



HAL
open science

Engagement and Discrete Emotions in Game Scenario: Is There a Relation Among Them?

Renan Vinicius Aranha, Leonardo Nogueira Cordeiro, Lucas Mendes Sales,
Fátima Nunes

► **To cite this version:**

Renan Vinicius Aranha, Leonardo Nogueira Cordeiro, Lucas Mendes Sales, Fátima Nunes. Engagement and Discrete Emotions in Game Scenario: Is There a Relation Among Them?. 18th IFIP Conference on Human-Computer Interaction (INTERACT), Aug 2021, Bari, Italy. pp.22-42, 10.1007/978-3-030-85613-7_3. hal-04292359

HAL Id: hal-04292359

<https://inria.hal.science/hal-04292359>

Submitted on 17 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

Engagement and discrete emotions in game scenario: is there a relation among them?

Renan Vinicius Aranha^{1,2}, Leonardo Nogueira Cordeiro², Lucas Mendes Sales², and Fátima L. S. Nunes^{1,2}

¹ Interactive Technologies Laboratory - Escola Politécnica

² Laboratory of Computer Applications for Health Care - School of Arts, Sciences and Humanities

Universidade de São Paulo - USP

{renanvinicius, leonogueirac, lucasm1sales@usp.br, fatima.nunes}@usp.br
<http://www.poli.usp.br> <http://lapis.each.usp.br>

Abstract. Analysis of user engagement is a factor of great importance for different computational applications, especially games. The knowledge of this information allows evaluating the software and making possible enhancing the software. However, analyzing user engagement is still a challenge. The most common approach in the literature has been the application of questionnaires at the end of the interaction process. Although some studies investigate the use of emotions to measure user involvement, the analysis of facial expressions has been little explored in this context. This technique has the main advantage of using a camera, dispensing other invasive sensors. This paper presents an investigation to evaluate whether discrete emotions can be related to engagement in game scenarios. Two studies were conducted with 48 volunteers playing games to discover if there is relation between emotions recognized by facial expressions and engagement measured in two different ways (by software and by user's self-assessment). Our analysis comprises about 38 hours of games and more than one million records of emotional states. The results indicate that there is no universal relation between discrete emotions and engagement, suggesting that the game context can influence users' engagement.

Keywords: Games · Engagement · Discrete emotions.

1 Introduction

The term “engagement” is widely used in the context of games to define an ideal state, in which the player is experiencing positive involvement with the game [30]. In a more practical sense, an engaged player is not bored when using the game and is willing to continue exploring the game and its challenges. Although the term engagement has been widely used, identifying if a person is engaged or not can be a complex task.

Firstly, there is no universal understanding from theorists about what this term represents [11, 12, 19]. Possibly, as a consequence, there are different approaches and methods to measure engagement, which can provide different results. Nowadays, a state-of-the-art approach to measure the user's engagement consists in applying questionnaires [6, 19, 30]. Some questionnaires are well-known and validated by other studies, as the User Engagement Scale (UES) [29] or the Game Engagement Questionnaire

(GEQ) [8]. Despite its several benefits, the application of questionnaires can imply some disadvantages, since it usually interrupts the interaction process, breaking the focused engagement [31]. To avoid the interruption of the experience, this approach is more recommended to be used at the end of the interaction.

A promising approach consists in analyzing the user's physiological responses, such as heart rate, in addition to the aforementioned techniques. This approach, sometimes providing data in real time, has been explored both in research and in the gaming industry [30]. An interesting point of this approach is the possibility to collect data without interrupting the interaction process, besides improving the analysis with a kind of information that probably not even the user is aware of.

However, most of the initiatives are based on sensors coupled to the user body [30], and some of them can cause discomfort to the user – a scenario that has changed in the last few years with the development of wearable sensors. Additionally, this kind of measure is usually adopted in laboratory-based studies, since some sensors are not common and may require training in handling and use.

Conversely, facial expression recognition (FER) can infer the users' emotions in real time without involving the use of specific sensors [3]. FER can use popular devices, such as webcam or a smartphone camera to acquire images, and be used to analyze the player's emotions even remotely. Despite these advantages, the relation of the data provided by this technique with engagement have been little explored when compared to the traditional physiological measures.

In this challenging scenario, a few companies and researchers have been investigating the automatic measurement of users' engagement by analyzing facial expressions. At least two commercial software, named Affdex SDK [1] and FaceReader [27], provide indicators of engagement in their interfaces. In both cases, the engagement is computed considering the user's facial expressions. However, the weight of each facial expression, emotion, or action unit (AU) [15] in the engagement measurement is not explained in detail in either case [1]. Additionally, both software imply paid licenses, which can reduce or make the adoption of this technique impracticable in many contexts.

In the academic scope, some studies explore machine learning techniques to predict user engagement from facial expressions. In general, these studies involve a step of manual classification of facial images, which are used to train Artificial Intelligence algorithms [9,23]. However, finding a relation between the discrete emotions and engagement remains a challenge, since there is no consensus about the impacts and relations of the universal emotions in the engagement. Although the measurement of engagement based on facial expressions has been explored by a few researchers, their findings can be quite controversial [19].

Aiming to contribute within this scenario, we here investigate if there is a correlation between emotions recognized by FER and engagement. The choice for this technique considers its effectiveness, already explored by previous studies [3, 20, 22], as well as the ease of its use without major impacts on the user's environment or setup. Since the engagement has different interpretations and measurement approaches, we present two studies: the first one, using emotional and engagement data provided by the Affdex SDK during the use of serious games; and the second one, concerning emo-

tional data collected during the use of two games, with the engagement obtained by a self-assessment questionnaire.

2 Emotion recognition

Just as engagement, understanding human emotions can be a complex task, since the literature presents different interpretations of emotional aspects. In this sense, few approaches to the classification of emotions are known [14], and the most used are the discrete approach proposed by Darwin [10] and the dimensional approach proposed by Wundt [35]. While Darwin [10] understands emotions as discrete (or modular), and represented by terms such as joy, sadness and disgust, for example, Wundt [35] classifies emotions using dimensions, such as pleasant/unpleasant and low/high intensity. In the Computer Science context, researches involving emotions consider both discrete and dimensional approaches [3].

In most cases, choosing an approach depends on the emotion recognition technique used. In facial expressions analysis, for example, it is usual to adopt the discrete approach [32, 33]. This is because studies that investigate the expression of emotions on the face consider the discrete emotions approach [15].

Facial expressions are recognizable using at least two techniques: facial electromyography (fEMG), which uses sensors coupled to the user's face, and image processing. In this study, we consider only the latter, since it only needs the use of a camera and software to process the images. Usually, software that recognize facial expressions can use at least two approaches.

