Proceedings of the Fourteenth Artificial Intelligence and
Interactive Digital Entertainment Conference (AIIDE 2018)

Affect-Based Early Prediction of Player Mental Demand and Engagement for Educational Games

Joseph B. Wiggins,[1] Mayank Kulkarni,[1] Wookhee Min,[2] Bradford Mott,[2] Kristy Elizabeth Boyer,[1] Eric Wiebe,[3] James Lester[2]

[1]Department of Computer & Information Science & Engineering, University of Florida, Gainesville, FL 32601
[2]Center for Educational Informatics, North Carolina State University, Raleigh, NC 27695
[3]Department of STEM Education, North Carolina State University, Raleigh, NC 27695
[1]{jbwiggi3, mayankk91, keboyer}@ufl.edu, [2]{wmin, bwmott, lester}@ncsu.edu, [3]wiebe@ncsu.edu

Abstract

Player affect is a central consideration in the design of game-based learning environments. Affective indicators such as facial expressions exhibited during gameplay may support building more robust player models and adaptation modules. In game-based learning, predicting player *mental demand* and *engagement* from player affect is a particularly promising approach to helping create more effective gameplay. This paper reports on a predictive player-modeling approach that observes player affect during early interactions with a game-based learning environment and predicts self-reports of mental demand and engagement at the conclusion of gameplay sessions. The findings show that automatically detected facial expressions such as those associated with joy, disgust, sadness, and surprise are significant predictors of players' self-reported engagement and mental demand at the end of gameplay interactions. The results suggest that it is possible to create affect-based predictive player models that can enable proactively tailored gameplay by anticipating player mental demand and engagement.

Introduction

Educational games are highly effective at engaging players as they learn (Sabourin and Lester 2014). These games also present challenges to players, who must master not only the game itself but achieve the learning objectives incorporated within it (Aleven et al. 2016). To support players through these challenges, accurate player and learner models are critical for providing effective and timely adaptive feedback (Aleven et al. 2016; Corrigan et al. 2015; Martinez, Bengio, and Yannakakis 2013). Work on learner modeling has long focused on *cognitive* aspects of learning (Aleven et al. 2016), recognizing player goals (Min et al. 2016) and leveraging information about learners' actions or interactions with resources in a virtual world in order to provide insight into player performance (Baker et al. 2014).

While cognitive aspects of learning are highly informative for adapting to players during game-based learning, *affective aspects* may be equally important. Incorporating affect into learner models, for example, can increase their robustness (Arroyo et al. 2007; D'Mello and Graesser 2012; Mudrick et al. 2017) by allowing them to better account for phenomena such as "guessing" and "slipping" (Corrigan et al. 2015). One particularly promising data channel to support real-time adaptation to learner affect is facial expression, which can leverage increasingly accurate real-time automatic facial expression recognition models (Aleven et al. 2016; Mudrick et al. 2017).

While increasingly sophisticated facial expression recognition and player models are emerging, relatively little is known about the ways in which we can leverage *early* displays of affect (e.g., within the first few minutes of gameplay) to adapt subsequent gameplay experience. Early recognition of affect allows for proactive interventions and resolving player frustrations, which can be particularly problematic in educational games. This paper focuses on this open area of research, addressing the following question: to what degree can facial expressions during early gameplay predict end-of-gameplay outcomes? This paper investigates this research question in the context of an educational game for microbiology and literacy, considering the specific outcomes of mental demand (i.e., how difficult and demanding the player found the experience to be) and engagement (i.e., feeling of immersion and interest in the game).

Using data from a study with undergraduate students playing the educational game, we constructed a set of models to predict self-reported mental demand and engagement. The results show that facial expressions traditionally linked to states of joy, sadness, disgust, and surprise that appeared during the game's introductory tutorial are significant predictors of learner engagement and mental demand reported after gameplay is complete. The findings suggest that it is possible to create affect-based predictive models that can enable proactively tailoring gameplay to player mental demand and engagement. This line of work provides initial insights into not only how to shape short-term adaptations, such as when to provide help to the player, but also how to shape long-term adaptations, such as selecting and sequencing challenges within the game.

