An evaluation of Heart Rate and Electrodermal Activity as an Objective QoE Evaluation method for Immersive Virtual Reality Environments

Darragh Egan[1,2], Sean Brennan[3], John Barrett[2], Yuansong Qiao[3], Christian Timmerer[4], Niall Murray[1,2,3]

TIIMED[1], Faculty of Engineering and Informatics[2],
Software Research Institute[3],
Athlone Institute of Technology,
Ireland
email: {d.egan, ysqiao, nmurray}@research.ait.ie,
jbarrett@ait.ie, brennas5@tcd.ie

Alpen-Adria-Universitat Klagenfurt[4],
Institute of Information Technology (ITEC)
Klagenfurt am Worthersee, Austria
christian.timmerer@itec.uni-klu.ac.at

Abstract—**Recently, we have seen an emergence of affordable Head Mounted Displays (HMD) such as the Oculus Rift, HTC Vive, and the PS4 Project Morpheus which allow users to experience 3D virtual reality (VR). These types of hardware aim to facilitate new and novel experiences for users above and beyond what is possible with traditional audiovisual displays. However, a very limited number of studies exist in the literature to determine the influence of these technologies on user Quality of Experience (QoE). In order to evaluate QoE as users consume VR content, this paper proposes the use of affordable consumer electronics to capture objective physiological metrics: Heart Rate (HR) and ElectroDermal Activity (EDA). Our findings indicate different HR and EDA dependent on VR and non-VR environments. Additionally, we examine the relationship between these objective metrics and user QoE captured via a post-test questionnaire. To the best of the authors knowledge, this is the first work which demonstrates a tangible relationship between the EDA/HR combination and user QoE of immersive VR environments.**

Keywords—Quality of Experience; virtual reality; physiological measures; subjective evaluations

I. Introduction

Traditionally, research on evaluating user QoE has primarily focused on the user perception of audiovisual components and the influence of each individually [1][2]. More recently, motivated by the need to enhance user QoE, research and industry have reported works with respect to sensory experiences [3] or multiple sensorial multimodal media content (mulsemedia) [4][5], which includes olfaction (sense of smell) [6][7] and tactile (sense of touch) [8]. Generally speaking, each of these modalities have been used to enhance the traditional 2D media components. It is notable that a lack of QoE research of Virtual Reality (VR) exists, particularly since the first VR hardware and experiences emerged in the 1970s. It is possibly due to the lack of hardware but also due to the difficulty in creating high quality virtual environments.

However, research and industry are slowly making progress in this field. Head-mounted displays like Oculus Rift [9], PlayStation VR [10], and even the low tech Google cardboard [11] among others are making the first steps. In terms of this technology reaching its aim, i.e., presentations that mimic real life experiences, researchers and industry have a long distance to travel despite some of the marketing excitement. A key aspect of this journey will be understanding how users perceive quality of these environments.

The QUALINET group defined a working definition of user Quality of Experience as being “the degree of delight or annoyance of the user of an application or service. It results from the fulfilment of his or her expectations with respect to the utility and / or enjoyment of the application or service in the light of the user’s personality and current state”[12]. The user QoE of a multimedia experience is complex and multidimensional as outlined by Ebrahimi et al. [13] and encompasses many influencing factors as per Fig. 1. Capturing QoE to date has been based on analyzing the user perception of an experience via post-test questionnaires [14]. The literature has identified numerous issues with this approach such as time consuming, expensiveness, and the fact that such approaches do not capture user quality during user consumption. Some initial works, such as [15], facilitated users to rate their quality via remote controls as they were consuming multimedia content. Other approaches to capture metrics have employed varying types of EEG headsets [16]. Whilst the EEG Avenue provides excellent data on neural activity, it is assumed that the level of intrusiveness with such devices impacts user QoE. A position paper by Timmerer et al. suggested an interesting approach to capturing user physiological data via smart watches and health bands as part of a new quality assessment model in [17].

