2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)
October 11-14, 2020. Toronto, Canada

Validity Testing the NeuLog Galvanic Skin Response Device

Theresa Flagler
Department of Psychology
Center for Vision Research
York University
Toronto, Canada
tflagler@yorku.ca

Jonathan Tong
Department of Psychology
Center for Vision Research
York University
Toronto, Canada
tongj86@yorku.ca

Robert S. Allison
Department of Computer Science and Engineering
Center for Vision Research
York University
Toronto, Canada
allison@eecs.yorku.ca

Laurie M. Wilcox
Department of Psychology
Center for Vision Research
York University
Toronto, Canada
lwilcox@yorku.ca

***Abstract*—This paper describes validity testing of the NeuLog NUL-217 GSR measurement device. This was accomplished by comparing the NeuLog device to readings from the Biopac Student Lab Systems EDA system. The results of this research found that measurements from the NeuLog device are correlated with and comparable to the Biopac system. The absolute skin conductance levels typically differed between the two systems. For most psychological and human factors research the dynamic skin conductance responses are more important than absolute levels. The timing and relative magnitude of changes typically tracked well across the two systems indicating that the Neulog system is suitable for the purposes of its intended use in psychological and technological research.**

Keywords—galvanic skin response, electrodermal activity, virtual reality, head-mounted devices, simulator sickness

I. INTRODUCTION

Electro-Dermal Activity (EDA) measuring devices, also known as Galvanic Skin Response (GSR) devices, are sensors which measure the level of electrical activity in the skin. Most typically, the current produced in response to a controlled voltage between the electrodes is measured to estimate the skin conductance, which directly relates to the amount of activity in the sweat glands. EDA measurement is often used in psychophysiological research, as changes in electrodermal responses have been shown to be indicative of changes in psychological states. For instance, studies suggest that EDA readings increase when participants view novel stimuli, suggesting an orienting effect, and then steadily decrease with habituation [1]. EDA measurement has also been used to successfully discriminate between attentional orientation vs. surprise/alarm and has been found to be related to information processing, memory, and hemispheric lateralization [2]. However, the most common application of EDA measurement is to assess changes in physiological arousal, which implies strong physical or emotional reaction to a given stimuli [3]. This arousal leads to the natural 'fight, flight, or freeze' response, one aspect of which includes increased sweat production [4].

This research was supported by the Natural Sciences and Engineering Research Council of Canada through the Collaborative Research and Development Grant CRDPJ 523632-18

The present validation study was designed to assess the suitability of an EDA measurement system for research on simulator sickness in response to optical distortion [6]. Thus, EDA in response to arousal is directly relevant to this paper as electrodermal responses are associated with symptoms of illness, such as nausea, dizziness, and disorientation. For instance, Tamura et al. [5] found EDA was significantly increased following the experience of a standard disorientation training simulation designed for pilots in training.

Traditionally, EDA has been recorded with electrodes connected to bench-top data acquisition, data logger, or chart recording systems. These systems are designed to record either a variety of physiological signals or EDA specifically. Due to the recent increase in interest in wearable sensor systems, inexpensive portable systems have become more widely available. An example of such a device is the NeuLog NUL-217 GSR (Fig. 1) measurement device. The NeuLog device is more convenient than traditional systems primarily because of its stand-alone nature, size, and portability. Traditional systems are typically larger benchtop devices connected to a host PC, this makes it more difficult to transport it from one location to another and limits mobility of the wearer during data collection.

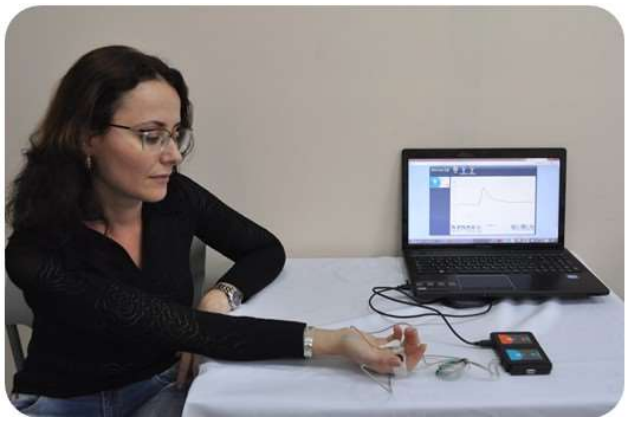

Fig. 1. Neulog device in use (https://neulog.com/wp-content/uploads/2014/11/gsr-under-1.jpg)

