Social Signal Processing in Affective Virtual Reality: Human Shaped Agents Increase Elecrodermal Activity in an Elicited Negative Environment

Sergio C. Torres, PhD, Desirée I. Garcia Laso, BsBME, Maria Eleonora Minissi, BAPsy, Luna Maddalon, BAPsy, Irene Alice Chicchi Giglioli, PhD, and Mariano Alcañiz, PhD

Introduction

Affect elicitation typically refers to the process of intentionally inducing positive or negative states to investigate various aspects of human emotions, including their nature, and associated behavioral and physiological responses[1]. A meta-analysis by Joseph et al.[2] supported the effectiveness of affect elicitation techniques, among which films, autobiographical recall, and written descriptions, are the most popular. A negativity bias was also confirmed so that negative stimuli show stronger elicitation effects than positive and neutral stimuli, particularly in women[3].

However, two aspects deserve special attention. On one hand, novel technologies such as virtual reality (VR) are becoming increasingly used for affect elicitation due to their capabilities for enhancing ecological validity as compared to traditional affect elicitation techniques[4,5]. Virtual reality is a technological system designed to simulate multimodal environments (e.g., visual, auditive, or kinaesthetic) that convey a high degree of sense of spatial presence (SoP[6]). For example, virtual environments (VEs) in form of “emotional parks” wherein characteristics such as luminosity, color, and sounds are manipulated, have been shown effective to induce positive and negative affective states[7–9]. In their review on VR for affect elicitation, Diniz et al.[10] highlight that, at an experiential level, affective responses are mostly evaluated through self-reports of emotional valence (the degree of pleasure or displeasure experienced in a specific situation) and arousal (high or low intensity). In addition, electrodermal activity (EDA) is the physiological measure of arousal most reported in VR contexts. EDA is defined as the changes in the electrical properties of the skin resulting from variations in the resistance of the epithelial tissue. This signal comprises the tonic and phasic components. The tonic activity represents slow changes produced over time and provides information about the basic skin conductance level (SCL). In contrast, the phasic activity reflects rapid variations that occur in response to a stimulus (skin conductance response, SCR[11]).

On the other hand, whereas stimuli used for affect elicitation are mostly designed and selected based on their valence/arousal qualities (or association with discrete emotions e.g., fear), other characteristics such as social content, are only scarcely investigated in this context. From a social signal processing account, that is highly relevant since affective states are at the core of "real world" social dynamics[12]. In this regard, previous research has reported, for instance, that films with social content are more arousing and elicit deeper visual attention than films without social content[13]. Pictures with social content (including people) have also been shown to elicit stronger EDA response than pictures with non-social content (without people[14]). Moreover, some findings suggest that this effect can be driven by pleasant rather unpleasant stimuli (i.e., *negativity EDA bias hypothesis*), due to the activation of the motivational appetitive system (i.e., *appetitive EDA bias hypothesis*[15]).

It is quite unclear, however, whether and how "social" VEs, particularly focused on affect elicitation, show this type of physiological arousal response-pattern too. The work here deals with this question from a "social-emotional elicitation" perspective by primarily comparing EDA responses of social versus non-social affective VEs. It can be expected that social contexts will trigger stronger physiological EDA responses, in general. Nonetheless, it is also investigated whether the potential effects support either the negativity or the appetitive EDA bias hypothesis.

Materials and methods

Participants

A total of n = 72 participants (55% women) aged 18-53 (M_{age} = 33.6 *SD* = 9.3) were recruited for the study. Inclusion criteria addressed being over 18 years old and right-handed. Exclusion criteria addressed: a) being pregnant or breastfeeding; b) having severe neurological or cardiac disorders, mental disorders, or severe psychotic symptoms (such as hallucinations or delusions); c) psychotropic drugs use (including antidepressants, antipsychotics, anxiolytics, and mood stabilizers); d) being undergoing psychotherapy, and e) scoring > 8 for the Beck Depression Inventory-II (BDI-II[16]) and > 20 for the Liebowitz Scale for Social Anxiety (LSAS[17]). Participants received written information about the study and signed informed consent before the experimental testing. The research was completed in accordance with the Declaration of Helsinki as revised in 2013, and was approved by the ethics committee of the Polytechnic University of Valencia (UPV; P2_18_06_19).