In this context, this paper explores to possibility of using two affordable consumer devices, the Fitbit heart rate monitor and the PIP biosensor (which monitors EDA) as an approach to objectively capturing user QoE. In other works, such combinations have been used with EEG analysis to evaluate user immersion, level of induced stress, etc. [18] during multimedia and immersive VR experiences. Assessors completed a post experience questionnaire and we examined the correlation between the objective and subjective metrics.

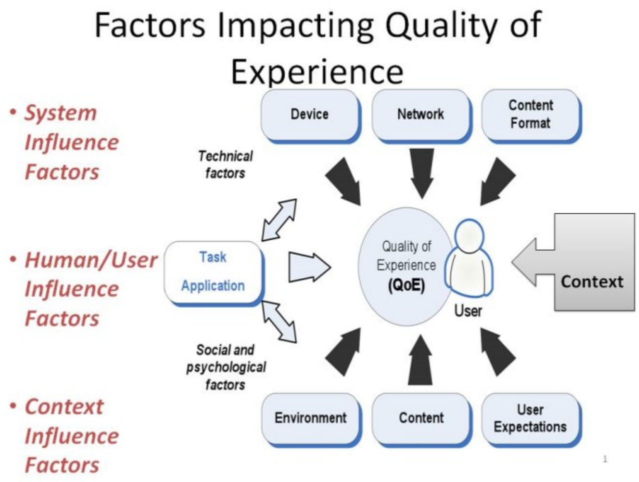

Fig. 1. Factors influencing user Quality of Experience, adapted from [13]

These three approaches to capturing user data was employed when a user experienced a VR (user consumed content whilst wearing a HMD) and non VR environment (user consumed content via a 2D display). We present the comparison of the EDA, HR, and self-reported measures for these two conditions with respect to user QoE.

II. Related Work

In [14], the authors aimed at measuring gaming QoE when users employed immersive HMD. They reported that the Oculus headset increased the sense of immersion in the 3D world as well as perceived game usability. They had 22 participants in the study and each participant was asked to perform a simple forklift driving task in a small virtual circuit. This was tested both with an Oculus headset and with a conventional 2D display monitor. No objective capture of physiological data was performed in this study.

In terms of EDA and heart rate monitoring, [19] aimed to induce a state of anxiety in the user by using aggressive outbursts in virtual and real world environments. They used EDA and heart rate to measure the effects of the aggressive outbursts on the assessor's emotional state and results showed an increase in EDA levels and heart rate during state of anxiety. Their sample size was 28 participants. Another work studied the correlation between HR, EDA, and player experiences while playing a first person shooter game [20]. They had a sample size of 16 participants where participants played three first person shooter games for 20 minutes each while their heart rate and EDA levels were being monitored. Every five minutes they completed an In-Game Experience Questionnaire (iGEQ). Results indicate a significant correlation between psychophysiological arousal (HR, EDA) and self-reported measures for capturing the player experience in the game. Our work differs from this in that [20] was a high speed interactive experience and secondly it was presented on a 2D display and not a VR environment with a HMD.

Fig. 2a. User during evaluation session of VR environment. Oculus HMD
Fig. 2b. User during evaluation session of non-VR environment (2D display)

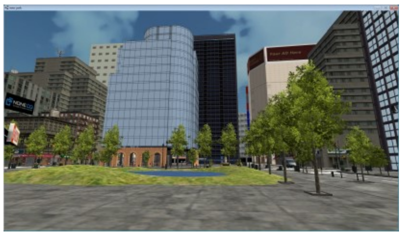

Fig. 3. First person perspective of the city landscape

III. Experimental Setup

This section outlines the immersive VR and non-VR display technologies, the equipment used to capture the objective metrics, the laboratory design, and the assessors who took part in the experiment as well as the screening process.

