Inferring Player Experiences Using Facial Expressions Analysis

Chek Tien Tan
Games Studio, Faculty of Engineering and IT, University of Technology, Sydney
chek@gamesstudio.org

Sander Bakkes
Intelligent Systems Laboratory, University of Amsterdam, Amsterdam
s.c.j.bakkes@uva.nl

Yusuf Pisan
Games Studio, Faculty of Engineering and IT, University of Technology, Sydney
yusuf.pisan@gamesstudio.org

ABSTRACT
Understanding player experiences is central to game design. Video captures of players is a common practice for obtaining rich reviewable data for analysing these experiences. However, not enough has been done in investigating ways of preprocessing the video for a more efficient analysis process. This paper consolidates and extends prior work on validating the feasibility of using automated facial expressions analysis as a natural quantitative method for evaluating player experiences. A study was performed on participants playing a first-person puzzle shooter game (Portal 2) and a social drawing trivia game (Draw My Thing), and results were shown to exhibit rich details for inferring player experiences from facial expressions. Significant correlations were also observed between facial expression intensities and self reports from the Game Experience Questionnaire. In particular, the challenge dimension consistently showed positive correlations with anger and joy. This paper eventually presents a case for increasing the application of computer vision in video analyses of gameplay.

Categories and Subject Descriptors
H.5.2 [**Information Interfaces and Presentation**]: User Interfaces: Evaluation/methodology; I.2.1 [**Applications and Expert Systems**]: Games.

Keywords
video games, game user research, facial expressions analysis, GEQ, player experience, game design

1. INTRODUCTION
The primary goal of most digital games is to provide players with appropriate and often positive overall experiences that are linked to concepts like flow [5] and immersion [19]. A game designer also often meticulously crafts different gameplay instances to hopefully provide appropriate short-term experiences like fear, anger and surprise. Hence it is essential in game design to be able to measure whether (and to which extent) these experiences are achieved. Research into methods to enable efficient and effective player experience evaluation is hence a key area in game design. Facial expression analysis is one way to provide a natural and non-invasive method to enable such evaluations.

1.1 Player Experience Evaluation
Collecting, detecting and analyzing player experiences are non-trivial tasks. This is because direct measurement methods are often disruptive and laborious, and affective states are complicated, derived concepts. Traditional approaches are often qualitative and includes collecting subjective data from direct observations, interviews and think-aloud protocols. These methods are most common amongst game practitioners and usually require formal playtest sessions in artificial play environments [11, 20]. Although these methods have been shown to reflect reasonably accurate states, they have several shortcomings. Firstly, they might inhibit true play experiences, as the players might not be totally at ease when someone is watching or questioning them. Players might not be able to properly self-articulate their play experiences concurrently during gameplay and might not even remember important details when post interviews are performed. Secondly, the sessions often require a lot of time and resources to conduct and analyze.

These shortcomings have driven much research towards quantitative methods that work on objective data. Quantitative methods have the potential to represent true player experiences in the game and are able to continuously capture a more diverse body of information. However, these quantitative methods currently do not serve to totally replace qualitative approaches, and many have utilized mixed methods in order to form more holistic analyses. Current quantitative work mostly fall within telemetry or psychophysiology approaches.

Telemetry primarily deals with the logging of player in-game interactions to build player models, and several studies have been performed [9, 14, 15, 26]. The advantage of Telemetry over qualitative methods is that it is non-disruptive and that it can continuously capture objective gameplay statistics in non-laboratory settings. However, the data is limited to the in-game actions available to the player and events in the game world. Hence these "virtual observations" do not capture full experiences and might not even represent the true experiences of the player in real life. For example, a player might take a long time to clear a level, but he might be having a high level of arousal in real life having fun exploring the level or simply stimulated by the aesthetics.

Psychophysiology is the other main branch of quantitative player experience research, which consists of methods to infer psychological states from physiological measurements, which commonly includes electrodermal activity (EDA), electromyography (EMG), electrocardiogram (ECG), electroencephalography (EEG), body temperature and pupil dilations. Current work [6, 12, 17, 18, 26] mostly involve inferring emotional valence and arousal by employing a combination of the measurements. Amongst them, EDA and facial EMG seems to be most popular as they are easily deployed and correspond well to emotional dimensions of arousal and valence respectively [22]. Similar to telemetry, physiological measurements are able to capture player experiences continuously in real-time. In addition, physiological data represent the real life experiences of the player. Unfortunately, most current approaches deal with expensive specialized equipment that are obtrusive, which are usually only viable in controlled laboratory settings.

1.2 Motivations and Approach

The reasons above have led to our investigation of a video-based approach based on facial expressions analysis to capture and analyze data in way that is more efficient, versatile, and has minimal disruptions to natural gameplay. Facial expressions analysis [10] is the use of automatically recognized facial expressions to infer affective states. Being a video-based approach, it is non-obtrusive compared to current physiological approaches. This allows for more authentic play experiences and enables data collection in non-laboratory settings.