The first one is to analyze facial landmarks directly. Facial landmarks are fiducial facial key points used in Computer Vision to extract and identify components of the face [34]. DLib, for example, a common library used to extract facial landmarks, recognizes 68 points in the human face. Considering these facial landmarks and previously labelled image datasets, it is possible to train machine learning algorithms to recognize facial expressions.

The second approach uses a coding system about muscle movements. In this scope, the FACS (Facial Action Coding System) is commonly adopted. FACS was proposed by Ekman and Friesen [15] to measure all visible facial behavior by analyzing muscle movements and their components, which receive the name of Action Units (AUs) [15, 16]. FACS contains the definition of 44 AUs, each one identified by a number. The analysis of movements of the AUs and their intensities allows mapping facial expressions to emotions through Emotional FACS (EMFACS) [17]. The emotion of joy, for example, involves AU6 (orbicularis oculi muscle) and AU12 (zygomaticus major muscle). Figure 1 represents neutral and facial expressions, with facial landmarks processed using DLib. From the neutral to the joy facial expression, it is possible to observe that the second case involves cheek and lip corners rising.

3 Related works

Although facial expressions have traditionally been used to recognize basic or universal emotions (anger, disgust, fear, joy, sadness and surprise), only few works have investi-

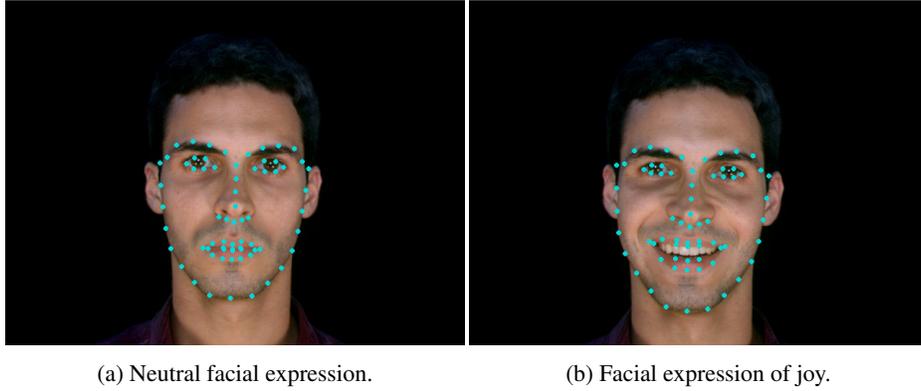


Fig. 1: Facial expressions with facial landmarks processed using the Dlib library. The facial expression of joy, compared to the neutral facial expression, involves raising the cheek and the lip corners. Images from FacesDB dataset [24].

gated their use for recognizing engagement. In a literature review about the measurement of engagement, Dewan, Murshed and Lin [11] indicated that only recently have FACS received attention for engagement recognition. In this review, the authors present some studies that investigated the recognition of discrete emotions associated with engagement in learning, such as “boredom”, “confusion” and “frustration” [11]. The authors identified no studies considering engagement as an emotion, but some authors related engagement to the combination of other emotional states, such as boredom and delight. However, even because these studies are relatively recent, there is not consensus about the findings. For example: in [23], the AUs related to confusion were AU1, AU4, AU7, and AU12. In [18], this emotion can be identified by AU4 and AU7. Finally, in a more recent work, confusion was associated with AU45, AU1, and AU4 [7].

In [9], researchers proposed FaceEngage, an engagement estimator that analyzes the facial expressions in gaming videos obtained from YouTube. Considering that engagement annotation is challenging both for computer engineers and psychologists, the authors analyzed the game events in the videos to classify if a user was active or not in the game. The authors extracted the eyes and the head motion as features, and trained Machine Learning algorithms. Despite the positive results, with an accuracy of 83.8%, the authors mentioned as further work the development of a Likert-scale engagement measurement tool, to enhance the analysis of engagement in videos initially classified as “uncertainty”.

Outside the FACS context, Nezami et al. [26] explored the use of deep learning to recognize user’s engagement through facial expressions. The authors collected a dataset containing about 20 hours of students interacting with an educational virtual environment. To classify the users engagement, Nezami et al. [26] considered two annotation steps, embracing the behavioral and emotional dimensions of engagement. The annotations were performed by Psychology students, who classified engagement in the behavioral dimension (on-task, off-task, cannot decide) and in the emotional dimension

(satisfied, confused, bored, cannot decide). Engagement was defined by mixing both dimensions labels. The results indicated that the approach proposed by the authors was more effective in recognizing non-engaged users, with accuracy of 60%.

Although these works show promising results, highlighting the possibility of analyzing users' engagement through their facial expressions, this field of study still has great open challenges. A relation between universal discrete emotions and user engagement, for example, was not addressed by these articles.

4 Methods

Given the challenging scenario about the automatic measurement of engagement and the feasibility of recognizing emotions by facial expressions, this paper analyzes if discrete emotions can indicate players' engagement in games. To offer a broader perspective about this topic, we describe two studies using different scenarios. The first study was based on data collected during the use of a serious game for motor rehabilitation, while the second study comprised two entertainment games. Considering these scenarios, our approach allows identifying whether there is any relation between emotions and engagement, even under different conditions, which could define the later using knowledge about the former. With both studies, we aimed to answer the following research questions (RQs):

- RQ1: Is there a relation between discrete emotions and users' engagement?
- RQ2: Can users experience different emotions when they are engaged and when they are not engaged?

We present the studies and how these questions are answered in each as follows.

4.1 Study 1: Emotions and engagement analyzed by software

In this first study, we analyze if there is a relation between discrete emotions and engagement when both are measured by software. For this, we use a previously collected dataset, which includes emotion and engagement records measured by software during the use of a serious game.

Software. Developed by Affectiva, the Affdex SDK is an emotion recognition software that analyzes the user's facial expressions considering FACS [1]. For each frame processed, in addition to the discrete categorization of seven emotions (joy, contempt, fear, disgust, anger, surprise and sadness), this software offers indices for dimensional affective states, such as valence and engagement [1]. This feature is a differential of *Affdex*, since most similar applications focus on the recognition of the six universal emotions.

The Affectiva website describes that engagement is "a measure of facial muscle activation that illustrates the subject's expressiveness" [1]. It is calculated from the weighted sum of some AUs, having a variable index from 0 to 100, where 0 indicates no engagement and 100 represents total engagement. The weight attributed to each AU in the calculation of engagement is not described by the company. However, a study

published by Kapoor and Picard [21] indicates the relation between some AUs and engagement. Although it is not possible to state that the engagement analysis adopted by *Affdex* is based on this study, there is a great similarity with the mentioned AUs in the *Affectiva* website.

