source: https://doi.org/10.57694/7276 | downloaded: 4.5.2024

Facial and Bodily Expressions of Emotional Engagement: How Dynamic Measures Reflect the Use of Game Elements and Subjective Experience of Emotions and Effort

SIMON GREIPL, Ludwig-Maximilians-University of Munich, Germany
KATHARINA BERNECKER, University of Zurich, Switzerland
MANUEL NINAUS, University of Innsbruck, Austria

Users' emotional engagement in a task is important for performance and motivation. Non-intrusive, computerized process measures of engagement have the potential to provide fine-grained access to underlying affective states and processes. Thus, the current work brings together subjective measures (questionnaires) and objective process measures (facial expressions and head movements) of emotions to examine users' emotional engagement with respect to the absence or presence of game-elements. In particular, we randomly assigned 156 adult participants to either a spatial working memory task with or without game elements present, while their faces and head movements were recorded with a webcam during task execution. Positive and negative emotions were assessed before the task and twice during task execution using conventional questionnaires. We additionally examined whether perceived subjective effort, assumed to inherit a substantial affective component, manifests at a bodily expressive level alongside positive and negative emotions. Importantly, we explored the relationship between subjective and objective measures of emotions across the two tasks versions. We found a series of action units and head movements associated with the subjective experience of emotions as well as to subjective effort. Impacted by game elements, these associations often fit intuitively or lined up with findings from literature. As did a linear increase of blink (action unit 45) intensity relate to participants performing the task without game elements, presumably indicating disengagement in the more tedious task variant. On other occasions, associations between subjective and objective measures seemed indiscriminative or even contraindicated. Additionally, facial and bodily reactions and the resulting subjective-objective correspondences were rather consistent within, but not between the two task versions. Our work therefore both gains detailed access to automated emotion recognition and promotes its feasibility within research of game elements while highlighting the individuality and context dependency of emotional expressions.

CCS Concepts: • **Human-centered computing → Empirical studies in HCI; Empirical studies in interaction design**; • **Applied computing →** *Psychology*; Interactive learning environments.

KEYWORDS: game elements; design; emotions; affective computing; working memory; gamification; game-based assessment

ACM Reference format:

Author's addresses: S. Greipl, Dept. of Media and Communication, Ludwig-Maximilians-University of Munich, Oettingenstraße 67, 80538 München, simon.greipl@ifkw.lmu.de; Katharina Bernecker, Departement of Psychology, University of Zurich, Binzmühlestrasse 14, Box 6, CH-8050 Zürich, Switzerland; Manuel Ninaus, Department of Psychology, University of Innsbruck, Innrain, 52, 6020 Innsbruck, Austria.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
Copyright © ACM 2021 2573-0142/2021/9 – Art240...$15.00
https://doi.org/10.1145/3474667

Simon Greipl, Katharina Bernecker, and Manuel Ninaus. 2021. Facial and bodily expressions of emotional engagement: How dynamic measures reflect the use of game elements and subjective experience of emotions and effort. In *Proceedings of the ACM on Human-Computer Interaction,* Vol. 5, CHI PLAY, Article 240 (September 2021), 25 pages, https://doi.org/10.1145/3474667

1 INTRODUCTION

Many of today's attempts to deepening the understanding how individuals perform tasks are concerned with engagement [45,52,64]. Such endeavours usually follow the rationale that engagement is a proxy for an ideal state characterised by active involvement in a task as opposed to disinterest, absence or surface attendance [58]. Engagement seems to be predictive of several positive educational parameters, including better learning outcomes, grades and increased ambition to higher education [83,19,84].

Game-based environments and gamification seem to be promising ways to promote engagement. In other words, the literature on game-based learning and gamification suggests that adding game elements increases the enjoyability of a (cognitive) task that is otherwise considered tedious or boring [3,59]. In particular, game elements can positively influence performance and learning indirectly through individuals changes in attitudes and behaviors [44,43], motivation, and affect [31,70]. However, underlying mechanisms and processes of game elements have not been studied sufficiently yet [6]. Recent studies aim at utilizing process measures, such as physiological [e.g. 50] and neurophysiological data [e.g. 88], to better understand these mechanisms on a fine-grained level [for an overview see 57]. For instance, Ninaus et al. [59] utilizd facial emotion classification and were able to predict whether participants engaged in a math learning task with or without game elements on facial emotion data alone. Like previous studies, the current work utilizes process measures and assumes that adding game elements leads to an increase in (emotional) task engagement. The resulting increase in task engagement should manifest in detectable bodily or facial expression alterations. Previous works have proven the feasibility of such attempts [for an overview see 57]. By comparing two variants of a working memory task, one with and one without game elements, the current work aims at broadening our knowledge on how game elements and their effects on emotional user engagement relate to dynamic, objective and fine-grained expressive measures of affect such as facial (action units) and head activity features (head translations and rotations). First of all, we generally examine which of these facial and head activity features predominantly reflect the presence (or absence) of game elements during the interaction with the task. Secondly, we sharpen these correspondences by examining the relationship between facial/bodily features and self-reported reported affective states (i.e. positive and negative emotions) as well as the subjective feeling of effort. Finally, we explore whether these results show any consistencies on a meta-level, such as whether identified facial features likewise reflect positive/negative emotions and the presence/absence game elements.

2 RELATED WORK

2.1 Game elements & (emotional) engagement

Engagement usually describes the active involvement in a given task as opposed to a lack of interest, apathy, or superficial participation [58]. Ample research evidence suggests that engagement is related to better educational outcomes and thus is of primary importance in education and training [19,83,84].

Despite an ongoing debate about its exact underpinnings, engagement can be conceptualized to consist of the dimensions behavior, cognition and emotion [20]. In the current study, we focus on aspects of emotional engagement. That is, for instance, learners' emotional reactions to instructional material [see 1,33 for a more comprehensive definition]. Research suggests that a lack of emotional engagement can also lead to behavioral and cognitive disengagement [34,30,2], emphasizing the importance of emotional factors.

Within the last decade, studies capturing any kind of engagement have become more and more digitally enhanced and technologically supported [71]. Studies using computer vision and machine learning techniques in the domain of affective computing showed promising results to differentiate between different emotional states and levels of engagement [49,59,74]. Automatic analysis of, for instance, images or videos of users faces during task interaction to estimate operator attention [67], emotion [46] or even engagement in the classroom [for a review see 25] has become a popular and efficient means.

What is sometimes referred to as advanced, analytic, and automated (AAA) approaches utilize heavily computerized examination of relevant, often affective learning states [12]. These AAA approaches include, amongst others, video analyses of facial expressions and bodily postures. In this context, facial expressions and action units are often used to classify discrete emotions such as happiness and sadness. For instance, Ninaus et al. [59] compared a game-based and a non-game-based math learning task directly to evaluate the affective impact on the player using discrete emotions and machine learning techniques. They found that game elements can exert an affective effect that also entails an expressive component measurable by automatic recognition of facial expressions and standard questionnaires. They concluded that adding game elements to a task leads to more emotional engagement in both positive and negative directions.

2.2 Game elements and subjective effort

Sustaining cognitive performance on a task is ubiquitous and critical in everyday life [38]. Subjective effort and its phenomenological qualities unify aspects of all three different levels of engagement: cognitive, behavioural and emotional. Importantly though, investing effort into a given task is often discussed on a cognitive or behavioural level [12,20].

Actual and subjective effort may often be congruent regarding their direction [see 77]. However, as this is not necessarily the case, subjective and actual effort are conceptually inequivalent [76,77]. An individual's perception of its capacity seems to determine the amount of effort that can be exerted in the first place [77]. Assuming that an individual's capacity may be grounded in his or her motivations [75], task engagement, and its accompanying motivational upheaval, should lead to an increase in the expended effort. Game elements potentially make a task more rewarding or enjoyable. As such, there is evidence of heightened objective effort expenditure in (cognitive) tasks with game elements present [55,61]. It is suggested that game elements might influence the willingness to invest (cognitive) effort as they alter the learner's perception of the task at hand [60,73]. That is, tasks with game elements can be perceived as less effortful or less strenuous than tasks without game elements present [3]. This argument relies on the assumption that cognitive resource allocation is flexible and can be influenced by ones motivational and emotional state [62,65,79]. This potential link between subjective experience of effort and emotions might further suggest that subjective effort might also be reflected in users' facial or even bodily features during task execution.

2.3 Face & head features and affective states:

Table 1: Action Units used in the current study (adapted from FACS)

AU			AU		
1	Inner Brow Raiser		14	Dimpler	
2	Outer Brow Raiser		15	Lip Corner Depressor	
4	Brow Lowerer		17	Chin Raiser	
5	Upper Lid Raiser		20	Lip stretcher	
6	Cheek Raiser		23	Lip Tightener	
7	Lid Tightener		25	Lips part	
9	Nose Wrinkler		26	Jaw Drop	
10	Upper Lip Raiser		28	Lip Suck	
11*	Nasolabial Deepener		45	Blink	
12	Lip Corner Puller				

*not included in the data collection

2.3.1 Mapping action units onto emotions. Interest in the reliable recognition of expressive facial features and the interpretation of their configuration continues unabated [46,51]. This is particularly interesting when investigating player experience, as players are not disturbed by, for instance, responding to a questionnaire during gaming, which might distract the players or even disrupt rather volatile user states such as flow or presence (e.g. Nebel & Ninaus, 2019). A substantial body of work used the facial action coding system (FACS) – a taxonomy of human facial 57 expressions. Originally developed in 1978 (Ekman & Friesen, 1978; revision: Ekman, 2002), the system specifies 32 distinct facial muscle actions called Action Units (AUs; see Table 1 for an overview of AUs). Due to their fine-grained partition into facial features, action units are still considered state-of-the-art for a detailed facial expression analysis.