Facial expressions analysis can be also viewed as a type of psychophysiological approach, which seemed to be underexplored in current player experience research. Prior psychophysiological approaches (e.g., [17]) have shown promising correlations between the measured physiological data (using EDA and facial EMG) and self-reported flow experience. To the best of our knowledge, no work has been performed to evaluate the feasibility of facial expressions analysis as a basis to infer gameplay experience dimensions such as challenge, flow and immersion. Moreover, a key insight for using facial expressions in this paper is that game personalisation techniques have been shown to be able to leverage facial expression analysis techniques to unobtrusively infer player experiences automatically to alter gameplay in real-time [21].

As further motivation, research in non-game domains have shown promising results for inferring other kinds of user experience metrics from facial expressions. For example, Branco [2] showed some encouraging results evaluating positive and negative expressions of users of an online shopping website. In other domains, general emotion detection based on facial expression recognition [1, 10] have also shown promising results.

The first step in any facial expressions analysis system is to recognize facial expressions, and facial expression recognition is a fairly mature domain in computer vision with techniques that boast a high level of accuracy and robustness [3, 13]. This allows us to build on a vast pool of well-developed work as a basis for our framework.

Current technological advancements in gaming technology also favors our video-based approach. The advent of motion detection game consoles like Microsoft's Kinect[1] and Nintendo's 3DS[2], video feeds are naturally incorporated into gameplay. For other games, webcams are also relatively cheap and prevalent in most mobile computing devices nowadays, especially when compared to specialized physiological equipment.

However, before we even venture into inferring these experiences, the question of whether games elicit enough facial expressions, and further more, whether these expressions can be captured robustly, needs to be answered. This paper hence extends prior work [25] to present a comprehensive mixed method analysis based on the above questions.

2. EVALUATION

From the above discussion based on related work, we can see that research into using facial expressions for understanding player experiences is still in very early stages. Our evaluation is hence planned as an exploratory study to uncover interesting opportunities for further research rather than make hypothesis-driven generalizable claims. The following subsections describe our experimental details.

2.1 Method

To evaluate the feasibility of using facial expressions to understand player experiences, we employ a mixed method approach that is primarily qualitative as this research is exploratory. A repeated-measures setup was used for the experiments. All participants had their on-screen actions and facial video captured in two play conditions. The two conditions are the two different games played - the first game being Portal 2 by Valve[3], and the second game Draw My Thing by OMGPOP[4]. Portal 2 is a story-based first-person puzzle shooter in which players shoot portals in order to solve spatial puzzles. In Portal 2, the participants played the single-player story mode from the start. Draw My Thing is a social word trivia game where players take turns to draw a picture using the mouse or track pad in a time-constrained setting, and lets the other player guess the correct word. In Draw My Thing, the participants played with a single human opponent.

After the end of the experiments, the facial videos were then fed through our facial expression recognizer [25] and graphs were generated for each player. These were then consolidated to produce the analyses as shown in Section 3.

2.2 Setup

The apparatus setup consists of a 15-inch Intel Core i7 Apple notebook with 8GB RAM, a three-button mouse, and a Logitech webcam capable to capturing full 1080p High Definition video. The on-screen actions were captured using the

[1]www.xbox.com/en-US/kinect/
[2]www.nintendo.com/3ds
[3]www.thinkwithportals.com
[4]www.omgpop.com/games/drawmything

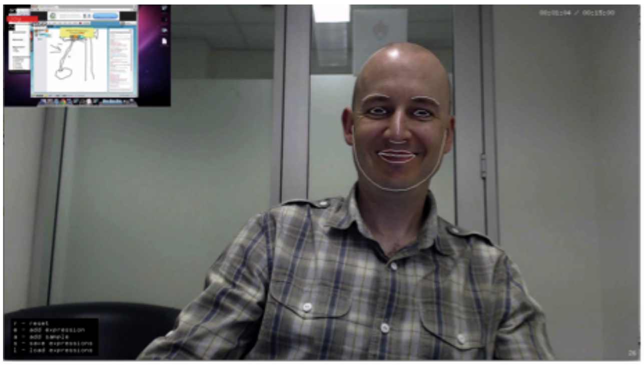

Figure 1: Screenshot of the Facial Expression Recognizer when a participant played Draw My Thing. The white curved lines on the face automatically track the facial expressions of the participant. The actual game screen is also shown in the top left subscreen.

Screenflow software[5]. The notebook was placed in an office with common fluorescent lighting.

The facial expression recognizer was built on the ofxFaceTracker add-on[6] in the openFrameworks C++ toolkit[7] using the OpenCV library[8] for the computer vision functions. A screenshot of the software is shown in Figure 1.