Psychological assessments

Depression Symptoms. Beck Depression Inventory-II[16] consists of 21 items (scored from 0 to 3) designed to assess the severity of depressive symptomatology.

Social Anxiety. Liebowitz Scale for Social Anxiety[17] assesses social anxiety through 24 items rated on a 4-point Likert Scale. It consists of 2 subscales: performance anxiety (13 items), and social anxiety (11 items).

Self-reports of emotional valence and arousal. Self-Assessment Manikin (SAM[18]). Participants reported emotional valence and arousal via two related 9-point Likert scales. Valence scores ranged from 1 (very unpleasant) to 9 (very pleasant). Arousal scores ranged from 1 (very relaxed) to 9 (very activated).

Sense of Presence (SoP). Inventory of Sense of Presence (ITC-SOPI[19]). The instrument assesses four dimensions: negative VR effects (6 questions), engagement with the content (13 questions), ecological validity (5 questions), and the one targeted in the study, the sense of spatial presence (SoP; 19 questions). A 5-point Likert scale is used for evaluation, with 1 ("Strongly disagree") and 5 ("Strongly agree").

Physiological arousal

Electrodermal activity (EDA) was recorded before and during the VR using the Shimmer3 GSR+ Unit sensor (Shimmer Sensing, Dublin, Ireland) recording at 128 Hz and 0,001 – 100 µS. The device was placed on the participants' non-dominant wrist. Its electrodes were placed on the second phalanx of the ring and middle fingers of the same hand. The data was preprocessed using the Ledalab tool (version 3.4.8) included in the Matlab R2011b software. The preprocessing was divided into two successive stages: initially, a low-pass Butterworth filtering with a cutoff frequency at 2.5 Hz[20] was performed and then artifacts were corrected through a Continuous Decomposition Analysis (CDA). Subsequently, the signal was divided into tonic and phasic components.

Social-Emotional Virtual Environments

The virtual environments (VE) in the study were developed with the software Unity 5.5. They built upon valence-laden scenarios (positive, negative, and neutral), which narrative developed within a city park[7]. The park was square-shaped and surrounded by buildings. A bandstand was displayed in the middle. The non-social parks included different elements (e.g., streetlamps, and plants) which features were manipulated according to the intended affective elicitation (e.g., light, colors, or sounds). Furthermore, three additional park versions were developed to explicitly convey social meaning by adding non-responsive virtual agents (human-shaped pre-programmed characters). Concretely, in the negative park, a mother showed up scolding his son because it was raining, and he didn't want to leave the park. Afterward, a man with a black hoodie, characterized by an unfriendly attitude, appeared abruptly, and sat on a bench (see Figure 1a). In the positive park, a couple of runners appeared reinforcing each other (e.g., "well done"). In addition, a father and his son appeared playing and laughing. They approached the screen to say "hello" (see Figure 1b). In the neutral park, a woman walked across the park. Then, two young men appeared and sat for a while on the bandstand stairs. One of them asked the other for the time (see Figure 1c; see details in Torres et al.[21]) Therefore, a total of six VEs (3 with virtual agents and 3 without them) were displayed.

(Figures 1a-1c)