Related Work

Designing a game experience involves managing the cognitive load being introduced to the player while attempting to design a smooth experience that will be consistent across users (e.g., Horn et al. 2017; Liapis et al. 2013). In addition to cognitive considerations, changes to game designs impact the emotional progression of a game experienced by players, for example, within the game's narrative (Hernandez et al. 2014). Real-time affective information has been used to improve user experience in games, such as verifying narrative events that induce the intended emotions (Wang and Marsella 2006). Blom and colleagues (2014) proposed an adaptive game difficulty that adjusted itself based on the player's facial expressions, increasing the difficulty when the player was happy or neutral and decreasing the difficulty when she was angry. They determined that this emotional adaptation improved the player's experience.

The affective, or emotional, experience of a player may have a particularly important role in educational games due to the well-established relationship between affect and learning. Learner emotional states such as boredom and confusion have been a big focus in learner modeling, using sources such as log data (Baker et al. 2014) and text the learner writes during interaction with the system (D'Mello and Graesser 2012). These emotional states also relate to facets of user engagement, such as felt involvement (i.e., flow) and aesthetic enjoyment (Wiebe et al. 2014). Player modeling has also begun to utilize diverse data streams from players, such as eye gaze to model cognitive-affective states (Hutt et al. 2017) and recognize player goals (Min et al. 2017), and galvanic skin response to cluster player reactions to game events (Nogueira et al. 2014) and to model player affective states (Martinez, Bengio, and Yannakakis 2013).

Among all sources of information about player affect, facial expression may be particularly important. Facial expressions have long been thought of as a rich source of affective information (Ekman and Friesen 1978), and mounting evidence is pointing to their importance in educational games. Facial expressions have been found to be predictive of learner outcomes while playing educational games (Sawyer et al. 2017), and adapting to facial expressions has been shown to increase on-task behavior and reduce hint abuse (Arroyo et al. 2007).

In contrast to the previous work, the present work aims to create early predictions of mental demand and engagement—two key factors in gameplay (Wiebe et al. 2014). Some prior work has suggested that as little as five minutes of a tutoring interaction help to predict the frustration and learning gains that will be observed at the end of the session (Grafsgaard et al. 2013), but it is unclear if these early predictive features also transfer to other domains such as educational games. This paper explores whether early prediction in a different context (i.e., educational games) is also feasible for other outcome metrics (i.e., engagement and mental demand).

Educational Game Environment

To investigate the predictive potential of players' affect exhibited early in interactions with educational games, we conducted a study to model player affect while players interacted with CRYSTAL ISLAND, an educational game for microbiology and literacy (Sabourin and Lester 2014). The game's premise is that an outbreak has occurred on a remote island research station, and the player, embodied as a medical field agent in the game, is given the responsibility of identifying the disease and its source. Players navigates the game from a first-person perspective (Figure 1). Their objective is to diagnose the outbreak and develop a recommendation for a plan of treatment for the infected people on the island. To do so, players must gather evidence from non-player characters, explore in-game educational resources, such as books and quizzes, and test their hypotheses using virtual lab equipment.

Figure 1: First-person perspective experienced by players.

To successfully complete the game, players must perform a variety of problem-solving actions. These actions include developing and testing hypotheses, gathering information from various sources, and successfully diagnosing the disease that has infected the island inhabitants. The game design allows learners to experience an ongoing series of successes and failures punctuated by emotions such as joy, sadness, and surprise over the course of learning interactions. Thus, these activities have a rich affective dimension.

Description of Gameplay

When players begin interacting with the game, they play through a tutorial that introduces the game mechanics (e.g., how to move the player's character within the virtual world), the narrative (e.g., the backstory that draws the player into the game world), and the learning tasks for the player (e.g., reading scientific texts to inform hypothesis generation and testing). This game tutorial proceeds in a linear fashion where learners explore key elements of the game world, meet characters, and interact with objects in the environment.

After completing the tutorial, players reach a milestone in the game by talking to a key non-player character (NPC) within the game world, and from that point players diverge and follow their own paths in the open-world game. Learners are free to approach and enter any of the buildings in the game world including an infirmary, a dining hall, a laboratory, and two residences. Learners can interact with the following game elements: NPCs that provide game-related information; virtual books, posters, and articles that teach microbiology concepts relevant to the game; and possibly contaminated objects which players must test in the virtual lab. It is the player's prerogative to determine what information obtained is useful and take notes using their in-game diagnosis worksheet. They submit the worksheet to Kim, the camp nurse NPC, who then assesses the findings. Solving the mystery results in completion of the game.