The implementation of facial expression recognition is based on deformable model fitting. It is principled on the concept of learning independent image patches centered on landmarks on the face and has shown superior performance to holistic approaches (refer to [23] for a details of the technique). An important advantage of using this method is that it requires no training and no user intervention throughout the whole tracking, which is aligned with our goals of providing a non-intrusive method of data collection. This implementation represents the state-of-the-art in facial feature tracking which leads to high recognition rates in unconstrained video environments [4].

In this study, we used three common player expressions, namely joy, surprise, and anger, with an additional neutral expression as the baseline. These are derived from the six universal expressions that has been shown to be a basis for emotions across diverse cultures [7]. The use of these six basic expressions, as opposed to the more detailed Facial Action Coding System (FACS) [8] is a conscious decision due to the fact that FACS action unit recognition being still an open problem [1, 13]. In the future, we will gradually investigate the feasibility of using FACS as action unit recognition improves. This approach also ensures that we have a comprehensible record of data for expert analyses of the data when needed, or as a complementary verification to the automatic analyses.

[5] www.telestream.net/screen-flow
[6] github.com/kylemcdonald/ofxFaceTracker
[7] www.openframeworks.cc
[8] sourceforge.net/projects/opencvlibrary

2.3 Participants

Participants were recruited via university mailing lists which includes university employees, undergraduates and alumni. 12 participants (4 females) took part in the study aged between 20 and 48 (mean = 34, SD = 8).

The participants represented a wide mix of player types. Four participants indicated that they play games for more than five hours per week, eight participants less than five hours per week and one participant did not play games at all. Ten participants indicated they enjoy playing first-person shooters, five participants enjoy role-playing games, five participants enjoy strategy games, three participants enjoy simulations, two participants enjoy puzzle games, and three participants enjoy playing social word and trivia games. Six participants indicated that they have played the Portal Series and three have played Drawing games by OMGPOP.

2.4 Procedure

After indicating their informed consent in the study, participants were asked to fill in a background questionnaire to determine player demographics (with the results as described in Section 2.3 above). They then proceed to play the two games for 15 minutes each, one after another in an enclosed room by themselves. The opponent in Draw My Thing played against the participant from a separate room over the Internet.

A short brief on the structure of the session was given to participants before starting the experiments. There were no tutorials or practice sessions prior to gameplay and they were left to figure out the games themselves so as to obtain those initial experiences as well. The participants were also told to play as they normally would, and not to think of this as an experiment. After each game, participants filled the full Game Experience Questionnaire (GEQ) [16].

After playing both games, participants were given a short interview on their experiences which includes describing whether the presence of the camera or other aspects of the experimental setup affected their play experience. No compensations were given to the participants at the end of the experiment.

3. RESULTS AND DISCUSSION

Our results consist of a qualitative analysis of the data collected augmented by a quantitative correlation analysis with a subjective self-report questionnaire. In the following subsections, we present these results with discussions.

3.1 Visual Analysis

Before venturing into the task of inferring gameplay experience metrics, the question of whether games elicit rich enough facial expressions, and further more, whether these expressions can be captured robustly, needs to be answered first. This initial visual analysis of the data is meant to provide a qualitative perspective of the feasibility of facial expressions. Sample plots of 2 participants for the discussion in this section are extracted in Figures 2 and 3, and summarized plots of the averages across all the participants are shown in Figures 4 and 5 for Portal 2 and Draw My Thing respectively. The full plots and detailed descriptions for each play session of Portal 2 and Draw My Thing for

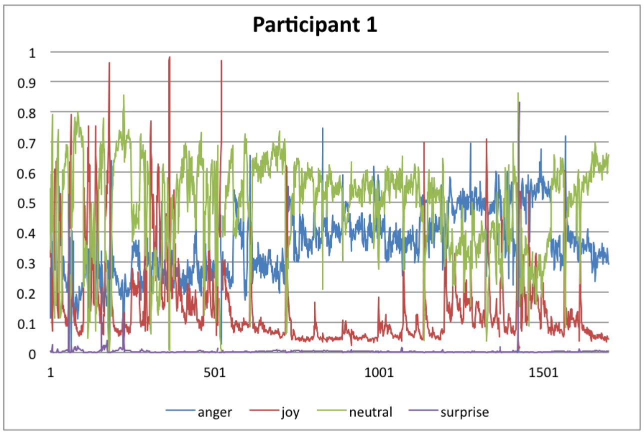

Figure 2: A sample plot with good quality readings: plot of expression intensities (y-axis) against frame count (x-axis) of participant 1 playing Portal 2.