A. *Immersive Virtual Reality presentation equipment*

The immersive VR environment equipment employed was The Oculus Rift (OR) [9] Head Mounted Display (HMD) Development Kit 2, with version 6 SDK as per Fig. 2a. The OR has a field of view (FoV) of 100 degrees (typical viewing angle of a human eye is approximately 120 degrees). In the experiments conducted, the frame rate was 75 frames per second (fps). The resolution of the OR is 960 x 1080 per eye. Finally, the OR supports user's head movement tracking. This facilitates the users "to look" around the virtual scene based on their head movement.

The non-VR environment as per Fig. 2b was a conventional 2D monitor (screen size is 22 inches and the virtual scene is played at a resolution of 1600 x 900). The PC used was Windows 7 Professional, Intel Core™ i5 – 4590 CPU @ 3.30GHz, 10.0 GB RAM with a 4GB nVidia GTX 970 Graphics Card. The application displayed at 60 fps. To navigate around the virtual city, the users used a mouse connected to the PC. The users sat approximately 50–75 centimetres from the monitor.

For the non-VR and VR environments, a virtual city environment was created and rendered using Unity Game Engine. The scene for both conditions lasts for 2 minutes. The assessor will view this scene from a first person perspective as per Fig 3. During the two minutes, the assessor will be brought around the city using a predefined path that was created for this test. During the scene, ambient city sounds will be played through a set of Afterglow wireless headphones.

B. Objective Metric Capture Equipment

Two consumer devices were used for the capture of the objective metrics, i.e., HR and EDA. HR was captured using the Fitbit heart rate monitor [21] and is presented in Fig. 4(A). This device costs approximately €120. Heart rate was used as a metric as it offered a method determine user emotional arousal whilst experiencing the immersive experience. The Fitbit uses optical heart rate sensors on its strap to detect blood volume changes. Internal Fitbit algorithms then measure the users heart rate as outlined in [22]. The Fitbit provides readings once per second. The PIP Biosensor was used to measure participant's electrodermal activity (EDA) [23] and is presented in Fig. 4(B). This device cost €179. Electrodermal Activity (EDA) also known as skin conductance levels (SCL) or galvanic skin response (GSR) has been used in many studies to measure the emotional responses of users to game events [20]. To capture the data, the PIP was held between the thumb and index finger of the users hand during testing. The PIP provides readings approximately eight times per second.

C. Assessor Screening

All assessors were screened for visual acuity and red-green colour deficiencies. In terms of visual acuity, the Snellen chart procedure was used [24]. This involved assessors being placed 10 feet away from the Snellen chart. Their visual acuity was evaluated per eye by assessors identifying letters on the different scales. In terms of the red-green colour deficiencies screening, the Ishihara Colour Blindness Test [25] was employed. Assessors evaluated colour discs which required them to identify numbers and patterns. Additionally, if assessors suffered from any illnesses such as epilepsy there were excluded from the testing. One assessor did not pass the screening stage due to epilepsy and another was omitted based on the colour deficiencies screening.

D. The Laboratory Design

The design of the test lab was inspired by [27] and is in accordance with ISO standard [27]. One of the key aspects of this standard is the design of test rooms which reduce the effects that psychological factors and physical conditions can have on human judgment.

E. Participants

A total of 33 assessors took part in this study. They came from a variety of backgrounds: students, post graduate researchers and academic staff. Out of the 33, 19 were male and 14 were female. 3 of the assessors had used VR headsets previously. The average age of the assessors was 23 years with a range from 19 years to 30 years. A convenience sampling approach was taken to assessor recruitment. All assessors were screened as per the assessor screening process outlined in Section III.C.