Dataset. *EasyAffecta* Dataset is a dataset of emotional data collected during a longitudinal experimental evaluation with 11 individuals (approved by the Ethics Committee of the School of Arts, Sciences and Humanities of the University of São Paulo). This dataset comprises more than one million records of emotional states obtained with *Affdex SDK* during the use of a serious game with four exercises in a motor rehabilitation process, encompassing about 33 hours of data collection. In addition to the percentages of seven discrete emotions, each record also includes the engagement and valence indices. An important feature of this dataset is that the individuals' facial expressions were not stimulated. Therefore, they tend to represent a situation closer to everyday life.

Data analysis. Considering the data from the *EasyAffecta* Dataset, the following indices were calculated to answer the questions RQ1 and RQ2: i) the Pearson correlation between emotions and engagement; ii) the average intensity of each emotion when users were engaged and when they were not engaged; and iii) the analysis of emotional intensity (AEI) that comprises, for each emotion, the percentage of records in which it was the most intense. At this point, we should highlight that for each frame, the ER software provides a probability index for each emotion. The AEI analysis focuses on the peaks, indicating the percentage of frames in which that emotion has the highest probability. The threshold to split the records between engaged and non-engaged is the engagement average of all players. The results are presented and discussed in Section 5.

4.2 Study 2: Emotions analyzed by software, and engagement self-assessed by users

Complementing the aforementioned approach, the second study consists of a hybrid analysis, with discrete emotions recognized by software being compared to the engagement self-assessed by the users. This analysis involved the preparation of a specific experimental evaluation with users, which is described below.

Evaluation design. Figure 2 presents an overview of the experimental evaluation, approved by the Ethics Committee of School of Arts, Sciences and Humanities of the University of São Paulo. A web platform was developed to conduct the experiment. When accessing the platform, the user found a text informing the invitation to participate in the experiment. Accepting the invitation, the user was redirected to a screen containing the Informed Consent Form. Agreeing to the term, the user made a registration and then was presented with the activities to be performed during the experiment. The stages of the experiment involved: activating the camera; interacting with two games and, after using each, self-evaluate the engagement using a questionnaire. In the last phase, we asked users to leave free text comments to researchers on their perceptions when participating in the experiment. The last phase was not mandatory.

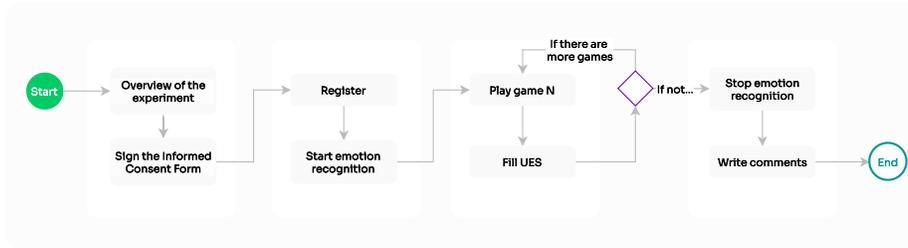
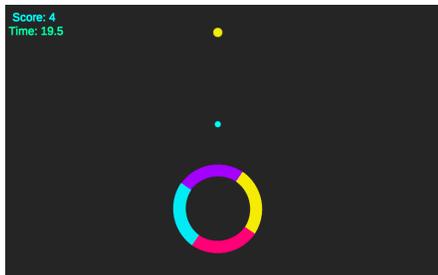


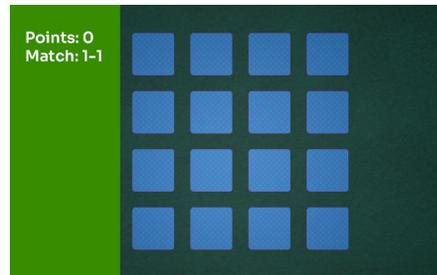
Fig. 2: Overview of the experimental evaluation.

Games. During the experimental evaluation, the volunteers can interact with two games, as shown in Figure 3. The first one is based on “Color Switch Replica”, a project available on GitHub [4]. In this game, illustrated in Figure 3a, there is a small circle with gravity in the center of the screen. There is also a greater and spinning circle around it drawn in different colors. The player should click on the screen to move up the small circle. However, when moving up, the player should take care to prevent the small circle from colliding with the part of the greater circle that has another color.

The second game consists of a game developed by the authors, based on the classic Memory Game, in which the player should find a pair of cards, as shown in Figure 3b. Both games were customized to communicate with the experimental platform created for this evaluation. Volunteers were asked to play each game as many times as wanted.



(a) Screenshot of the Color Switch Replica.



(b) Screenshot of the Memory Game.

Fig. 3: Screenshots of the games used in this experimental evaluation.

Self-assessment engagement. To assess the user engagement, we chose the short form of User Engagement Scale (UES) [28], which contains 12 items with statements about the experience of using software. For each statement, the volunteer should classify his/her agreement in a Likert scale, from 1 to 5. Besides a global engagement measure, UES also provides information about four dimensions of engagement: aesthetic appeal (AE), focused attention (FA), perceived usability (PU), and reward (RW) [28].

Table 1 shows the items that composes this scale. Following the instructions described in [28], the items order was randomized in each application of the UES.

Table 1: Items that composes the short form of UES.

Dimension	Item
FA-1	I lost myself in this experience.
FA-2	The time I spent using this game just slipped away.
FA-3	I was absorbed in this experience.
PU-1	I felt frustrated while using this game.
PU-2	I found this game confusing to use.
PU-3	Using this game was taxing.
AE-1	This game was attractive.
AE-2	This game was aesthetically appealing.
AE-3	This game appealed to my senses.
RW-1	Using this game was worthwhile.
RW-2	My experience was rewarding.
RW-3	I felt interested in this experience.

Facial expressions. Unlike the first study, we chose to use another facial expression recognition software. *Face-api.js* is an open-source software that recognizes facial expressions, but does not provide engagement levels as *Affdex SDK* does. For each processed frame, this software offers the probability of the presence of each one of these facial expressions: neutral, anger, disgust, fear, joy, sadness, and surprise. In addition to its open source feature, which allows its adoption without financial impact, the accuracy of *Face-api.js* in recognizing emotions was investigated in [2]. Besides the aforementioned advantages, the fact that this tool works in the browser without the need to install components was the main factor that contributed to its choice. The emotional data generated was stored in a database.