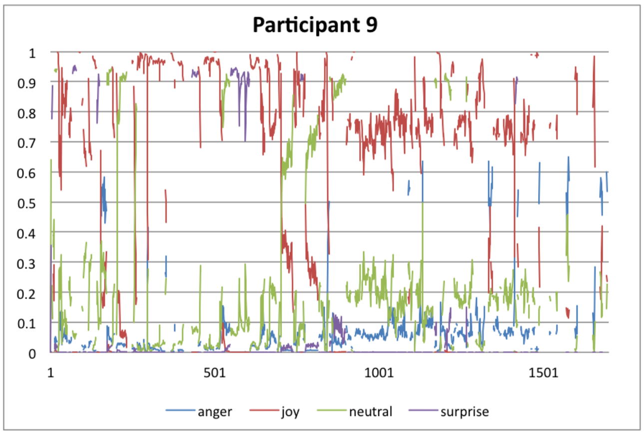

Figure 3: A sample plot with poor quality readings (due to occlusions): plot of expression intensities (y-axis) against frame count (x-axis) of participant 9 playing Portal 2.

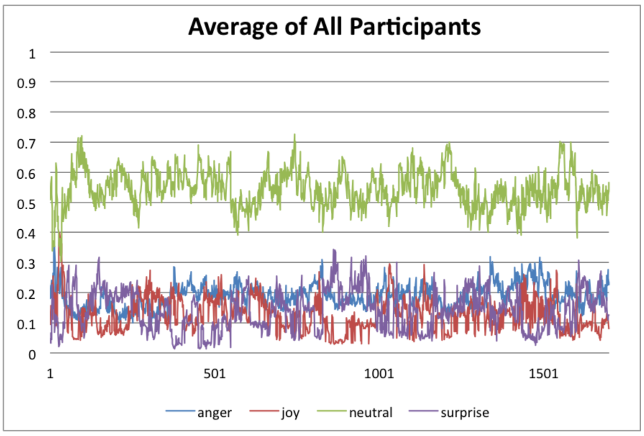

Figure 4: Average expressions of all participants over time for Portal 2: plot of expression intensities (y-axis) against frame count (x-axis).

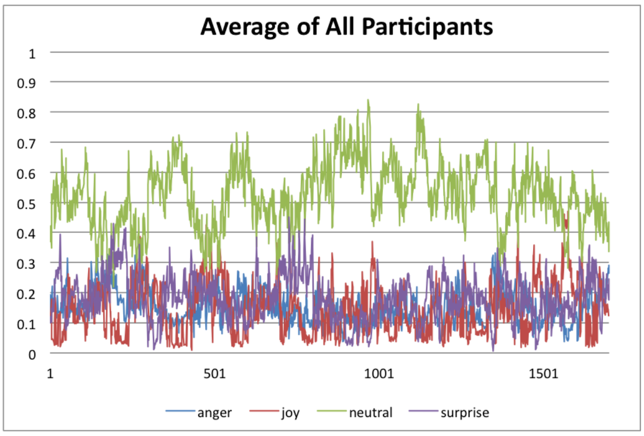

Figure 5: Average expressions of all participants over time for Draw My Thing: plot of expression intensities (y-axis) against frame count (x-axis).

each player can be found in a prior paper [25]. This section is a summarized analysis in order to form a complete picture of the study together with the correlation analysis presented in this paper. As can be seen from the plots, a variety of rich facial expressions were exhibited (other than neutral).

Strong associations were found when comparing the automatically captured expressions with the self-reported experiences from our interviews. For example in the plot of Participant 1 in Figure 2, it can be seen that he/she starts the game with primarily neutral expressions (green line) with spikes of joy (red line) being exhibited, and then anger (blue line) gradually builds up over time. This corresponds well to his/her own account of the experience where he/she mentioned being periodically amused by the opening sequence of the game, but became increasingly frustrated when he/she could not solve even the first puzzle in the game. This correlation can also be observed when performing a visual inspection of the recorded video. Similar rich experiences can be inferred from the plots of the other users.

Another observation was the difference in the expression fluctuations between the 2 games being played. Portal 2 is an immersive single-player first-person puzzle shooter game whilst Draw My Thing is a casual, social drawing game played with friends over the Internet. The individual plots of Draw My Thing exhibited larger fluctuations over the plots of Portal 2, which can also be verified from the summary plots in in Figures 4 and 5. A visual inspection of the videos also showed that participants were more expressive in Draw My Thing than Portal 2. This observation might be attributed to the additional social element in playing Draw My Thing. This implies that the automatically recorded expressions captured the different qualities between the two game genres.