IV. ASSESMENT METHODOLOGY

A. Test Protocol

On arrival to the building, the assessors were brought to the

Fig 4. (A). Fitbit Charge HR (B): PIP Biosensor.

waiting room where they provided with an information sheet on the assessment. Any questions were addressed and assessors were required to sign a consent form. Once the consent form was signed, assessors were screened according to Section III.C. In order to capture "resting" state in terms of the objective metrics captured, the Fitbit was placed on the assessor's and assessors were asked to hold the PIP biosensor. The assessor heart rate and EDA levels were measured for 5 minutes. They were then brought to the experimentation room where the testing was carried out. They were then asked to sit and watch the virtual environment that was presented to them, both on the conventional 2D display and also through the Oculus HMD (the ordering was randomized). The entire time for one interaction with the environment was 2 minutes. Upon completion, assessors completed the questionnaire based on the scales provided in Section IV.B. The entire testing time it took for a single subject was approximately 25 minutes. This consisted of information/screening (10mins), a period to obtain resting state (5mins), training (2mins) and evaluations (8mins (4*2mins (for VR and non-VR))).

B. Questionnaire and Rating Scale

The questionnaire used in the experiment had nine questions. The aim of the questionnaire was to evaluated two key aspects of user QoE for an immersive VR environment: immersion and usability. We considered the model proposed by Ebrahimi et al. [13] in definition of our model for immersive VR experiences. Each of the nine questions and associated answers are presented in Table I and Table II. Question 1 analysed how immersed the user felt in the two environments evaluated. This fits within the immersion factor of our model and we propose sits within content format, device and content influencing factors of Fig 1. Question 2 queried the users sense of presence in the environments evaluated. Again, this sits within the immersion aspect of our model and content format, device and content influencing factors of [13]. Question 3 asked assessors to rate their level of enjoyment of the scene. This transcended both the immersion and usability categories in our model and the content and device influencing factors in the Ebrahimi model. Question 4 requested assessor to rate the level of realism in the environments they experienced. This question feeds into both of our factors for QoE and into the content, device, and user expectation factors in [13]. Questions 5-7 were derived from [28] and are relevant to the usability categorization in our model. Question 8 asked the assessor to rate their experience based on their

TABLE I RATING SCALE FOR EACH OF THE IMMERSION QUESTIONS (LIKERT SCALE)

Questions	Q1 I was immersed in the environment		Q3 I enjoyed experiencing the virtual environment	Q4 The virtual environment was realistic	
Answers	• Strongly Disagree	• Disagree	• Neither Agree nor Disagree	• Agree	• Strongly Agree
Scale	1	2	3	4	5

Questions	Q2 I did not feel a strong sense of presence whilst experiencing the system				
Answers	• Strongly Disagree	• Disagree	• Neither Agree nor Disagree	• Agree	• Strongly Agree
Scale	5	4	3	2	1

TABLE II: RATING SCALE FOR EACH OF THE USABILITY QUESTIONS (LIKERT SCALE)

Questions	Q5 The system was easy to use		Q7 I would have liked more time in the virtual environment	Q9 I did not feel any discomfort while using the application	
Answers	• Strongly Disagree	• Disagree	• Neither Agree nor Disagree	• Agree	• Strongly Agree
Scale	1	2	3	4	5
Questions	Q6 I needed to learn a lot of things before I could get going with this system			Q8 My experience did not meet my expectations	
Answers	• Strongly Disagree	• Disagree	• Neither Agree nor Disagree	• Agree	• Strongly Agree
Scale	5	4	3	2	1

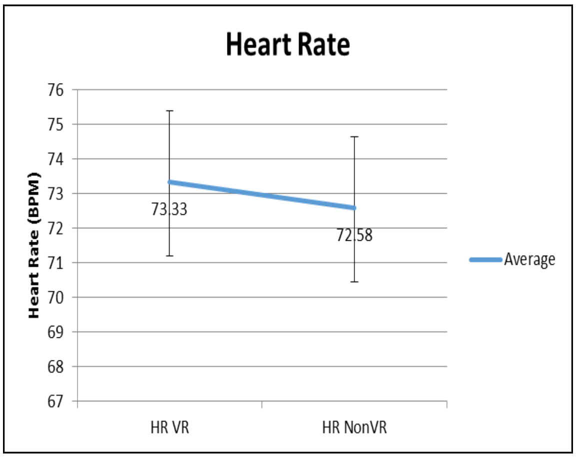

Fig. 5. Users Average Heart Rate Data during VR and Non-VR User QoE (with standard error)

