FacePsy: A Deployable Remote Sensing Framework to Understand Affective Learning States

No Author Given

No Institute Given

Abstract. As online learning becomes more prominent, it is increasingly difficult for instructors to gauge the learning states of their students in real time, thus missing the opportunity to evaluate their teaching effects. While existing approaches suggested systems that require high computing power in detecting student engagement, there is a gap of broad adoption of systems that meet the needs of both students and teachers. Therefore, in this paper, we propose an easily-deployable sensing framework, FacePsy, which unobtrusively senses students' behavioral signals to see if they are fully immersed in online learning. In this proof of concept study, we collected behavior markers from 12 students to validate the feasibility of FacePsy in online learning. We achieved an accuracy of 93.5% (F1 = 0.96) in detecting low versus high flow using a LightGBM binary classifier. The implications of these results allow us to move on to the next stage of expanding the framework's scalability with larger data sets by validating its performance throughout a variety of contexts and tasks.

Keywords: Feasibility · Deployability · Student Engagement · States of Flow · Remote Sensing Framework · Online Learning.

1 Introduction

The worldwide adoption of online learning during the COVID-19 pandemic has demonstrated unprecedented opportunities and challenges for equitable learning and teaching across all levels of education. Reduced engagement of students has been a fundamental setback for online learning compared to face-to-face learning. Some key factors that impact students' engagement include screen fatigue and multitasking [23], a lack of suitable learning environments at home, a lack of feedback and technical skills from both teachers and students, as well as a lack of parental support [24][25]. The advancement of affective computing has laid out the algorithmic foundations towards intelligent learning environments with real-time pedagogical feedback in response to students' learning states such as engagement and flow [13][6][3][4][7][14][11].

In reality, the evident digital divide and privacy concerns have been major barriers to a broad adoption of intelligent learning environments described above, which further exacerbates inequality in education [26]. First, automatic detection technologies of learning states are inaccessible for the majority of students and teachers. In spite of the flourishing of technologies for modeling student learning states in the past decade [4], they are mostly confined to small scale research studies, and are thus out of reach for students and teachers with limited technical skills. Second, existing technologies are impractical for the existing technology infrastructure of learners outside of schools. These detection technologies often require high-speed, high-capacity internet connectivity and significant computing power in order to transmit and process data, which is not the reality for all students, especially those in developing countries or low-income households [27][28]. Third, video recording of minors and in-home environments raises various privacy concerns as family members share the living space and sensitive information may be revealed [29].

To address the above challenges, we propose the FacePsy framework, an easily-deployable, lightweight, privacy-preserving sensing framework designed for students and teachers with limited technical skills and various internet connectivity and computing devices. Individual students can easily install the program on their personal computer and use one-click button to turn on automatic detection at their own choice. To constrain data transmission and preserve privacy, the FacePsy framework directly extracts facial behaviors using a built-in or external camera on the local computer, and only transmits the facial behaviors information for cloud-based processing to detect learning states, without sharing any video recording. To validate the efficiency of FacePsy, we conducted a proof-of-concept user study with 12 participants to detect a student's flow state, a critical learning state of full immersion linked to learning outcomes [30], during online programming activities using personal computers at home.

Our main contributions are as follows:

1. We propose an easily-deployable learning state sensing framework, FacePsy, to tackle the digital divide for equitable online learning at scale by promoting the wide adoption of affect-sensitive learning environments that meet the practical needs of learners and teachers with heterogeneous technology access and skills.

2. We adopt the framework in sensing the state of flow during online at-home learning, and reached a high sensing accuracy at 93.5% and the promising effectiveness of FacePsy to leverage low granularity data, computational power and storage, and privacy preservation for sensing learner states in the wild.

2 Related Works

2.1 Privacy-Preserving Frameworks

There have been several studies that incorporate frameworks of feature extraction processes of learner behaviors similar to FacePsy to detect specific learning states, of which the most similar is Bosch and D'Mello's mind wandering study [3]. Similarly to FacePsy's privacy-preserving features, Bosch and D'Mello incorporate real-time feature (including Head Pose and AUs) extraction due to privacy concerns. However, although their classroom study achieved promising results, it is unclear whether validated studies in lab settings, in-person, using standardized, school-issued equipment are applicable in online learning contexts. It might be difficult to gauge success of deploying the same prediction methods in everyday distance learning. Students may undergo specific distance-induced distractions not seen in classroom-based online learning or a lack of accessibility to the hardware required to run these state of the art technologies. Bosch and D'Mello pointed out that their feature extraction process was "computationally expensive," while using standardized computers [3]. It is unclear if deploying such heavy processes to personal devices with a variety of computational capabilities would elicit the same level of success.

