Toward Computer-Mediated Emotional Monitoring and Burnout Mitigation for University STEM Students

AISHWARYA CHANDRASEKARAN, University of Delaware, USA
LANA MAI HUYNH, California Polytechnic State University San Luis Obispo, USA
LUCY ZHANG BENCHARIT, California Polytechnic State University San Luis Obispo, USA
MATTHEW LOUIS MAURIELLO, University of Delaware, USA

Increasing levels of stress and burnout are pervasive among university STEM students leading to decreased well-being, lower productivity during coursework-related tasks, and negative long-term mental health outcomes. Monitoring physiological cues such as emotional expressions during computing tasks opens up numerous opportunities to support their well-being and productivity while also potentially mitigating burnout. However, researchers have raised concerns about the accuracy of vision-based emotional state recognition tools. In this workshop paper, we explore using off-the-shelf tools to model and code for the emotional states of users. Our goal is to evaluate these tools for use during just-in-time interventions designed to help students perform better in coursework while maintaining positive mental health.

CCS Concepts: • **Human-centered computing → Human computer interaction (HCI)**;

Additional Key Words and Phrases: affective computing, emotion coding, productivity, well-being, burnout, university students

ACM Reference Format:
Aishwarya Chandrasekaran, Lana Mai Huynh, Lucy Zhang Bencharit, and Matthew Louis Mauriello. 2022. Toward Computer-Mediated Emotional Monitoring and Burnout Mitigation for University STEM Students. The Future of Emotion in Human-Computer Interaction (CHI'22 Workshop), (April 2022), 7 pages.

1 INTRODUCTION

Despite increasing enrollment numbers in STEM fields, the U.S. Department of Defense recently warned of a shortage of graduates to meet the future needs of the domestic Science, Technology, Engineering, and Mathematics (STEM) workforce [24]. Indeed, the National Center for Education Statistics estimates that over 48% of those enrolled in a STEM bachelor degree program will leave their major for a non-STEM field or, worse, the university entirely [7]. One reason for these high attrition rates is the increasing levels of stress and burnout which are pervasive among university students leading to decreased well-being, lower productivity during coursework-related tasks, and negative long-term health outcomes. For example, a 2014 report from U.C. Berkeley suggests that an estimated 45% of their STEM graduate students were living with depression [2]. As a result, there have been numerous calls to address the mental health crisis in university STEM majors [1, 38]. Furthermore, the COVID-19 pandemic has exacerbated these conditions and introduced a range of potentially new social and environmental stressors in the lives of students resulting in renewed calls for longitudinal research and multi-pronged public health interventions to address anxiety and depression during and after the pandemic [27].

Recent work has approached student burnout in a variety of ways. For example, developing predictive models to identify students who might be at risk of dropping out based on performance indices have been explored towards creating an early intervention system [34]. Others have explored developing new psychological scales to measure burnout. Adapted from the job demand-resource model of burnout, the University Demand-Resource Questionnaire (UDRQ) has been shown to be a reliable measure of burnout in university students [19]. Another

Part of the Future of Emotion in Human-Computer Interaction Workshop (Virtual)
at CHI'22, April 30 - May 5, 2022, New Orleans, LA, USA

promising approach is to explore different kinds of digital phenotyping to predict mental illness from user mobile device traces and other computing hardware [36]. However, these approaches tend to be reactive based on the time it takes to collect enough data for inferences to be made.

Our long-term goal is to develop new, real-time sensing techniques and just-in-time interventions that would allow STEM university students to self-monitor, develop skills, and better manage their personal well-being to avoid burnout. As this population spends significant time using computers, one potential way of monitoring these students while they work is through their front-facing camera using vision-based tools to code for emotions they display *in-situ* in conjunction with other meta-data. In this workshop paper, we explore how such tools might allow us to monitor and study the emotional state of students while performing cognitively heavy tasks. In addition to evaluating vision-based emotional recognition, the proposed system would use multimodal emotion detection to generate features and mitigate the limitations of vision-based methods. As our project develops, we expect to answer several research questions including: *How effective are existing emotional state coding tools? How can they be used to design, build, and evaluate just-in-time interventions for students to help them manage time and mitigate academic burnout? And, what effects do these interventions have on students' understanding of a subject?*