Further, to use many traditional devices for frequent sessions, a researcher must have access to a single room in which both the system and the computer used to run the corresponding program can permanently reside, or it needs to be mounted on a portable cart. The Biopac Student Labs system (Fig. 2) is a flexible and powerful multi-signal recording device and is representative of traditional EDA measuring devices. However, this device requires training for appropriate configuration for a given use case. In practice this means that for optimal use researchers would require a technician available on site to troubleshoot any issues. Further, traditional benchtop systems are much more expensive, as they require that the researcher purchase a software license as well as the system, electrode leads, electrodes, and electrode gel. In comparison, the NeuLog system is small, easily transported, user friendly, and can be purchased

Fig. 2. Biopac data acquisition system (top) and leads with electrodes (bottom; https://www.biopac.com/wp-content/uploads/EDA-Guide.pdf)

for a modest price which includes the software, system, and reusable electrode leads (with attached dry electrodes).

Although the Biopac system has some drawbacks it is an extensively validated laboratory-grade system [7]. Specifically, the Biopac has a high sampling rate (with hundreds of samples per second typical for EDA measurements); further, the use of gelled electrodes can be an advantage. For instance, research has suggested that gelled or wet electrodes provide more accurate measurements than dry electrodes, as they provide more consistent contact with the skin [8].

For clinical measurement or medical research in a laboratory such expense and complexity may be justified. Important factors to consider before using any device are the existence of validated data collection protocols and the flexibility of the system to record a range of other physiological signals, which will allow the device to be used in a wide range of psychophysical research designs – criteria to which the Biopac system adheres. For ambulatory applications or commercial applications the lack of portability and cost make such traditional EDA systems less practical. For research which does not have direct implications for the clinical well-being of a population (such as the social sciences, forensic science, and technological development research), particularly for research performed in real world rather than laboratory settings, the disadvantages of traditional systems such as the Biopac system might outweigh their benefits.

If a cheaper, more convenient device such as the NeuLog system can be shown to have sufficient validity, these findings would be useful for researchers in fields that do not rely on strict levels of accuracy. Further, if the NeuLog readings are not valid as compared to laboratory grade devices, this is also important for researchers who intend to use the NeuLog device in their data collection. Therefore, the purpose of the present study was to confirm the validity of the NeuLog device by comparing it directly to a Biopac system. To this end we recorded EDA/GSR measurements from participants using both devices simultaneously and comparing the outputs of each.

II. Method

A. Participants

Participants consisted of eight people from neighbouring laboratories ranging in age from 25 to 55 (M = 37.75, SD = 11.23), with equal numbers of males and females. All participants had normal or corrected-to-normal vision and normal stereoacuity (< 40 arcsec). There was no compensation provided for participating.

B. Materials and Apparatus

Electrodermal activity was recorded using two systems. (1) The NeuLog NUL-217 GSR device [9] with reusable dry electrodes and Velcro straps to secure them to the fingers. (2) The Biopac MP36 DAQ data acquisition system with compatible electrode leads (SS57LA) and self-adhesive, pre-gelled, disposable electrodes (EL507) using the BSL 4 data acquisition and analysis software (packaged as the Biopac Student lab system) [10]. The Biopac software was run on a Dell desktop computer (Intel Core i7-4790 CPU @ 3.6 GHz, 16GB RAM, Windows 10 OS) connected to the MP36 hardware. A separate Dell laptop computer was used to upload the NeuLog data recordings offline. The Biopac sampling rate was set to 250 Hz, whereas the maximum setting for the session duration on the NeuLog allowed for a rate of only 50 Hz. For data analysis, every fifth sample from the Biopac reading was selected for comparison to the complete NeuLog dataset.

Visual displays were presented using a phone-based virtual reality headset. A Google Pixel 2 XL (Octa-core CPU, 6” 1440 x 2880 screen size, 2160 pixel resolution, 4 GB RAM, Android 8.0 OS) was placed in a Google Daydream (2017) viewer providing an immersive virtual reality display. Participants were shown both a blank screen (no visual stimulus) and a video displayed in immersive stereoscopic Virtual Reality using the YouTube VR app [11] depicting a first-person experience of 5 different 3D roller coaster models (based on real roller coasters in existing theme parks).

C. Procedure

Participants were brought to a dedicated Medical Devices Laboratory at York University, Toronto. Upon arrival, they were connected to both the NeuLog and the Biopac devices, donned the HMD, and seated at a desk. The devices were placed on opposite hands on the volar middle phalanx (between the top and middle joints), with one lead placed on the index finger and the other placed on the middle finger. The specific hand upon

which each device was placed was counterbalanced across participants.