Participants. The volunteers were invited by e-mail and by using social networks. In general, 37 volunteers concluded all the steps. In this scope, 12 are female and 25 are male. No volunteer declared not to identify with one of these binary genres. The average age of the volunteers was about 34 years. Then 80 activities were recorded, 39 for the Memory Game and 41 for the Color Game. The term “activity” is used to define the process that involves the use of a game, with the subsequent completion of the UES. Two volunteers participated in the experiment more than once, which justifies the greater number of computed activities. Additionally, 107876 records of emotional states were stored, encompassing about 5 hours of data collection.

Data analysis. Similarly to the first study, to answer question RQ1 we considered the Pearson correlation between the engagement measured by UES with the average

intensity of each emotional state, both computed for each activity. To answer RQ2, we analyzed the average intensity of each emotional state when users were engaged and when they were not engaged. In this case, since two games were addressed in the experiment, the analysis was separated for each game. Also similarly to the previous experiment, the AEI was also computed, indicating, for each emotion, the percentage of records in that emotion was the most intense.

For these last two analyses, two thresholds were calculated, which consist of: i) the average engagement of all players in activities related to the Memory Game (MG); and ii) the average engagement of all players in activities related to the Color Switch Replica (CSR). The results are presented and discussed in Section 5.

5 Results and discussion

We next present the results and discussions from the two studies aforementioned. In the first study (described in Section 4.1), which comprises emotions and engagement measured by software in a rehabilitation scenario, we observed that: i) players did not experience all the discrete emotions while using the game; ii) there was great variability in the engagement average of the players; iii) only facial expression of joy had a moderate correlation with engagement; and iv) when players are engaged, they do not necessarily feel different emotions from when they are not engaged. This discussion will be detailed in Section 5.1.

Concerning the second study (described in Section 4.2), which mixed emotions measured by software and engagement measured using a questionnaire in entertainment games, we can point out that: i) similarly to the previous study, players did not experience all the discrete emotions when playing the games; ii) players themselves indicate high levels of engagement by filling the questionnaire; iii) for both games, no emotion has a moderate correlation with engagement; and iv) again, similarly to the previous study, there was not a difference between the emotions experienced when users are engaged or not. The discussion will be presented in Section 5.2.

5.1 Study 1: Emotions and engagement analyzed by software

Overview. Figure 4 shows the distribution of each facial expression of emotion in the dataset. In general, there is a greater density of records with low intensity (in the range of 0 to 20%). Only the facial expressions of joy, contempt and disgust showed significant densities at high intensity (in the range of 80 to 100%).

These data can indicate that users may not experience all seven facial expressions of emotions while using games, and reaffirm that people do not necessarily express emotions all the time, as mentioned by Paul Ekman [13]. In the context of games, this aspect may be relevant, since the facial expression of an emotion tends to indicate a reaction to an event of the game - which arouses an emotion in the user.

Concerning engagement, Figure 5 shows an overview about the engagement levels experienced by users in this dataset, considering the minimum, the maximum, the median and the first and third quartiles of each volunteer's engagement. Aiming to improve the understanding, the data were sorted by the median value. This figure indicates

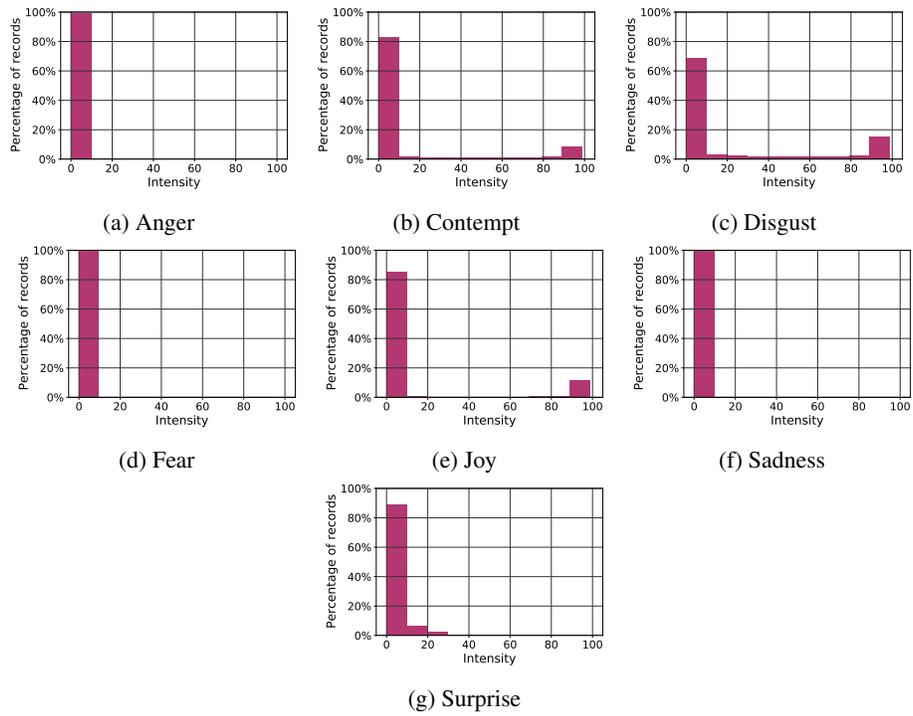


Fig. 4: Distribution of emotional states in the first experiment.

that there was a great variation in the engagement of the participants, which had low indexes. A great part of the volunteers (77%) showed median engagement under the average, as indicated by the red line.

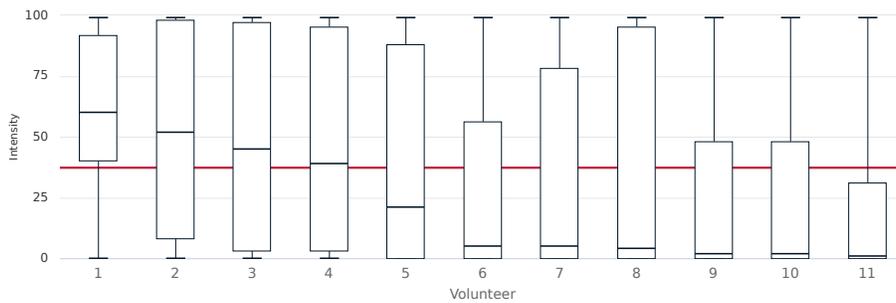


Fig. 5: Overview about user’s engagement in the first experiment.

