined as the sum of the power spectral density (psd) of the IBI signal in the 0.04 to 0.15 Hz range, over high frequency (HF) HRV, defined as the sum of the psd from 0.15 to 0.4 Hz range. Features derived from the eye tracker included the mean fixation number and mean fixation duration, defined as maintenance of visual gaze on a single location, defined as 3 visual degrees and >100 ms [22].

Once the features were selected, two separate models were trained: one in which the participants' data were broken down by task into 2 states (low and high engagement) corresponding to the video and simulation tasks due to the lack of eye tracking data in the live task; and one in which the participant's data were broken down by tasks into 2 states (low and high engagement) corresponding to the video and live tasks for situations in which eye-tracking is not feasible. To predict the engagement state of the participants, the extracted features were subjected to a logistic regression analysis. A series of models were trained to map the 5 selected continuous physiological variables to two discrete states (low/high engagement) using 5 fold cross-validation, iteratively repeated 100 times, to minimize bias and variance. Final model parameters were taken from the model with the highest classification accuracy across both testing and training sets.

2.5. Data Analysis and Statistics

Participant performance was assessed using two measures. First, outcome performance was measured based on the number of challenge levels that a participant successfully completed, out of a maximum of 10, with participants receiving 10 points per level completed. Second, participant errors were assessed using a count of the number of times that the UAS bumped into an object or crashed.

Differences between training task self-reported engagement were evaluated using repeated measures ANOVA with $\alpha = 0.05$. Differences in performance between the video and simulation tasks were analyzed using repeated measures t tests with $\alpha = 0.05$. Statistical analyses were performed in IBM SPSS Statistics version 23.

3. Results

The sociodemographic factors in the study sample are listed in Table 1. Four participants were removed from the study due to missing physiological data, leaving forty-five participants for data analysis and modeling. Most participants were male, of an average age 25.7 ± 9.9 years, and had some previous UAS experience, defined as UAS usage 1 - 3 times per year.

Users evaluated each training task immediately following exposure using FSS scale (Figure 2). Differences were observed in self-reported engagement [$F(2, 88)$ = 13.697, p = 0.0000066]. Post-hoc testing indicated significant differences between video and simulation training (p = 0.003), between video and live training (p = 0.0001), and between simulation and live training (p = 0.007).

Table 1. List of sociodemographic factors in the study sample.

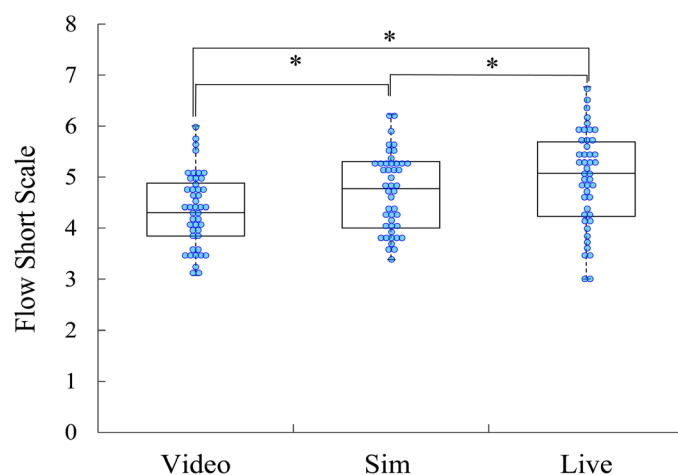|	Study sample % (*n*)
Gender	
Male	82 (37)
Female	18 (8)
Age group	
17 - 20	31 (14)
21 - 25	42 (19)
26 - 30	11 (5)
>30	16 (7)
UAS Experience	
Never flown a UAS	33 (15)
Sometimes (1 - 3 times/year)	40 (18)
Often (monthly)	27 (12)

Figure 2. Boxplots overlaid with raw data of self-reported engagement using the FSS survey. Participants rated the simulated training session significantly higher than the video training, and the live training significantly higher than both the simulated and video training sessions. *p < 0.05.