Furthermore, although it is possible to extract features directly from participants, many existing engagement detection studies extracted facial features from user recordings, as seen in such engagement detection in online learning review [4]. Extracting features from videos of participants after the experimental portion of a study has concluded not only adds a privacy risk to participants, but the entire processes of recording, storing, and analyzing videos of participant faces increases the risks of privacy violations. The direct extraction of features removes the need to store videos of the participants themselves, easily identifiable information that is susceptible to potential outside influence.

2.2 Existing Methods in Flow and Engagement Detection

Disregarding the process used for feature extraction, existing learning state detection studies have a variety of characteristics depending on researchers' preferences or goals. When deciding on a ground-truth platform, they may use either a participant's self-report [8][10][6][11] or through an individual's externals observation [13][14]. Existing studies have also used some form of log-file analysis [7][9][10] or sensor data analysis [6][8][11] for learning state detection, as opposed to computer-vision based methods, as seen in this study.

Similarly to these engagement detection studies, flow, a state of total immersion, detection studies have incorporated both a variety ground-truth methods as well as a variety of methods for analysis. When it comes to flow detection in an online learning context, log-file analysis methods that look at a user's interaction with a specific learning systems are widely used [31][32]. This study, however, is more similarly structured to computer-vision based engagement detection studies, which measure participant engagement by analyzing unobtrusively detected facial and positional cues [4]. As mentioned in the previous section, the majority of these studies extract features from pictures or video recordings of participants, which both increases the risk for participant privacy violations as well as requires a large amount of computational power. Furthermore, similarly to Bosch and D'Mello's study [3], it is also unclear if the methods and results of these studies can be replicated in a distance learning setting.

3 Designing a Deployable Remote Sensing Framework

In this section we introduce and describe the creation process and characteristics of the FacePsy framework. We focus on its lightweight - *low use of system components* - computation, its easy deployability, - *how easily it can be adopted by students and instructors* - and its privacy preserving - *minimizing the risk of identifying users* - processes.

3.1 Easy Deployability

Student Interface. The initial version of the app was designed to get objective data from users automatically (24/7) whenever they participated in the class, using their laptops as long as the camera is tuned in. An approach of this nature enables to lower human burden for both instructors and students. However, after a usability test, we found that the privacy concerns made students less inclined to participate. In response to these concerns, we then decided to give users 'full control' of the app by implementing "start and stop buttons" to enhance user controllability, making sure that the app is runs only when users allow the app to collect their behavior signals when doing a specific task. We also provided the data collection status (either active or inactive) and uploading progress to ensure that the app only collects data when prompted by users. Therefore, we implemented three controls for users (Figure 1) in the GUI: (a) the Start button: to start data collection, (b) the Stop button: to stop data collection, and (c) the Survey button: to open the survey link in a browser.

User Identifiers and Data Transmission. Upon installation, the FacePsy framework generates a unique participation identifier for each individual. The unique identifier, with a length of 8 random characters, is generated using shortuuid [17] and does not include any identifying participant markers, such as their name, contact information, or email. Upon clicking the 'Start button' on the FacePsy framework GUI, the application loads feature extraction modules and initiates a connection to the computer's webcam. Each camera frame from the webcam is then processed through the pipeline to extract low level facial behavior features, which are stored in temporary local storage through a SQLite database. Upon clicking the 'Stop button,' the database is automatically synced to Google Cloud Storage for further processing and research purposes over a secured connection.The participants are provided with a 'Survey button' following stopping the data collection, which, upon clicking it, opens a Google form with study survey instruments. The responses to the survey are stored in Google Forms and Google Sheets.