2 BACKGROUND AND RELATED WORK

2.1 Stress and Burnout

As noted in the introduction, research has shown that peak levels of stress, psychological distress, and burnout in university students have become increasingly prevalent on college campuses [4, 31]. The term "burnout" was first coined by Freudenberger (1974) when he described it to be exhaustion caused by excessive working hours and intensity among social workers and medical personnel [14]. It was defined as a work-related mental health impairment comprising three dimensions namely emotional exhaustion, depersonalization, and reduced personal accomplishment [3]. Burnout has been studied in different individual work professions (e.g., nurses, teachers, etc.) [22, 35]. In our work, we focus on student population to improve their immediate well-being while, in the long-term, reducing attrition in STEM majors. Burnout can oftentimes be indicated by the prevalence of negative emotions and the absence of positive emotions. Facial expressions might act as a window to one's emotional experience even before they are aware of it. Therefore, it is important to classify these emotions accurately.

2.2 Multimodal Emotion Recognition

Numerous methods have been explored to determine a user's emotional state. Early work by psychologists Paul Ekman and Wallace Friesen [1] theorized that certain facial expressions and features correspond to a set of basic emotions including anger, disgust, fear, happiness, sadness, and surprise. Ekman later added contempt to this list. These expressions were theorized to form a universally understood emotional language and continue to be the basis of state-of-the-art facial emotion recognition algorithms available today. Cloud services like Amazon Web Services (AWS)[2], for example, have begun to offer Emotion Recognition services which, in addition to emotional coding, attempts to extract a number of additional features such as gender, the presence of a beard or mustache, whether the person is smiling or not, etc. However, a recent report [11] highlighted that some companies have been employing these technologies to analyze candidates' feelings during job interviews and in public spaces despite the software being prone to various biases such as those related to race, culture, and gender. One of the reasons for these biases is that the training sets are not representative of the vast diversity that is increasingly present in the wider US population (and this problem compounds when global representation is considered). This suggests that there is likely a difference between how software codes an emotion and what is going on in someone's mind–particularly for diverse groups, which are known to value different emotional states [5].

[1]https://www.paulekman.com/facial-action-coding-system/
[2]https://aws.amazon.com/

Facial expressions, gestures, and other signals are not only a product of the body and brain but also of context, of what is happening around a person (i.e., interpreting ambiguous context). Thus, in addition to problems with underlying training data sets, limitations related to the task include the variability in the way people express different emotions (i.e., one individual's happiness expression might be different in intensity and/or facial features compared to that of another), facial configuration indicative of multiple distinct emotional states, and the same facial expression being interpreted differently in different contexts. As a result, there has been increasing calls to address these limitations prior to continuing application in the wild (e.g., to screen job candidates).

To address this concern, research into affective computing (i.e., Emotion AI) have proposed various methods for the measurement of affective states using multimodal input including- assessing emotional state via physiological signals (e.g., heart rate, EEG, blood pressure) [8, 36], wearable sensor data via (e.g., Microsoft Band or mobile phone) [30, 41], environmental data (e.g., lighting condition of the room, weather) [20], data directly reported by the user (via prompts or lifelogging) [15], and sometimes data reported by caretakers (e.g., parents in the case of infants) [10]. Our work draws on this research and explores how to to develop acceptable monitoring technologies for STEM university students using multimodal inputs including camera-based technologies.

2.3 Just-In-Time Interventions

Interventions for burnout have been shown to be effective [6]. In addition to our our algorithmic work, our project also explores Just-In-Time Adaptive Interventions (JITAI) which are a class of digital interventions designed to appear at the right moment in time for a target user by combining data from numerous sources with predictive algorithms. JITAIs have been used for skill building (e.g., coping strategies, decision-making, planning behavior), emotional support (e.g., encouragement, empathy), and instrumental support (e.g., feedback, reminders) [16, 25, 26]. JITAI systems have already been implemented to prevent lapses in addiction behavior [12, 23], dietary lapses [13, 17], and to increase physical activity and reduce sedentary behavior [33, 39, 40]. Our work will explore designing, building, and evaluating similar interventions to mitigate stress and burnout in STEM university students.