Real-Time Affective Data Collection

As players interact with CRYSTAL ISLAND, their facial expressions are logged and interpreted in real time. Logging is performed at one-second intervals, using the webcam on the players' laptops. Images of learners' faces are processed using Affectiva's AFFDEX SDK (McDuff et al. 2016). This system has the ability to monitor the muscle movements of the face from the video feed and, using these low-level facial action unit-based features, estimate the likelihoods of higher-level affect features. For this work, we consider the higher-level features of Joy, Sadness, Surprise, Fear, Anger, Disgust, and Contempt (Ekman and Friesen 1978). In this work, only the players' affective states from the tutorial phase of gameplay were considered. For each feature, the Affectiva AFFDEX SDK computes an estimate of the likelihood that the state is present in a given one-second window. These numbers are positive and scaled to a range from 0 to 100.

The intervals at which to sample affective data are an important consideration, and the chosen sampling interval was selected because it is sufficiently sparse to support the educational game running on laptops in real time.

Methodology

To explore the research question of how learners' facial expressions during early gameplay can inform player models that predict end-of-gameplay player outcomes, we collected data of undergraduate students playing CRYSTAL ISLAND. We describe the participants, the study design, the instruments utilized, and the collection of facial expression data.

Participants

Participants were undergraduate students at the University of Florida, enrolled in an introductory computer science course in which the instructor offered extra credit for participating in research studies. The original sample consisted of 34 participants. From these participants, we obtained 31 complete facial expression data sets which were used in this analysis. Table 1 shows descriptive statistics for these 31 participants.

Characteristic	Details
Gender	12 Female; 18 Male; 1 Other
Age	18.97 (SD = 0.84)
Semesters Completed	1.95 (SD = 1.75)
Game Play Frequency	2 Not at all; 7 Rarely; 10 Occasionally; 7 Frequently; 5 Very frequently

Table 1: Descriptive statistics for study participants.

Study Design

The study proceeded by administering a pre-survey, providing learners with up to one hour of gameplay, and administering a post-survey. All learners were given a laptop and mouse and asked to use headphones. After learners completed the pre-survey, they were provided a brief overview of the game and were informed that they could play for up to one hour. At the conclusion of gameplay, either by completing the game or reaching the time limit, learners completed the post-survey. Of the 31 participants, 17 completed the game within the hour.

Survey Instruments

After playing the game, players self-reported their affective outcomes and were asked for the demographic information reported in Table 1. In this paper, we focus on engagement as calculated using the User Engagement Scale (UES) (O'Brien and Toms 2010; Wiebe et al. 2014) and the learners' mental demand as calculated using the NASA-TLX (Hart and Staveland 1988; Rubio et al. 2004).

Example Items from User Engagement Scale:

I lost myself in this gaming experience. I blocked out things around me when I was playing the game. I felt involved in the game. I would continue to play the game out of curiosity. I would recommend this game to my friends and family.

Mental Demand Survey Item:

(*mental demand can mean things like thinking, deciding, calculating, remembering, looking, searching*) How mentally demanding was playing CRYSTAL ISLAND today?

Facial Expression Logs

Players were in a well-lit room and positioned away from one another to prevent issues with facial tracking picking up other players. The facial expression data analyzed in this work was collected from the time players first began their gameplay session until they finished the tutorial phase.

Player Outcome	Mean (Standard Deviation)
User Engagement (out of 5)	3.51 (0.34)
Mental Demand (out of 100)	45.35 (20.31)
Correlation	-0.37 (p-value = 0.04)

Table 2: Descriptive statistics on the players' self-reported experiences while playing.

Data

This work focuses on predicting player outcomes from affect data generated from players' affective states exhibited during gameplay. The player outcomes are user engagement and mental demand. Ideally, a gaming experience would be deeply engaging without causing excessive mental demand (Horn et al. 2017). In educational games, the learning components risk becoming too mentally demanding and negatively impacting the user experience.