3.2 Lightweight and Privacy Preserving

Feature Extraction To circumvent traditional study requirements, our application is designed to run on commodity hardware with limited computation and memory allocation, which advance the other studies need high computing power

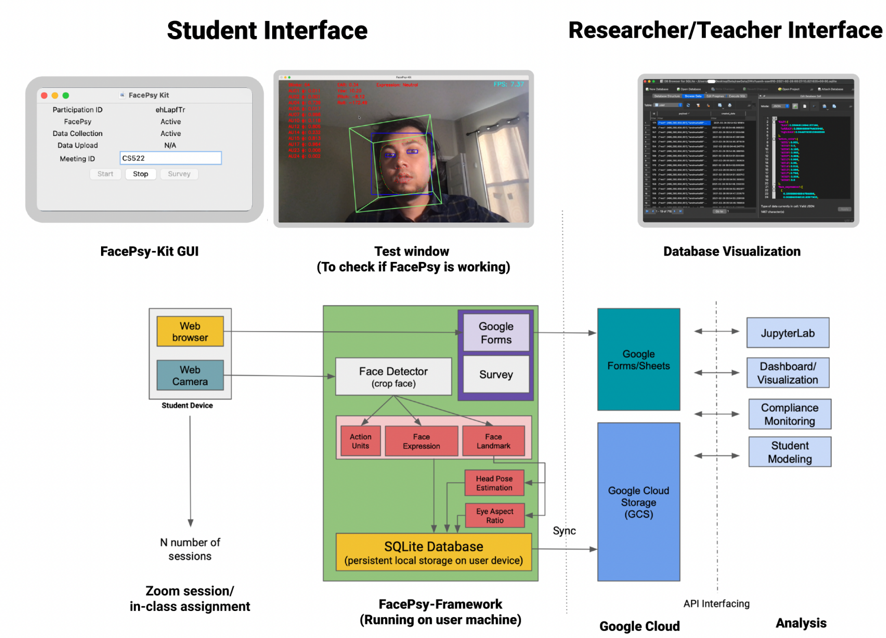

Fig. 1. System Architecture for the FacePsy Framework

[3]. The framework is available on both macOS and Windows 10 operating systems, with a minimum requirement of an Intel Core i5 processor and 8GB of ram for both systems. It incorporates state-of-the-art technologies, including the dlib face and shape detector for face detection and 68 landmark detection tasks [16]. We adopted facial behavior detection modules including Facial Action Unit (AU) detection, Facial expression recognition, Facial Landmark detection, and Head pose estimation to provide enough features in sensing different learning states [1]. The specific modules incorporated into the framework include:

(a) Action Units (AUs), which are used to describe the movement of individual muscles or groups of muscles by their appearance in the human face, detailed in the Facial Action Coding System [18]. The FacePsy framework implements an existing model architecture and weights from the BP4D+ dataset [19], which predicts the probability of any AU occurring in a frame with a face. Our system uses the following Action Units: *AU1, 2, 4, 6-7, 10, 12, 14, 17, 23-24*. The module achieves an average AUC of 0.866 (F1=0.599) across all 12 AUs. Aus have been used to predict an individual's anxiety or stress levels [33].

(b) Facial Expression, which have been shown to be a proxy for communicating feelings and cognitive learning states [1]. For facial expression recognition, we implement a CNN network [15] and train it with the FER 2013 dataset to infer a person's discrete facial expression in seven categories: neutral, anger, happy,

surprise, sad, disgust, and fear, with an accuracy of 66.37%. Facial expressions are useful in indicating an individual's emotional states [34]

(c) Head Pose, which is defined in the Euler angles (pitch, yaw, and roll) which parameterize head configurations with respect to the webcam. The yaw angle represents head movement left or right, pitch represents head movement up or down, and roll represents head tilts left or right towards the shoulder. Head Pose can indicate if an individual is concentrating or thinking [36]

And *(d) Eye Aspect Ratio (EAR)*, which is an estimate of the eye-opening state, it is a measure of the aspect ratio of the eye region. An individual's EAR was has been associated with fatigue and inattention [35]

Privacy-Preserving Approach The combined pipeline of all these tasks builds to an in-device feature extraction module, which runs at 7 frames per second (FPS). For collecting participant surveys, we use Google forms. Our application generates a unique survey link for each session, with pre-filled unique identifiers which is automatically created and assigned to each participant and session identifiers. Like the Bosch and D'Mello study, these frames are directly extracted from the users, instead of extracted from images or videos of the users, which both erases the need to upload large files of data to cloud storage as well as ensures that no distinguishing images or characteristics of the participants are saved [3]. The synced low level facial feature frames in the database and survey responses are then consumed by other applications, such as JupyterLab, for analysis, building dashboards, study compliance monitoring, and student modeling through Google Cloud API.