Experimental procedure

Participants were invited to a three-screen CAVE™ (Cave Assisted Virtual Environment) room. This room conformed to a semi-immersive three wall setting (4x4x3 meters) with three video projectors with ultra-short-throw lens (Optoma HD35UST; Optoma, Alicante, Spain) connected to a computer (Intel Core i7-7700 CPU @3.6GHZ with dual DVI output Nvidia Geforce GTX 1060 6GB). The Logitech Z906 500W 5.1 THX (Logitech, Canton Vaud, Switzerland) was used as the sound system. Participants sat on a chair at the center of the room and were equipped with the EDA sensors. Before experiencing the VE, EDA was recorded at the baseline for each participant. To do so, they looked at to the (empty) central wall with their eyes closed for 60 seconds and subsequently with their eyes open for another 60 seconds. Then, participants filled demographic information and read the instructions before experiencing the VE. The VEs presentation order was randomized according to their valence category (positive, negative, and neutral). The social and non-social VE versions were displayed sequentially within each affective category, but their order was counterbalanced across participants (e.g., positive social – positive non-social, or vice versa). Each version lasted 30 seconds with an inter-stimulus interval of 1 second with a black screen. The EDA was recorded without interruption to minimize artifacts and enable observing more accurately, potential signal changes between the two versions. After each affective VE block, including social and non-social, participants rated their felt emotional valence and arousal with the SAM scales. EDA was recorded across the whole experimental session. At the end of the experiment, participants answered the SOPI-ITC questionnaire.

Analyses

Data analyses were performed by means of the Statistical Package for Social Sciences (SPSS v5). Specifically, self-reports of valence, arousal, and EDA were analyzed from a generalized linear mixed model (GLMM) framework. Specifically, the data was analyzed with a log-link function and a robust standard errors (SE) estimator. Subjects' id was used as a random factor with random intercepts. For valence reports, the analysis targeted the fixed factor VE (positive, negative, and neutral). This analysis served a manipulation check of affect elicitation. Arousal reports were analyzed as for emotional valence. EDA analyses were submitted to a 3 x 2 design with VE and the factor Context (social vs. non-social) manipulated within participants. Gender was included as a control covariate in all analyses together with standardized age and SoP scores.

Results

Self-reported emotional valence

The covariates gender, $F(1, 210) = 2.79$, $p = .097$, and age $F(1, 210) = .92$, $p = .34$, did not show statistical significance. In contrast, there was a general incremental SoP effect on valence $F(1, 210) = 15.55$, $p < .001$ (b = .082, CI[.041, .122]; $\eta_p^2 = .07$). As expected, the factor VE $F(2, 210) = 86.39$, $p <.001$ ($\eta_p^2 = .45$) was highly significant. On average, the positive VE elicited higher valence ratings, $M_{\text{positive}} = 7.45$; CI (7.11, 7.81), than the neutral VE, $M_{\text{neutral}} = 5.81$, CI (5.44, 6.20), and the negative VE, $M_{\text{negative}} = 2.80$, CI (2.43, 3.24; see Figure 2a).

Self-reported arousal

Self-reported arousal was analyzed as for emotional valence. Again, gender $F(1, 210) = .19$, $p = .66$, and age $F(1, 210) = .15$, $p = .70$, did not show statistical significance. On the other hand, SoP showed again an incremental effect on arousal. SoP $F(1, 210) = 4.14$, $p = .043$ (b = .101, CI[.003, .198]; $\eta_p^2 = .02$). Finally, the factor VE F (2, 210) = 23.40, $p = .043$ ($\eta_p^2 = .18$) was significant, too. On average, the negative VE elicited higher arousal ratings, $M_{\text{negative}} = 5.66$; CI (5.17, 6.20), followed by the positive, $M_{\text{positive}} = 4.25$, CI (3.73, 4.85), and the neutral, $M_{\text{neutral}} = 3.24$, CI (2.81, 3.74; see Figure 2b).

(Figure 2a-2b)

Electrodermal Activity (EDA)

Data inspection revealed some inconsistencies in the registered data (due to a bad connection or failure to record) leading to the data exclusion of n = 14 participants. After being processed, the EDA data of the remaining n = 58 participants were segmented in the time intervals corresponding to the baseline (120s) and each VE (6x30s; 180s in total). The analyses primarily targeted individual SCR (phasic signal) means (baseline-corrected)[22]. The model included the fixed factors VE, Context, and their interaction. These factors were modeled with a diagonal covariance type for repeated measures. The results did not show a significant main effect neither of the covariates gender $F(1, 299) = 1.35$, $p = .24$, age $F(1, 299) = 1.85$, $p = .17$, and SoP $F(1, 299) = 1.04$, $p = .31$ nor the VE factor $F(2, 299) = 2.41$, $p = .091$.