Traditionally, those small units were combined and related to larger and compound concepts of emotional expressions (e.g. discrete emotions). Basic emotions can in this system be represented as a combination of action units [80] (see Table 2). Happiness, for instance, is frequently represented as a combination of cheek raiser (AU6) and lip corner puller (AU12) [51]. Besides this widely applied approach, other works established associations between single AUs and basic emotions. The current study similarly relates single action units to a negative and a positive continuum of emotion. For instance, using happy again as an example, this emotion has cheek raiser (AU 6), lid tightener (AU 7), lip corner puller (AU 12) and jaw drop (AU 26) as positive associations but inner and outer brow raiser (AU 1, AU 2), upper lid raiser (AU 5) and nose wrinkle (AU 9) as negative associations [81].

Another popular approach utilizes the two-dimensional circumplex model of affect [69] that arranges emotional experiences within the spectrum of arousal and valence. Accordingly, one can map discrete positive and negative emotions onto the valence dimension. That is, happy and surprise are, for instance, usually located in a positive valence space [48]. A more direct inclination of the valence dimension and action units has shown the discriminative power of single action units [54]. As such, lip corner puller (AU12) seems to be highly indicative for positive valence, whereas nasolabial deepener (AU 11) is strongly indicative for negative valence (see also Table 2).

Table 2: Overview of Action Units and their association to basic emotions

[adapted from 80]	AUs	[adapted from 51]
{4,5,7,10,22,23,25∨26} {4,5,7,10,23,25∨26} {4,5,7,17,23∨24} {4,5,7,23∨24} {4,5∨7} {17,24}	anger	4, 5, 7, 10, 17, 22-26
{9∨10,17} {9∨10,16,25∨26} {9∨10}	disgust	9, 10, 16, 17, 25, 26
{1,2,4} {1,2,4,5,20,25∨26∨27} {1,2,4,5,25∨26∨27} {1,2,4,5} {1,2,5,25∨26∨27} 5,20,25∨26∨27} {5,20} {20}	fear	1,2,4,5,20,25,26,27
{12} {6,12}	happiness	6, 12, 25
{1,4} {1,4,11∨15} {1,4,15,17} {6,15} {11,17} {1}	sadness	1,4,6,11,15,17
{1,2,5,26∨27} {1,2,5} {1,2,26∨27} {5,26∨27}	surprise	1,2,5,26,27
-	pain	4,6,7,9,10,12,20,25,26,27,43
	cluelessness	1,2,5,15,17,22

Additional information stems from attempts to map action units onto more complex emotions than the basic ones. Researcher examined so-called learner-centered affective states such as boredom [10,56], frustration [41,see also 13] and confusion [41,26], using facial expressions. McDaniel and colleagues manually coded action units of facial video recordings from participants interacting with the system AutoTutor [53]. They identified several action units that, when present (+) or absent (-), distinguished emotions from neutral states. Confusion, for instance, was indicated by brow lowerer [AU4(+)], lid tightener [AU7(+)] and lip corner puller [AU12(-)]. Frustration was indicated only by lip corner puller [AU12(+)]. Another work correlated task difficulty with activation of action units, and revealed that only lower blink rate (AU45) was consistently indicative [86]. An approach of automated frustration and confusion classification based on action units likewise mainly relied on blink (AU45) [4]. Brow lowerer (AU4) or the *corrugator* is the only one mentioned to be linked to mental effort, prominently termed as "the muscle of concentration" by Darwin [47].

2.3.2 Mapping head movements: Examining "bodily expressions" [single body parts or the whole body [82]] as the expressive component of emotions has a substantial body of research that, however, primarily focused on controlled expressions from actors in specific situations [11,32]. However, literature's focus is very limited with respect to effort, which is why we will focus on related emotional expressions such as boredom. For instance, it has been shown that increased upper body shifts are associated with boredom. In contrast, positive emotions, like joy and happiness, are often expressed by an elevated upper body posture [32] or a head leaning upwards [11,82]. In a different study, increased head activity was related to frustration during a computer-supported tutoring lesson [27]. Moreover, positioning the upper body closer to the screen was also indicative of self-reported frustration. Accordingly, bodily expressions might provide information about a users' disposition to engage in a particular task [22,23]. More specifically related to gaming, differently engaging tasks (i.e., a "boring" vs. "interesting" reading task vs. playing a shooting game on a computer) lead to different head activity with highest activity in the "boring reading task" condition [87]. In contrast, van den Hoogen et al. [35] showed that boredom can also lead to decreased upper body activity. Riemer et al. [68] recently used a Kinect sensor to capture head and upper body movements from participants interacting with a computer platform game. Positioning the upper body towards the gaming screen was associated with self-reported frustrations, just like keeping the head turned to the right (while keeping it turned to the left indicated enjoyment). Boredom was most significantly indicated by increased head rolling activity (tilting to the left or right shoulder). Additionally, keeping the head turned to the right was a sign of frustration, keeping it turned to the left however was linked to enjoyment [see also 4].

2.4 The current study

We augmented a classic working memory paradigm, the spatial n-back task [39], with game elements to evaluate their effect on processes of engagement. In particular, we compare two digital versions of the n-back task that are equivalent as regards their basic task characteristics. However, one version utilized various game elements, e.g. a narrative with matching visual aesthetics as well as enhanced feedback. Importantly, game elements did not change the basic mechanics of the spatial n-back task so that maximal comparability with the version without game elements was given.

In contrast to most previous work, we perform a fine-grained analysis of the influence of game elements on expressive (emotional) facial and bodily features. Because previous research indicated that certain facial and bodily features can act as objective proxies for subjective (emotional) experiences, we aim to explore the potential of these dynamic measures to study the impact of game elements on (emotional) engagement. All measures used are part of a larger project from which results of questionnaire ratings regarding emotions and subjective effort have been already published elsewhere [3]. Drawing upon this previous analysis, it revealed that positive emotions decreased over time in the working memory task without game elements (no-game condition), but remained at baseline levels in the working memory task with game elements present (game condition). In this latter condition, participants had a higher baseline level of negative emotions but experienced fewer negative emotions than the non-game condition afterwards. For subjective effort, participants in the game condition experienced the task as less effortful throughout the experiment. No performance differences were found between participants in the game or no-game condition. The current work extends this data by linking it to objective facial and body movement features such as action units and head movement and rotations.

Action units are usually examined using rough intensity change estimates [see 47 for another use of de-/increase measures of AUs] or mere frequency [29]. In the present work, we take a different approach using two process parameters describing AU changes across conditions and blocks of the experiment. One describes the (latent) intensity de- or increase over time. The second is a variability parameter, indicating fluctuations over time. Both measures overcome the threshold problem of counting significant activation of action units, meaning that they are only considered if a human judge or a machine classifies them as significant. These process measures should allow us highly detailed access to the properties of single AUs over the course of the experiment.
Further, we establish a relation between different movement and orientation parameters of the head and game elements as well as emotional engagement. Variants of the same movement and orientation parameters have been proven useful in the past [68]. In particular, we try to establish indicators of head movements such as range, velocity and stability, and link them to the presence or absence of game elements as well as participants' subjective (emotional) experience.

In summary, the current work focuses on two research questions (RQs). Assuming an underlying affective component, can facial and head activity features reflect the use and impact of game elements in a cognitive task (RQ1)? Can subjective effort, alongside positive and negative emotions, be meaningfully mapped onto objective, expressive components of emotions such as facial and head activity features (RQ2)? Beyond these research questions, we will also consider and discuss potential overarching patterns identified by our analyses. This will shed light on the stability of the relationship between objective, expressive components of emotions and established subjective measures over the course of the experiment (e.g., are subjective negative ratings accompanied by similar negative expressions across time and experimental conditions?).

3 Method

The current study is part of a larger project investigating the effects of game elements in cognitive training. As such, part of the data acquired in the project have been published elsewhere [3]. However, data, analyses, and results presented in the current study have not been published and focus on facial and head movement data.

3.1 Design

The study was designed with one between-subjects factor (game vs. non-game task condition). Participants in both conditions worked on spatial 2-back tasks. In the game condition, the 2-back task was augmented with typical game elements, i.e. game-like visual design, game narrative, individual score, progress bar, and streaks (extra points awarded for 5 correct responses in a row). In the non-game condition, these game elements were absent in the 2-back task.

3.2 Participants and procedure

In total, 190 adult participants were recruited (42 male, 147 female, 1 na, mean age= 22.47 years, SD age = 2.46) via the research institutions participant tool in a city in Germany, which contains mainly university students. All participants were reimbursed participating for participating in the study with 8 EUR, which took about 1 hour. Unfortunately, complications (image quality or face recognition problems) reduced the sample of participants to a total of 156, equally split to both conditions (N=78 each).

After providing informed consent, participants were randomly assigned to one of the two conditions (game vs. non-game task version) and received the corresponding instruction for the spatial 2-back task. First, participants completed 50 practice trials, where all participants received trial-based feedback on whether their response was correct or incorrect. This should ensure that every participant understood the objective of the spatial 2-back task. Afterwards, participants worked in two blocks, consisting of 120 trials each. After each of the two spatial 2-back task blocks, participants' positive and negative affect as well as subjective experience of effort were assessed with questionnaires. During the spatial 2-back task, participants faces were recorded with a conventional webcam. For more details on the procedure unrelated to the current study see [3].

3.3 Measures and materials

3.3.1. Game and non-game spatial 2-back task.

Non-game task: The spatial 2-back task was realized using PsychoPy [63]. During each trial of the task, a blue square was randomly presented in one of four locations (i.e., grey squares) for precisely 900 ms (see Figure 1 left). Each of the four locations was assigned to a key on a QWERTZ keyboard (i.e., Q, A, P, L). After stimulus presentation, participants had 900 ms to press the corresponding key on the keyboard to indicate the location of the stimulus presented 2 trials ago to the current trial. That is, for the first two trials of each block no response was possible or necessary. A new stimulus would appear after 500 ms independent of whether a participant pressed a key or not within the 900 ms time window. Accordingly, the interval between any two stimuli was 1400 ms.