4 Methods

As a first step of the study, we used the FacePsy framework to obtain its feasibility in detecting a low versus high state of flow, by conducting a proof-of-concept study in an online programming course.

Write and submit a script that asks for a text file and then uses a regular expression to find all of the capitalized words in that text file. (For this exercise, capitalized words are those that begin with a capital letter.) Sort and print the words alphabetically, removing any duplicated entries. Download and use Juliet.txt to check and debug your code.

Fig. 2. A Sample Programming Assignment

4.1 Study Deployment

This study was approved by the university's Institutional Review Board. Students were provided with study guidelines before participation, and the study was introduced during online classes through Zoom in the beginning of the semester. Students were required to provide informed consent to join. The informed consent included an overview of the type of data collected, privacy preservation and confidently policies during data collection, study compliance, and course completion. Although the students were encouraged to participate with an extra credit incentive, participation in the study was voluntary, and students were able to drop out any time. Those who did not choose to participate in the study were given an alternative extra credit assignment. Students were also informed that their participation would not impact their grade or course completion.

Participants and Tasks. The participants were all graduate students, between the ages of 22 to 30 years old, including 25% female students and 75% male students. 88% of the participants reported their ethnicity as Asian, 6% reported their ethnicity as Hispanic, and 6% reported their ethnicity as Caucasian. All participants had at least three months of programming experience. The collected data was recorded over 8 weeks in the semester, with a new programming assignment given each week. The assignments tested a variety of topics seen in Python programming, including conditionals, imported modules, and file input/output. An example of one assignment given to the students can be seen in Figure 2.

Measurements. As part of the study participation, students were asked to install the FacePsy framework on their device and take an online survey at the end of a programming session. After completing their programming task, participants filled out an end-of-programming-session survey, which included three questions [30], and participants rated to what degree they agreed with statements either directly (e.g. "I was fully immersed during the programming session") or indirectly (e.g. "The programming task challenges my capabilities to their limits") mentioning a specific item. Final scores for each flow dimension were computed by taking the average rating over each item (n = 3), which were ranked on a scale from 1-10, where 1 corresponded with strongly disagree and 10 corresponded with strongly agree.

4.2 Dataset Preparation

In total, we collected 54 end of session survey reports from 19 graduate students, with each session representing a student working on the programming task. We found 19 total surveys without any accompanying sensor data from 3 students, some of which were relatively new to programming. We excluded 4 sessions with a duration of less than 5 minutes, from 4 students. In total, we had 31 sessions from 12 students in our dataset. The average number of session across all participants was 2.5 sessions.

Before developing the flow detection model, we first needed to classify sessions as either 'high' or 'low' flow. The median was selected to distinguish flow levels as well as an absolute value at 5 out of 10 because of the high vs. low flow distributions. Using the median as a cutoff, of the 31 sessions in the dataset, 15 were classified as having a low flow level, and 16 were classified as having a high flow level, as opposed to 5 low vs 26 high flow level using 5 as a cutoff.

4.3 Data Pre-processing

In the final dataset, 398,126 student facial feature-frames were sampled every second. Because the FacePsy framework extracts and processes facial images at a rate of 7 FPS and the median session length was approximately 30 minutes, we found it impractical to incorporate all of the raw feature data into the dataset. We decided that one minute was the best unit of analysis, as it has the ability to showcase how a participant's facial features change over the course of a session without an abundance of data to analyze. So, we built and applied an aggregation pipeline, with a frequency of one minute, that computes statistical summaries (sum, minimum, median, maximum, mean, standard deviation, first quartile, and third quartile) for each individual participant and session, for a total of 200 features. Each entry of the dataset consists of the statistical summaries of every feature, calculated over one minute, and they were labeled with ground truth surveys by mapping an entire session to its respective survey. We replaced the Null values in the dataset with 0. Further, the dataset was pre-processed using the MinMax scalar normalization and fed to the machine learning model during Leave-One-Session-Out cross-validation.