3 SYSTEM DESIGN

To begin exploring our research questions, we plan to design and build an emotional coding and burnout tracking platform for STEM university students and other heavy computer users. We will conduct iterative deployments of this tool with the goal of gaining a better understanding of trends in students' emotional states and how these might correlate with their performance on cognitively demanding tasks (e.g., programming, mathematics, composition). We plan to validate our resulting models by using *in-situ* prompts based on the Experience Sampling Method (or Ecological Momentary Assessment) [9, 21, 37]. Once validated, we will then investigate designing personal informatics dashboards and different interventions (i.e., reflective and in-situ) given the detection of mood changes, prolonged bouts of negative emotions, and indications of burnout conditions. In our context, it is important for our platform to effectively determine the most effective intervention if there is a range of potential options (e.g.,[28]) and determine whether intervention at a particular moment might be counterproductive. Therefore, the user's context must be understood to avoid disruptive interventions and increase acceptability.

4 PRELIMINARY WORK

As a preliminary step, we have been investigating emotion recognition tools using data captured from front-facing cameras. Each video in our dataset is approximately 60 seconds long and features users completing a formal interview-esque task where participants are asked to describe projects they have worked on, personal weaknesses, and other challenges. We have used an algorithm based on Residual Masking Networks [29] and software from

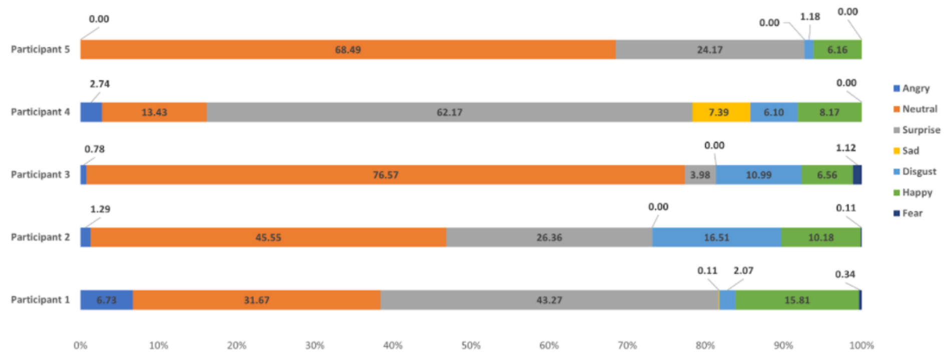

Fig. 1. Results of proportions of different emotions using Residual Masking Network