Emotions intensity and engagement. The correlation between discrete facial expression of emotions and engagement in the scope of EasyAffecta Dataset is shown in Table 2. The main finding is that the facial expression of joy had a moderate correlation ($\rho = 0.6$) with engagement. When considering the AUs involved in engagement described in [21], this correlation is expected, since the AUs referring to the emotional state of joy, AU6 and AU12, are positive indicators of engagement. Even moderate, this correlation can reinforce a common feeling that, if the user is experiencing joy or a positive emotion, he/she is engaged in the task.

Table 2: Correlation between engagement and discrete emotions.

Emotion	Anger	Contempt	Disgust	Fear	Joy	Sadness	Surprise
Correlation (ρ)	0.09	-0.05	0.19	0.08	0.6	0.06	0.32

Complementarily, although it presented a weak correlation ($\rho = 0.32$), the facial expression of surprise is also an expected correlation, considering that some of its action units (AUs 1,2 and 5) are considered positive indicators of engagement by Kapoor and Picard [21]. Since the emotion of surprise indicates a reaction to an unexpected event, its correlation with engagement is likely, indicating that the user was probably paying attention to the games.

Finally, the facial expression of disgust had a weak correlation with engagement ($\rho = 0.19$). However, this is an unexpected positive correlation, considering that AUs 9 and 15, which make up this facial expression together with AU16, are presented as negative indicators of engagement [21]. We can next observe, respectively, the following facial expressions of emotions: anger ($\rho = 0.09$), fear ($\rho = 0.08$) and sadness ($\rho = 0.06$), all with low indices.

Emotions experienced. Another observed aspect, related with the RQ2 question, aims to identify whether the volunteers experienced facial expression of emotions while they were engaged differently from when they were not engaged. Figure 6 presents a comparative analysis of the average and AEI in the two engagement situations.

Considering the average of the intensity, Figure 6a reveals that, while the volunteers were engaged in the task, they mainly experienced four emotional states: joy, disgust, contempt and surprise. When users were not engaged, the facial expressions of joy and surprise had a very low intensity average. Only the facial expression of contempt showed a higher intensity average in not engaged situations.

In turn, although it looks similar to Figure 6a, Figure 6b reveals an important fact regarding the facial expression of disgust. Although this facial expression of emotion had a higher average of intensity when users were engaged, it was not necessarily the predominant emotion most of the time. Figure 6b shows, in reality, that there is a greater number of records of the facial expression of disgust with higher intensity in situations of non-engagement. This conflicting result can suggest that the duration of the facial expression of emotion in higher intensities may be a factor that impacts engagement.

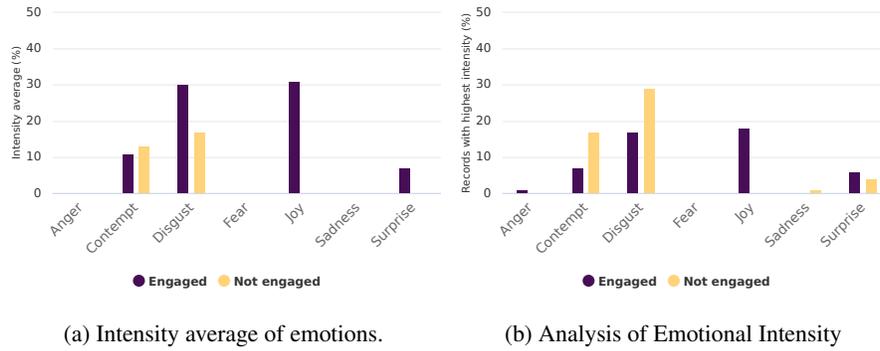


Fig. 6: Facial expressions of emotions experienced by users.

5.2 Study 2: Emotions measured by software and engagement measured by self-assessment

Overview. The distribution of each facial expression of emotion in the dataset obtained in this experiment is shown in Figure 7. Similarly to the previous experiment, there was a greater density of facial expression of emotions with records in low intensity (0 to 20%). Besides the neutral facial expression, only the facial expressions of joy and sadness got significant densities at high intensity (in the range of 80 to 100%). Although the facial expressions of anger and surprise had high intensity emotional records, a minor volume of data is related with these records.

The data shows that not all universal facial expressions of emotions were experienced by the users; likewise, users will not necessarily facially express an emotion all the time. This fact is even more evident, since the software used recognizes neutral facial expression.

Figure 8 shows the engagement obtained by the volunteer's self-assessment in this experimental evaluation, showing the minimum, the maximum, the median and the first and third quartiles of engagement for each volunteer. To enhance the understanding, the data were sorted by the median value. Users' engagement in activities was high. As indicated by the figure, the rates of most users exceed the average engagement, indicated by the red line.

Emotion intensity and engagement. The correlation between the average intensity of facial expression of emotions experienced by players and the engagement shown in Table 3 reveal a remarkable variation in this metric from one game to another. Despite this variation, facial expression of emotion did not have a strong correlation with the engagement measured by UES in either of the cases. In a way, a low correlation between these indices can be expected, since it involves two different approaches to analyze the player's experience, and relating them is a complex task [25].

In this experiment, the highest positive index identified indicates correlation of $\rho = 0.22$ of the facial expression of joy with engagement when users were playing the Memory Game. In this game, joy was the only facial expression of emotion with a

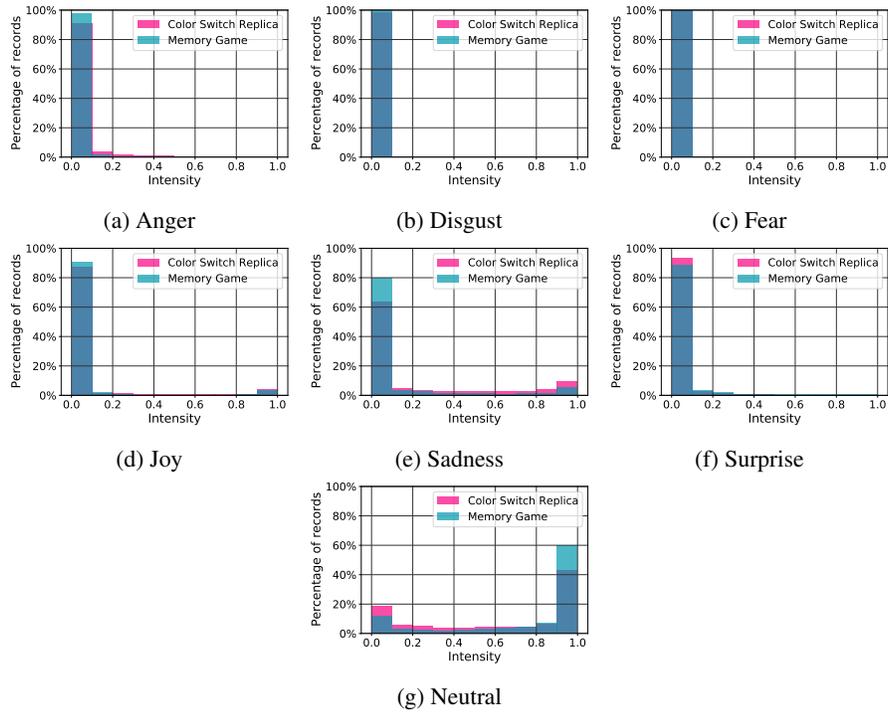


Fig. 7: Distribution of emotional states in the second experiment. In each histogram, the darker blue color indicates the overlap of registers from the two games.