4.4 Modeling

After we mapped the data from each participant to their flow ground truth values, a LightGBM Binary Classification, an implementation of boosted trees, was to trained to predict flow levels for a given session. We used the hyperparameter tuning framework, Optuna, to optimize our LightGBM binary classifier with an objective function of maximizing Area under the ROC Curve (AUC) and boosting type as Dropouts meet Multiple Additive Regression Trees (DART).

5 Results and Discussion

In this section, we introduce our findings in detecting a state of flow and discuss how the FacePsy framework was implemented and adopted by students and instructors.

5.1 Feasibility of Flow Modeling

It appears that our low-cost remote sensing framework is applicable to deploy in the real-world online learning setting. Using the data collected by the FacePsy framework, we obtained the feasibility of the computational model that predicts a students aggregated flow level when taking online programming sessions. Our best performing model had an accuracy of 93.5%, an F1 of 0.96, a precision of 0.96, a recall of 0.96, and AUC of 0.89. These results were achieved using a high/low flow cutoff of 5. However, these results are only a marginal improvement on the second model, which used a median cutoff (accuracy = 87.1%, F1 = 0.86, precision = 0.86, recall = 0.86, and AUC = 0.83). Both models results can be seen in Table 1. These results indicate that the traditional threshold in binary classifications was the best choice for our dataset. The model with a halfway cutoff (5 out of 10 scale) also outperformed its baseline model in detecting the majority class (High flow, n=26) (84%) by 9.6%.

Model	Baseline	Accuracy	F1	Precision	Recall	AUC
LightGBM w/Halfway Cutoff	84% (26/31)	93.5%	0.96	0.96	0.96	0.89
LightGBM w/Median Cutoff	52% (16/31)	87.1%	0.86	0.86	0.86	0.83

Table 1. Model Comparison in Detecting a Low versus High Flow State

5.2 Easy Deployability

We found that all (N=19) who joined the study were able to install the framework, which incorporates 5 short, easy to follow steps. Approximately 63% (12/19) of the students completed a sufficiently long recording session paired with a survey during their programming assignments. 3 students submitted multiple surveys without accompanying sensor data, and 4 students recorded sessions with a duration of less than 5 minutes. At least 1 of the 3 students who only submitted surveys informed their instructor that they forgot to start the system due to assignment-related anxiety. The other students may have been either confused by the instructions given or the interface, or also forgot to initialize the recording. It is likely that the 4 students who recorded sessions with durations of less than five minutes were testing the system.

5.3 Lightweight and Privacy Preserving

We found the optimal rate (7 FPS) that balances the framework's sampling and the user's hardware performance, and, although the bandwidth of the framework is relatively low, the data collected can feed machine learning modeling. Traditional real-time automatic flow detection studies incorporate user self-reports at a high granularity, probing for them in multiple small intervals [11][10][12], while we used one end-of-programming-session report. Our feature extraction window of minute was also less granular than other seconds long feature extraction windows seen in other studies [13][6][7][8]. The framework was designed to passively collect the users' facial behavioral signals once initiated without any video recordings, so the students were less likely to perceive risks towards their privacy but were satisfied with the framework's unobtrusiveness.

6 Limitations and Future Work

To improve generalization, we plan to extend the data collection and model training to multiple learning contexts. This is well recognized practices in the learning state detection community [13][5]. It is common to combine vision features with the learning contexts data such as synchronized log of learning platform and video recording of the learning platform in predicting learning outcomes [13]. Therefore, adding a mechanism for incorporating learning context information will be very helpful for learning analytics in the future. Furthermore, because our study only required one self-reported ground truth survey, its granularity was significantly lower than other real-time learning state detection models, who typically record ground-truths in regular intervals [11][10][12]. Increasing the framework's scalability and confirming its validity in an expanded context is essential before eventually introducing it to instructors on a large scale.

7 Conclusion

We introduced a new facial behavioral sensing framework, the FacePsy framework, which extracts facial features in real-time, without the need to save recordings of participants for later extraction. In this proof-of-concept study, features detected using the FacePsy framework were used to identify college students' subjectively reported aggregated state of flow after a naturalistic online programming session. With the data extracted from students by the framework, we created a machine learning model with a maximum accuracy of 93.5% (F1 = 0.96). These initial results provide a good basis for future implementations of the FacePsy framework. Expanding our study to a variety of contexts with a larger dataset is the next step in giving instructors a non-invasive look into the learning states of their students without the need for increased computation power or privacy risks.