In contrast, the factor Context showed statistical significance $F(1,299) = 5.57$, $p = .019$, $\eta_p^2 = .02$. Specifically, social VEs elicited higher phasic activation, $M = 1.90$, CI(1.68, 2.16), on average, than non-social VEs, $M = 1.55$, CI(1.40, 1.72). In contrast, the factor Context showed statistical significance $F(1,299) = 5.57$, $p = .019$, $\eta_p^2 = .02$. Specifically, social VEs elicited higher phasic activation, $M = 1.90$, CI(1.68, 2.16), on average, than non-social VEs, $M = 1.55$, CI(1.40, 1.72). The result-pattern of the social context resembled the one of the self-reported arousal (see Figure 3). Interestingly, the most revealing finding was characterized by the interaction between VE and Context, $F(2, 299) = 4.87$, $p = .008$; $\eta_p^2 = .03$ (see Figure 3). Post-hoc comparisons (Bonferroni-Holm corrected) indicated that, VE did not reach statistically significant differences within the social and non-social contexts. Only the non-social neutral VE showed higher EDA than non-social positive, $t(299) = 2.63$, $p = .03$. However, when comparing the social (vs. non-social) VEs, the negative revealed the stronger differences $F(1, 299) = 12.26$, $p <.001$ ($\eta_p^2 = .04$). Positive social (vs. non-social) VEs showed a higher activation tendency, $F(1, 299) = 3.11$, $p =.078$ ($\eta_p^2 = .01$), and neutral social (vs. non-social)

VEs did not show significant differences, $F(1, 299) = .91, p = .34$ (see Table 1). Supplementary material shows analyses before exclusions.

(Figure 3)

(Table 1)

Discussion

This study primarily investigated differential patterns of physiological arousal by skin conductance response (SCR) within the framework of social-emotional elicitation through virtual reality (VR[5]). Specifically, positive, negative, and neutral "emotional parks" were presented with human-shaped virtual agents (social context) and without them (non-social context). In general, the VEs elicited the intended affective response in terms of self-reports of emotional valence. Moreover, negative VEs elicited higher ratings of emotional arousal than positive and negative. This finding is congruent with a negativity (self-report) bias observed in affect research whereby negative stimuli typically trigger higher perceived arousal[3,21]. Noteworthy, the most important finding in the study addresses physiological arousal. Social VEs showed, on average, stronger EDA responses than non-social VEs. Prior research using affective pictures in clinical contexts reported similar effects[14]. According to literature tapping on social signal processing, a potential mechanism behind this effect refers to the fact that social content triggers deeper visual attention than non-social content. For example, Rubo & Gamer[13] reported evidence in this direction using eye-tracking on affective films.

However, the main effect of social content found in the study was qualified by an interaction with the VEs affective category. Specifically, the effect of social context was mainly driven by differences in the negative VEs, thus, supporting a *negativity EDA bias hypothesis.* Social positive VEs only showed a higher EDA tendency than non-social positive VE, and neutral VEs did not show significant EDA differences. These findings partially contrast with Kosonogov et al.[15]. Their study using affective pictures showed a similar result-pattern regarding self-reports of emotional valence and arousal as for the study here. However, their reported EDA effect was driven only by social positive pictures. That was explained in the light of motivational systems wherein positive stimuli activate appetitive tendencies, increasing the level of arousal (*appetitive EDA bias hypothesis*).

We argue that these contrasting findings might be explained by at least two aspects. Contrary to static pictures, the VEs in our study were characterized by first, acoustic and visual features, and second, dynamic scenes. These aspects might have stimulated a deeper sense of presence (SoP) when experiencing the VEs, which, in turn, increased the negativity (attentional) bias. In a situation of stress, tension, or anxiety, as when experiencing a negative scene, sweat secretion increases, and it becomes a conductor, reducing the skin resistance[23]. The novelty in our study is that, in addition, social stimuli, particularly negative, increased that physiological response.