On average, players took approximately 18.48 minutes (SD = 7.12) to complete the tutorial, approximately 31% of their overall gameplay time. As seen in Table 2, the average user engagement is above neutral and mental demand is near the middle of the scale. The distributions of user engagement and mental demand can be seen in Figures 2 and 3 respectively.

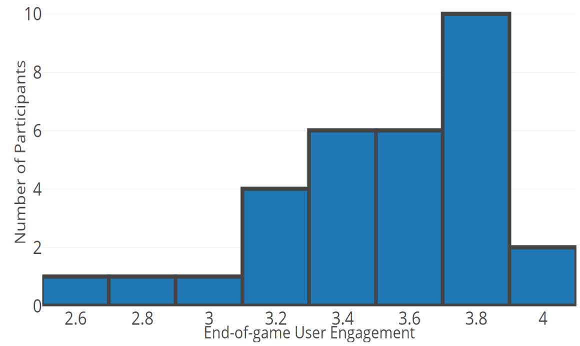


Figure 2: Distribution of end-of-gameplay user engagement, possible scores ranging from 1 to 5.

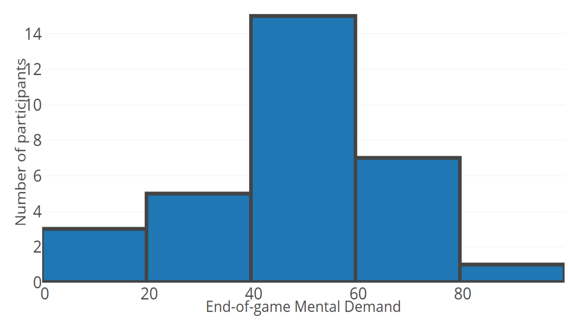


Figure 3: Distribution of end-of-gameplay mental demand, possible scores ranging from 0 to 100.

Analysis and Results

To predict players' self-reported engagement and mental demand, we constructed a stepwise linear regression model that used the average facial expression confidences across the emotion features during the game tutorial for each player. The models used minimum Bayesian Information Criterion (BIC) as their stopping rule for adding features, building models with high predictive power while applying a penalty for including more features.

Predicting Engagement

User engagement is measured by the UES described earlier. The UES includes subscales of *felt involvement* and *aesthetic enjoyment*, which are related to both engagement and affective state (Wiebe et al. 2014). As shown in Table 3, the model includes significant predictors of Surprise, Disgust, Sadness, and Joy. Disgust, Sadness, and Joy all have a positive correlation with user engagement, but Surprise is negatively correlated.

User Engagement =	R^2	p
- 0.0322 * Surprise	0.1844	< 0.001*
+0.0175 * Disgust	0.1777	0.043*
+0.0348 * Sadness	0.0825	0.022*
+0.0519 * Joy	0.0954	0.028*
+3.5545 (Intercept)		1.000

Table 3: Predictive model for self-reported user engagement at end of gameplay (R^2 = 0.54). Each predictor represents that affective feature as observed within the game tutorial.

Predicting Mental Demand

Learners reported their sense of mental demand upon the completion of their gameplay sessions. Mental demand was presented to the learners as tasks such as thinking, deciding, or calculating. Table 4 shows the single predictor selected for mental demand, Surprise. Surprise had a positive correlation with mental demand.

Mental Demand =	R^2	p
+0.7558 * Surprise	0.1282	0.048*
+39.399 (Intercept)		1.000

Table 4: Predictive model for self-reported mental demand at end of gameplay (R^2 = 0.13). Each predictor represents that affective feature as observed within the game tutorial.

Discussion

The results suggest that automatically recognized player facial expressions exhibited during early interactions with an educational game can predict learners' self-reported engagement and mental demand at the conclusion of gameplay. Below, we discuss the features that were found to be most informative for the successful models for engagement and mental demand prediction.