To investigate the true feasibility of a video-based approach, we instructed the participants to play naturally and not be conscious of staying within the camera's vision. When visually inspecting the encoding process, we observed a number of participants either moved out of the camera or placed their hands on their faces while playing Draw My Thing

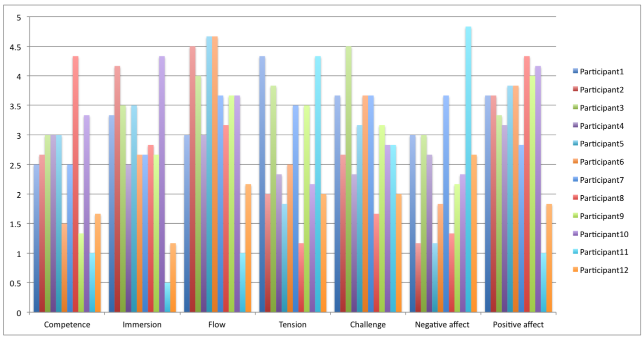

Figure 6: GEQ scores for all participants in Portal 2.

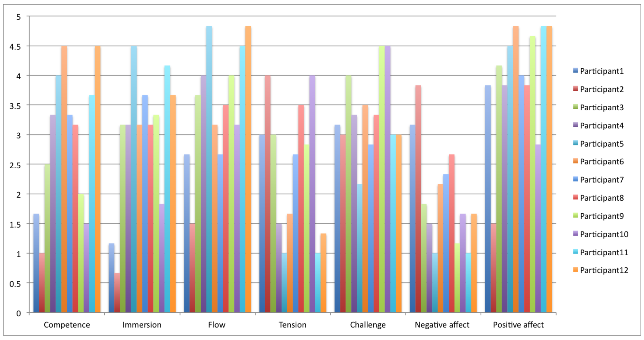

Figure 7: GEQ scores for all participants in Draw My Thing.

(which requires only one hand when drawing), or when watching cut scenes in Portal 2. Minor occlusions did not affect the readings but exaggerated ones resulted in empty readings during these instances. We have recorded all the empty readings as empty plots in the graph as illustrated in Figure 3, where participant 9 had a number of empty readings which can be seen from the disjoint lines on the graph. Upon a visual inspection of participant 9's video, he/she sometimes slouched very low when playing, with a section of his/her face outside of the camera. Fortunately, these occurrences occurred in minority.

Overall, participants felt that the presence of the facial expression detection system was generally not obtrusive to their experiences. At the end of each session, the participants were asked about whether aspects of the experimental setup affected their play. Responses were generally positive:

> "i didn't think about being recorded...it was unobtrusive" (P12)

> "Forgot all about the video recording!!!" (P10)

> "Not really - only when I switched between games, or was waiting for a game. " (P3)

However, some did express a small amount of discomfort due to the recording equipment.

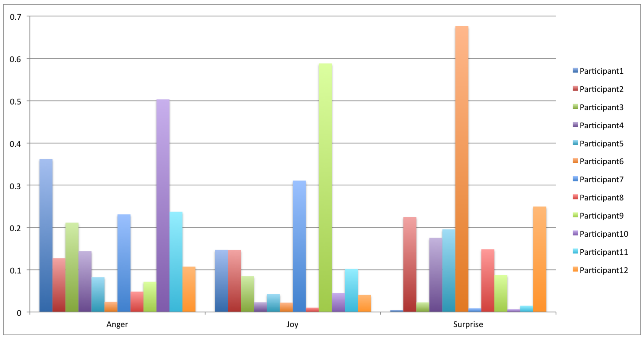

Figure 8: Average facial expression intensities over the game for each participant in Portal 2.

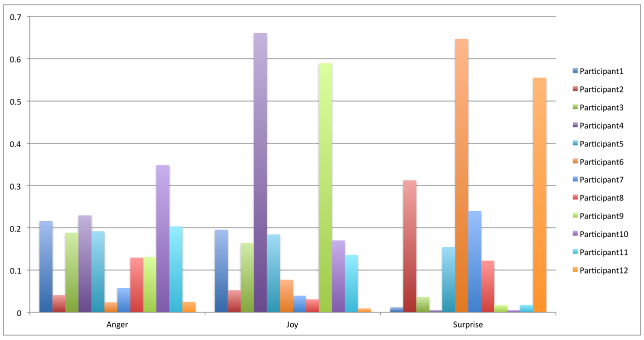

Figure 9: Average facial expression intensities over the game for each participant in Draw My Thing.

> "the video recording wasn't affecting me too much. However, I guess I would have shown more frustration/anger if the video recording was not present." (P2)

> "The knowledge that I am being recorded and that it is based upon my expressions, sometimes make me realize that I'm not just playing and I exit the state of mind that I am usually in when playing games." (P11)

These responses show promise that a video-based approach is indeed largely unobtrusive to gameplay. Perhaps alterations to the way the recording hardware was presented might improve this even further. A thorough study on obtrusiveness might be explored in future research.