Nevertheless, some limitations should be pointed out. Experiential valence wasn't assessed neither at the baseline nor after each emotional park version (social and non-social). These assessments would be necessary to enable proper affect or mood induction protocols (MIP) aimed at comparing changes in affective states[10]. Moreover, the study cannot derive conclusions regarding high vs low arousing social content. For example, virtual environments intended to elicit fearful vs sadness or excited vs calm states[7]. In this line of reasoning, future studies should include assessments of discrete emotions. On the other hand, our study was based on visual and acoustic modalities. Interestingly, from the perspective of social neuroscience, additional sensory modalities should be considered in future studies. For example, the sense of olfaction is particularly connected to the amygdala and the orbitofrontal cortex, thus providing a close link with the limbic system, which is directly related to emotion processing[24]. Therefore, investigating the role of body odors in emotional responses to affective social stimuli may be of great value in the current framework.

Conclusions

The study sheds light on the potential of VR for investigating human emotions and the impact of social content on physiological EDA responses. Understanding the interplay between social-emotional stimuli and physiological arousal in virtual environments can contribute to a better understanding of social-emotional processing.

References

1. Coan JA, Allen JJB. Handbook of Emotion Elicitation and Assessment. Oxford University Press, USA; 2007. 492 p.
2. Joseph DL, Chan MY, Heintzelman SJ, Tay L, Diener E, Scotney VS. The manipulation of affect: A meta-analysis of affect induction procedures. Psychological Bulletin. 2020;146:355-75.
3. Cacioppo JT, Berntson GG. The Affect System: Architecture and Operating Characteristics. Curr Dir Psychol Sci. octubre de 1999;8(5):133-7.
4. Parsons TD, Gaggioli A, Riva G. Virtual Reality for Research in Social Neuroscience. Brain Sciences. abril de 2017;7(4):42.
5. Somarathna R, Bednarz T, Mohammadi G. Virtual Reality for Emotion Elicitation – A Review. IEEE Transactions on Affective Computing. 2022;1-21.
6. Cipresso P, Giglioli IAC, Raya MA, Riva G. The Past, Present, and Future of Virtual and Augmented Reality Research: A Network and Cluster Analysis of the Literature.

Frontiers in Psychology [Internet]. 2018 [citado 5 de junio de 2023];9. Disponible en: https://www.frontiersin.org/articles/10.3389/fpsyg.2018.02086

7. Baños RM, Liaño V, Botella C, Alcañiz M, Guerrero B, Rey B. Changing Induced Moods Via Virtual Reality. En: IJsselsteijn WA, de Kort YAW, Midden C, Eggen B, van den Hoven E, editores. Persuasive Technology. Berlin, Heidelberg: Springer; 2006. p. 7-15. (Lecture Notes in Computer Science).

8. Felnhofer A, Kothgassner OD, Schmidt M, Heinzle AK, Beutl L, Hlavacs H, et al. Is virtual reality emotionally arousing? Investigating five emotion inducing virtual park scenarios. International Journal of Human-Computer Studies. 1 de octubre de 2015;82:48-56.

9. Jicol C, Wan CH, Doling B, Illingworth CH, Yoon J, Headey C, et al. Effects of Emotion and Agency on Presence in Virtual Reality. En: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems [Internet]. New York, NY, USA: Association for Computing Machinery; 2021 [citado 5 de junio de 2023]. p. 1-13. (CHI '21). Disponible en: https://dl.acm.org/doi/10.1145/3411764.3445588

10. Diniz Bernardo P, Bains A, Westwood S, Mograbi DC. Mood Induction Using Virtual Reality: a Systematic Review of Recent Findings. J technol behav sci. 1 de marzo de 2021;6(1):3-24.

11. Vahey R, Becerra R. Galvanic skin response in mood disorders: A critical review. Research outputs 2014 to 2021 [Internet]. 1 de enero de 2015; Disponible en: https://ro.ecu.edu.au/ecuworkspost2013/1350

12. Cacioppo JT, Cacioppo S. Social Neuroscience. Perspect Psychol Sci. noviembre de 2013;8(6):667-9.

13. Rubo M, Gamer M. Social content and emotional valence modulate gaze fixations in dynamic scenes. Sci Rep. 28 de febrero de 2018;8(1):3804.