Users performed significantly better in simulated UAS flying compared to live [t(44) = 6.961, p = 0.000000013], and made similar errors in simulation as compared to live [t(44) = −1.539, p = 0.131] (**Figure 3**).

Average, session-level physiological and behavioral data is shown in **Table 2**. Features were extracted from individual, baseline-normalized physiological and behavioral data and used to train an algorithm of engagement state.

All data used to train the model was labeled either as high (sim or live) or low (video) engagement. This data was then preprocessed to ensure the model: 1) generalizes across individuals and 2) varies on a time scale useful for learning interventions. This was accomplished by training the model with feature data from 45 different subjects and breaking the data into 30 second windows. A logistic regression model with both behavioral (eye tracker) and physiological (Equivital) features was capable of predicting whether a 30 second physiological data set

Table 2. Average (SD) session level physiological and behavioral (eye-tracking) data from the cohort.

|	Baseline	Video	Simulation	Live
HR (bpm)	86.03 (14.98)	90.86 (12.60)	90.22 (19.32)	92.88 (17.52)
QRS dur (ms)	125.65 (13.08)	124.52 (14.15)	129.64 (14.85)	137.43 (15.83)
QRS sum (ms)	126.12 (13.36)	123.18 (13.92)	129.82 (14.79)	138.92 (16.32)
LF/HF HRV	4.45 (3.94)	5.65 (4.69)	4.39 (3.37)	3.60 (2.64)
EDA (μS)	6.70 (4.57)	6.95 (5.56)	7.24 (4.40)	7.44 (4.44)
RR (bpm)	18.95 (3.35)	16.45 (6.11)	17.89 (3.47)	17.03 (3.30)
SMA (mg)	19.00 (1.38)	19.08 (1.63)	19.02 (1.41)	19.33 (1.41)
Fixation Number		16.13 (3.61)	7.58 (2.23)	
Fixation Duration (s)		2.09 (1.18)	6.98 (4.18)	
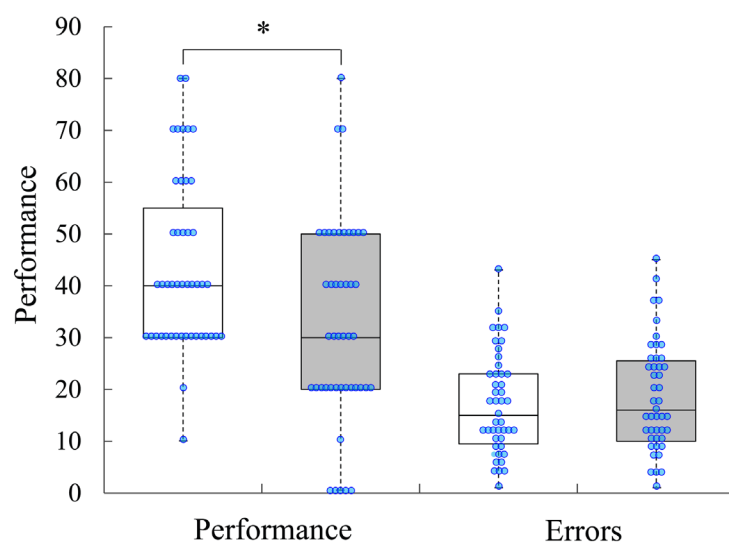
Figure 3. Boxplots overlaid with raw data of performance in simulated (white) and live (gray) UAS flight. *p < 0.05.

came from a high or low engagement condition with a training accuracy of 79% and a testing accuracy of 85% (Figure 4). Parameters included fixation number (w = 0.002), EDA (w = 0.063), fixation duration (w = −0.030), QRS duration (w = −0.205), and heart rate (w = 1.125), with an intercept of 0.0164. An additional logistic regression classifier was developed without eye-tracking features to account for situations when eye-tracking is not available, and was capable of predicting whether a 30 second physiological data set came from the high or low engagement condition with a training accuracy of 76% and a testing accuracy of 81% (Figure 4). Parameters included heart rate (w = −0.021), QRS sum (w = 0.041); LF/HF HRV (w = 0.074, SMA (w = 0.047), EDA (w = 0.050), and respiration rate (w = 0.091), with an intercept of 7.163.