Before the participant arrived, the researcher set the video to the correct time stamp. The participant was asked to sit as still as possible for a duration of 5 minutes viewing a blank display while both devices recorded baseline measurements. The participant was then asked to verbally direct the researcher to press the play button on the VR video using the controller. Once the video had been initiated, both the NeuLog and the Biopac computer software were manually switched by the researchers to begin recording, using a verbal countdown to ensure the recordings began as close in time to each other as possible. After 5 minutes had elapsed, the EDA recordings were stopped, and the participant was once again asked to guide the researcher in stopping the video. The participant was then asked to sit for an additional 5 minutes to record a post-baseline while

Fig. 3. Timeline of data collection process

viewing a blank display. In three cases, the participants were then asked to complete a series of hand gestures consisting of two repetitions of: palms facing up, palms facing down, wrist bending up and down, wrist bending up and down (Fig. 3).

The baseline condition was conducted to determine if there was a delay between onset of recording and the achievement of stable readings in the case of the NeuLog device. A delay was expected, based on existing studies of the efficiency and stability of wet versus dry electrodes [8]. We chose a 3D roller coaster video because it was recorded in stereoscopic VR and elicited strong vection illusions, both of which will be present in the future study for which we intend to use the NeuLog device. Like a real rollercoaster, the virtual rides varied between intense and less intense phases of the ride providing opportunity for eliciting electrodermal responses. The post-baseline condition was added to get a second comparison of no stimuli against visual stimuli. Finally, the hand-sequence condition was used to determine if accuracy was impacted by motion of the hands upon which the devices had been placed. A normalized cross-correlation was conducted using the MATLAB R2019a software and its *xcorr* function (Mathworks, Natick MA; [12]) to estimate the correlation and time delay between the recordings from the two instruments. For this analysis the first 30 s and any samples beyond the 5 min duration were removed from each recording and the data was detrended to remove slow drifts typical of EDA recordings.

III. Results and Discussion

The analysis revealed a significant positive correlation for all eight subjects between the normalized recordings obtained from the two systems in each of the four recording intervals: pre-baseline, video, post-baseline, and hand-sequence (Table 1). Although all correlations were significant, some participants displayed lower than average cross-correlation levels in some conditions. For instance, subjects 2, 3, and 5 all had cross-correlation levels below 0.50 during the pre-baseline condition. Although this could be due to individual differences in participants, the fact that these low correlations were only seen in one condition suggests a difference in the NeuLog and Biopac devices in the amount of time the electrodes need to be in contact with the skin before they are able to give accurate readings. It

TABLE I. Correlation between instruments, for each participant and mean, with [95% confidence interval]

Cross-Correlations	**Condition**			
Participant	***Pre-Baseline***	***Video***	***Post-Baseline***	***Hand-Sequence***
1	0.98 [0.96,1.00]	0.87 [0.85,0.89]	0.92 [0.90,0.94]	
2	0.33 [0.31,0.35]	0.96 [0.94,0.98]	0.95 [0.93,0.97]	
3	0.28 [0.26,0.30]	0.53 [0.51,0.55]	0.75 [0.73,0.75]	
4	0.98 [0.96,1.00]	0.94 [0.92,0.96]	0.97 [0.95,0.99]	0.29 [0.27,0.31]
5	0.27 [0.25,0.29]	0.86 [0.84,0.88]		
6	0.51 [0.49,0.53]	0.93 [0.91,0.95]	0.96 [0.94,0.98]	0.87 [0.85,0.89]
7	0.96 [0.94,0.98]	0.83 [0.81,0.85]	0.81 [0.79,0.83]	0.83 [0.81,0.85]
8	0.73 [0.71,0.75]	0.93 [0.91,0.95]	0.97 [0.95.0.99]	
Average overall R	0.63 [0.61,0.65]	0.86 [0.84,0.88]	0.90 [0.88,0.92]	0.66 [0.64,0.68]
Average n	15,001	14,979	15,001	923

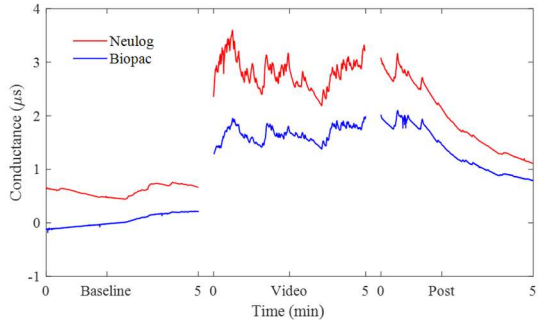

Fig. 4. Raw EDA data points presented as a time series for a single observer (Subject 2) for all conditions (baseline, video, posttest). The readings have been temporally aligned to correct for a delay between devices in recording start time and show a strong agreement between the two instrument readings.

should be noted that there was a small variable delay across individuals between the electrode placement and initiation of trials which may have been responsible for some of this variability. Between-subjects cross-correlations (first 30 s removed and lag-corrected) averaged across the video, baseline, and post-baseline conditions were also significant. An example of a representative raw data plot during the video session can be found in Fig. 4, showing the level of fluctuation in EDA response, as well as the level of correspondence between the two measurements.