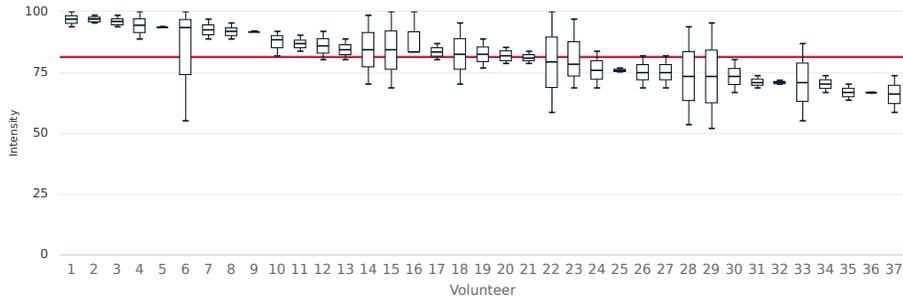


Fig. 8: Overview about user’s engagement in the second experiment.

positive correlation with engagement. Curiously, this facial expression had a negative correlation ($\rho = -0.05$) with engagement in the Color Switch Replica.

Still regarding the differences, while the facial expression of fear showed a negative correlation ($\rho = -0.28$) with engagement in the Memory Game, this facial expression showed a positive correlation ($\rho = 0.15$) with engagement in the Color Switch Replica. Also in the latter game, the facial expression of disgust showed a similar correlation

($\rho = 0.18$). However, the facial expressions of fear and disgust had lower intensities, with their values usually equal to zero.

The variation in the correlation from one game to another can be justified by the difference in the context of the games. While the Memory Game requires less agility from the user, the Color Switch Replica demands constant agility and attention, which may have made users tenser, albeit engaged.

Table 3: Correlation between engagement and discrete emotions for each game.

Emotion	Anger	Disgust	Fear	Joy	Sadness	Surprise	Neutral
CSR (ρ)	-0.25	0.18	0.15	-0.05	0.06	0.04	-0.02
MG (ρ)	-0.03	0	-0.28	0.22	0	-0.12	0

Emotions experienced. Figure 9 shows the average intensity of facial expression of emotions experienced by users when using the two games, also as the AEI (metric explained in Section 4). This analysis allows us to identify the facial expressions of emotions experienced by users when engaged and not engaged. As described in Section 4, the data were grouped into two categories, engaged (E) and non-engaged (NE).

Figure 9 enriches the discussion by reinforcing that the results vary from game to game. While in the Memory Game higher intensities of sadness may indicate that the user is not engaged, the opposite is observed in relation to the Color Switch Replica.

Another curious aspect can be observed analyzing data from both games. One of the most intense facial expression of emotions experienced by users while engaged was sadness, an emotion traditionally considered negative. In the Color Switch Replica, for example, there is a curious fact: the volunteers experienced the facial expressions of emotions joy and surprise at a greater intensity when they were not engaged. Even limited to this scope, this fact indicates that there is not necessarily a relation between positive emotions and engagement, which is often suggested or expected.

No significant differences were found in the metrics of medium intensity and AEI. In this case, this may be due to a characteristic of the software used. The software Face-api.js does not recognize that more than one facial expression of emotion can happen simultaneously with maximum intensity, as occurs in Affdex. For this reason, in most cases, when an emotion is identified by Face-api.js, this emotion tends to be the most intense at that moment.

Comments from volunteers. In the last stage of the experiment, the volunteers were asked to write their suggestions, criticisms or considerations. Only 10 volunteers wrote their considerations. In general, the comments involve the users’ opinions about the game based on the Color Switch Replica and considerations about the questionnaire.

Concerning the game, some volunteers considered the game complex to use and, in some cases, expressed their preference for the Memory Game. One volunteer wrote that he/she “was a little upset with the first game (Color Switch Replica) for not being able

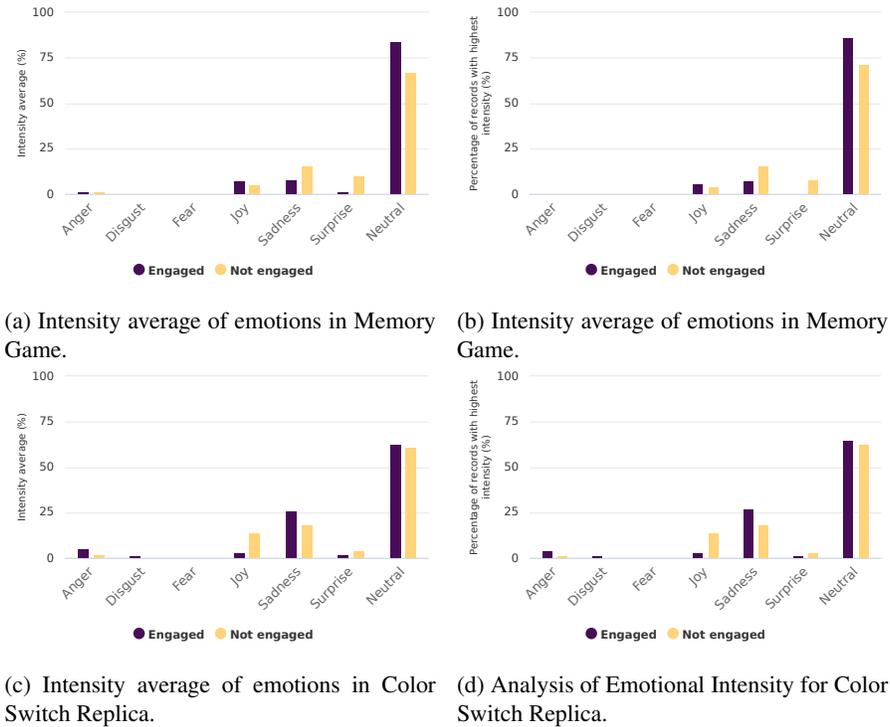


Fig. 9: Facial expressions of emotions experienced by users in both games.

to keep the ball going up”, and “really liked it (Memory Game), because it was much more interesting”. In fact, the data collected reinforce the preference of this volunteer: while in Color Switch Replica the engagement was 2.75, the engagement was 4.33 in the Memory Game. Regarding facial expressions of emotions, in the Color Switch Replica, the most frequent facial expressions identified were joy and anger (besides the neutral facial expression), while in the Memory Game the neutral facial expression and the facial expression of joy were identified.