4. Discussion

The ability to measure and ultimately optimize learner engagement is considered to be one of the most effective countermeasures in combating student dropout, disaffection, and low performance [11]. Our results suggest that utilizing low non-invasive physiological measures that capture electrodermal and cardiovascular measures can provide a window into the real-time state of learner engagement. As predicted, results indicated that learner engagement levels significantly increased with increasing levels of interactivity, challenge, goal clarity, feedback and immersion. In addition, features derived from physiological sensing and eye-tracking equipment were able to successfully classify engagement with high accuracy.

In the current effort we developed a two class model of engagement using physiological and behavioral features, which classified engagement as high or low, with 85% accuracy with eye-tracking features included, and 81% accuracy without eye-tracking features. Both models are based on significant differences observed in self-reported engagement, user physiology, and task performance between
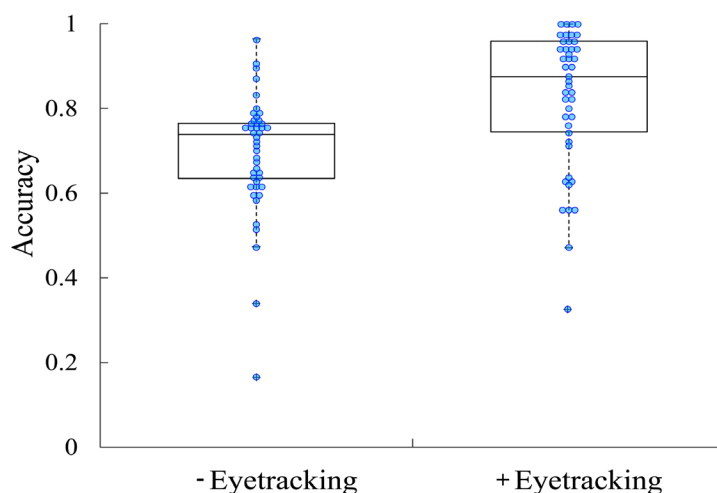

Figure 4. Logistic regression model accuracy. The model average accuracy was 81% without eyetracking features, and 85% with eyetracking features.

video, simulated, and live environments. Both simulated and live environments indicating significant increases in engagement as compared to video, and were labeled as high engagement in data modeling. Similar levels of accuracy have been reported recently in measuring engagement using features derived from wearable physiological sensors, touchscreens, and optical biomechanics [8], or using EDA only [11]. Given the physiological and behavioral differences observed between scenarios of increasing engagement, we leveraged a random forest classifier to select classifier features. As compared to other feature selection techniques, random forest approaches have shown the highest accuracy in modeling user state including engagement [8].

The engagement classifier reported here is associated with features derived from physiological and behavioral data. Previous research has indicated that various physiological metrics are associated with engagement, including features derived from EDA [23] [24] [25], ECG [24] [26], and from eye-tracking [27]. Our results suggest a decrease in heart rate and an increase in HRV occurs during increasingly engaging tasks, likely associated with increasing parasympathetic activity [24]. Our results also indicate a decrease in fixations and increasing fixation duration with increasing engagement. Such results agree with theory that posits that individuals will direct visual attention to salient images under engaging conditions because they are thinking more deeply about those images [27].