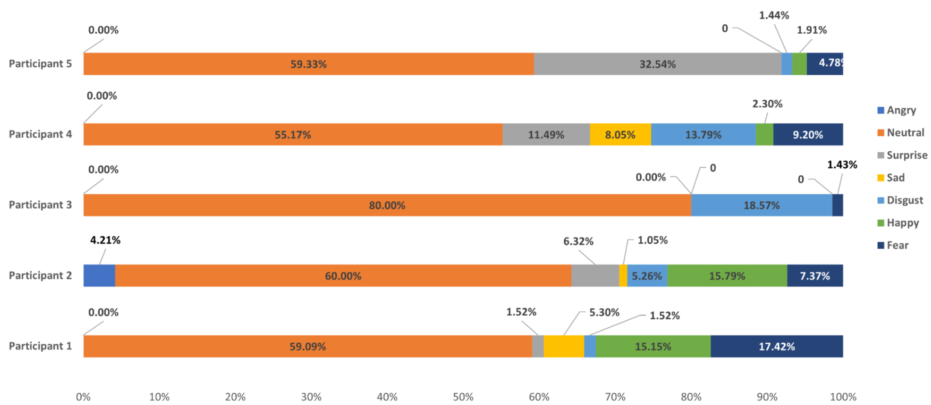

Fig. 2. Results of proportions of different emotions using Noldus FaceReader

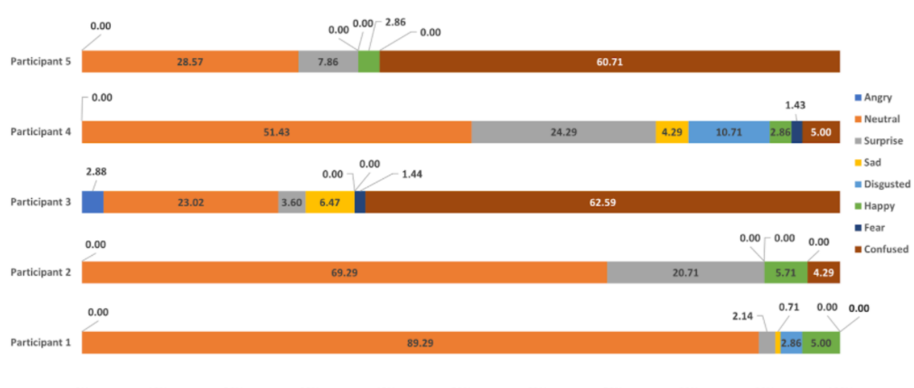

Fig. 3. Results of proportions of different emotions using Amazon Rekognition

Noldus, FaceReader,[3] based on the Facial Action Coding System to compare the results of their predictions. In the former, we see that the predictions are distributed as shown in Figure 1, while the Noldus results are as shown in Figure 2. The Residual Masking Networks algorithm is based on a segmentation network to refine feature maps, enabling the network to focus on relevant information to make correct decisions. It uses a UNet-based [32] architecture in combination with Deep Residual Networks [18] to create a Residual Masking Network to generate the emotional predictions while Noldus determines facial expressions in 3 steps, i.e., Face finding using a deep learning based face-finding algorithm which searches for areas in the image having the appearance of a face at different scales, deep neural networks based Face Modeling which synthesizes an artificial face model to describe the location of 468 key points in the face, and Face Classification using a trained deep artificial neural network to recognize patterns in the face and directly classify the facial expressions from image pixels. Finally, Figure 3 shows results from Amazon Rekognition which allows us to upload frames and receive a JSON packet of predictions as described in the related work section.

4.1 Observations

As can be seen from Fig 1, Fig 2 and Fig 3, the most dominating emotion in all three predictions is "Neutral". Furthermore, the predictions from Residual Masking Network and AWS Rekognition indicate that "Surprised" and "Confused" are their second most dominant emotions, respectively. Making comparison between algorithms more difficult, different algorithms provide different sets of emotional classes, different confidence values, and can't always be separated into singular emotions (making our results an approximation in the case of FaceReader). For example, we observed for coding one frame, Rekognition responded with "Sad" (.62 confidence) whereas the Residual Masking Network responded with "Surprise" (.88 confidence). Thus, it is difficult to conclude from the predictions which of these should be used without a ground truth. Therefore, these solutions may present problems if used to bootstrap an adaptive recognition system. However, an interesting note about FaceReader is that these predictions are based on a human-in-the-loop calibration process which suggests these results may be more reliable. That said, future work might consider exploring the difference in predictions between algorithms to determine how much value the human-in-the-loop process adds to the emotional tagging process compared to a fully autonomous algorithm and using these outputs for ensemble learning alongside other multimodal inputs.

5 CONCLUSION

The long-term goal of this project is to develop new sensing techniques and just-in-time intervention that allow students to self-monitor, develop skills, and better manage their personal well-being towards improving productivity and avoiding burnout. Our work probes into the different responses returned from several advanced tools for emotional coding of video data. Similar to concerns raised in [11], we notice discrepancies between these approaches suggesting further audits of these tools, new and more representative training data, and multimodal input are likely necessary to achieve our goals. However, even if such methods proved highly effective off-the-shelf, we would still need a better understanding of students' work context as well as the types of interventions they find effective and adoptable. As a result, we plan to: (i) continue our technical work exploring multimodal algorithms for emotional detection including using both front-facing cameras and computer peripherals with *in-situ* ESM surveys to improve coding accuracy, (ii) audit models for bias data and explore mitigation techniques, and (iii) interview and survey students to begin soliciting feedback about their work habits and willingness to adopt digital interventions. As a result, some expected contributions of this work include new data sets, novel multimodal algorithms for emotional detection, and design guidelines for interventions aimed at university student populations.