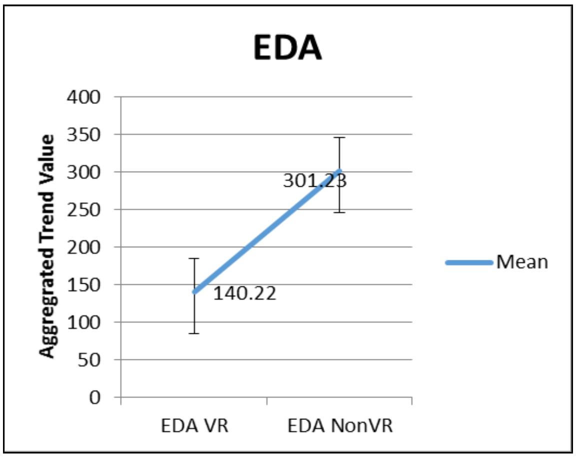

Fig. 6. Users Mean EDA Data during VR and Non-VR User QoE (with standard error)

expectations and this is a factor in both immersion and usability in QoE. For the model shown in Fig. 1 this question refers to user expectation, content, device, content format. Lastly, question 9 asks the assessor if they felt any discomfort while using the application. This is currently the key issue for HMDs and this is part of usability in our model for QoE. The content used in this experiment was designed with this factor in mind, i.e., users were taken slowly through the virtual city.

For all the questions, assessors were required to state their level of rating (ACR) method proposed in ITU-T P.910 [29]. They did this after observing each sample agreement on a 5 point Likert scale as per the absolute category rating.

V. RESULTS AND DISCUSSION

In this section of the paper we present and discuss our findings on the objective and subjective data captured.

A. *Objective Metrics: Heart Rate & EDA*

As outlined above, the physiological metrics employed during this study were heart rate captured by the Fitbit and EDA captured by the Pip Biosensor. Fig. 5 and Fig. 6 present the assessors average HR and EDA readings during their experience of the VR and non-VR environments.

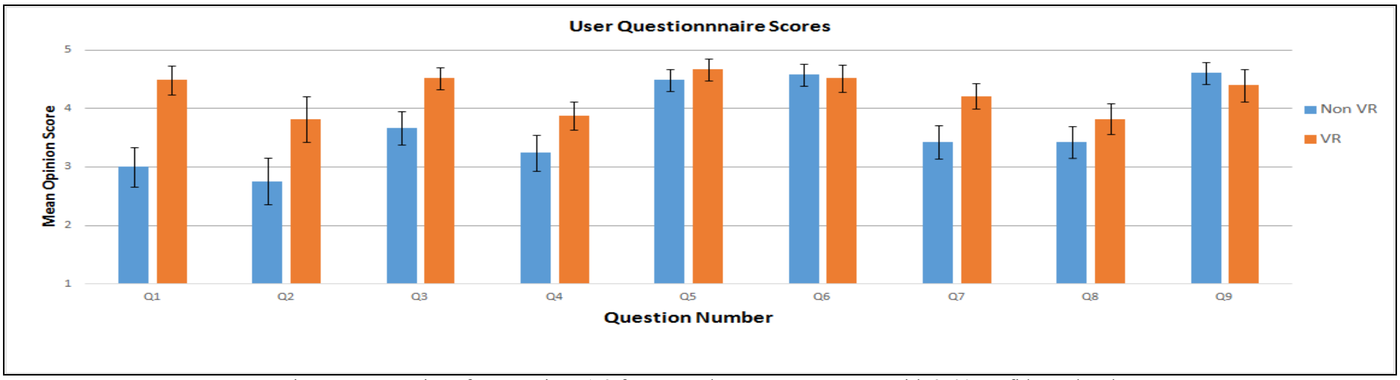

Fig. 7. MOS Ratings for Questions 1-9 for VR and Non-VR User QoE with 95% confidence level

