Reliability of the Empatica E4 wristband to measure electrodermal activity to emotional stimuli.

Adrian Borrego[1], Jorge Latorre[1, 2], Mariano Alcañiz[1], Roberto Llorens[1, 2]

[1]*Neurorehabilitation and Brain Research Group, Instituto de Investigación e Innovación en Bioingeniería*
Universitat Politècnica de València
Valencia, Spain.

[2]*NEURORHB.*
Servicio de Neurorrehabilitación de Hospitales Vithas
Valencia, Spain.

***Abstract*— The electrodermal activity has been used in research as a physiological measure of the activity of the autonomic nervous system when patients face emotional stimuli. Laboratory-grade systems that register the galvanic skin response include sensors that are wired to expensive amplifiers, which can limit its portability and accessibility for non-research centers. In the last years, different devices, mostly wireless wristbands, have come into the market at lower prices. While their portability and lower cost are interesting features for virtual reality applications, their reliability still remains unclear. The objective of this study was to determine the reliability of a wireless wristband, the Empatica E4, to register changes in the galvanic skin response when facing emotionally-valenced images in comparison to a laboratory-grade system. Our data showed that correlation between both systems were low to moderate for negative and positive-valenced images, and non-existent for neutral images. These results suggest that the E4 should be used with caution when assessing changes in emotionally-valenced images and should be avoided for neutral images.**

Keywords*— *electrodermal activity, galvanic skin response, emotional valence, sensors, virtual reality.

I. INTRODUCTION

Electrodermal activity (EDA), also referred to as galvanic skin response or skin conductance [1], describes all electrical phenomena in skin, including active and passive activity, which depends on the amount of secretion that is brought to the ducts and on the number of sweat glands that are activated. These responses are mediated by the autonomic nervous system. For this reason, EDA has been traditionally used to measure emotional valence, it is, the intrinsic attractiveness (positive valence) or aversiveness (negative valence) of a stimulus [2], and arousal. Immersive stimulation, as that facilitated by virtual reality (VR) could increase emotional perceptions [3] and, consequently, the interest of investigating EDA under this condition. Acuallly, use of EDA is not strange in VR, and has involved different domains, such as cyberpsychology [4].

Common instrumentation to register the EDA signal uses Ag-AgCl electrodes attached to the palmar medial phalanges of the index and middle fingers wired to amplifiers [1]. Laboratory-grade systems as the BIOPAC (BIOPAC Systems, Inc., Goleta, CA, USA) or the Refa system (Twente Medical Systems International B.V., Zutphenstraat, Nederland), use this configuration. However, their high cost, low portability, and restriction of movements limit their use in real-life situations, difficult to replicate in research laboratories. In the last years, an increasing number of portable wireless devices have come into the market at lower prices. These new devices usually consist on armband or bracelets that combine different sensors that measure not only EDA, but also the skin temperature, heart rate, and movement. The portability and reduced cost of these devices could spread their use outside the laboratory allowing unobtrusive assessment and releasing the fingers, which is particularly interesting in VR, where fingers are usually needed to deal with controllers. However, the reliability of these devices remains unclear [5].

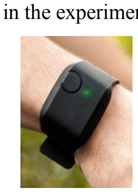
The aim of this study was to determine the reliability of a wireless armband to detect changes in EDA due to emotional valence in comparison to a laboratory-grade system.

II. METHODS

A total of 59 individuals with a mean age of 33.3±7.5 years old (35 men and 24 women) participated in the study. All of the participants provided written consent to participate in the study.

A wireless wristband, the E4 (Empatica, Milan, Italy), and a laboratory-grade device, the Refa system, were used to register the EDA (Fig.1). The E4 consists of two sensors located in the strap of the wristband and allows acquisition of data at 4 Hz. The Refa system consists of two Ag-AgCl sensors that are wired to an external amplifier and allows acquisition at 256 kHz.

Fig. 1. Devices used in the experiment: a) E4; and b) Refa system



a)

b)

Stimuli consisted of a series of pictures extracted from the IAPS, images whose emotional valence has been categorized database [6]. Four negative, four positive, and eight neutral images were used in this study. A sequence of 16 images was designed alternating negative and positive images, separated by neutral images. Duration of each image was 12 s. A black image of the same duration was included at the beginning so that the total duration of the sequence was 204 s (3 minutes and 24 s) (Fig.2).

The experiment was conducted in a dedicated temperature-controlled room free of distractors in our laboratory. The setup

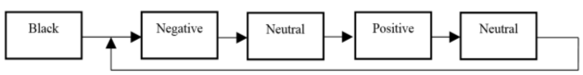
Note: This is not the official copyright released version of the IEEE proceedings paper. When citing this paper, use the following format: Borrego A, Latorre J, Raya Alcaniz M and Llorens R, "Reliability of the Empatica E4 wristband to measure electrodermal activity to emotional stimuli", Proc. 13th Int'l Conf. on Virtual Rehab., WG Wright, S Subramanian, G Fluet, M Agmon, RM Proffitt, M Roberts (Eds), Tel Aviv, Israel, 21-24 July 2019.

consisted of a conventional table with a 55” TV screen on it, and a chair with armrests in front of the table. Participants were asked to comfortably sit in the chair facing the screen. Participants were equipped with the sensors and were briefly introduced to the objective of the experiment. After this, participants were asked to rest their arms on the table and remain still for 5 minutes. After this time, the experiment started and participants were asked to “look at the pictures”.