The Color Switch Replica attracted another volunteer, who wrote: “In the Color Switch Replica I had more difficulty, but it is challenging and holds my attention; I want to overcome at each restart”. Although this volunteer played the Color Switch Replica more than once, which tends to reinforce the interest in this game, the engagement levels obtained on both occasions (4.17) were inferior to the Memory Game (5). Regarding facial expressions of emotions, in one play of Color Switch Replica, this volunteer has only the neutral facial expression identified. For the other activities, from Color Switch Replica and Memory Game, neutral facial expression and joy were identified.

These comments from the volunteers reinforce the fact that the individual characteristics of the players can have a great influence on their perceptions and on their involvement with the game.

Another aspect addressed in the comments involved the engagement analysis questionnaire. Some volunteers reported difficulty in understanding the questions, while one said that he felt interest in the games only for the context of the experiment, but that he would not want to use such games in the future. For this last volunteer, the writing of the question did not make this difference clear.

The considerations of the volunteers in relation to the questionnaire is an important aspect to be considered for the assessment of engagement. Although the UES is widely used and validated, the difficulty in understanding certain statements in the questionnaire identified in our study may indicate the need for better clarification of the items. Additionally, this is already a challenge known in the literature [11]. In further investigations, it is interesting to assess whether simpler metrics, such as “like” or “don’t like”, can offer users a better understanding when evaluating their experiences.

5.3 Overall analysis

In the first study, the evaluation indicated a correlation between facial expressions of discrete emotions, especially joy and surprise, with engagement. However, some emotions – such as fear and sadness – were little experienced by the volunteers. Therefore, the correlation of these facial expressions of emotions with engagement may not have been identified. Additionally, an important observation regarding this study is that the data collection occurred in a process of motor rehabilitation, with the game used reproducing real rehabilitation tasks. Since the motor rehabilitation tasks can result in pain or physical discomfort in patients, this may also have influenced the users’ facial expressions and, consequently, the correlation of facial expressions of emotions with engagement.

Although our findings in the first study did not fully agree with that stated in [21], it was possible to confirm some promises about facial expressions of discrete emotions and engagement using the Affdex SDK. At this point, we can conclude that the facial expressions of emotions of joy and surprise can be considered indicative – even if slight – that the user is engaged in the task.

However, the second study improved this discussion, indicating that the correlation of engagement with facial expressions of discrete emotions is variable to the context and the design of the game. The great difference regarding the correlation presented in the games of the second study and the game of the first study is a factor that deserves to be highlighted, suggesting that there is no universal relation between facial expressions of emotions and engagement.

Another contribution from the second study is that there is not necessarily a relation between positive emotions and engagement (which can be understood in the first study). Interestingly, the facial expressions of emotions identified when users were engaged at using Color Switch Replica encompass emotions usually considered to be negative. These findings reinforce the discussion presented by Bopp, Mekler and Opwis [5], who observed that negative affect can be related with a positive experience in games. Similarly to movies, games can arouse emotions in players that are considered negative, but the arousal of these emotions in the player/viewer can be an indication that the user is involved with the multimedia content.

As already discussed the facial expression of fear was rarely identified in both studies. Some facial expressions of emotions, such as disgust and anger, were identified in one study, but not in the other. Although these emotions were not widely experienced, they can play an important role in engagement depending on the context of the game.

In general, the volunteers' engagement levels in the second study were higher than in the first study. This difference is understandable and justified for two reasons: i) first, the volunteers from the first study were in physiotherapy sessions, while the volunteers from the second study participated in the research freely; and ii) the metrics used to measure engagement are considerably different, which justifies the wide variation in this metric.

Although the results of our studies are not conclusive, the facts discussed here make the analysis of the relation between facial expressions of discrete emotions and engagement even more challenging. The differences identified in the relation of engagement and facial expressions of discrete emotions depend on the goal and the characteristics of each game.

In a game for motor rehabilitation, for example, it is important to motivate the player to carry out the activities – which may not occur with such intensity in traditional games. These aspects certainly influence the player's reaction to the events of the game. In addition, motor rehabilitation patients, for example, tend to deal with pain and motor difficulties, unlike typical users of entertainment games. Therefore, it is essential that the characteristics of each game are considered when investigating the correlation between engagement and emotions.

6 Final remarks

Although there are several techniques for evaluating user's engagement, most of them are focused on face-to-face evaluations, with a high degree of involvement of specialists in the collection and analysis of data. In order to contribute to this scenario, we here investigated if facial expressions of discrete emotions can indicate player's engagement. Our findings from two studies involving the use of three games and involving different contexts initially revealed that users do not experience universal emotions constantly. Although it seems strange, this is in agreement with the literature of the area and, in the case of games, it can indicate an emotional reaction of the user to specific events of the game.

Both studies indicated that the correlation between discrete emotions identified by facial expressions and engagement is weak. In the first experiment, joy and surprise could be considered positive indicators of engagement. In the second experiment, although joy also showed a positive correlation with engagement in the Memory Game, this emotion has a negative correlation with engagement in the Color Switch Replica.

Another important finding in this article is that there is not necessarily a relation between positive emotions and engagement, as evidenced in the aforementioned case of Color Switch Replica. In other words, our findings indicated that there is not necessarily a global relation of emotions or group emotions with engagement, which can be used or exchanged to different types of games.

Finally, even though the results of this study do not provide an equation for calculating engagement from facial expressions, for example, our findings reinforce the need for further investigation in this regard, also including the use of other techniques for analyzing the player's engagement in conjunction with facial expressions. In future work, we intend to explore the correlation of engagement with discrete emotions and other metrics extracted from non-invasive methods, such as gaze and game logs. Additionally, this should involve the use of Machine Learning techniques to identify possible correlations between these data, also including a temporal analysis considering the use of the game at different moments.