Game task: In the game version of the spatial 2-back task entitled "Brains vs. Zombies", several game elements augmented the conventional/non-game task version. A game narrative was introduced before the instructions and stated that a cemetery in the city of the study is haunted and zombies appear on every gravestone. The zombies need to be eliminated before they reach the city and can only be eliminated with a remote-controlled weapon. The weapon must be aimed at the exact position of the zombies, but they are very fast and visible only for a brief period of time. After this short game narrative statement, the identical instructions as in the non-game version of the task were presented to the participants. The only exception was that a graveyard design was presented (see Figure 1 right). Moreover, instead of a blue square, zombies were used as stimuli to further enhance the credibility of the narrative. Finally, in the game version of the task we implemented and individual score (i.e. 10 points for each correct response), a progress bar showing the progress within each block in the form of a brain, and performance streaks (i.e. 10 extra points were awarded if participants gave 5 correct responses in a row).

In [3] the number of correct responses served as dependent measure. The authors did not find a difference in number of correct responses between conditions, but participants increased their performance from practice to first block and second block of the spatial 2-back task in both conditions.

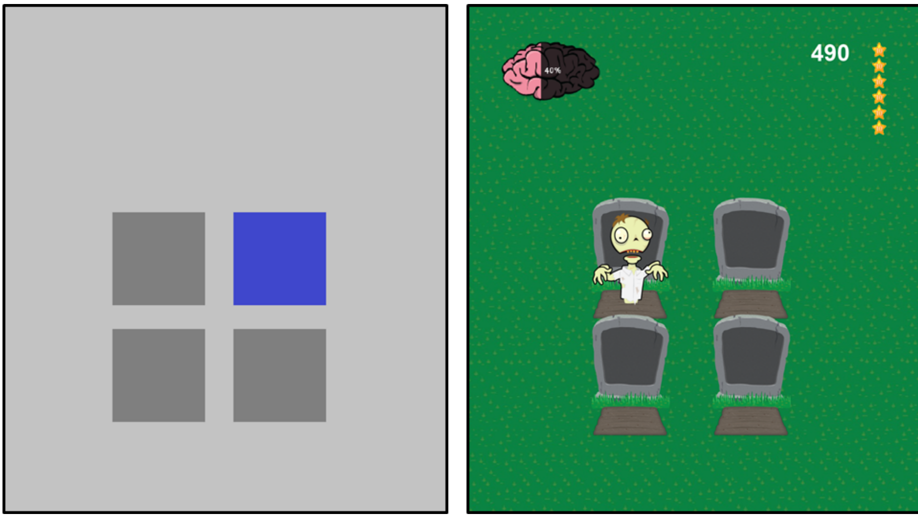


Fig. 1: Non-game (left) and game (right) version of the spatial 2-back task

3.3.2 Facial expressions. Video data were analysed using OpenFace 2.0 (Baltrusaitis et al., 2018). We obtained intensity values for a total of 18 action Units (see Table 1) provided by the system. Importantly, for every frame of the video, an intensity value (range from 0 to 5) is provided by the software, indicating to which extent an action unit was present in a single frame. We used linear regression models for every AU and subject with time (frames of the video) as the dependent variable to obtain slopes and residuals out of these separate models. Positive and negative slopes describe an increase or a decrease of the intensity of the AU over time, respectively. Residual values, on the other hand, describe the amount of deviation from the linear function and, therefore, the intensity's variability of a given AU across time, or, in the present case, across a block of the experiment.

3.3.3 Head movement. Estimation of head pose was obtained using the same software and returned process values on two different levels, translation, and rotation (for detailed information see https://github.com/TadasBaltrusaitis/OpenFace/wiki/Output-Format).

Translation: Head translation refers to the heads motions in space (visible by the camera) along the three axes (X,Y,Z), while the camera represents the origin of the space. For instance, alterations along the Z-axis describe movements *away* or *towards* camera (distal, the camera is the point of origin), Y indicates the head moving up or down (vertical) and X describes movements to the left or right (horizontal).

Rotation: We obtained head rotations (also known as pitch, yaw and roll) around the three axes (X,Y,Z; relative to the camera as point zero). For instance, X means *nodding* the head to look up or downwards, Y is *turning* the head to the left or to the right and Z describes *rolling* (or tilting) the head left or right towards the shoulders.

For both rotation and translation parameters, we obtained range, velocity and stability metrics for every subject n. Range (d_n) refers to the sum of absolute differences of head movements for frames k:

$$d_n = \sum_{i=1}^{k} |x_{i+1} - x_i|$$

Velocity was calculated as the average movement speed (millimeters for translation and radiants for rotation) per second per subject (v_n) as follows:

$$v_n = \frac{\sum_{i=1}^{k} \sqrt{\left(\frac{x_{i+1} - x_i}{t}\right)^2}}{k}$$

with t=1/30 being the timespan of a single frame in seconds and k the number of frames. To calculate stability we used the homonymous function from the *tsfeatures* package in R [36] which calculates variance of means within user defined timeframes, which were set to 30 seconds in the current analysis. Higher values of the parameters range and velocity, but lower values of the parameter stability would thus refer to more movement and therefore more activity (vice versa for less movement and thus less activity).

3.3.4 Subjective Emotions. Positive and negative emotions were assessed using the German short version of the Positive and Negative Affect Schedule [PANAS; 42,85]. The scale consists of 6 items for positive affect (PA; e.g. "interested", "excited") and 6 items for negative affect (NA; e.g. "upset", "irritable"). Participants rated each item to what extend they felt the way described on a scale from 0-100 (not at all – very much). PA and NA were assessed before the spatial 2-back task as a baseline and after the first and second block of the task to assess whether affect changed over time and between task versions.

3.3.5 Subjective Effort. Subjective task effort was assessed using three items (i.e. "How effortful/difficult/strenuous has the task been so far?"[adapted from 40]. We again used a scale from 0-100 (not at all – very much) to assess how much participants agreed with each item.

3.4 Analysis

Data analysis was carried out using R version 3.6.3 [66]. We calculated batches of logistic as well as linear regression models to predict either condition (game/non-game, logistic regression, RQ1) or affective states (pos/neg emotions and subjective effort, linear regression, RQ2).

To address RQ1, two batches of models were established where condition (game/non-game) was predicted by either action units or head pose features in separate logistic elastic net regression models (see below for further explanation on the regression technique). In the first batch with facial features as the predictor, all 18 action units were simultaneously entered in one model. However, models were separately calculated for the two parameters of action units, slopes and residuals, as well as block 1 and 2, totaling in 4 models for this batch.

The second batch addressed head pose features predicting condition. The actual predictors in this model batch were the axis values X,Y, and Z. Again, separate models were calculated for two head pose geometry parameters, movement and orientation and, like before, separately for blocks 1 and 2. In addition however, we also calculated separate models for each of the three head pose metrics, range, velocity and stability, totaling in 12 regression models for this batch.

To address RQ2, two further batches of models were calculated where one of the three subjective outcomes (pos./neg. emotions, subjective effort) was implemented as the dependent variable. This third and fourth batch is similar to previous batches, as slopes and residuals of action units as well as orientation and movement of head poses were again used as predictors in separate models. Unlike the previous batches, not only separate models for blocks as well as head pose metrics (range, velocity and stability, batch 4 only) were calculated, but also separate models for both conditions (game/non-game). Because our three subjective measures (i.e., PA, NA, and effort) of affect also required separate models, batch 3 comprised 24 models regarding the prediction of affective states by means of action units. Batch 4, in which affective states were predicted by head pose features, contained 72 models, resulting in a total of 112 regression models in our analysis (see https://osf.io/jz9sr/ for a tabular overview of models calculated).

Because there are 18 action units as predictors in our model, they were considered high dimensional and therefore potentially suffering from intercorrelations. To prevent such shortcomings and to more clearly identify relevant variables and possibly reducing the dimensionality, we calculated penalized logistic elastic net regressions for the models examining action units. This hybrid between ridge and lasso regression combines advantages of both L1 and L2 regularization by determining optimal alpha and lambda parameters. Lambda was evaluated using a cross-fold technique from R's Glmnet package [21]. Alpha was iteratively determined by repeating the procedure for a coarse sequence of possible alpha values between 0 and 1. The final model was selected using the lowest AIC value, indicating the best model fit. Estimation and hypothesis testing was done using induced smoothing as implemented by the *islasso* package [9].

4 RESULTS

4.1 Facial action units and head movements reflecting game-elements (RQ1)

4.1.1 Action units & game/non-game task. To investigate whether the use of game elements in a cognitive task has a measurable impact on (emotionally) expressive facial features, we analyzed individual slopes of facial AU intensity (i.e. increase or a decrease of the intensity of the AU over time) as well as the residual values from the individual linear function (i.e. intensity's variability of a given AU across time; for more details see also 3.3.2 and 3.4). Concerning the slopes, a negative/positive beta value indicates that an increasing intensity over a block is predictive of the lower/higher conditional value, i.e. the non-game/game condition. Increasing slopes of outer brow raiser (AU02), block 2 only), lid tightener (AU07, block 1 only) and blink (AU45, block 1 only) were predictive of the non-game condition. This means, for instance, that eye blink (AU 45) intensity increased over the course of a block which might either be caused by an increase in blink rate or the eyelid opening decreased over the course of the experiment predominantly in the non-game condition. Only increasing slopes of upper lid raiser (AU05) were predictive of the game condition (block 1 only).

Higher variability (see Table 3, residuals) in nose wrinkler (AU09, block 2 only) was indicative for the non-game condition. In contrast, higher variability in lips part (AU25) was significantly associated with the game condition (see Table 3 for statistical details). No other AUs were significant (all ps>.06). Overall, we found slopes of four AUs and residuals (variability) of two action units to be indicative for game or non-game condition. Further, in most cases effects were significant for the first block.