3.2 Correlation Analysis

Pearson's correlation coefficients (chosen as both facial expression and GEQ data are parametric) between each game experience dimension in the GEQ and each average facial expression intensity were calculated for participants playing Portal 2 and Draw My Thing, as shown in Tables 1 and 2 respectively. The detailed GEQ scores are shown in Figures 6 and 7, and the average facial expression intensities are shown in Figures 8 and 9.

GEQ Dimension	Anger	Joy	Surprise
Competence	0.15	0.41	-0.24
Immersion	0.29	-0.01	-0.07
Flow	-0.22	0.07	0.40
Tension	0.38	0.44	-0.44
Challenge	0.24	0.26	-0.08
Negative affect	0.43	0.13	-0.47
Positive affect	-0.04	0.09	0.16

Table 1: Pearson's correlation coefficient between average facial expression intensities and average scores from each GEQ dimension, for participants playing Portal 2. All correlations observed are statistically significant ($p < 0.001$).

GEQ Dimension	Anger	Joy	Surprise
Competence	-0.37	-0.18	0.53
Immersion	-0.10	0.07	0.08
Flow	0.18	0.25	-0.07
Tension	0.14	-0.14	-0.25
Challenge	0.35	0.38	-0.30
Negative affect	-0.35	-0.42	0.22
Positive affect	-0.12	0.12	0.15

Table 2: Pearson's correlation coefficient between average facial expression intensities and average scores from each GEQ dimension, for participants playing Draw My Thing. All correlations observed are statistically significant ($p < .001$).

Portal 2

For the anger expression, it can be seen in Table 1 that it was significantly positively correlated with moderate to large effect sizes ($0.30 < r < 0.50$, $p < 0.001$) with GEQ dimensions tension and negative affect. Immersion was significantly positively correlated with an almost moderate effect size ($r = 0.29$, $p < 0.001$). Flow was significantly negatively correlated with a near moderate effect size ($r = -0.22$, $p < 0.001$). Significant positive correlations with smaller effect sizes ($0.10 < r < 0.30$, $p < 0.001$) were exhibited in the GEQ dimensions competence and challenge. Positive affect had a significantly negligible correlation ($r = -0.04$, $p < 0.001$).

For the joy expression, significant positive correlations with moderate to large effect sizes ($0.30 < r < 0.50$, $p < 0.001$) were observed with the GEQ dimensions competence and tension. Significant positive correlations with smaller effect sizes ($0.10 < r < 0.30$, $p < 0.001$) were observed with the GEQ dimensions challenge and negative affect. Immersion, flow and positive affect had significantly negligible correlations ($-0.10 < r < 0.10$, $p < 0.001$).

For the surprise expression, flow was significantly positively correlated with a large effect size ($r = 0.40$, $p < 0.001$). Tension and negative affect were significantly negatively correlated with a large effect size ($r < -0.30$, $p < 0.001$). Competence was significantly negatively correlated with a near moderate effect size ($r = -0.22$, $p < 0.001$). Positive affect was significantly positively correlated with a small effect size ($r = 0.16$, $p < 0.001$). Immersion and challenge had significantly negligible correlations ($-0.10 < r < 0.10$, $p < 0.001$).

From the above correlation findings in Portal 2, it shows that the facial expressions were significantly correlated to a majority of GEQ dimensions. Anger and joy had primarily positive correlations with the GEQ dimensions except for flow, whilst the surprise expression had primarily negative correlations except for flow. Hence an initial implication is that the average facial expression intensities over a certain play period can be used, with relative confidence, to infer these gameplay dimensions with larger affect sizes. For example, large amounts of anger corresponds with higher tension and negative affect.

Draw My Thing

In Draw My Thing, as shown in Table 2, the anger expression only showed a significant positive correlation with a moderate effect size ($r = 0.35$, $p < 0.001$) with the GEQ dimension challenge. Competence and negative affect exhibited significant negative correlations with moderate to large effect sizes ($-0.50 < r < -0.30$, $p < 0.001$). Flow and tension exhibited significant positive correlations with smaller effect sizes ($0.10 < r < 0.30$, $p < 0.001$). Immersion and positive affect were negatively correlated with smaller effect sizes ($-0.30 < r < -0.10$, $p < 0.001$).

For the joy expression, challenge similarly exhibited a significant positive correlation with a moderate effect size ($r = 0.38$, $p < 0.001$). Negative affect was shown to significantly negatively correlate with a moderately large effect size ($r = -0.42$, $p < 0.001$). Positive affect exhibited a significant positive correlation with a rather small effect size ($r = 0.12$, $p < 0.001$). Competence and tension were shown to exhibit significant negative correlations with a smaller effect size ($-0.30 < r < -0.10$, $p < 0.001$). Immersion had a significantly negligible correlation ($r = 0.07$, $p < 0.001$).