The signals provided by both the E4 and the Refa systems were low-pass filtered at 1.99 Hz to satisfy the Nyquist frequency imposed by the 4 Hz sampling frequency of the E4. Both tonic and phasic components of the EDA were taken into account [5]. Signals were divided into time slots of 12 s according to the duration of the images. For all the time slots, the change in the EDA was estimated as the difference between the mean amplitude of the signal in the last two seconds and the mean amplitude of the signal in the first two seconds. The change in the EDA of negative, positive, and neutral images was estimated averaging the differences in those images.

The correlation between the detected change by both systems was computed using the Spearman rank correlation coefficient. The α level was set at 0.05 for all analyses (two-sided).

Fig. 2. Flowchart of the sequence. The loop was repeated four times.



III. RESULTS

The change detected by both systems showed low to moderate correlations for both positive (r=0.554, p=0.000) and negative images (r=0.449, p=0.000) while no significant correlation was found for neutral stimuli (r=0.062, p=0.639) (Table I).

TABLE I. MEAN CHANGE IN THE GALVANIC SKIN RESPONSE

	Negative	**Positive**	**Neutral**
E4 (µS)	0.009	0.007	0.000
Refa (µS)	-0.009	0.016	-0.078

IV. DISCUSSION

Previous research on EDA has identified an increase in this response for positive and negative stimuli, and a decrease for neutral stimuli [8]. In our study, the average changes detected by both devices for positive images are in line with these reports. However, while the E4 also detected an average increase of the EDA for negative images, thus supporting previous reports, the Refa system registered the opposite tendency. The biggest differences between systems were detected for neutral stimuli. As expected from previous studies, an average decrease in the EDA was detected by the Refa system. However, the E4 failed to detect these changes and remained unaltered.

Even though both devices showed variable performance in average, low to moderate correlations were detected for positive and negative images. These results seem to depict weaker relationships than those reported by the only previous study that analyzed the correlation between EDA records retrieved by different devices [8]. These differences might be due, in part, to the different sampling rates of the devices. It is important to highlight that the sampling rate of the E4 could be too low to register the phasic component of the EDA, which encodes changes related to arousal. This component of the EDA has been suggested to range from 0.05-1.5 Hz [9] or even 0.05-5 Hz [10], which would lead to a minimum sampling rate of 3 Hz or 10 Hz, respectively. According to this latter case, the sampling rate of the E4 (4 Hz) would not be sufficient to capture the phasic component of the EDA, which would restrict the E4 to measure changes in the tonic component.

Although wearable wireless devices are particularly interesting in VR, the low to moderate correlations detected between the E4 and the laboratory-grade Refa system, and the low sampling rate of the E4 support caution in using this device to measure changes in EDA to emotionally valence-stimuli.

FUNDING

This study was partially funded by project VALORA, grant 201701-10 of the Fundació la Marató de la TV3 (Barcelona, Spain), grant BES-2014-068218 of the Ministerio de Economía y Competitividad of Spain (Madrid, Spain), and grant PAID-06-18 of Universitat Politècnica de València (Valencia, Spain).

REFERENCES

[1] W. Boucsein *et al.*, “Publication recommendations for electrodermal measurements,” *Psychophysiology*, vol. 49, no. 8, pp. 1017–1034, 2012.

[2] W. M. Winton, L. E. Putnam, and R. M. Krauss, “Facial and autonomic manifestations of the dimensional structure of emotion,” *J. Exp. Soc. Psychol.*, vol. 20, no. 3, pp. 195–216, 1984.

[3] S. Estupiñán, F. Rebelo, P. Noriega, C. Ferreira, and E. Duarte, “Can virtual reality increase emotional responses (Arousal and Valence)? a pilot study,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2014.

[4] T. D. Parsons, *Cyberpsychology and the Brain: The Interaction of Neuroscience and Affective Computing*. 2017.

[5] B. Figner and R. O. Murphy, “Using skin conductance in judgment and decision making research,” *A Handb. Process Tracing Methods Decis. Res. A Crit. Rev. User’s Guid.*, pp. 163–184, 2010.

[6] B. N. Lang, P.J., Bradley, M.M., & Cuthbert, “International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8.,” 2008.

[7] G. K. Levenston, C. J. Patrick, M. M. Bradley, and P. J. Lang, “The psychopath as observer: emotion and attention in picture processing.,” *J. Abnorm. Psychol.*, vol. 109, no. JULY 2000, pp. 373–385, 2000.

[8] D. Estepa Sanz, “Desarrollo de un sistema de detección de espurios a partir del hardware del anillo de Usenns,” 2015.

[9] J. Schumm, M. Bächlin, C. Setz, B. Arnrich, D. Roggen, and G. Tröster, “Effect of movements on the electrodermal response after a startle event,” in *Proceedings of the 2nd International Conference on Pervasive Computing Technologies for Healthcare 2008, PervasiveHealth*, 2008, pp. 315–318.

[10] C. Kappeler-Setz, F. Gravenhorst, J. Schumm, B. Arnrich, and G. Tröster, “Towards long term monitoring of electrodermal activity in daily life,” *Pers. Ubiquitous Comput.*, vol. 17, no. 2, pp. 261–271, 2013.