Table 3: action units and condition (game = 1, non-game = 0, Df[a] = 155)

		Estimate	Std. Error	Df	z value	p value	Estimate	Std. Error	Df	z value	p value
	Block	*Slopes*					*Residuals*				
AU02 	1										
	2	-11507.33	5881.37	0.73	-1.96	.050					
AU05 	1	122271.15	30385.78	0.96	4.02	<.001					
	2										
AU07 	1	-23703.62	9685.96	0.96	-2.45	.014					
	2										
AU09 	1										
	2						-6.33	3.23	0.83	-1.96	.050
AU25 	1						1.80	0.92	0.74	1.97	.049
	2										
AU45 	1	-52306.78	13973.85	0.96	-3.74	<.001					
	2										

a Degree of freedom (null model, constant)
b values of the predictor variable slopes were extremely low (range ~-0.0001 to 0.0001), leading to very high beta estimates associated to unit changes in slopes.

4.1.2 Head movement and game/non-game task. To investigate whether the use of game elements in a cognitive task has a measurable impact on head movement features, we analyzed range, velocity and stability of head translations and rotations along and around the three axes X, Yand Z, respectively. All significant predictors occurred exclusively in block 2.

Translation: The only predictive variable in this measure was vertical head movement stability around (Y) in block 2. Higher stability was associated with the presence of game elements (see Table 4 for statistical details).

Rotation: As in the analysis before, positive beta values indicate a positive increase significantly related to the presence of game elements, whereas negative values indicate an increase significantly associated with the absence of game elements. With respect to range, higher levels of head turning (rotations around Y) significantly predicted the game condition, whereas higher head tilting (rotations around Z) was related to the non-game condition (both block 2). The same pattern occurred with respect to velocity: head turning significantly predicted the game condition, whereas higher head tilting was related to the non-game condition. Additionally, stability of head rotations (Z) was associated with the game condition. All other head features were non-significant (ps>.05). In sum, also head movements were quite predictive of game vs. non-game condition, especially head turning was related to game condition in the second block and tilting to the nongame condition also in the second block.

Table 4: head movements and condition (game = 1, non-game = 0, Dfa =155/152)

			Estimate	95% CI	z	p value	Estimate	95% CI	z	p value
Parameters	*Axis*	*Block*	*Translation*				*Rotation*			
Range										
	Y	1								
		2					0.15	[0.04, 0.31]	2.25	.025
	Z	1								
		2					-0.21	[-0.38, -0.08]	-2.84	.005
Velocity										
	Y	1								
		2					40.41	[9.83, 81.68]	2.23	.026
	Z	1								
		2					-56.85	[-100.48, -21.49]	-2.83	.005
Stability										
	Y	1								
		2	1.31	[0.08, 2.59]	2.05	.040				
	Z	1								
		2					1.43	[0.16, 2.76]	2.16	.031

[a] Degrees of freedom (null model, constant/residual model, respectively)

4.2 Mapping facial action units & head movements onto subjective scales of emotions (RQ2)

To investigate whether subjective ratings of emotions and effort align with certain features of facial expressions and head movements, we ran similar analyses as above. However, instead of using experimental group membership (condition) as the dependent variable, we used subjective ratings of positive and negative emotions and subjective effort and analyzed conditions separately.

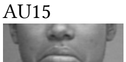
4.2.1 Action units, subjective emotions and subjective effort

Positive emotions: Only for dimpler (AU 14) in non-game condition an intensity increase (positive beta estimate of slope) during the second block was present (see Table 5). Other AUs were not significant (all ps>.07).

Negative emotions: In contrast, increasing dimpler (AU14) variability in the non-game condition was associated with negative emotions in the second block. In block 2 and in the non-game condition, nose wrinkle (AU09), and lip stretcher (AU20) had also positively associated variability, while lip stretcher (AU20) variability was also positively associated with negative emotions in block 1. In the game condition, only upper lip raiser (AU10) had a positive association to negative emotions and only in the first block (see also Table 5).

No other AUs were significant (all ps>.09). In sum, residuals (variability in AUs) was associated to subjective ratings of negative emotions predominantly in the non-game condition.

Table 5: Action Units and subjective scales (Df[a] = 77)

	Block	Parameter	Estimate	Std. Error	Df	z	p value	Estimate[b]	Std. Error	Df	z	p value
			Game					Non-game				
Action Units and positive emotions												
AU14 	2	Slopes						60949.80	28343.22	0.50	2.15	.032
		Residuals										
Action Units and negative emotions												
AU09 	2	Slopes										
		Residuals						18.11	9.14	0.01	1.98	.047
AU10 	1	Slopes										
		Residuals	34.48	4.79	0.05	7.20	<.001					
AU14 	2	Slopes										
		Residuals						8.83	4.24	0.01	2.08	.037
AU20 	1	Slopes										
		Residuals						43.31	15.39	0.58	2.81	.005
	2	Slopes										
		Residuals						25.72	7.69	0.01	3.34	.001
Action Units and subjective effort												
AU05 	1	Slopes										
		Residuals	-31.95	9.28	0.02	-3.45	.001					
	2	Slopes										
		Residuals	-27.20	9.02	0.06	-3.02	.003					
AU06 	1	Slopes										
		Residuals	26.36	5.57	0.02	4.74	<.001	58.35	7.12	0.007	8.20	<.001
AU10 	1	Slopes										
		Residuals	30.26	4.76	0.02	6.35	<.001					
AU12 	1	Slopes										
		Residuals	16.54	5.30	0.02	3.12	.002					
AU14 	1	Slopes										
		Residuals	10.13	3.86	0.02	2.63	.009					
	2	Slopes										
		Residuals	19.99	3.86	0.06	5.23	<.001					
AU15 	1	Slopes										
		Residuals						12.66	5.39	0.010	2.35	.019

[a] Degrees of freedom (null model, constant)
[b] values of the predictor variable slopes were extremely low (range ~-0.0001 to 0.0001), leading to very high beta estimates associated to unit changes in slopes

Subjective effort: No slopes were significantly related to subjective effort across conditions. With respect to residuals in the game condition, upper lid raiser (AU05) showed a negative association in both blocks. Further, cheek raiser (AU06, block 1 only), upper lid raiser (AU10, block1 only), lip corner puller (AU12 ,block 1 only) and dimpler (AU14, both blocks) were significantly positively predictive of subjective effort. In the non-game condition, only cheek raiser (AU06, block 1 only) and lip corner depressor (AU15, block 1 only) were positively associated with subjective effort (see Table 5). None of the other AUs were significant (all ps>.13). Especially residuals (variability in AUs) was correlated with effort ratings in the game condition but not in the non-game condition.

4.2.2 Head movements, emotions and subjective effort

Positive emotions: All significant relationships were established in the game condition (see Table 6). Increased nodding range and velocity (X) positively predicted positive emotions in the game condition (block 2). Lower stability was also significantly related to positive emotions with respect to horizontal movements (X, block 1), nodding (up/down rotations, X, block 1) as well as rolling (tilting left/right, rotation around Z, block 2). None of the other head features were significant (all ps>.08).

In sum, measures with respect to increased activity and lower stability indicate positive emotions, however in the game-condition only.

Negative emotions: In the game condition, increased vertical head movements showed a significant positive relation to negative emotions with respect to range and velocity in block 1. However, in block 2, these relationships were also significant but reversed: decreased vertical head movements showed a significant positive relation to negative emotions with respect to range and velocity. Stability of head nodding was significantly positively related to negative emotions. In the non-game condition, the same positive relationships between increased vertical head movements and negative emotions as in the game-condition were established for block 1, but not for block 2, i.e., block 2 showed no significant relationships at all.

Further, increased nodding velocity and range (both block 1) were positively related to negative emotions. None of the other head features were significant (all ps>.06). In both conditions, increased vertical movements covaried significantly with negative emotions in block 1. In contrast, in the game condition in block 2, negative emotions were indicated by decreased vertical movements. Interestingly, increased nodding activity further indicated higher negative emotions in the non-game condition but increased nodding stability related to negative emotions in the game-condition.

Subjective effort: In the game condition, range and velocity of vertical head movements showed a significant positive, while distal movements (along the Z-axis) showed a significant negative relationship with subjective effort (block 2 only). Lastly, a positive association betweeen horizontal head movement stability and subjective effort was significant in block 1. In the non-game condition, horizontal head movement with respect to range and velocity were significantly negatively related to subjective effort (both block 2 only). Additionally, head movement around Z was significantly positively related to subjective effort (block 2), whereas head rotation stability around Z showed a significant negative relationship to subjective effort (block 2, see Table 6 for statistical details). None of the other head features were significant (all ps>.09). That is, in the gaming condition, both increased (e.g. vertical) and decreased (e.g. distal) head movements were associated with subjective effort. In the non-game condition, decreased horizontal head movements indicated higher subjective effort. Additionally, higher distal movement but lower tilting stability indicated higher subjective effort.

Table 6: head movements and subjective scales

			Estimate	**95% CI**	**t(74)**	**p value**	**Estimate**	**95% CI**	**t(74)**	**p value**
Parameters	*Axis*	*Block*	*Game*				*Non-game*			
Positive emotions										
Range										
rotation	X	1								
		2	0.32	[0.03, 0.61]	2.21	.030				
Velocity										
rotation	X	1								
		2	87.23	[8.18, 166.29]	2.20	.031				
Stability										
translation	X	1	-20.50	[-37.11, -3.89]	-2.46	.016				
		2								
rotation	X	1	-19.49	[-35.66, -3.31]	-2.40	.019				
		2								
	Z	1								
		2	-19.40	[-34.44, -4.35]	-2.57	.012				
Negative emotions										
Range										
translation	Y	1	0.01	[0.00, 0.01]	2.61	.011	0.01	[0.00, 0.02]	2.04	.045
		2	0.00	[-0.01, 0.00]	-2.36	.021				
rotation	X	1					0.63	[0.08, 1.17]	2.30	.025
		2								
Velocity										
translation	Y	1	2.82	[0.55, 5.09]	2.48	.015	3.68	[0.04, 7.33]	2.01	.048
		2	-1.28	[-2.40, -0.17]	-2.29	.025				
rotation	X	1					255.06	[36.23, 473.89]	2.32	.023
		2								
Stability										
rotation	X	1								
		2	13.74	[0.76, 26.72]	2.11	.038				
Subjective effort										
Range										
translation	X	1								
		2					-0.02	[-0.03, 0.00]	-2.01	.048
	Y	1								
		2	1.81	[0.65, 2.97]	3.11	.003				
	Z	1								
		2	-2.54	[-4.34, -0.74]	-2.81	.006				
Velocity										
translation	X	1								
		2					-4.42	[-8.78, -0.06]	-2.02	.047
	Y	1								
		2	479.34	[170.10, 788.58]	3.09	.003				
		1								
	Z	2	-670.12	[-1,148.16, -192.07]	-2.79	.007				
Stability										
translation	X	1	20.47	[2.72, 38.22]	2.30	.024				
		2								
	Z	1								
		2					26.95	[5.52, 48.38]	2.51	.014
rotation	Z	1								
		2					-28.08	[-52.15, -4.01]	-2.32	.023

5 DISCUSSION

The current study showed that certain facial and head movement features might be indicative of participants being either engaged in a cognitive task with or without game elements present. Further, we found a range of facial and head movement features to be related to subjectively reported positive emotion, negative emotion, and effort. In the following, we will discuss the results in greater detail.