A Repeated Measures Analysis of Variance (ANOVA) revealed at least one significant difference in the correlation coefficients of the two instruments between conditions (baseline, video, post-baseline, and hand-sequence) with $F(3, 18) = 6.27$, $p = 0.004$. However, post-hoc paired samples t-tests (using a Bonferonni corrected p-value of 0.0125) conducted between the four groups showed no significant differences, perhaps due to low statistical power. Although it appears at first that the correlation levels in the hand-sequence condition are

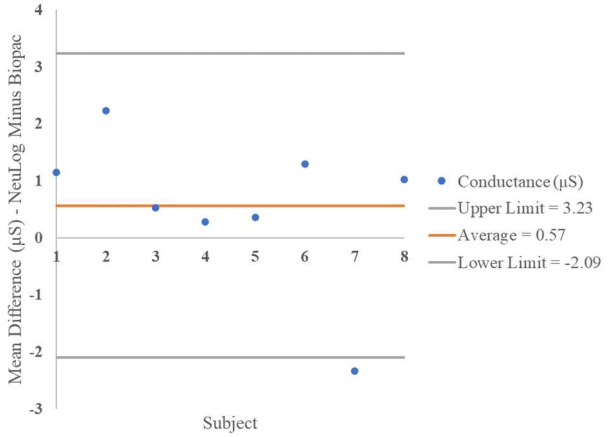

Fig. 5. Average of raw mean differences in micro-Siemens for participants in the video condition. Level of mean difference varies widely between participants and the Neulog readings are typically higher than those obtained using the Biopac.

lower than those of the other three conditions, this analysis may suggest that minor movement in the form of gesture does not affect the reliability of the NeuLog device readings. However, since only three subjects participated in this condition, further research should be conducted in order to confirm these results.

Fig. 4 shows a typical recording with larger skin conductance in the NeuLog device, both in terms of average value and amplitude of fluctuations. Generally, the timing and shape of the fluctuations was similar for the two instruments although the absolute values differed. Fig. 5 shows the mean differences between the two instruments by observer (with average mean difference across conditions ranging from 0.24 to 0.57 μS). For the Neulog, the readings ranged from 0.51 to 1.34 μS on average, and for the Biopac, the readings ranged from 0.54 to 1.06 μS on average. Absolute values in EDA measurements exhibit bias due to variability in the electrode skin

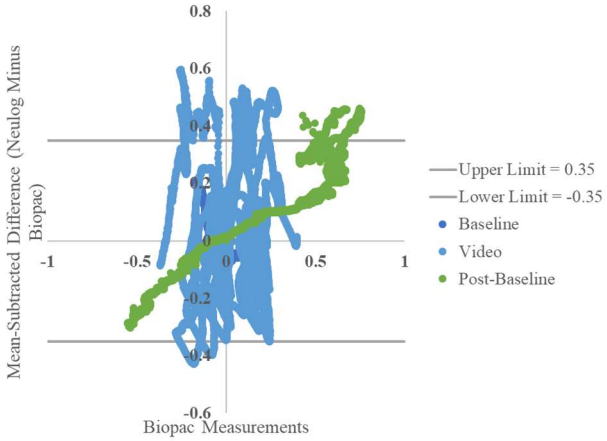

Fig. 6. Differences between mean-subtracted Neulog and Biopac data (μS) from a representative observer (Subject 1), as function of mean-subtracted Biopac signal level. The data are colour-coded by condition, with limits specified as two standard deviations from the mean. This example shows a weak tendency of the Neulog to show a larger amplitude in readings (mean difference tends to increase with signal level). Note: large differences in scale were seen between subjects, but overall trends were similar.

interface and is not surprising given that the Neulog uses dry electrodes while the Biopac uses isotonic gel electrodes. Measured skin conductance level also varies between participants, over time for a given participant, and shows dependences on skin site, skin potential, and electrode potentials. Absolute skin conductance levels are of limited interest for behavioural, psychological or human-computer interaction studies where phasic changes in EDA (skin conductance responses) corresponding to events or experimental conditions are typically more important. Thus, to compare the two instruments in terms of the measurement of phasic skin conductance responses we compared mean-subtracted signals.