An ANOVA with 95% confidence level was conducted to compare mean HR and EDA scores in VR and Non VR conditions. A significant effect was found for EDA $F(1, 30) = 12.04$, $p < .002$. However no significant effect was found for HR $F(1,30) = .16$, $p = .70$. With respect to EDA, the lower the value the higher the arousal level (EDA rating) of assessors. In this particular VR environment, these results suggests that EDA is more representative of user QoE than HR. The EDA analysis more closely reflected changes in QoE as represented by the self reported measures (presented in Section V.B). However further work is necessary to understand correlated patterns in the objective data and self-reported measures and is discussed in future work.

B. *Self Reported Questionnaire Results*

Fig. 7 presents the MOS ratings for each of the 9 questions and Table III presents the results of the statistical analysis employed using IBM SPSS. Since all assessors experienced all conditions, i.e., non-VR and VR, we employed an ANOVA with 95% confidence interval. The statistical analysis indicates that in seven of the nine questions, there were statistically significant between the user ratings of VR and non-VR experiences. In question 6 and 9 the results were not statistically significant. Question 6 aimed to evaluate the amount effort that was required from assessors to interact and use the VR and non-VR systems. Considering question 9, significant reports have highlighted issues with virtual reality HMDs with respect to them causing nausea and dizziness. As outlined above, the virtual city was specifically designed to ensure users were brought through the environment at a slow pace and, thus, reducing the cybersickness factor which causes the dizziness/nausea effects [30]. It has also been demonstrated in related works that the more immersed or higher their sense of presence users experience in VR environments, the less the nausea/dizziness effects exist.

In terms of the two questions where statistically significant differences exist but just so, question 5 and 8. Question 5 relates to ease of use of the VR environment w here assessor could look around the scene with their head movement. For the non-VR environment, assessors used a computer mouse. Both are very easy systems to use and therefore a big statistical significance was not expected. Question 8 relates the assessors experience to their expectations. This was a surprise as we would have suggested VR would have been much higher but maybe this is due to the fact this was most assessors first time ever using a VR device so they did not know what to expect.

For the remaining questions (S1, S2, S3, S4 S7), the results were emphatically significant and were still significant with 99% confidence levels. It is notable that four of these reside in the user level of immersion aspect of our model. Assessors clearly reported being more immersed and felt a greater sense of presence in the VR environment. Interestingly, assessors reported higher level of enjoyment wearing the HMD as opposed to the 2D display in particular since, they did not report from a usability perspective (S5 & S6) that managing their viewing angles required the use of the mouse interface in the non-VR environment.

The strength of the relationship between HR/EDA measures and questionnaire experience ratings was examined using bivariate correlations. The results found a significant medium strength positive correlation between EDA in Virtual Reality (VR) and Question 6, $r = .407$, $n = 33$, $p = .019$, indicating that high EDA values (low physiological response) were associated with less difficulty getting used to the system.

VI. Conclusion

This paper has presented a subjective and objective QoE evaluation of Immersive VR and non-VR environments. It is the first work which analysed the correlation between objective metrics of heart rate and electrodermal activity to user QoE of immersive virtual reality environments. These initial findings indicate a correlation between EDA and some of the self-reported measures. The elevated HR and EDA are correlated with physiological arousal and as such with some of the influencing factors associated with user QoE. The results also indicate higher QoE ratings whilst experiencing the environments with the HMD as opposed to the 2D display screen. Further subjective testing is required to facilitate regression analysis to analyse in greater depth the relationship between the subjective and objective data but also to consider order effects and gender influences. As such, we can further hypothesize which of the influencing factors enhance user QoE. Future work will also extend the immersive virtual environments to include mulsemedia components such as olfaction and haptic effects as a step towards understanding the effect on user QoE of a truly immersive multisensory multimedia system.