For the surprise expression, competence exhibited a significant positive correlation with a large effect size ($r = 0.53$, $p < 0.001$). Challenge exhibited a significantly negative correlation with a moderate effect size ($r = -0.30$, $p < 0.001$). Tension similarly exhibited a significant negative correlation with a near moderate effect size ($r = -0.25$, $p < 0.001$). Negative affect and positive affect were shown to exhibit significant positive correlations with smaller effect sizes ($0.10 < r < 0.30$, $p < 0.001$). Immersion and flow had significantly negligible correlations ($-0.10 < r < 0.10$, $p < 0.001$).

Although the correlations for Draw My Thing were mostly different from those seen in Portal 2, there were several alignments. Focusing on only the larger effect sizes, only challenge showed consistent significant correlations in the same directions. Challenge were both positively correlated with anger and joy, and negatively correlated with surprise. Other than the challenge dimension, the other correlations for Draw My Thing were either in the opposite direction or had vastly different effect sizes, when compared to Portal 2.

4. CONCLUSION

The results discussed above indicate that automated facial expressions analysis can indeed be used to infer meaningful player experiences. Our approach builds on existing research aimed at devising a non-obtrusive player experience analysis method that can be employed to infer player experience metrics.

The first key finding is that each participant's graph represents a rich body of player experience data that often relates to participants' self reports. This implies that there is potentially ample information from automatically captured facial expressions to infer player experiences. We also observed clear visual differences in the graphs between the two game genres played, which means genre differences are being reflected in the expressions as well. We showed that the results here enable a human to make meaningful analyses, which provides a foundational confidence that a machine learning algorithm might be possible for automatically inferring experience dimensions. This will be the next major step in this research.

Another key finding is that significant correlations were exhibited between the facial expressions and the GEQ dimensions. In this study, only challenge showed consistency of a large effect size and direction across both games played. This correlation implies that challenge intensities might be automatically inferred from facial expressions across different game genres. There were also large correlations exhibited in the rest of the GEQ dimensions but they did not align in both games, which means that different games might induce different bindings of facial expressions with gameplay experiences. Moreover, adequacy of only comparing these results to a single instrument (i.e., the GEQ) is also a limitation of our study. This calls for further experiments of larger scopes in the future. Nevertheless we believe the current findings can serve as an initial basis for the implementation of an automated play classifier as part of a supervised learning system (see Figure 10) we have proposed in prior work [24].

In terms of limitations of our framework, missed readings were abundant in some graphs due to occlusions. This would be rather inevitable in natural gameplay for players with a lot of movement around the face. Future work might be to investigate the use of webcams with a larger field of view coupled with motion sensors to rotate accordingly; another approach is to investigate computer vision methods that can handle occlusions better.

In terms of limitations of our experimental setup, one aspect that can be improved is the correlation test. Our current setup used a single GEQ score for the entire gameplay whereas facial expressions were captured 2 times per second. This is still good for our current aim in getting an aggregated correlation for the entire game but it would be useful to also find correlations at a finer granularity during gameplay. However, as detailed gameplay experience metrics were intended for our current study, we chose to use the full GEQ (as opposed to the shorter iGEQ [16]), which means that repeatedly filling out the questionnaire would cause too much participant fatigue. Hence for future experiments, it might be more feasible to evaluate the correlations of the facial expressions with other physiological signals like EDA and EMG, in order to perform a finer-grained analysis.

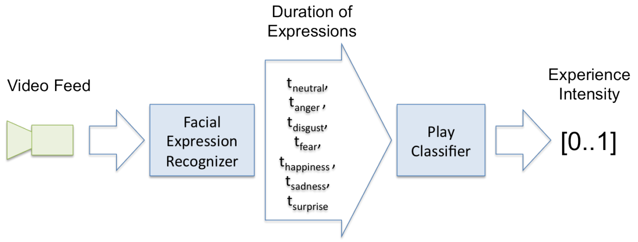

Figure 10: Overview of the automated player experience detector system. The player's captured video is fed into the Facial Expression Recognizer which outputs six expressions based on Ekman's basic emotions [7] and their durations. These features are then input into the Play Classifier which determines the intensity of a gameplay experience metric.

Acknowledgements

This research was kindly supported by the Games Studio, the Centre for Human Centred Technology Design and the School of Software at the University of Technology, Sydney. The authors also thank all reviewers for their constructive feedback.

5. REFERENCES

[1] T. Baltrusaitis, D. McDuff, N. Banda, M. Mahmoud, R. E. Kaliouby, P. Robinson, and R. Picard. Real-time inference of mental states from facial expressions and upper body gestures. *Face and Gesture 2011*, pages 909–914, Mar. 2011.

[2] P. Branco. *Computer-based facial expression analysis for assessing user experience*. PhD thesis, 2006.