Separate Bland-Altman plots were generated using the mean-subtracted data to show the average differences in micro-Siemens per participant for the baseline, video, and post-baseline conditions – an example of which can be found in Fig. 6. Although the scale varied largely across participants (with the

smallest normalized differences ranging from -0.08 to 0.08 μS and the largest ranging from -1 to 1 μS), there were some general trends which became clear. Overall, the NeuLog device tended to produce larger values than the Biopac when conductance was low, but lower values when conductance was high. This suggests that as well as an offset difference between the two devices, there is typically a difference in gain. Further, when traces from both devices are plotted on a similar scale, the NeuLog sequences seemed to have higher amplitude compared to the Biopac readings. Although these findings appear to suggest systematic differences between the readings, the fact that these differences are relatively stable across participants suggests that they are easy to predict and correct for. In particular, the high normalized cross-correlation in Table 1 shows that these gain differences can be normalized. Finally, in most research examining EDA measurements, it is not the exact value in μS which is of interest, but the moment-to-moment change in response to a given stimulus. Our data show these changes should be consistent for the two instruments for a given subject. This is important because there will naturally be some variability in EDA readings between subjects due to individual biological differences.

For practical usage, there are some general methods that future researchers would be advised to follow based on our results. First, due to the low correspondence rate between devices in the pre-baseline condition, it is recommended that the Neulog be attached to the participant for a specified period (5 min) before recording begins, so as to ensure the device readings have stabilized. Second, it is recommended that the device not be used if the research design requires that the participant generate movements or gestures with the hand upon which the device will be attached. Given the high level of agreement in the moment to moment differences between the Biopac and NeuLog readings, we conclude that the NeuLog is a sufficiently accurate and convenient option for psychophysiological and commercial EDA measurement.

REFERENCES

[1] L. M. Williams, M. J. Brammer, D. Skerrett, J. Lagopolous, C. Rennie, K. Kozek, G. Olivieri, T. Peduto, and E. Gordon, "The neural corrleates of orienting: An integration of fMRI and skin conductance orienting," NeuroReport, vol. 11, pp. 3011-3015, September 2000.

[2] G. Turpin, F. Schaefer, and W. Boucsein, "Effects of stimulus intensity, risetime, and duration on autonomic and behavioural responding: Implications for the differentiation of orienting, startle, and defense responses," Psychophysiology, vol. 36, pp. 453-463, July 1999.

[3] W. Boucsein, Electrodermal Activity. Wuppertal, New York, NY: Springer, 2012.

[4] P. Walschburger, "Psychophysiological Activation Research," in The Individual Subject and Scientific Psychology, J. Valsiner, Ed., Boston, MA: Springer, 1986, pp. 311-345.

[5] A. Tamura, T. Iwamoto, H. Ozaki, M. Kimura, Y. Tsujimoto, and Y. Wada, "Wrist-worn electrodermal activity as a novel neurophysiological biomarker of autonomic symptoms in spatial disorientation," Front. Neurol., vol. 9, article 1056, 2018.

[6] J. C. Miller, T. J. Sharkey, G. A. Graham, and M. E. Mccauley, "Autonomic physiological data associated with simulator discomfort," Aviat. Space Environ. Med., vol. 64, pp. 813-819, 1993.

[7] J. J. Braithwaite, D. G. Watson, R. Jones, and M. Rowe, "A guide for analysing Electrodermal Activity (EDA) & Skin Conductance Responses (SCR) for psychological experiments," Psychophysiology, vol. 49, pp. 1017-1034, 2013.

[8] Y. Higashi, Y. Yokota, and Y. Naruse, "Signal correlation between wet and original dry electrodes in electroencephalogram according to the contact impedence of dry electrodes," Proceedings 39th Anuual IEEE EMBC, Seogwipo, pp. 1062-1065, 2017.

[9] NeuLog, GSR Logger sensor NUL-217, 2017, https://neulog.com/gsr/.

[10] Biopac Systems Inc., EDA: Electrodermal Activity, 2020, https://www.biopac.com/application/electrodermal-activity/.

[11] 3D-VR-360 VIDEOS, VR 360 video of top 5 roller coaster rides 4K virtual reality [Video], Youtube, December 2019.

[11] The MathWorks Inc, MATLAB, 2020, https://www.mathworks.com/products/matlab.html?s_tid=hp_products_matlab.