TABLE III STATISTICAL ANALYSIS OF SELF REPORTED MEASURES WITH 95% CONFIDENCE LEVEL

	NON VR	NON VR	VR	VR			
	MEAN	SD	MEAN	SD	F	df	Sig. (2-tailed)
Q1	3.000	1.000	4.4848	.71244	76.986	32	0.000
Q2	2.7576	1.17341	3.8182	1.13067	22.897	32	0.000
Q3	3.6667	0.81650	4.5152	0.56575	28.970	32	0.000
Q4	3.2424	0.90244	3.8788	0.69631	15.474	32	0.000
Q5	4.4848	0.56575	4.6667	0.54006	5.053	32	0.032
Q6	4.5758	0.56071	4.5152	0.66714	.000	32	1.000
Q7	3.4242	0.83030	4.2121	0.64988	42.250	32	0.000
Q8	3.4242	0.79177	3.8182	0.76871	5.485	32	0.026
Q9	4.6061	0.55562	4.3939	0.82687	2.435	32	0.129

VII. ACKNOWLEDGEMENTS

This research was supported by the Athlone Institute of Technology Presidents Seed fund (grant no. P221-037)

References

[1] M. Vaalgamaa amd B. Belmudez, "AudioVisual Communication", in "Quality of Experience: Advanced Concepts, Applications and Methods" (S. Muller and A. Raake, eds.), Springer, Heidelberg, Germany, pp. 195-121, 2014

[2] C. Alberti, D. Renzi, C. Timmerer, C. Mueller, S. Lederer, S. Battista and Marco Mattavelli. “Automated QoE evaluation of dynamic adaptive streaming over HTTP” In Fifth International Workshop on Quality of Multimedia Experience (QoMEX) pp. 58-63, 2013.

[3] C. Timmerer, M. Waltl, B. Rainer, and N. Murray, "Sensory Experience: Quality of Experience Beyond Audio-Visual", in "Quality of Experience: Advanced Concepts, Applications and Methods" (S. Muller and A. Raake, eds.), Springer, Heidelberg, Germany, pp. 351-365, 2014 DOI= 10.1007/978-3-319-02681-7_24

[4] G. Ghinea, S.R. Gulliver, and F. Andres (eds.), “Multiple Sensorial Media Advances and Applications: New Developments in MulSeMedia”, IGI Global, 2011.

[5] G. Ghinea, C. Timmerer, W. Lin and SR Gulliver. 2014. “Mulsemedia: State of the Art, Perspectives, and Challenges”. ACM Trans. Multimedia Comput. Commun. Appl. 11, 1s, Article 17 (October 2014), 23 pages. DOI=10.1145/2617994 http://doi.acm.org/10.1145/2617994

[6] N. Murray, B. Lee, Y. Qiao, G.M. Muntean. “Olfaction enhanced multimedia: A survey of application domains, displays and research challenges". In ACM Computing Surveys 48:4, 2016 DOI: http://dx.doi.org/10.1145/2816454.

[7] N. Murray, Y. Qiao, B. Lee, AK. Karunakar, G.M. Muntean, “Multiple-Scent Enhanced Multimedia Synchronization” In *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM),* vol. 11, Issue 1s, September 2014 Article No. 12 DOI=10.1145/2637293 http://dx.doi.org/10.1145/2637293

[8] F. Danieau, A. Lecuyer, P. Guillotel, J. Fleureau, N. Mollet, M. Christie, “Toward Haptic Cinematography: Enhancing Movie Experience with Haptic Effects based on Cinematographic Camera Motions”. In In *IEEE TRANSACTIONS ON Multimedia,* vol . 21, Issue 2, pp. 11-21, 2014

[9] https://www.oculus.com/en-us/ accessed on 10.03.2016

[10] https://www.playstation.com/en-us/explore/playstation-vr/ accessed on 10.03.2016

[11] https://www.google.com/get/cardboard/ accessed on 10.03.2016

[12] P. Le Callet, S. Möller, A. Perkis, “Qualinet White Paper on Definitions of Quality of Experience“. European Network on Quality of Experience in Multimedia Systems and Services (COST Action IC 1003), 2012

[13] T. Ebrahimi, “Quality of Multimedia Experience: Past, Present and Future”, ACM Multimedia Conference (MM’09), pp. 3-4, 2009.