Individuals reported increasing levels of engagement across video, simulation and live tasks, as measured using the FSS. Underlying these results are aspects of interactivity, task challenge, goal clarity, feedback and immersion. Interactivity has been shown to impact engagement by providing the individual with control in the learning task through technology, leading to increased learner motivation, interest, and learning gains [28] [29]. The level of challenge associated with a learning task is related to engagement as the presence, level, and appropriateness of a challenge can lead to intrinsic motivation, flow and micro-engagement [30] [31] [32]. However, the level of challenge must be appropriate to an individual's capabilities to increase engagement. A task where the challenge is too advanced leads to a stress response which can reduce the motivation of the individual [26], [33] and reduce motor and cognitive performance [34]. If a task is not challenging enough or too easy it may produce apathy and a reduced attention due to a lack of motivation [33] [35]. An important facet in challenge is that as the individual learns, the challenge generally becomes easier, thus requiring the challenge to grow with the individual's progress [36] [37] [38]. Increasing challenge between the simulation and live flight tasks was associated with decreasing performance. Further, engagement levels can also be positively influenced by presenting clear goals and performance feedback as these elements can improve learner motivation and strategy use [39]. Finally, immersion has been shown to influence engagement through increased interest and motivation towards learning tasks, ultimately leading to higher engagement [40] [41].

In the current effort, we were able to achieve high accuracy classification of engagement similar to recent reports [8] [11]. However, results of this study and

the implications for use in educational settings should be interpreted with caution. The inclusion of eye-tracking features in this modeling effort are difficult to scale, and are associated with necessary calibration and restriction of user head movements [10]. Similarly, physiological sensors add cost, are difficult to scale, have varying levels of data quality [42], and may be associated with privacy issues [43].

In summary, this research demonstrated the capability of monitoring, and assessing individualized learner engagement across learning situations and contexts using physiological and behavioral inputs. Application of such an approach can facilitate to support them in determining where they need to adjust training to optimize learner engagement.

Acknowledgements

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Curran, V., *et al.* (2017) A Review of Digital, Social, and Mobile Technologies in Health Professional Education. *Journal of Continuing Education in the Health Professions*, **37**, 195-206. https://doi.org/10.1097/CEH.0000000000000168

[2] Cannon-Bowers, J. and Bowers, C. (2010) Synthetic Learning Environments: On Developing a Science of Simulation, Games, and Virtual Worlds for Training. In: Kozlowski, S.W.J. and Salas, E., Eds., *Learning, Training, and Development in Organizations*, Routledge, New York, 229-261.

[3] Rangel, B., *et al.* (2015) Rules of Engagement: The Joint Influence of Trainer Expressiveness and Trainee Experiential Learning Style on Engagement and Training Transfer. *International Journal of Training and Development*, **19**, 18-31. https://doi.org/10.1111/ijtd.12045

[4] de Manzano, O., *et al.* (2010) The Psychophysiology of Flow during Piano Playing. *Emotion*, **10**, 301-311. https://doi.org/10.1037/a0018432

[5] Appleton, J.J., *et al.* (2006) Measuring Cognitive and Psychological Engagement: Validation of the Student Engagement Instrument. *Journal of School Psychology*, **44**, 427-445. https://doi.org/10.1016/j.jsp.2006.04.002

[6] Engeser, S. (2012) Advances in Flow Research. Springer-Verlag, New York. https://doi.org/10.1007/978-1-4614-2359-1

[7] Carroll, M., *et al.* (2019) An Applied Model of Learner Engagement and Strategies for Increasing Learner Engagement in the Modern Educational Environment. *Interactive Learning Environments*. https://doi.org/10.1080/10494820.2019.1636083

[8] Huynh, S., *et al.* (2018) EngageMon: Multi-Modal Engagement Sensing for Mobile Games. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, **2**, 13. https://doi.org/10.1145/3191745

[9] Althubaiti, A. (2016) Information Bias in Health Research: Definition, Pitfalls, and Adjustment Methods. *Journal of Multidisciplinary Healthcare*, **9**, 211. https://doi.org/10.2147/JMDH.S104807

[10] Dewan, M.A.A., Murshed, M. and Lin, F. (2019) Engagement Detection in Online Learning: A Review. *Smart Learning Environments*, **6**, Article No. 1. https://doi.org/10.1186/s40561-018-0080-z

[11] Di Lascio, E., Gashi, S. and Santini, S. (2018) Unobtrusive Assessment of Students' Emotional Engagement during Lectures Using Electrodermal Activity Sensors. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, **2**, 103. https://doi.org/10.1145/3264913

[12] Berka, C., *et al.* (2007) EEG Correlates of Task Engagement and Mental Workload in Vigilance, Learning, and Memory Tasks. *Aviation, Space, and Environmental Medicine*, **78**, B231-B244.