5.1 From body features to game elements (RQ1)

On a general level, we found that at least some facial action units helped to disambiguate between participants engaged in the game or non-game version of the spatial working memory task. We found intensity of upper lid raiser (AU5) and variability of lips part (AU25) to be indicative for the game condition – AUs that reflect both negative and positive valence [54]. Both, however, are also characterized by high arousal [5]. AUs associated with the task without game elements tend to range in the negative to non-discriminative spectrum (AU02,07,09,45). Outer brow raiser (AU02) and lid tightener (AU07) have only small discriminative validity in terms of valence [54]. But nose wrinkle (AU09) is indicating negative valence [48]. Blink (AU45) strongly correlates with boredom [53] and is considered to be an indicator for confusion and frustration [4]. In the non-game condition blink (AU45) intensity increased over the course of the first block of the experiment, which might indicate fatigue or boredom.

While the pattern of results of action units is not straightforward, with high arousal positive as well as negative AUs associated with the game condition, we assume that game elements facilitated a more emotionally engaging experience overall – in a positive as well as negative direction. Even though these results have to be treated with caution, they seem to be in line with previous research of increased emotional engagement when using game elements in cognitive tasks [59].

The non-game condition seems to be primarily associated with negative or neutral AUs (AU02,07,09,45) and indicators of boredom (i.e. increased blinks). In line with this, we found that higher stability in head rolling (leaning the head left/right) related to the game condition, whereas increased velocity and range of head rolling indicates the non-game condition. This seems to corroborate findings from literature in which such heightened activity was related to the experience of increasing boredom [68], which might one assume was more prevalent in the non-game condition. While rotations seemed to be more indicative than movements in general, stability might be more consistently related to the game condition than other variables. This may generally indicate that head rolling stability might be associated with the user's overall state of engagement.

5.2 Mapping objective to subjective scales (RQ2)

Positive emotions: Only one action unit, the dimpler (AU14), in the non-game condition (slope in block 2) was positively related to positive emotions. Literature, in this case, is mixed. AU14 is sometimes linked to negative emotions [29] but also to positive pleasure [5]. Considering that subjective positive emotions were more prevalent in the game condition ([3], see also section "Present study"), current results seem to be not in line with subjective reports. It might be the case that positive emotions were simply not pronounced enough to trigger emotionally expressive behavior in participants faces. Subjective reports indicated ([3], see also section "Present study") that participants in the game condition maintained the level of positive emotions throughout the experiment. In contrast, participants' ratings of positive emotions dropped over the course of the experiment in the non-game condition. This might also be the reason for the absence of explanatory AUs with respect to ratings of positive emotions in the game condition, as no linear change was detectable.

In turn, a series of head pose parameters was associated with the prevalence of positive emotions. Two patterns emerge here. First and most apparently, none of our head pose features were able to explain differential positive emotional ratings in the non-game condition, possibly reflecting the overall decline in positive emotions in this condition. Second, heightened nodding velocity and range as well as lower nodding stability were linked to positive emotions in the game condition. This suggests that head activity associated with positive emotions seemed to increased rather than decreased (at least around the X-axis). This outcome partly opposes results from literature, in which head and upper body activity is usually associated with negative states like boredom [27,32,35]. However, van den Hoogen et al. [35] also found increased head activity to be related to enjoyment.

Negative emotions: In the game condition, higher variability of upper lid raiser (AU10) indicated higher negative emotions (block 1 only). AU10 is usually discriminative for negative valence in emotions [54], suggesting a linear overlap with our results, but also correlates with positive discrete emotions like happiness [81]. Participants in the non-game condition showed increased variability in AU09, AU14, and AU20, which are usually indicative of negative valence [28,48]. Accordingly, as regards the match between subjectively reported negative emotions and AUs usually associated with negative valence, there seems to be considerable overlap with the literature and results reported above. That is, more negative emotions in the non-game condition elicited more facial movements associated with negative valence.

Previous works related increased head activity to more negative affective states [68]. We identified similar patterns because head activity was partly increased in the non-game condition. Increased movement range and velocity along the y-axis (up/down movements) might indicate efforts to correct a slumped sitting position, possibly reflecting the tediousness of the task. In the game condition, relations between head activity and negative emotions were mixed. We found congruent but also opposing indicators compared to the non-game condition.

Subjective effort: Our analysis revealed several action units related to subjective effort, such as AU05, AU06, AU10, AU12, AU14, and AU15, representing a relatively broad spectrum of positive and negative valence as well as high and low arousal. Noteworthy within the game condition, AU14 showed a positive and AU05 showed a negative relation to subjective effort across blocks. The first is sometimes indiscriminative in terms of emotional valence, the latter is closely tied to arousal [54,81]. Under the premise that subjective effort unifies both positive and negative emotional components, it seems plausible that facial expressions indicating ambivalent valence and are tied to arousal best indicate subjective effort. Interestingly, relationships between AUs and subjective effort were predominantly found in the game condition as well as in block 1. Both these findings are intriguing because, according to participants' ratings, subjective effort seemed to be stronger in the non-game condition and, in both conditions, increasing over time [see 3] and expressions might therefore be more prevalent in the later course of the task.

Significant head activity features were predominantly found in block 2 (except for left/right movement stability). This seems plausible because subjective effort was increasing over time and corresponding expressions may be stronger. However, this does not explain the differential pattern between conditions: For instance, increased up/down movements in the game-condition are opposed to decreased left-right movements in the non-game condition, both indicating subjective effort. Also contradicting this account is the absence of significant relationships between action units and subjective effort in block 2 (but their presence in block 1). There seems to be no intuitive access to these results in the first place, and literature also provides mixed results. Increased bodily movements may also indicate low attention [17] or disengagement [89]. It is assumed further that head activity is more generally an indicator of boredom [68]. In contrast, though, other results showed that boredom (but also enjoyment) co-occurred with decreased amounts of upper body activity [35]. Subjective effort thus seems to be - at least partially - tangible through facial and head activity features.

5.3 Implications – (Un)steady patterns

In our comprehensive analysis of face and head activities, various (in-)congruities emerged with respect to theoretical assumptions (e.g. increased emotional engagement in the game version) and subjectively reported emotions. Nevertheless, current results indicate the potential of such objective metrics and continuous measurements of subjective experiences and provide important theoretical as well as practical implications.

On a *theoretical level*, de-/increases in AU intensities (slopes) were more prevalent in predicting the presence or absence of game elements than in predicting emotions or subjective effort, where almost exclusively variability measures (residuals) were indicative. This seems reasonable because emotions, at least compared to overall mood, tend to be event-triggered and time-limited affective reactions [e.g. 7,14]. These affective reactions might be manipulated by game elements and are therefore reflected by short-lived fluctuations in facial expressions. On the other hand, modelling conditions with and without game elements based on facial activity may involve more abstract concepts of engagement that are better represented by latent changes in facial expressions.

Congruencies in relevant face and head movement features across conditions were rare. This suggests that facial and head activity features may not be readily comparable across different contexts (i.e. game vs. non-game). Moreover, correspondences between facial/head activity features and subjective ratings were not straightforward either. However, this is not unique to this study. When studying the literature on AUs, it becomes apparent that the same AU in different studies (see also Table 2), for instance, the upper lip raiser (AU10), can be indicative for positive [54] and negative valence [81]. Similar inconsistencies across studies emerge when looking at head and upper body activity [e.g. 27,32,35].

Apart from these inconsistencies, AUs and head activity did contribute to a better understanding of the users in the game and non-game condition. For instance, blinks (AU45) increased in the non-game condition across a block in the experiment, which is usually linked to boredom [53]. However, blinks (AU45) did not equivalently indicate any of the allegedly negative subjective scales (e.g. negative emotions or subjective effort). This raises the fundamental question of how to best validate objective and continuous measures of subjective experiences. The intuitive answer and often employed strategy seems to be the comparison between subjective (post-hoc) ratings and the objective metrics obtained [for a related but different question on how to validate emotion recognition algorithms, see 90] . While this can work well when subjective measures and objective measures are compared in close temporal proximity, it can be suboptimal when a single subjective estimate is compared to continuous data acquired over several minutes or even longer. In particular, subjective post-hoc ratings are affected by cognitive biases and memory effects, such as the peak-end rule, which states that experiences in the last moments of an episode are strongly emphasised by participants' hindsight. [e.g. 24]. In a similar vein, primacy and recency effects could distort subjective ratings [91] and might explain the potential mismatch between objective, continuous measures and subjective post-hoc ratings. However, this needs to be systematically investigated in future studies with optimised experimental designs. A fruitful approach in this regard might be to experimentally induce subjective experiences such as boredom and effort using experimental manipulations and link them to face/head/body movements. In a second step, identified indicators for these experiences can be compared between game and non-game contexts.