14. Martínez-Velázquez ES, Honoré J, de Zorzi L, Ramos-Loyo J, Sequeira H.

Autonomic Reactivity to Arousing Stimuli with Social and Non-social Relevance in Alexithymia. Frontiers in Psychology [Internet]. 2017 [citado 5 de junio de 2023];8. Disponible en: https://www.frontiersin.org/articles/10.3389/fpsyg.2017.00361

15. Kosonogov V, Sanchez-Navarro JP, Martinez-Selva JM, Torrente G, Carrillo-Verdejo E. Social stimuli increase physiological reactivity but not defensive responses. Scandinavian Journal of Psychology. 2016;57(5):393-8.

16. Beck AT, Steer RA, Brown G. Beck Depression Inventory–II [Internet]. American Psychological Association; 2011 [citado 5 de junio de 2023]. Disponible en: http://doi.apa.org/getdoi.cfm?doi=10.1037/t00742-000

17. Liebowitz MR. Social phobia. Modern Problems of Pharmacopsychiatry. 1987;22:141-73.

18. Bradley MM, Lang PJ. Measuring emotion: The self-assessment manikin and the semantic differential. Journal of Behavior Therapy and Experimental Psychiatry. 1 de marzo de 1994;25(1):49-59.

19. Lessiter J, Freeman J, Keogh E, Davidoff J. A Cross-Media Presence Questionnaire: The ITC-Sense of Presence Inventory. Presence: Teleoperators and Virtual Environments. 1 de junio de 2001;10(3):282-97.

20. Dessai A, Virani H. Emotion Detection Using Physiological Signals. En: 2021 International Conference on Electrical, Computer and Energy Technologies (ICECET). 2021. p. 1-4.

21. Torres SC, Gracia Laso D, Minissi ME et al. Modulating Virtual Affective Elicitation by Human Body Odors: Advancing Research on Social Signal Processing. D. D. Schmorrow and C. M. Fidopiastis (Eds.): HCII 2023, LNAI 14019, pp. 1–11, 2023.

22. Horvers A, Tombeng N, Bosse T, Lazonder AW, Molenaar I. Detecting Emotions through Electrodermal Activity in Learning Contexts: A Systematic Review. Sensors. enero

de 2021;21(23):7869.

23. Sanchez-Comas A, Synnes K, Molina-Estren D, Troncoso-Palacio A, Comas-González Z. Correlation Analysis of Different Measurement Places of Galvanic Skin Response in Test Groups Facing Pleasant and Unpleasant Stimuli. Sensors. enero de 2021;21(12):4210.

24. de Groot JHB, Semin GR, Smeets MAM. On the Communicative Function of Body Odors: A Theoretical Integration and Review. Perspect Psychol Sci. 1 de marzo de 2017;12(2):306-24.

25. Leys C, Ley C, Klein O, Bernard P, Licata L. Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. Journal of Experimental Social Psychology. 1 de julio de 2013;49(4):764-6.

Table 1. Estimated marginal means of phasic EDA signal linked to social and non-social affective VEs.

				95% CI	
VR	Context	M	SE	lower	upper
Negative	Social	2.34	.22	1.93	2.82
	Non-Social	1.47	.13	1.23	1.76
Positive	Social	1.74	.23	1.34	2.25
	Non-Social	1.30	.10	1.12	1.51
Neutral	Social	1.70	.17	1.40	2.07
	Non-Social	1.96	.22	1.56	2.45

Note. Sensitivity analyses revealed a rather homoscedastic residual distribution, which may bias the results interpretation. To test the model robustness, a residual exclusion criterion was set of 2.5 Median absolut deviations[25]. Supplemental Table S1 shows the results before the exclusion.

Figure 1a-1c. Negative social park (1a left side). Positive social park (1b centre). Neutral social park (1c right side). Non-social parks omit the virtual agents.

Figure 2a-2b. Marginal means of self-reported emotional valence (left part) and arousal (right part) regarding the three affective blocks. VE = Virtual Environment

Figure 3. Interaction between VE affective categories and social Context. SCR = Skin Conductance Response; VE = Virtual Environment