[13] Hardy, M., *et al.* (2013) Physiological Responses to Events during Training: Use of Skin Conductance to Inform Future Adaptive Learning Systems. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, **57**, 2101-2105. https://doi.org/10.1177/1541931213571468

[14] Murata, A. (2005) An Attempt to Evaluate Mental Workload Using Wavelet Transform of EEG. *Human Factors*, **47**, 498-508. https://doi.org/10.1518/001872005774860096

[15] Witchel, H.J., *et al.* (2012) Comparing Four Technologies for Measuring Postural Micromovements during Monitor Engagement. *Proceedings of the 30th European Conference on Cognitive Ergonomics*, Edinburgh, 28-31 August 2012, 189-192. https://doi.org/10.1145/2448136.2448178

[16] Renshaw, T., Stevens, R. and Denton, P.D. (2009) Towards Understanding Engagement in Games: An Eye-Tracking Study. *On the Horizon*, **17**, 408-420. https://doi.org/10.1108/10748120910998425

[17] Lorigo, L., *et al.* (2008) Eye Tracking and Online Search: Lessons Learned and Challenges Ahead. *Journal of the Association for Information Science and Technology*, **59**, 1041-1052. https://doi.org/10.1002/asi.20794

[18] Nakano, Y.I. and Ishii, R. (2010) Estimating User's Engagement from Eye-Gaze Behaviors in Human-Agent Conversations. *Proceedings of the 15th International Conference on Intelligent User Interfaces*, Hong Kong, 7-10 February 2010, 139-148. https://doi.org/10.1145/1719970.1719990

[19] Engeser, S. and Rheinberg, F. (2008) Flow, Performance and Moderators of Challenge-Skill Balance. *Motivation and Emotion*, **32**, 158-172. https://doi.org/10.1007/s11031-008-9102-4

[20] Pedregosa, F., *et al.* (2011) Scikit-Learn: Machine Learning in Python. *Journal of Machine Learning Research*, **12**, 2825-2830.

[21] Zong, W., Moody, G. and Jiang, D. (2003) A Robust Open-Source Algorithm to Detect Onset and Duration of QRS Complexes. *Computers in Cardiology*, **30**, 737-740. https://doi.org/10.1109/CIC.2003.1291261

[22] Winslow, B.D., *et al.* (2013) Combining EEG and Eye Tracking: Using Fixation-Locked Potentials in Visual Search. *Journal of Eye Movement Research*, **6**, 5.

[23] Kim, P.W. (2018) Real-Time Bio-Signal-Processing of Students Based on an Intelligent Algorithm for Internet of Things to Assess Engagement Levels in a Classroom. *Future Generation Computer Systems*, **86**, 716-722. https://doi.org/10.1016/j.future.2018.04.093

[24] Cheng, W.K.R. (2014) Relationship between Visual Attention and Flow Experience in a Serious Educational Game: An Eye Tracking Analysis. George Mason University, Fairfax.

[25] Richardson, D.C., *et al.* (2018) Measuring Narrative Engagement: The Heart Tells the Story. https://doi.org/10.1101/351148

[26] Peifer, C., *et al.* (2014) The Relation of Flow-Experience and Physiological Arousal under Stress—Can U Shape It? *Journal of Experimental Social Psychology*, **53**, 62-29. https://doi.org/10.1016/j.jesp.2014.01.009

[27] Plochl, M., Ossandon, J.P. and Konig, P. (2012) Combining EEG and Eye Tracking: Identification, Characterization, and Correction of Eye Movement Artifacts in Electroencephalographic Data. *Frontiers in Human Neuroscience*, **6**, 278. https://doi.org/10.3389/fnhum.2012.00278

[28] Adams, W.K., *et al.* (2008) A Study of Educational Simulations Part I Engagement and Learning. *Journal of Interactive Learning Research*, **19**, 397-419.