On a more *practical level*, the current study provides an example of the importance of investigating subjective experiences with different measures. Although matching between face/head movements and subjectively reported (emotional) experiences varied and relevant features were not constant across conditions, current data helped, for instance, to differentiate between participants engaged in the game or non-game condition. Of course, more research is needed, which is particularly warranted in the field of games research, as emotions are often seen to be at the core of gaming [e.g. 72,31]. The use of continuous objective measures of subjective experience might be particularly useful for better understanding temporal dynamics between ordinary (e.g., emotionally moderate) and extraordinary (e.g., emotionally intense) player experience, as recently termed by Tyack & Meckler [78]. In this context, the current study, went beyond the majority of previous studies by not only analysing the frequency of certain (emotional) facial and head movement features [e.g. 29], but tried to utilize the obtained continuous data more comprehensively by analysing changes across time (i.e. slopes) and their variability (i.e. residuals). However, this could also explain why the match between objective, continuous measurements and subjective assessments in the current study could be considered unsatisfactory (see also above). This further underlines the difficulty of implementing such measurement methods and therefore also represents a limitation for the present results: the implementation of detailed subjective and objective measurement methods is complex and the interaction of subjective and objective measurement methods in relation to emotional engagement over time still needs to be researched and understood in much more detail. Emotions are central to learning games, so inaccuracies in measurement or interpretation can quickly lead to effects that are detrimental in any learning process, such as demotivation or frustration.

Nevertheless, obtaining objective and continuous measures to assess subjective experiences, e.g. through webcam recordings, as was the case in the current study, offers advantages for basic and applied game-related research [for a more comprehensive overview see e.g. ,57]. First, subjective experiences can be measured continuously during playing without disturbing users with prompts and questionnaires. Second, continuous and objective measures are not affected by cognitive biases and memory effects [e.g. 24,91]. However, current results need to be treated with caution as the overall pattern of the results was quite diverse. Even though measures derived from simple face recordings can have advantages, previous studies indicated that other variables, such as emotion regulation, context, gender, age, and personality, can influence how and if emotions are expressed [e.g. 8,37].

6 CONCLUSIONS

Regarding our initial research questions we can conclude, that facial and head movement features can reflect the use and impact of game elements in a cognitive task (RQ1). Results indicated that game elements emotionally engaged participants, which is reflected by bodily features such as action units, especially with respect to intensity changes across time. For instance, an increase of blink rate (AU45) intensity – an indicator of boredom – was related to playing without game elements. Head movements were also able to distinguish between the absence or presence of game-elements, while head rotations (rolling/nodding) seemed to be more significant than head translations (vertical or horizontal etc.).

Positive and negative emotions could be mapped onto a series of facial as well as head pose features (RQ2). Our assumption that subjective effort manifests at levels of expressive facial and head movements was also supported. However, we identified considerable inconsistencies across conditions (with or without game elements) and (partly) to the literature. For instance, several significant AUs were either indiscriminative or even contraindicated and game-elements lead to substantial differences in the visual displays of emotions in relation to subjective ratings. This could be because our manipulation of affect through game elements was not impactful enough, or that our subjective measurements employed were too superficial and coarse to identify appropriate correspondences. Our diverse results on the expressive component of emotions and subjective effort best fit a conclusion from related work, stating that feelings that arise while or because of doing something (not) worth the cost critically depend upon the individual and the context [18]. Nonethless, action units and head movement provide valuable metrics that augment conventional approaches to assessing subjective experiences such as post-hoc ratings, particularly in the context of gaming.

7 AUTHOR CONTRIBUTION

KB and MN conceptualized, designed and constructed the study and recuited participants. SG conducted the analysis. SG and MN wrote the first draft. All authors contributed to writing the final version.

ACKNOWLEDGMENTS

This research was funded by the Postdoc Network "Cognitive Conflicts During Media Use" funded by the competition fund of the Leibniz Association (SAW16) within the framework of the "Pact for Research and Innovation".

REFERENCES

[1] James J. Appleton, Sandra L. Christenson, and Michael J. Furlong. 2008. Student engagement with school: Critical conceptual and methodological issues of the construct. Psychol. Sch. 45, 5 (May 2008), 369–386. DOI: https://doi.org/10.1002/pits.20303

[2] Isabelle Archambault, Michel Janosz, Jean-Sébastien Fallu, and Linda S. Pagani. 2009. Student engagement and its relationship with early high school dropout. J. Adolesc. 32, 3 (June 2009), 651–670. DOI: https://doi.org/10.1016/j.adolescence.2008.06.007

[3] Katharina Bernecker and Manuel Ninaus. 2021. No Pain, no Gain? Investigating motivational mechanisms of game elements in cognitive tasks. Comput. Hum. Behav. 114, (January 2021), 106542. DOI: https://doi.org/10.1016/j.chb.2020.106542

[4] Nigel Bosch, Yuxuan Chen, and Sidney D'Mello. 2014. It's Written on Your Face: Detecting Affective States from Facial Expressions while Learning Computer Programming. In Intelligent Tutoring Systems (Lecture Notes in Computer Science), Springer Int'l Publishing, Cham, 39–44. DOI: https://doi.org/10.1007/978-3-319-07221-0_5

[5] H. Boukricha, I. Wachsmuth, A. Hofstätter, and K. Grammer. 2009. Pleasure-arousal-dominance driven facial expression simulation. In 2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, 1–7. DOI: https://doi.org/10.1109/ACII.2009.5349579

[6] Elizabeth A. Boyle, Thomas Hainey, Thomas M. Connolly, Grant Gray, Jeffrey Earp, Michela Ott, Theodore Lim, Manuel Ninaus, Claudia Ribeiro, and João Pereira. 2016. An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. Comput. Educ. 94, (2016), 178–192.

[7] John T. Cacioppo and Wendi L. Gardner. 1999. EMOTION. Annu. Rev. Psychol. 50, 1 (February 1999), 191–214. DOI: https://doi.org/10.1146/annurev.psych.50.1.191

[8] Tara M. Chaplin and Amelia Aldao. 2013. Gender differences in emotion expression in children: A meta-analytic review. Psychol. Bull. 139, 4 (2013), 735–765. DOI: https://doi.org/10.1037/a0030737

[9] Giovanna Cilluffo, Gianluca Sottile, Stefania La Grutta, and Vito MR Muggeo. 2020. The Induced Smoothed lasso: A practical framework for hypothesis testing in high dimensional regression. Stat. Methods Med. Res. 29, 3 (March 2020), 765–777. DOI: https://doi.org/10.1177/0962280219842890

[10] Mihaly Csikszentmihalyi. 1990. The psychology of optimal experience New York. Harper & Row.

[11] Nele Dael, Marcello Mortillaro, and Klaus R. Scherer. 2012. Emotion expression in body action and posture. Emotion 12, 5 (2012), 1085–1101. DOI: https://doi.org/10.1037/a0025737

[12] Sidney D'Mello, Ed Dieterle, and Angela Duckworth. 2017. Advanced, Analytic, Automated (AAA) Measurement of Engagement During Learning. Educ. Psychol. 52, 2 (April 2017), 104–123. DOI: https://doi.org/10.1080/00461520.2017.1281747

[13] Sidney D'Mello and Art Graesser. 2012. Dynamics of affective states during complex learning. Learn. Instr. 22, 2 (April 2012), 145–157. DOI: https://doi.org/10.1016/j.learninstruc.2011.10.001

[14] Paul Ekman. 1992. An argument for basic emotions. Cogn. Emot. 6, 3–4 (May 1992), 169–200. DOI: https://doi.org/10.1080/02699939208411068

[15] Paul Ekman. 2002. Facial action coding system (FACS). Hum. Face (2002).

[16] Paul Ekman and E. Friesen. 1978. Facial action coding system: a technique for the measurement of facial movement. Palo Alto 3, 2 (1978), 5.

[17] Anna Marie Farrace-Di Zinno, Graham Douglas, Stephen Houghton, Vivienne Lawrence, John West, and Ken Whiting. 2001. Body movements of boys with attention deficit hyperactivity disorder (ADHD) during computer video game play. Br. J. Educ. Technol. 32, 5 (2001), 607–618.

[18] Alexander L. Francis and Jordan Love. 2020. Listening effort: Are we measuring cognition or affect, or both? WIREs Cogn. Sci. 11, 1 (January 2020). DOI: https://doi.org/10.1002/wcs.1514

[19] Jennifer Fredricks, Phyllis C Blumenfeld, and Alison H Paris. 2004. School Engagement: Potential of the Concept, State of the Evidence. Rev. Educ. Res. 74, 1 (March 2004), 59–109. DOI: https://doi.org/10.3102/00346543074001059

[20] Jennifer Fredricks, Michael Filsecker, and Michael A. Lawson. 2016. Student engagement, context, and adjustment: Addressing definitional, measurement, and methodological issues. Learn. Instr. 43, (June 2016), 1–4. DOI: https://doi.org/10.1016/j.learninstruc.2016.02.002

[21] Jerome Friedman, Trevor Hastie, and Robert Tibshirani. 2010. Regularization Paths for Generalized Linear Models via Coordinate Descent. J. Stat. Softw. 33, 1 (2010), 1–22.