Facial Expressions of Joy

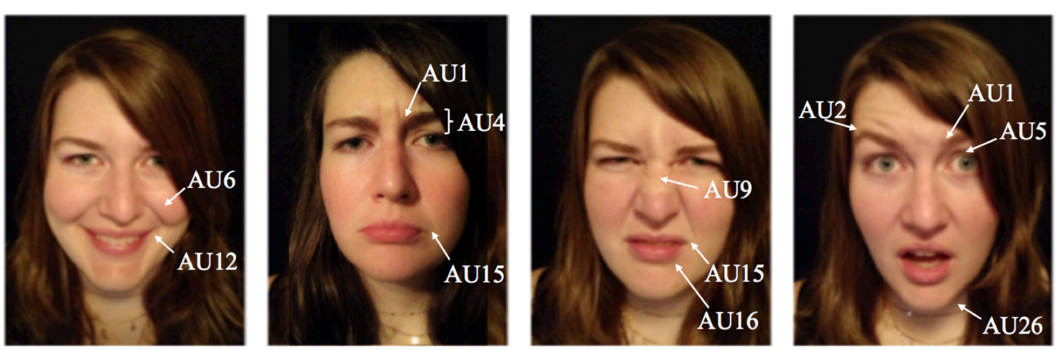
In the learner engagement model, facial expressions of Joy were selected as a predictive feature, as seen in Table 3. Joy is characterized by cheek raiser (AU6) and lip corner puller (AU12) as shown in Figure 4 (Ekman and Friesen 1978).

Joy in the game tutorial had a positive correlation with user engagement, as reported at the end of the gameplay sessions. Facial expressions of Joy have traditionally been associated with enjoyment and pleasure (Ekman and Friesen 1978). In learning contexts, it has also been shown to have some associations with positive learning gains (Vail et al. 2016). In our case, Joy may signal an early affective manifestation of the player engagement.

Facial Expressions of Sadness and Disgust

The next two features selected for the user engagement model, Sadness and Disgust, perhaps seem counterintuitive because they are typically associated with negative affective states. However, displays of Sadness and Disgust early in the game tutorial were positively correlated with user engagement. To consider why Sadness and Disgust may be important signs of user engagement in educational games, it is helpful to understand what features characterize these emotions. Sadness is characterized by the activation of inner brow raiser (AU1), brow lowerer (AU4), and lip corner depressor (AU15) as shown in Figure 4 (Ekman and Friesen 1978). Disgust has some similarities to Sadness, such as the activation of lip corner depressor. Disgust requires the activation of nose wrinkler (AU9), lip corner depressor (AU15), and lower lip depressor (AU16) as shown in (Ekman and Friesen 1978).

Sadness and Disgust, unlike Joy, are traditionally associated with negative affective states and their connection with user engagement may seem unclear. However, prior



Joy Sadness Disgust Surprise

Figure 4: Example prototypical facial expressions of Joy, Sadness, Disgust, and Surprise.

work has found that facial expressions of Sadness relate to a higher sense of presence (a related concept to engagement) in educational games, which has also been associated with player empathy (Sawyer et al. 2017). Features of Disgust have also been shown to have an association with positive outcomes, such as work that found a relationship between nose wrinkler (AU9) and learning gain in females, where it may be indicative of mental effort (Vail et al. 2016).

Facial Expressions of Surprise

In both the model of user engagement (Table 3) and the model of mental demand (Table 4), the first and most predictive feature selected was Surprise, although the model of mental demand was only able to achieve a low R^2 of 0.13. Surprise is characterized by the activation of inner brow raiser (AU1), outer brow raiser (AU2), upper lid raiser (AU5), and jaw drop (AU26) as shown in Figure 4 (Ekman and Friesen 1978).

In the models, Surprise in the game tutorial had a negative correlation with the user engagement reported at the end of the learners' gameplay sessions, and it had a positive relationship with their mental demand. Surprise has been shown to have a negative relationship with learning gain (Sawyer et al. 2017; Vail et al. 2016). Prior work has also seen facial features of Surprise to be indicative of mental effort (Vail et al. 2016) and confusion (Grafsgaard et al. 2011). While confusion can be beneficial in learning, if left unmanaged it may transition into detrimental states, such as Frustration and Boredom (D'Mello et al. 2014), which in turn are related to lower feelings of felt involvement (Wiebe et al. 2014). The models suggest that seeing early displays of Surprise may in fact be related to less desirable learner outcomes.