[3] https://www.noldus.com/facereader/new

REFERENCES

[1] Jennfier Akullian, Adam Blank, Lauren Bricker, Linda DuHadway, and Christian Murphy. *Supporting Mental Health in Computer Science Students and Professionals*, page 958–959. Association for Computing Machinery, New York, NY, USA, 2020.
[2] UC Berkeley Graduate Assembly. *2014*. Graduate student happiness and well-being report, 2014.
[3] Wendy L Awa, Martina Plaumann, and Ulla Walter. Burnout prevention: A review of intervention programs. *Patient education and counseling*, 78(2):184–190, 2010.
[4] Ellen Bara Stolzenberg, Melissa C Aragon, Edgar Romo, Victoria Couch, Destiny McLennan, M Kevin Eagan, and Nathaniel Kang. The american freshman: National norms fall 2019. 2020.
[5] Lucy Zhang Bencharit, Yuen Wan Ho, Helene H Fung, Dannii Y Yeung, Nicole M Stephens, Rainer Romero-Canyas, and Jeanne L Tsai. Should job applicants be excited or calm? the role of culture and ideal affect in employment settings. *Emotion*, 19(3):377, 2019.
[6] Edgar Bresó, Wilmar B Schaufeli, and Marisa Salanova. Can a self-efficacy-based intervention decrease burnout, increase engagement, and enhance performance? a quasi-experimental study. *Higher Education*, 61(4):339–355, 2011.
[7] Xianglei Chen. Stem attrition: College students' paths into and out of stem fields. statistical analysis report. nces 2014-001. *National Center for Education Statistics*, 2013.
[8] Joanne Wai Yee Chung, Henry Chi Fuk So, Marcy Ming Tak Choi, Vincent Chun Man Yan, and Thomas Kwok Shing Wong. Artificial intelligence in education: Using heart rate variability (hrv) as a biomarker to assess emotions objectively. *Computers and Education: Artificial Intelligence*, 2:100011, 2021.
[9] Mihaly Csikszentmihalyi and Reed Larson. Validity and reliability of the experience-sampling method. In *Flow and the foundations of positive psychology*, pages 35–54. Springer, 2014.
[10] C Emily Durbin and Sylia Wilson. Convergent validity of and bias in maternal reports of child emotion. *Psychological assessment*, 24(3):647, 2012.
[11] Rose Eveleth. Your employer may be spying on you—and wasting its time. *Scientific American*, 16, 2019.
[12] Florian Ferreri, Alexis Bourla, Stephane Mouchabac, and Laurent Karila. e-addictology: an overview of new technologies for assessing and intervening in addictive behaviors. *Frontiers in psychiatry*, page 51, 2018.
[13] Evan M Forman, Stephanie P Goldstein, Fengqing Zhang, Brittney C Evans, Stephanie M Manasse, Meghan L Butryn, Adrienne S Juarascio, Pramod Abichandani, Gerald J Martin, and Gary D Foster. Ontrack: development and feasibility of a smartphone app designed to predict and prevent dietary lapses. *Translational behavioral medicine*, 9(2):236–245, 2019.
[14] Herbert J Freudenberger. Staff burn-out. *Journal of social issues*, 30(1):159–165, 1974.
[15] Kurara Fukumoto, Tsutomu Terada, and Masahiko Tsukamoto. A smile/laughter recognition mechanism for smile-based life logging. In *Proceedings of the 4th Augmented Human International Conference*, pages 213–220, 2013.
[16] Stephanie P Goldstein, Brittney C Evans, Daniel Flack, Adrienne Juarascio, Stephanie Manasse, Fengqing Zhang, and Evan M Forman. Return of the jitai: applying a just-in-time adaptive intervention framework to the development of m-health solutions for addictive behaviors. *International journal of behavioral medicine*, 24(5):673–682, 2017.
[17] Stephanie P Goldstein, Fengqing Zhang, John G Thomas, Meghan L Butryn, James D Herbert, and Evan M Forman. Application of machine learning to predict dietary lapses during weight loss. *Journal of diabetes science and technology*, 12(5):1045–1052, 2018.
[18] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
[19] Balázs Jagodics and Éva Szabó. Student burnout in higher education: A demand-resource model approach. *Trends in Psychology*, pages 1–20, 2022.
[20] Nobuyoshi Komuro, Tomoki Hashiguchi, Keita Hirai, and Makoto Ichikawa. Predicting individual emotion from perception-based non-contact sensor big data. *Scientific reports*, 11(1):1–9, 2021.
[21] Reed Larson and Mihaly Csikszentmihalyi. The experience sampling method. In *Flow and the foundations of positive psychology*, pages 21–34. Springer, 2014.
[22] Daniel J Madigan and Lisa E Kim. Towards an understanding of teacher attrition: A meta-analysis of burnout, job satisfaction, and teachers' intentions to quit. *Teaching and teacher education*, 105:103425, 2021.
[23] Lisa A Marsch, Aimee Campbell, Cynthia Campbell, Ching-Hua Chen, Emre Ertin, Udi Ghitza, Chantal Lambert-Harris, Saeed Hassanpour, August F Holtyn, Yih-Ing Hser, et al. The application of digital health to the assessment and treatment of substance use disorders: The past, current, and future role of the national drug abuse treatment clinical trials network. *Journal of substance abuse treatment*, 112:4–11, 2020.
[24] Amy Y Murray. Fy2019 industrial capabilities report to congress. Technical report, OFFICE OF THE UNDER SECRETARY OF DEFENSE (ACQUISITION) WASHINGTON DC …, 2020.
[25] Inbal Nahum-Shani, Eric B Hekler, and Donna Spruijt-Metz. Building health behavior models to guide the development of just-in-time adaptive interventions: A pragmatic framework. *Health psychology*, 34(S):1209, 2015.