[22] Nico H. Frijda. 1988. The laws of emotion. Am. Psychol. 43, 5 (1988), 349–358. DOI: https://doi.org/10.1037/0003-066X.43.5.349

[23] Beatrice de Gelder. 2006. Towards the neurobiology of emotional body language. Nat. Rev. Neurosci. 7, 3 (March 2006), 242–249. DOI: https://doi.org/10.1038/nrn1872

[24] Xiaowei Geng, Ziguang Chen, Wing Lam, and Quanquan Zheng. 2013. Hedonic Evaluation over Short and Long Retention Intervals: The Mechanism of the Peak-End Rule: Hedonic Evaluation over Retention Intervals. J. Behav. Decis. Mak. 26, 3 (July 2013), 225–236. DOI: https://doi.org/10.1002/bdm.1755

[25] Patricia Goldberg, Ömer Sümer, Kathleen Stürmer, Wolfgang Wagner, Richard Göllner, Peter Gerjets, Enkelejda Kasneci, and Ulrich Trautwein. 2019. Attentive or Not? Toward a Machine Learning Approach to Assessing Students' Visible Engagement in Classroom Instruction. Educ. Psychol. Rev. (December 2019). DOI: https://doi.org/10.1007/s10648-019-09514-z

[26] Arthur C. Graesser and Brent A. Olde. 2003. How does one know whether a person understands a device? The quality of the questions the person asks when the device breaks down. J. Educ. Psychol. 95, 3 (2003), 524–536. DOI: https://doi.org/10.1037/0022-0663.95.3.524

[27] Joseph F. Grafsgaard, Kristy Elizabeth Boyer, Eric N. Wiebe, and James C. Lester. 2012. Analyzing Posture and Affect in Task-Oriented Tutoring. In FLAIRS Conference, 438–443.

[28] Joseph F. Grafsgaard, Joseph B. Wiggins, Kristy Elizabeth Boyer, Eric N. Wiebe, and James C. Lester. 2013. Automatically Recognizing Facial Indicators of Frustration: A Learning-centric Analysis. In 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction, IEEE, Geneva, Switzerland, 159–165. DOI: https://doi.org/10.1109/ACII.2013.33

[29] Joseph Grafsgaard, Joseph B. Wiggins, Kristy Elizabeth Boyer, Eric N. Wiebe, & James Lester. 2013. Automatically recognizing facial expression: Predicting engagement and frustration. In Educational Data Mining 2013.

[30] Gillian Green, Jean Rhodes, Abigail Heitler Hirsch, Carola Suárez-Orozco, and Paul M. Camic. 2008. Supportive adult relationships and the academic engagement of Latin American immigrant youth. J. Sch. Psychol. 46, 4 (August 2008), 393–412. DOI: https://doi.org/10.1016/j.jsp.2007.07.001

[31] Simon Greipl, Manuel Ninaus, and Korbinian Moeller. 2020. Potential and limits of game-based learning. Int. J. Technol. Enhanc. Learn. 12, 4 (2020), 363. DOI: https://doi.org/10.1504/IJTEL.2020.10028417

[32] Hatice Gunes, Caifeng Shan, Shizhi Chen, and YingLi Tian. 2015. Bodily Expression for Automatic Affect Recognition. In Emotion Recognition, Amit Konar and Aruna Chakraborty (eds.). John Wiley & Sons, Inc., Hoboken, NJ, USA, 343–377. DOI: https://doi.org/10.1002/9781118910566.ch14

[33] Curtis R. Henrie, Lisa R. Halverson, and Charles R. Graham. 2015. Measuring student engagement in technology-mediated learning: A review. Comput. Educ. 90, (December 2015), 36–53. DOI: https://doi.org/10.1016/j.compedu.2015.09.005

[34] Paul J. Hirschfield and Joseph Gasper. 2011. The Relationship Between School Engagement and Delinquency in Late Childhood and Early Adolescence. J. Youth Adolesc. 40, 1 (January 2011), 3–22. DOI: https://doi.org/10.1007/s10964-010-9579-5

[35] Wouter M. van den Hoogen, Wijnand A. IJsselsteijn, and Yvonne AW de Kort. 2008. Exploring behavioral expressions of player experience in digital games. In Proceedings of the workshop on Facial and Bodily Expression for Control and Adaptation of Games ECAG, Citeseer, 11–19.

[36] Rob Hyndman, Yanfei Kang, Pablo Montero-Manso, Thiyanga Talagala, Earo Wang, Yangzhuoran Yang, and Mitchell O'Hara-Wild. 2020. tsfeatures: Time Series Feature Extraction. Retrieved from https://CRAN.R-project.org/package=tsfeatures

[37] Herman Ilgen, Jacob Israelashvili, and Agneta Fischer. 2021. Personal Nonverbal Repertoires in facial displays and their relation to individual differences in social and emotional styles. Cogn. Emot. (January 2021), 1–10. DOI: https://doi.org/10.1080/02699931.2021.1877118

[38] Michael Inzlicht, Amitai Shenhav, and Christopher Y. Olivola. 2018. The Effort Paradox: Effort Is Both Costly and Valued. Trends Cogn. Sci. 22, 4 (April 2018), 337–349. DOI: https://doi.org/10.1016/j.tics.2018.01.007

[39] Susanne M. Jaeggi, Martin Buschkuehl, John Jonides, and Walter J. Perrig. 2008. Improving fluid intelligence with training on working memory. Proc. Natl. Acad. Sci. 105, 19 (May 2008), 6829–6833. DOI: https://doi.org/10.1073/pnas.0801268105

[40] Veronika Job, Carol S. Dweck, and Gregory M. Walton. 2010. Ego Depletion—Is It All in Your Head?: Implicit Theories About Willpower Affect Self-Regulation. Psychol. Sci. 21, 11 (November 2010), 1686–1693. DOI: https://doi.org/10.1177/0956797610384745

[41] B. Kort, R. Reilly, and R. W. Picard. 2001. An affective model of interplay between emotions and learning: reengineering educational pedagogy-building a learning companion. In Proceedings IEEE International Conference on Advanced Learning Technologies, 43–46. DOI: https://doi.org/10.1109/ICALT.2001.943850

[42] Heinz W. Krohne, Boris Egloff, Carl-Walter Kohlmann, and Anja Tausch. 1996. Untersuchungen mit einer deutschen version der" positive and negative affect schedule"(PANAS). Diagn.-Gottingen- 42, (1996), 139–156.

[43] Richard N. Landers. 2014. Developing a Theory of Gamified Learning: Linking Serious Games and Gamification of Learning. Simul. Gaming 45, 6 (December 2014), 752–768. DOI: https://doi.org/10.1177/1046878114563660

[44] Richard N. Landers, Elena M. Auer, Andrew B. Collmus, and Michael B. Armstrong. 2018. Gamification Science, Its History and Future: Definitions and a Research Agenda. Simul. Gaming 49, 3 (June 2018), 315–337. DOI: https://doi.org/10.1177/1046878118774385

[45] Hao Lei, Yunhuo Cui, and Wenye Zhou. 2018. Relationships between student engagement and academic achievement: A meta-analysis. Soc. Behav. Personal. Int. J. 46, 3 (March 2018), 517–528. DOI: https://doi.org/10.2224/sbp.7054

[46] Shan Li and Weihong Deng. 2018. Deep Facial Expression Recognition: A Survey. ArXiv180408348 Cs (April 2018). Retrieved January 24, 2019 from http://arxiv.org/abs/1804.08348

[47] Gwen C. Littlewort, Marian S. Bartlett, Linda P. Salamanca, and Judy Reilly. 2011. Automated measurement of children's facial expressions during problem solving tasks. In Face and Gesture 2011, IEEE, Santa Barbara, CA, USA, 30–35. DOI: https://doi.org/10.1109/FG.2011.5771418

[48] Meng Liu, Yaocong Duan, Robin A. A. Ince, Chaona Chen, Oliver G. B. Garrod, Philippe Schyns, and Rachael E. Jack. 2020. Facial Expressions of Emotion Categories are Embedded within a Dimensional Space of Valence-arousal. DOI: https://doi.org/10.31234/osf.io/pw5uh

[49] Konstantinos Makantasis, Antonios Liapis, and Georgios N. Yannakakis. 2021. The Pixels and Sounds of Emotion: General-Purpose Representations of Arousal in Games. ArXiv210110706 Cs (February 2021). Retrieved February 17, 2021 from http://arxiv.org/abs/2101.10706

[50] Regan L. Mandryk and M. Stella Atkins. 2007. A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. Int. J. Hum.-Comput. Stud. 65, 4 (April 2007), 329–347. DOI: https://doi.org/10.1016/j.ijhcs.2006.11.011

[51] Brais Martinez, Michel F. Valstar, Bihan Jiang, and Maja Pantic. 2019. Automatic Analysis of Facial Actions: A Survey. IEEE Trans. Affect. Comput. 10, 3 (July 2019), 325–347. DOI: https://doi.org/10.1109/TAFFC.2017.2731763

[52] Gerald Matthews, Joel S. Warm, and Andrew P. Smith. 2017. Task Engagement and Attentional Resources: Multivariate Models for Individual Differences and Stress Factors in Vigilance. Hum. Factors J. Hum. Factors Ergon. Soc. 59, 1 (February 2017), 44–61. DOI: https://doi.org/10.1177/0018720816673782

[53] Bethany McDaniel, Sidney D'Mello, Brandon King, Patrick Chipman, Kristy Tapp, and Art Graesser. 2007. Facial features for affective state detection in learning environments. In Proceedings of the Annual Meeting of the Cognitive Science Society.

[54] Daniel McDuff, Rana El Kaliouby, Karim Kassam, and Rosalind Picard. 2010. Affect valence inference from facial action unit spectrograms. In 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, IEEE, San Francisco, CA, USA, 17–24. DOI: https://doi.org/10.1109/CVPRW.2010.5543833

[55] Elisa D. Mekler, Florian Brühlmann, Alexandre N. Tuch, and Klaus Opwis. 2017. Towards understanding the effects of individual gamification elements on intrinsic motivation and performance. Comput. Hum. Behav. 71, (June 2017), 525–534. DOI: https://doi.org/10.1016/j.chb.2015.08.048

[56] Marianne Miserandino. 1996. Children who do well in school: Individual differences in perceived competence and autonomy in above-average children. J. Educ. Psychol. 88, 2 (1996), 203–214. DOI: https://doi.org/10.1037/0022-0663.88.2.203

[57] Steve Nebel and Manuel Ninaus. 2019. New Perspectives on Game-Based Assessment with Process Data and Physiological Signals. In Game-Based Assessment Revisited, Dirk Ifenthaler and Yoon Jeon Kim (eds.). Springer International Publishing, Cham, 141–161. DOI: https://doi.org/10.1007/978-3-030-15569-8_8

[58] Fred M. Newmann (Ed.). 1992. Student engagement and achievement in American secondary schools. Teachers College Press, New York.