[14] I Hupont, J. Gracia, L. Sanagustin, and M. A. Gracia, “How do new visual immersive systems influence gaming QoE? A use case of serious gaming with Oculus Rift,” in *2015 Seventh International Workshop on Quality of Multimedia Experience (QoMEX)*, 2015, pp. 1–6.

[15] M. Waltl, C. Timmerer, and H. Hellwagner, “Improving the Quality of multimedia Experience through sensory effects,” in 2010 Second International Workshop on Quality of Multimedia Experience (QoMEX), 2010, pp. 124–129.

[16] E. Kroupi, P. Hanhart, JS Lee, M. Rerabek, T. Ebrahimi. “Modeling Immersive Media Experiences by Sensing Impact on Subjects” Multimedia Tools and Applications doi>10.1007/s11042-015-2980-z , 2015.

[17] C. Timmerer, T. Ebrahimi and F. Pereira. “Toward a new assessment of quality” In IEEE Computer Society, Issue 3, vol. 48, pp. 108-110, 2015. doi.ieeecomputersociety.org/10.1109/MC.2015.89

[18] A.-M. Brouwer, M. A. Neerincx, V. L. Kallen, L. van der Leer, and M. ten Brinke, “EEG alpha asymmetry, heart rate variability and cortisol in response to virtual reality induced stress,” J. Cybertherapy Rehabil., vol. 4, no. 1, pp. 21–34, 2011.

[19] R. Blankendaal, T. Bosse, C. Gerritsen, T. de Jong, and J. de Man, “Are Aggressive Agents As Scary As Aggressive Humans?,” in *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, Richland, SC, 2015, pp. 553–561

[20] A Drachen, L.E. Nacke, G. Yannakakis, and AL Pedersen. “Correlation between Heart Rate, Electrodermal Activity, and Player Experience in FirstPerson Shooter games”. Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games, ACM New York, NY, USA (2010), 49-54.

[21] https://www.fitbit.com/chargehr accessed on 10.03.2016

[22] https://help.fitbit.com/articles/en_US/Help_article/Heart-rate-FAQs#How

[23] https://thepip.com/ accessed on 10.03.2016

[24] H. Snellen, Probebuchstaben zur Bestmmung der Sehschärfe P.W. vander Weijer, Utrecht, cited in L.A. Riggs, Visual Acuity, in: C.H. Graham (Ed.), Vision and Visual Perception. Wiley, New York, 1862

[25] J. Pokorny, B. Collins, G. Howett, R. Lakowski, M. Lewis, J. Moreland, H. Paulson, VC Smith, S. Shevell *Procedures for Testing Color Vision: Report of Working Group 41*. National Academies Press, 1981

[26] N. Murray, Y. Qiao, B. Lee, AK Karunakar, G.-M. Muntean, “Subjective Evaluation of Olfactory and Visual Media Synchronization” In Proceedings of *ACM Multimedia Systems conference*. Feb 26 - March 1, Oslo, Norway. 2013.

[27] ISO/IEC 8589 Sensory analysis – General guidance for the design of test rooms.

[28] P. W. Jordan, B. Thomas, I. L. McClelland, and B. Weerdmeester, *Usability Evaluation In Industry*. CRC Press, 1996. ISBN 9780748404605

[29] ITU-T P.910. Subjective video quality assessment methods for multimedia applications, 2008.

[30] M. E. McCauley and T. J. Sharkey, “Cybersickness: Perception of self-motion in virtual environments,” Presence Teleoperators Virtual Environ., vol. 1, no. 3, pp. 311–318, 1992.