References

1. El Kaliouby, Rana, and Peter Robinson. "Real-time inference of complex mental states from facial expressions and head gestures." In Real-time vision for human-computer interaction, pp. 181-200. Springer, Boston, MA, 2005.
2. Martin, Florence, and Doris U. Bolliger. "Engagement matters: Student perceptions on the importance of engagement strategies in the online learning environment." Online Learning 22, no. 1 (2018): 205-222.
3. Bosch, Nigel, and Sidney D'Mello. "Automatic detection of mind wandering from video in the lab and in the classroom." IEEE Transactions on Affective Computing (2019).
4. Dewan, M., Mahbub Murshed, and Fuhua Lin. "Engagement detection in online learning: a review." Smart Learning Environments 6, no. 1 (2019): 1-20.
5. D'Mello, Sidney K., and Caitlin S. Mills. "Mind wandering during reading: An interdisciplinary and integrative review of psychological, computing, and intervention research and theory." Language and Linguistics Compass 15, no. 4 (2021): e12412.
6. Pham, Phuong, and Jingtao Wang. "AttentiveLearner 2: a multimodal approach for improving MOOC learning on mobile devices." In International Conference on Artificial Intelligence in Education, pp. 561-564. Springer, Cham, 2017.

7. Yang, Tsung-Yen, Ryan S. Baker, Christoph Studer, Neil Heffernan, and Andrew S. Lan. "Active learning for student affect detection." In International Conference on Educational Data Mining (EDM), pp. 208-217. 2019.
8. Dhamija, Svati. "Learning based visual engagement and self-efficacy." In 2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII), pp. 581-585. IEEE, 2017.
9. Bosch, Nigel, Sidney D'Mello, and Caitlin Mills. "What emotions do novices experience during their first computer programming learning session?." In International Conference on Artificial Intelligence in Education, pp. 11-20. Springer, Berlin, Heidelberg, 2013.
10. Lee, Po-Ming, Sin-Yu Jheng, and Tzu-Chien Hsiao. "Towards automatically detecting whether student is in flow." In International Conference on Intelligent Tutoring Systems, pp. 11-18. Springer, Cham, 2014.
11. Carroll, Meredith, Mitchell Ruble, Mark Dranias, Summer Rebensky, Maria Chaparro, Joanna Chiang, and Brent Winslow. "Automatic Detection of Learner Engagement Using Machine Learning and Wearable Sensors." Journal of Behavioral and Brain Science 10, no. 3 (2020): 165-178.
12. Faber, Myrthe, Robert Bixler, and Sidney K. D'Mello. "An automated behavioral measure of mind wandering during computerized reading." Behavior Research Methods 50, no. 1 (2018): 134-150.
13. Okur, Eda, Nese Alyuz, Sinem Aslan, Utku Genc, Cagri Tanriover, and Asli Arslan Esme. "Behavioral engagement detection of students in the wild." In International Conference on Artificial Intelligence in Education, pp. 250-261. Springer, Cham, 2017.
14. Booth, Brandon M., Asem M. Ali, Shrikanth S. Narayanan, Ian Bennett, and Aly A. Farag. "Toward active and unobtrusive engagement assessment of distance learners." In 2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII), pp. 470-476. IEEE, 2017.
15. Tümen, Vedat, Ömer Faruk Söylemez, and Burhan Ergen. "Facial emotion recognition on a dataset using convolutional neural network." In 2017 International Artificial Intelligence and Data Processing Symposium (IDAP), pp. 1-5. IEEE, 2017.
16. King, Davis E. "Dlib-ml: A machine learning toolkit." The Journal of Machine Learning Research 10 (2009): 1755-1758.
17. Sharma, Nishank. "A Generator Library for Concise, Unambiguous and URL-Safe UUIDs." PyPI, 2020. https://pypi.org/project/shortuuid/.
18. Cohn, Jeffrey F., and Paul Ekman. "Measuring facial action. The new handbook of methods in nonverbal behavior research." The new handbook of methods in nonverbal behavior research (2005): 9-64.
19. Zhang, Zheng, Jeff M. Girard, Yue Wu, Xing Zhang, Peng Liu, Umur Ciftci, Shaun Canavan et al. "Multimodal spontaneous emotion corpus for human behavior analysis." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 3438-3446. 2016.
20. McCuaig, Judi, Mike Pearlstein, and Andrew Judd. "Detecting learner frustration: towards mainstream use cases." In International Conference on Intelligent Tutoring Systems, pp. 21-30. Springer, Berlin, Heidelberg, 2010.
21. Tiede, Mark, Christine Mooshammer, and Louis Goldstein. "Noggin nodding: Head movement correlates with increased effort in accelerating speech production tasks." Frontiers in psychology (2019): 2459.
22. Ooko, Ryota, Ryo Ishii, and Yukiko I. Nakano. "Estimating a user's conversational engagement based on head pose information." In International Workshop on Intelligent Virtual Agents, pp. 262-268. Springer, Berlin, Heidelberg, 2011.