Acknowledgments

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001; Brazilian National Council of Scientific and Technological Development (CNPq) - grant 309030/2019-6); São Paulo Research Foundation (FAPESP) – National Institute of Science and Technology – Medicine Assisted by Scientific Computing (INCT-MACC) – grant 2014/50889-7; PIBIC/CNPq/USP - grant 160013/2020-8), and PUB program of Provost for Undergraduate affairs of Universidade de São Paulo (USP).

References

1. Affectiva. Emotion AI 101: All About Emotion Detection and Affectiva's Emotion Metrics, August 2020.
2. Renan Vinicius Aranha, André Biondi Casaes, and Fátima L. S. Nunes. Influence of environmental conditions in the performance of open-source software for facial expression recognition. In *Proceedings of the 19th Brazilian Symposium on Human Factors in Computing Systems*. ACM, October 2020.
3. Renan Vinicius Aranha, Cleber Gimenez Correa, and Fátima L. S. Nunes. Adapting software with affective computing: a systematic review. *IEEE Transactions on Affective Computing*, pages 1–1, 2019.
4. Oussama Bonnor. Color switch replica - github, january 2020.
5. Julia Ayumi Bopp, Elisa D. Mekler, and Klaus Opwis. Negative emotion, positive experience? In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, May 2016.
6. João Bosco Borges, Carmen Li Juy, Izac Sidarta De Andrade Matos, Paulo Vitor Angelo Silveira, and Ticianne De Gois Ribeiro Darin. Player experience evaluation: Which instrument should i use? *Journal on Interactive Systems*, 11(1):74–91, December 2020.
7. Nigel Bosch. Detecting student engagement: Human versus machine. In *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization, UMAP '16*, page 317–320, New York, NY, USA, 2016. Association for Computing Machinery.
8. Jeanne H. Brockmyer, Christine M. Fox, Kathleen A. Curtiss, Evan McBroom, Kimberly M. Burkhart, and Jacquelyn N. Pidruzny. The development of the game engagement questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology*, 45(4):624–634, July 2009.
9. Xu Chen, Li Niu, Ashok Veeraraghavan, and Ashutosh Sabharwal. FaceEngage: Robust estimation of gameplay engagement from user-contributed (YouTube) videos. *IEEE Transactions on Affective Computing*, pages 1–1, 2019.

10. Charles Darwin. The expression of the emotions in man and animals. Oxford University Press, USA, 1998.
11. M. Ali Akber Dewan, Mahbub Murshed, and Fuhua Lin. Engagement detection in online learning: a review. Smart Learning Environments, 6(1), January 2019.
12. Kevin Doherty and Gavin Doherty. Engagement in HCI. ACM Computing Surveys, 51(5):1–39, January 2019.
13. Paul Ekman. Emotions Revealed: Understanding Faces and Feelings. A Phoenix paperback. Phoenix, 2004.
14. Paul Ekman. What scientists who study emotion agree about. Perspectives on Psychological Science, 11(1):31–34, jan 2016.
15. Paul Ekman and Wallace V. Friesen. Measuring facial movement. Environmental Psychology and Nonverbal Behavior, 1(1):56–75, 1976.
16. Paul Ekman and Harriet Oster. Facial expressions of emotion. Annual Review of Psychology, 30(1):527–554, January 1979.
17. Paul Ekman and Erika L. Rosenberg. What the Face Reveals Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS). Oxford University Press, April 2005.
18. Joseph F. Grafsgaard, Joseph B. Wiggins, Kristy Elizabeth Boyer, Eric N. Wiebe, and James C. Lester. Automatically recognizing facial indicators of frustration: A learning-centric analysis. In 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction, pages 159–165, 2013.
19. Geoffrey Hookham and Keith Nesbitt. A systematic review of the definition and measurement of engagement in serious games. In Proceedings of the Australasian Computer Science Week Multiconference, ACSW 2019, New York, NY, USA, 2019. Association for Computing Machinery.
20. Yunxin Huang, Fei Chen, Shaohe Lv, and Xiaodong Wang. Facial expression recognition: A survey. Symmetry, 11(10), 2019.
21. Ashish Kapoor, Selene Mota, and Rosalind W. Picard. Towards a learning companion that recognizes affect. In Proceedings from Emotional and Intelligent II: The Tangled Knot of Social Cognition, AAAI Fall Symposium, number 543 in 1, pages 2–4, 2001.
22. Shan Li and Weihong Deng. Deep facial expression recognition: A survey. IEEE Transactions on Affective Computing, pages 1–1, 2020.
23. Bethany McDaniel, Sidney D’Mello, Brandon King, Patrick Chipman, Kristy Tapp, and Art Graesser. Facial features for affective state detection in learning environments. In Proceedings of the Annual Meeting of the Cognitive Science Society, 2007.
24. Jesus Mena-Chalco, Roberto Marcondes, and Luiz Velho. Banco de dados de faces 3d: Impa-face3d. Technical report, IMPA - VISGRAF Laboratory, nov 2008.
25. Lennart E. Nacke. Introduction to biometric measures for Games User Research. Oxford University Press, March 2018.
26. Omid Mohamad Nezami, Mark Dras, Len Hamey, Deborah Richards, Stephen Wan, and Cécile Paris. Automatic recognition of student engagement using deep learning and facial expression. In Machine Learning and Knowledge Discovery in Databases, pages 273–289. Springer International Publishing, 2020.
27. Noldus. Facial expression recognition software — FaceReader, June 2020.
28. Heather L. O’Brien, Paul Cairns, and Mark Hall. A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. International Journal of Human-Computer Studies, 112:28–39, April 2018.
29. Heather L. O’Brien and Elaine G. Toms. 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, oct 2009.

30. Raquel Robinson, Katelyn Wiley, Amir Rezaeivahdati, Madison Klarkowski, and Regan L. Mandryk. "let's get physiological, physiological!". In Proceedings of the Annual Symposium on Computer-Human Interaction in Play. ACM, November 2020.
31. David Sharek and Eric Wiebe. Measuring video game engagement through the cognitive and affective dimensions. Simulation & Gaming, 45(4-5):569–592, August 2014.
32. Sabrina Stöckli, Michael Schulte-Mecklenbeck, Stefan Borer, and Andrea C. Samson. Facial expression analysis with AFFDEX and FACET: A validation study. Behavior Research Methods, 50(4):1446–1460, dec 2017.
33. Jordi Vallverdú, editor. Handbook of Research on Synthesizing Human Emotion in Intelligent Systems and Robotics. IGI Global, 2015.
34. Yue Wu and Qiang Ji. Facial landmark detection: A literature survey. International Journal of Computer Vision, 127(2):115–142, May 2018.
35. Wilhelm Wundt. Emotions. Grundriss der Psychologie, 13, 1896.