[59] Manuel Ninaus, Simon Greipl, Kristian Kiili, Antero Lindstedt, Stefan Huber, Elise Klein, Hans-Otto Karnath, and Korbinian Moeller. 2019. Increased emotional engagement in game-based learning – A machine learning approach on facial emotion detection data. Comput. Educ. 142, (December 2019), 103641. DOI: https://doi.org/10.1016/j.compedu.2019.103641

[60] Manuel Ninaus, K. Kiili, G. Wood, K. Moeller, and S. E. Kober. 2020. To Add or Not to Add Game Elements? Exploring the Effects of Different Cognitive Task Designs Using Eye Tracking. IEEE Trans. Learn. Technol. 13, 4 (October 2020), 847–860. DOI: https://doi.org/10.1109/TLT.2020.3031644

[61] Manuel Ninaus, Gonçalo Pereira, René Stefitz, Rui Prada, Ana Paiva, Christa Neuper, and Guilherme Wood. 2015. Game elements improve performance in a working memory training task. Int. J. Serious Games 2, 1 (February 2015). DOI: https://doi.org/10.17083/ijsg.v2i1.60

[62] Babette Park, Lisa Knörzer, Jan L. Plass, and Roland Brünken. 2015. Emotional design and positive emotions in multimedia learning: An eyetracking study on the use of anthropomorphisms. Comput. Educ. 86, (August 2015), 30–42. DOI: https://doi.org/10.1016/j.compedu.2015.02.016

[63] Jonathan Peirce, Jeremy R. Gray, Sol Simpson, Michael MacAskill, Richard Höchenberger, Hiroyuki Sogo, Erik Kastman, and Jonas Kristoffer Lindeløv. 2019. PsychoPy2: Experiments in behavior made easy. Behav. Res. Methods 51, 1 (February 2019), 195–203. DOI: https://doi.org/10.3758/s13428-018-01193-y

[64] Reinhard Pekrun and Lisa Linnenbrink-Garcia. 2012. Academic Emotions and Student Engagement. In Handbook of Research on Student Engagement, Sandra L. Christenson, Amy L. Reschly and Cathy Wylie (eds.). Springer US, Boston, MA, 259–282. DOI: https://doi.org/10.1007/978-1-4614-2018-7_12

[65] Jan L. Plass, Steffi Heidig, Elizabeth O. Hayward, Bruce D. Homer, and Enjoon Um. 2014. Emotional design in multimedia learning: Effects of shape and color on affect and learning. Learn. Instr. 29, (February 2014), 128–140. DOI: https://doi.org/10.1016/j.learninstruc.2013.02.006

[66] R Core Team. 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/

[67] M. Ramzan, H. U. Khan, S. M. Awan, A. Ismail, M. Ilyas, and A. Mahmood. 2019. A Survey on State-of-the-Art Drowsiness Detection Techniques. IEEE Access 7, (2019), 61904–61919. DOI: https://doi.org/10.1109/ACCESS.2019.2914373

[68] Valentin Riemer, Julian Frommel, Georg Layher, Heiko Neumann, and Claudia Schrader. 2017. Identifying Features of Bodily Expression As Indicators of Emotional Experience during Multimedia Learning. Front. Psychol. 8, (July 2017), 1303. DOI: https://doi.org/10.3389/fpsyg.2017.01303

[69] James A. Russell. 1980. A circumplex model of affect. J. Pers. Soc. Psychol. 39, 6 (1980), 1161–1178. DOI: https://doi.org/10.1037/h0077714

[70] Michael Sailer and Lisa Homner. 2020. The Gamification of Learning: a Meta-analysis. Educ. Psychol. Rev. 32, 1 (March 2020), 77–112. DOI: https://doi.org/10.1007/s10648-019-09498-w

[71] Andrew Schall. 2014. New Methods for Measuring Emotional Engagement. In Design, User Experience, and Usability. User Experience Design Practice (Lecture Notes in Computer Science), Springer International Publishing, Cham, 347–357. DOI: https://doi.org/10.1007/978-3-319-07638-6_34

[72] Jesse Schell. 2015. The art of game design: a book of lenses (Second edition ed.). CRC Press, Boca Raton.

[73] Wolfgang Schnotz, Stefan Fries, and Holger Horz. 2009. Motivational aspects of cognitive load theory. In Contemporary motivation research: From global to local perspectives. Hogrefe & Huber Publishers, Ashland, OH, US, 69–96.

[74] Y. A. Sekhavat, S. Roohi, H. Sakian Mohammadi, and G. N. Yannakakis. 2020. Play with One's Feelings: A Study on Emotion Awareness for Player Experience. IEEE Trans. Games (2020), 1–1. DOI: https://doi.org/10.1109/TG.2020.3003324

[75] Stuart G. Shanker and Devin M. Casenhiser. 2013. Reducing the effort in effortful control. In A Wittgensteinian Perspective on the use of Conceptual Analysis in Psychology. Springer, 214–232.

[76] Nicolas Silvestrini and Guido H.E. Gendolla. 2019. Affect and cognitive control: Insights from research on effort mobilization. Int. J. Psychophysiol. 143, (September 2019), 116–125. DOI: https://doi.org/10.1016/j.ijpsycho.2019.07.003

[77] James Steele. 2020. What is (perception of) effort? Objective and subjective effort during task performance. PsyArXiv (June 2020). Retrieved January 30, 2021 from https://pure.solent.ac.uk/en/publications/what-is-perception-of-effort-objective-and-subjective-effort-duri

[78] April Tyack and Elisa D. Mekler. 2021. Off-Peak: An Examination of Ordinary Player Experience. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, ACM, Yokohama Japan, 1–12. DOI: https://doi.org/10.1145/3411764.3445230

[79] Eunjoon "Rachel" Um, Jan L. Plass, Elizabeth O. Hayward, and Bruce D. Homer. 2012. Emotional design in multimedia learning. J. Educ. Psychol. 104, 2 (May 2012), 485–498. DOI: https://doi.org/10.1037/a0026609

[80] M.F. Valstar and M. Pantic. 2006. Biologically vs. Logic Inspired Encoding of Facial Actions and Emotions in Video. In 2006 IEEE International Conference on Multimedia and Expo, IEEE, Toronto, ON, Canada, 325–328. DOI: https://doi.org/10.1109/ICME.2006.262464

[81] S. Velusamy, H. Kannan, B. Anand, A. Sharma, and B. Navathe. 2011. A method to infer emotions from facial Action Units. In 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2028–2031. DOI: https://doi.org/10.1109/ICASSP.2011.5946910

[82] Harald G. Wallbott. 1998. Bodily expression of emotion. Eur. J. Soc. Psychol. 28, 6 (1998), 879–896. DOI: https://doi.org/10.1002/(SICI)1099-0992(1998110)28:6<879::AID-EJSP901>3.0.CO;2-W

[83] Ming-Te Wang and Jennifer Fredricks. 2014. The Reciprocal Links between School Engagement, Youth Problem Behaviors, and School Dropout during Adolescence. Child Dev. 85, 2 (March 2014), 722–737. DOI: https://doi.org/10.1111/cdev.12138

[84] Ming-Te Wang and Rebecca Holcombe. 2010. Adolescents' Perceptions of School Environment, Engagement, and Academic Achievement in Middle School. Am. Educ. Res. J. 47, 3 (September 2010), 633–662. DOI: https://doi.org/10.3102/0002831209361209

[85] David Watson, Lee Anna Clark, and Auke Tellegen. 1988. Development and validation of brief measures of positive and negative affect: The PANAS scales. J. Pers. Soc. Psychol. 54, 6 (1988), 1063–1070. DOI: https://doi.org/10.1037/0022-3514.54.6.1063

[86] Jacob Whitehill, Marian Bartlett, and Javier Movellan. 2008. Automatic facial expression recognition for intelligent tutoring systems. In 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, IEEE, Anchorage, AK, USA, 1–6. DOI: https://doi.org/10.1109/CVPRW.2008.4563182

[87] Harry J. Witchel, Carlos P. Santos, James K. Ackah, Carina E. I. Westling, and Nachiappan Chockalingam. 2016. Non-Instrumental Movement Inhibition (NIMI) Differentially Suppresses Head and Thigh Movements during Screenic Engagement: Dependence on Interaction. Front. Psychol. 7, (February 2016). DOI: https://doi.org/10.3389/fpsyg.2016.00157

[88] Matthias Witte, Manuel Ninaus, Silvia Erika Kober, Christa Neuper, and Guilherme Wood. 2015. Neuronal Correlates of Cognitive Control during Gaming Revealed by Near-Infrared Spectroscopy. PLOS ONE 10, 8 (August 2015), e0134816. DOI: https://doi.org/10.1371/journal.pone.0134816

[89] Beverly Woolf, Winslow Burleson, Ivon Arroyo, Toby Dragon, David Cooper, and Rosalind Picard. 2009. Affect-aware tutors: recognising and responding to student affect. Int. J. Learn. Technol. 4, 3/4 (2009), 129. DOI: https://doi.org/10.1504/IJLT.2009.028804

[90] Georgios N. Yannakakis, Roddy Cowie, and Carlos Busso. 2021. The Ordinal Nature of Emotions: An Emerging Approach. IEEE Trans. Affect. Comput. 12, 1 (Jan. 2021), 16–35. DOI: https://doi.org/10.1109/TAFFC.2018.2879512

[91] Yei-Yu Yen, Christopher D. Wickens, and Sandra G. Hart. 1985. The Effect of Varying Task Difficulty on Subjective Workload. Proc. Hum. Factors Soc. Annu. Meet. 29, 8 (October 1985), 765–769. DOI: https://doi.org/10.1177/154193128502900808

Received February 2021; revised June 2021; accepted July 2021.