[26] Inbal Nahum-Shani, Shawna N Smith, Bonnie J Spring, Linda M Collins, Katie Witkiewitz, Ambuj Tewari, and Susan A Murphy. Just-in-time adaptive interventions (jitais) in mobile health: key components and design principles for ongoing health behavior support. *Annals of Behavioral Medicine*, 52(6):446–462, 2018.

[27] Hans Oh, Caitlin Marinovich, Ravi Rajkumar, Megan Besecker, Sasha Zhou, Louis Jacob, Ai Koyanagi, and Lee Smith. Covid-19 dimensions are related to depression and anxiety among us college students: Findings from the healthy minds survey 2020. *Journal of Affective Disorders*, 292:270–275, 2021.

[28] Pablo Paredes, Ran Gilad-Bachrach, Mary Czerwinski, Asta Roseway, Kael Rowan, and Javier Hernandez. Poptherapy: Coping with stress through pop-culture. In *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare*, pages 109–117, 2014.

[29] Luan Pham, The Huynh Vu, and Tuan Anh Tran. Facial expression recognition using residual masking network. In *2020 25th International Conference on Pattern Recognition (ICPR)*, pages 4513–4519. IEEE, 2021.

[30] Kiran K Rachuri, Mirco Musolesi, Cecilia Mascolo, Peter J Rentfrow, Chris Longworth, and Andrius Aucinas. Emotionsense: a mobile phones based adaptive platform for experimental social psychology research. In *Proceedings of the 12th ACM international conference on Ubiquitous computing*, pages 281–290, 2010.

[31] David Robotham. Stress among higher education students: Towards a research agenda. *Higher education*, 56(6):735–746, 2008.

[32] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[33] Matthew Saponaro, Ajith Vemuri, Greg Dominick, and Keith Decker. Contextualization and individualization for just-in-time adaptive interventions to reduce sedentary behavior. In *Proceedings of the Conference on Health, Inference, and Learning*, pages 246–256, 2021.

[34] Annalina Sarra, Lara Fontanella, and Simone Di Zio. Identifying students at risk of academic failure within the educational data mining framework. *Social Indicators Research*, 146(1):41–60, 2019.

[35] Megha K Shah, Nikhila Gandrakota, Jeannie P Cimiotti, Neena Ghose, Miranda Moore, and Mohammed K Ali. Prevalence of and factors associated with nurse burnout in the us. *JAMA network open*, 4(2):e2036469–e2036469, 2021.

[36] Astha Sharma and Shaun Canavan. Multimodal physiological-based emotion recognition. In *International Conference on Pattern Recognition*, pages 101–113. Springer, 2021.

[37] Saul Shiffman, Arthur A Stone, and Michael R Hufford. Ecological momentary assessment. *Annu. Rev. Clin. Psychol.*, 4:1–32, 2008.

[38] Lígia Maria Soares Passos, Christian Murphy, Rita Zhen Chen, Marcos Gonçalves de Santana, and Giselle Soares Passos. *The Prevalence of Anxiety and Depression Symptoms among Brazilian Computer Science Students*, page 316–322. Association for Computing Machinery, New York, NY, USA, 2020.

[39] J Graham Thomas and Dale S Bond. Behavioral response to a just-in-time adaptive intervention (jitai) to reduce sedentary behavior in obese adults: Implications for jitai optimization. *Health Psychology*, 34(S):1261, 2015.

[40] Ajith Vemuri, Keith Decker, Mathew Saponaro, and Gregory Dominick. Multi agent architecture for automated health coaching. *Journal of Medical Systems*, 45(11):1–7, 2021.

[41] Kangning Yang, Chaofan Wang, Yue Gu, Zhanna Sarsenbayeva, Benjamin Tag, Tilman Dingler, Greg Wadley, and Jorge Goncalves. Behavioral and physiological signals-based deep multimodal approach for mobile emotion recognition. *IEEE Transactions on Affective Computing*, 2021.