Game Design Implications

This work explored the occurrence of facial indicators of Joy, Sadness, Disgust, and Surprise during the tutorial of an educational game and found relationships to key end-of-gameplay outcomes. Tutorials are key components of games that are designed to acclimate players to their new environment, and identifying facial expression indicators that players are struggling at this phase can help avoid negative outcomes such as lower engagement or higher mental demand.

Limitations

A limitation of this work is that it considers the domain of an educational game specifically focused on microbiology, which involved disease and sickness. As such, the domain of the game may heavily influence important affective states, such as disgust, which may not transfer well to other domains for the same predictive task. It is also the case that players spent different amounts of time in the game tutorial, and some players spent less overall time playing the game because they solved the game's science mystery more quickly than others. Players also moved and shifted their bodies while they played the game, so there were moments of facial tracking loss or partial occlusion. More accurate facial expression detection could be achieved if the learners had limited movement, but we opted for more naturalistic conditions in which players were free to move unconstrained. Finally, our sampling of facial expression image collection frequency is once per second, so if a player were to very briefly exhibit a facial expression, it may fall between samples and not be recorded. Exploring the utility of higher frequency sampling is a promising direction for future work.

Conclusion

Educational games hold great potential to promote strong cognitive and affective outcomes for players, but difficulties can arise during the process of learning while playing. It is important to enable educational games to accurately predict when learners are likely to experience negative affective learning outcomes such as low levels of engagement or harmfully high mental demand. This paper presents a predictive learner modeling approach that observes players in early interactions with an educational game. Using facial expressions displayed by players during the tutorial of educational games, models can predict perceived mental demand and engagement at the conclusion of gameplay. Facial indicators of surprise during early phases of gameplay were not only significant for predicting both user engagement and mental demand, but also explained the greatest variance for both models. Moments of surprise may hold potential for effective points of intervention, and should be investigated in future work. It will also be instructive to consider additional windows of observation that are complementary to affect observations in the tutorials of games. Future efforts should examine alternate windows in gameplay as indicated by game narrative episodic framing and task structure to investigate which families of emotions are most predictive of learning outcomes.

Acknowledgements

This work is supported by the National Science Foundation through grant CHS-1409639. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

Aleven, V., McLaughlin, E. A., Glenn, R. A., and Koedinger, K. R. 2016. Instruction based on adaptive learning technologies. *Handbook of Research on Learning and Instruction*. Routledge, 1–79.

Arroyo, I., Ferguson, K., Johns, J., Dragon, T., Meheranian, H., Fisher, D., Barto, A. Mahadevan, S., and Woolf, B. P. 2007. Repairing disengagement with non-invasive interventions. *In Proceedings of the 2007 conference on Artificial Intelligence in Education*, 195–202.

Baker, R. S., Ocumpaugh, J., Gowda, S. M., Kamarainen, A. M., and Metcalf, S. J. 2014. Extending log-based affect detection to a multi-user virtual environment for science. *In International Conference on User Modeling, Adaptation, and Personalization,* 290–300.

Blom, P. M., Bakkes, S., Tan, C. T., Whiteson, S., Roijers, D. M., Valenti, R., and Gevers, T. 2014. Towards personalised gaming via facial expression recognition. *In Proceedings of the Tenth Annual AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 30–36.

Corrigan, S., Barkley, T., and Pardos, Z. 2015. Dynamic approaches to modeling student affect and its changing role in learning and performance. *In International Conference on User Modeling, Adaptation, and Personalization*, 92–103.

D'Mello, S., and Graesser, A. 2012. AutoTutor and affective AutoTutor: Learning by talking with cognitively and emotionally intelligent computers that talk back. *ACM Transactions on Interactive Intelligent Systems,* 2(4): 23.

D'Mello, S., Lehman, B., Pekrun, R., and Graesser, A. 2014. Confusion can be beneficial for learning. *Learning and Instruction*, 29: 153–170.

Ekman, P., and Friesen, W. V. 1978. Manual for the facial action coding system. *Consulting Psychologists Press*.

Grafsgaard, J. F., Boyer, K. E., and Lester, J. C. 2011. Predicting facial indicators of confusion with hidden Markov models. *In International Conference on Affective Computing and Intelligent Interaction*, 97–106.