[3] J. M. Buenaposada, E. Muñoz, and L. Baumela. Recognising facial expressions in video sequences. *Pattern Analysis and Applications*, 11(1):101–116, Oct. 2007.

[4] J. F. Cohn and M. A. Sayette. Automated Facial Coding. *Behavioral Research Methods*, pages 1–33, 2010.

[5] M. Csikszentmihalyi. *Flow: The Psychology of Optimal Experience*, volume 54 of *P. S. Series*. Harper & Row, 1990.

[6] A. Drachen, L. Nacke, G. Yannakakis, and A. Pedersen. Correlation between heart rate, electrodermal activity and player experience in First-Person Shooter games. In *Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games*, pages 49–54. ACM, 2010.

[7] P. Ekman and W. V. Friesen. The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica*, 1(1):49–98, 1969.

[8] P. Ekman, W. V. Friesen, and J. C. Hager. *Facial Action Coding System*, volume 48. A Human Face, 2002.

[9] A. Gagne, M. Seif El-Nasr, and C. Shaw. A Deeper Look at the use of Telemetry for Analysis of Player Behavior in RTS Games. In *International Conference on Entertainment Computing ICEC*, 2011.

[10] M. Ghijsen. Facial Expression Analysis for Human Computer Interaction. *Affective Computing, IEEE Transactions on*, 2(3):147–161, 2004.

[11] K. Isbister and N. Schaffer. *Game Usability: Advancing the Player Experience*. CRC Press, 1 edition, 2008.

[12] R. L. Mandryk, M. S. Atkins, and K. M. Inkpen. A continuous and objective evaluation of emotional experience with interactive play environments. *Proceedings of the SIGCHI conference on Human Factors in computing systems - CHI '06*, page 1027, 2006.

[13] D. McDuff, R. el Kaliouby, K. Kassam, and R. Picard. Acume: A new visualization tool for understanding facial expression and gesture data. In *Face and Gesture 2011*, pages 591–596. IEEE, Mar. 2011.

[14] B. Medler, M. John, and J. Lane. Data cracker: developing a visual game analytic tool for analyzing online gameplay. In *Proceedings of the 2011 annual conference on Human factors in computing systems*, pages 2365–2374. ACM, 2011.

[15] D. Moura, M. Seif El-Nasr, and C. D. Shaw. Visualizing and Understanding Players âĂŹ Behavior in Video Games : Discovering Patterns and Supporting Aggregation and Comparison. In *ACM SIGGRAPH Symposium on Video Games*, pages 2–7, 2011.

[16] L. E. Nacke. *Affective Ludology: Scientific Measurement of User Experience in Interactive Entertainment*. PhD thesis, 2009.

[17] L. E. Nacke, M. N. Grimshaw, and C. a. Lindley. More than a feeling: Measurement of sonic user experience and psychophysiology in a first-person shooter game. *Interacting with Computers*, 22(5):336–343, Sept. 2010.

[18] L. E. Nacke and C. Lindley. Flow and immersion in first-person shooters: measuring the player's gameplay experience. *Proceedings of the 2008 Conference on Future*, pages 81–88, 2008.

[19] D. Örtqvist and M. Liljedahl. Immersion and Gameplay Experience: A Contingency Framework. *International Journal of Computer Games Technology*, 2010:1–11, 2010.

[20] R. J. Pagulayan, K. Keeker, D. Wixon, R. L. Romero, and T. Fuller. User-centered design in games. In *Design*, volume 28, pages 883–906. 2003.

[21] Paris Mavromoustakos Blom, S. Bakkes, C. T. Tan, S. Whiteson, D. Roijers, R. Valenti, and T. Gevers. Towards Personalised Gaming via Facial Expression Recognition. In *Proceedings of Tenth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, Palo, 2014. AAAI Press.

[22] J. A. Russell. A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6):1161–1178, 1980.

[23] J. M. Saragih, S. Lucey, and J. F. Cohn. Deformable Model Fitting by Regularized Landmark Mean-Shift. *International Journal of Computer Vision*, 91(2):200–215, Sept. 2010.

[24] C. Tan and A. Johnston. Towards a Nondisruptive, Practical, and Objective Automated Playtesting Process. In *Workshops at the Seventh Artificial Intelligence and Interactive Digital Entertainment Conference*, pages 25–28, 2011.

[25] C. T. Tan, D. Rosser, S. Bakkes, and Y. Pisan. A feasibility study in using facial expressions analysis to evaluate player experiences. In *Proceedings of The 8th Australasian Conference on Interactive Entertainment*, page 5. ACM Press, 2012.

[26] V. Zammitto, M. Seif El-Nasr, and P. Newton. Exploring Quantitative Methods for Evaluating Sports Games. *Signals*, pages 8–11, 2010.