[29] Rodgers, D.L. and Withrow-Thorton, B.J. (2005) The Effect of Instructional Media on Learner Motivation. *International Journal of Instructional Media*, **32**, 333.

[30] Bauer, K.N., Brusso, R.C. and Orvis, K.A. (2012) Using Adaptive Difficulty to Optimize Videogame-Based Training Performance: The Moderating Role of Personality. *Military Psychology*, **24**, 148. https://doi.org/10.1080/08995605.2012.672908

[31] Orvis, K.A., Horn, D.B. and Belanich, J. (2008) The Roles of Task Difficulty and Prior Videogame Experience on Performance and Motivation in Instructional Videogames. *Computers in Human Behavior*, **24**, 2415-2433. https://doi.org/10.1016/j.chb.2008.02.016

[32] Sampayo-Vargas, S., *et al.* (2013) The Effectiveness of Adaptive Difficulty Adjustments on Students' Motivation and Learning in an Educational Computer Game. *Computers & Education*, **69**, 452-462. https://doi.org/10.1016/j.compedu.2013.07.004

[33] Kahn, W.A. (1990) Psychological Conditions of Personal Engagement and Disengagement at Work. *Academy of Management Journal*, **33**, 692-724. https://doi.org/10.5465/256287

[34] McEwen, B.S. and Sapolsky, R.M. (1995) Stress and Cognitive Function. *Current Opinion in Neurobiology*, **5**, 205-216. https://doi.org/10.1016/0959-4388(95)80028-X

[35] Nakamura, J. and Csikszentmihalyi, M. (2014) The Concept of Flow. In: *Flow and the Foundations of Positive Psychology*, Springer, Berlin, 239-263. https://doi.org/10.1007/978-94-017-9088-8_16

[36] Dweck, C.S. (1986) Motivational Processes Affecting Learning. *American Psychologist*, **41**, 1040. https://doi.org/10.1037/0003-066X.41.10.1040

[37] Miele, D.B. and Scholer, A.A. (2017) The Role of Metamotivational Monitoring in Motivation Regulation. *Educational Psychologist*, **53**, 1-21. https://doi.org/10.1080/00461520.2017.1371601

[38] Nakamura, J. and Csikszentmihalyi, M. (2002) The Concept of Flow. In: Snyder, C.R. and Lopez, S.J., Eds., *Handbook of Positive Psychology*, Oxford University Press, New York, 89-105.

[39] Gully, S. and Chen, G. (2010) Individual Differences, Attribute-Treatment Interactions, and Training Outcomes. In: Kozlowski, S.W.J. and Salas, E., Eds., *Learning, Training, and Development in Organizations*, Routledge, New York, 3-64.

[40] Deci, E.L. and Ryan, R.M. (2000) The "What" and "Why" of Goal Pursuits: Human

Needs and the Self-Determination of Behavior. *Psychological Inquiry*, **11**, 227-268.
https://doi.org/10.1207/S15327965PLI1104_01

[41] Shapley, K., *et al.* (2011) Effects of Technology Immersion on Middle School Students' Learning Opportunities and Achievement. *The Journal of Educational Research*, **104**, 299-315. https://doi.org/10.1080/00220671003767615

[42] Hernández, N., *et al.* (2017) Data Quality in Mobile Sensing Datasets for Pervasive Healthcare. In: *Handbook of Large-Scale Distributed Computing in Smart Healthcare*, Springer, Berlin, 217-238. https://doi.org/10.1007/978-3-319-58280-1_9

[43] Kreitmair, K.V., Cho, M.K. and Magnus, D.C. (2017) Consent and Engagement, Security, and Authentic Living Using Wearable and Mobile Health Technology. *Nature Biotechnology*, **35**, 617-620. https://doi.org/10.1038/nbt.3887