Grafsgaard, J., Wiggins, J. B., Boyer, K. E., Wiebe, E. N., and Lester, J. C. 2013. Automatically recognizing facial indicators of frustration: A learning-centric analysis. *In Proceedings of the International Conference on Affective Computing and Intelligent Interaction*, 159–165.

Hart, S. G., and Staveland, L. E. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, 52: 139–183.

Hernandez, S. P., Bulitko, V., and Hilaire, E. S. 2014. Emotion-based interactive storytelling with artificial intelligence. *In Proceedings of the Tenth Annual AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 146–152.

Horn, B., Cooper, S., and Deterding, S. 2017. Adapting cognitive task analysis to elicit the skill chain of a game. *In Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, 277–289.

Hutt, S., Mills, C., Bosch, N., Krasich, K., Brockmole, J., and D'Mello, S. 2017. Out of the fr-eye-ing pan: Towards gaze-based models of attention during learning with technology in the classroom. *In Proceedings of the 25th Conference on User Modeling, Adaptation, and Personalization*, 94–103.

Liapis, A., Yannakakis, G. N., and Togelius, J. 2013. Towards a generic method of evaluating game levels. *In Proceedings of the Ninth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 30–36.

Martinez, H. P., Bengio, Y., and Yannakakis, G. N. 2013. Learning deep physiological models of affect. *IEEE Computational Intelligence Magazine*, 8(2): 20–33.

McDuff, D., Mahmoud, A., Mavadati, M., Amr, M., Turcot, J., and Kaliouby, R. el. 2016. AFFDEX SDK: a cross-platform real-time multi-face expression recognition toolkit. *In Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems,* 3723–3726.

Min, W., Mott, B. W., Rowe, J. P., Liu, B., and Lester, J. C. 2016. Player goal recognition in open-world digital games with long short-term memory networks. *In Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI 2016)*, 2590–2596.

Min, W., Mott, B., Rowe, J., Taylor, R., Wiebe, E., Boyer, K. E., and Lester, J. 2017. Multimodal goal recognition in open-world digital games. *In Proceedings of the Thirteenth Annual AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 80–86.

Mudrick, N. V, Taub, M., Azevedo, R., Rowe, J., and Lester, J. 2017. Toward affect-sensitive virtual human tutors: The influence of facial expressions on learning and emotion. *In Seventh International Conference on Affective Computing and Intelligent Interaction*, 184–189.

Nogueira, P. A., Aguiar, R., Rodrigues, R. A., Oliveira, E. C., and Nacke, L. 2014. Fuzzy affective player models: A physiology-based hierarchical clustering method. *In Proceedings of the Tenth Annual AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 132–138.

O'Brien, H. L., and Toms, E. G. 2010. The development and evaluation of a survey to measure user engagement. *Journal of the American Society for Information Science and Technology*, 61(1): 50–69.

Rubio, S., Díaz, E., Martín, J., and Puente, J. M. 2004. Evaluation of subjective mental workload: A comparison of SWAT, NASA-TLX, and workload profile methods. *Applied Psychology*, 53(1): 61–86.

Sabourin, J. L., and Lester, J. C. 2014. Affect and engagement in game-based learning environments. *IEEE Transactions on Affective Computing*, 5(1): 45–56.

Sawyer, R., Smith, A., Rowe, J., Azevedo, R., and Lester, J. 2017. Enhancing student models in game-based learning with facial expression recognition. *In Proceedings of the 25th Conference on User Modeling, Adaptation, and Personalization*, 192–201.

Vail, A. K., Grafsgaard, J. F., Boyer, K. E., Wiebe, E. N., and Lester, J. C. 2016. Gender differences in facial expressions of affect during learning. *In Proceedings of the 2016 Conference on User Modeling Adaptation, and Personalization*, 65–73.

Wang, N., and Marsella, S. 2006. "Introducing EVG: an emotion evoking game." *International Workshop on Intelligent Virtual Agents*, 282-291.

Wiebe, E. N., Lamb, A., Hardy, M., and Sharek, D. 2014. Measuring engagement in video game-based environments: Investigation of the User Engagement Scale. *Computers in Human Behavior, 32*(3): 123-132.