23. Peper, Erik, Vietta Wilson, Marc Martin, Erik Rosegard, and Richard Harvey. "Avoid Zoom fatigue, be present and learn." NeuroRegulation 8, no. 1 (2021): 47.
24. Ferri, Fernando, Patrizia Grifoni, and Tiziana Guzzo. "Online learning and emergency remote teaching: Opportunities and challenges in emergency situations." Societies 10, no. 4 (2020): 86.
25. Ahshan, Razzaqul. "A framework of implementing strategies for active student engagement in remote/online teaching and learning during the COVID-19 pandemic." Education Sciences 11, no. 9 (2021): 483.
26. The World Bank, UNESCO and UNICEF (2021). The State of the Global Education Crisis: A Path to Recovery. Washington D.C., Paris, New York: The World Bank, UNESCO, and UNICEF.
27. Blaskó, Z., and S. V. Schnepf. "Educational inequalities in Europe and physical school closures during Covid-19." Fairness Policy Brief Series 4 (2020): 2020.
28. Mascheroni, Giovanna, Marium Saeed, Marco Valenza, Davide Cino, Thomas Dreesen, Lorenzo Giuseppe Zaffaroni, and Daniel Kardefelt Winther. Learning at a Distance: Children's remote learning experiences in Italy during the COVID-19 pandemic. No. inorer1182. 2021.
29. Washington, Peter, Qandeel Tariq, Emilie Leblanc, Brianna Chrisman, Kaitlyn Dunlap, Aaron Kline, Haik Kalantarian et al. "Crowdsourced privacy-preserved feature tagging of short home videos for machine learning ASD detection." Scientific reports 11, no. 1 (2021): 1-11.
30. Csikszentmihalyi, Mihaly, and Mihaly Csikzentmihaly. Flow: The psychology of optimal experience. Vol. 1990. New York: Harper Row, 1990.
31. Semerci, Yusuf Can, and Dionysis Goularas. "Evaluation of students' flow state in an e-learning environment through activity and performance using deep learning techniques." Journal of Educational Computing Research 59, no. 5 (2021): 960-987.
32. Lynch, Tiina, and Ioana Ghergulescu. "Large scale evaluation of learning flow." In 2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT), pp. 62-64. IEEE, 2017.
33. Gavrilescu, Mihai, and Nicolae Vizireanu. "Predicting depression, anxiety, and stress levels from videos using the facial action coding system." Sensors 19, no. 17 (2019): 3693.
34. Laksana, Eugene, Tadas Baltrušaitis, Louis-Philippe Morency, and John P. Pestian. "Investigating facial behavior indicators of suicidal ideation." In 2017 12th IEEE International Conference on Automatic Face Gesture Recognition (FG 2017), pp. 770-777. IEEE, 2017.
35. Benoit, Alexandre, and Alice Caplier. "Hypovigilence analysis: open or closed eye or mouth? Blinking or yawning frequency?." In IEEE Conference on Advanced Video and Signal Based Surveillance, 2005., pp. 207-212. IEEE, 2005.
36. Kaliouby, Rana el, and Peter Robinson. "Real-time inference of complex mental states from facial expressions and head gestures." In Real-time vision for human-computer interaction, pp. 181-200. Springer, Boston, MA, 2005.