Preprints.org

Article

Not peer-reviewed version

# PhysioKit: Open-source, Low-cost Physiological Computing Toolkit for Single and Multi-user Studies

Jitesh Joshi , Katherine Wang , Youngjun Cho *

Posted Date: 7 August 2023

doi: 10.20944/preprints202308.0545.v1

Keywords: Physiological Computing; Data Acquisition Toolkit; Multi-user HCI studies; Biofeedback; Signal Quality Assessment

Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

*Preprints* (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 August 2023 doi:10.20944/preprints202308.0545.v1

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

*Article*

# PhysioKit: Open-Source, Low-Cost Physiological Computing Toolkit for Single and Multi-User Studies

**Jitesh Joshi [1,†], Kathernine Wang [1,†] and Youngjun Cho [1,*,†]**

[1] Computer Science Department, University College London
* Correspondence: youngjun.cho@ucl.ac.uk
† Current address: Room: 403, UCLIC and Computer Science, 169 Euston Road, London, NW1 2AE, United Kingdom

**Abstract:** The proliferation of physiological sensors opens new opportunities to explore interactions, conduct experiments and evaluate the user experience with continuous monitoring of bodily functions. Commercial devices, however, can be costly or limit access to raw waveform data, while low-cost sensors are efforts-intensive to setup. To address these challenges, we introduce *PhysioKit*, an open-source, low-cost physiological computing toolkit. PhysioKit consists of (i) a sensor and hardware layer that can be configured in a modular manner per research needs, (ii) a software application layer that enables data acquisition, real-time visualization and machine learning (ML)-enabled signal quality assessment. This also supports basic visual biofeedback configurations and synchronized acquisition for co-located or remote multi-user settings. In a validation study for photo-plethysmography (PPG) sensor with 15 participants, PhysioKit shows strong agreement with research-grade PPG sensors. Furthermore, we report usability survey results from 10 micro-project teams (44 individual members in total) who use PhysioKit for 4-6 weeks, providing insights into its use cases and research benefits. We conclude by discussing extensibility and potential impact of the toolkit on the research community.

**Keywords:** physiological computing; data acquisition toolkit; multi-user hci studies; biofeedback; signal quality assessment

## 1. Introduction

Physiological signals and metrics, such as heart rate variability, have been actively explored in a wide range of contexts given their possible interpretation into both physical and psychological states. Particularly in human computer interaction (HCI), much attention has recently been paid to physiological computing research where it focuses on how to evaluate user responses to interventions, enhance user interactions with technology, or investigate how to support interpersonal interactions. Physiological computing has also enabled to create novel ways to measure engagement in virtual experiences [1] or to monitor affective states [2]. Researchers have also highlighted the potential of using physiological sensors in educational settings to help students interactively learn about human anatomy and physiology [3–6].

Physiological computing involves detecting, acquiring, and processing various biological data, such as heart rate (HR), skin conductance, blood-oxygen saturation, body temperature, blood glucose levels, muscle activity, and neural activity. Sensors commonly used for physiological data acquisition include: electrocardiogram (ECG), photoplethysmography (PPG), respiratory (RSP), electrodermal activity (EDA), electromyography (EMG) and electroencephalogram (EEG) [7]. The development of physiological computing systems can include systems that sense, interpret, react and interface biosignals. These can help us track health, design adaptive technology interactions, and offer personalized interventions [8–10].

PPG, EDA, and RSP are among the most prominent physiological sensing channels used in HCI research. Cardiovascular signals from PPG and perspiratory data from EDA have been explored to capture physiological states of students in remote classrooms [11,12] and to understand engagement

© 2023 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 August 2023 doi:10.20944/preprints202308.0545.v1



levels of game players [13]. PPG signals have also been exploited to communicate the emotional states of group members to one another [14], or to connect people in public spaces [15], while EDA signals have been used to evoke empathy through shared biofeedback [16]. Respiration cues have been used to help people understand and manage their stress [17,18], or to feel connected to others through sharing breathing signals [19].

With a wide range of commercially available research-grade physiological sensors (e.g., Procomp Infiniti System [1] by Thought Technology), BIOPAC [2], Empatica [3], Biofourmis [4]) and more affordable physiological sensing solutions (e.g., BITalino [5], and OpenBCI [6]), researchers have explored how to understand and improve user experiences with technology [20] or create novel ways to support social connections between people[21–23].

While the proliferation of low-cost commercial physiological sensing solutions and fitness trackers (e.g., Apple Watch [7], Fitbit [8], Garmin [9]) has been boosting the interest in physiological sensing forward, there is a need to pay attention to the challenges of using such devices, including limitations in robustness associated with sensor misplacement, body movements and ambient noise, reduced flexibility in adapting an acquisition interface to the study protocol, and relatively high costs of hardware or software [24–27]. The majority of these commercial sensors provide limited access to continuous physiological signals and metrics (e.g. inter-beat interval time-series) which can yield considerable insight into our bodily functions and psychological states [28–31]. HCI researchers often face difficulties in utilizing raw physiological data due to their complexity [32]. Furthermore, the storage of user data by consumer-grade devices on remote company servers for commercial purposes [33] raises data-privacy and ethical concerns.

Given the aforementioned limitations of commercial biosensors, researchers started turning to low-cost, open-source prototyping platforms (e.g. Arduino with physiological sensing nodes [10]). These platforms have been deployed in different contexts, such as personalizing user experiences [34] (e.g., adaptive interfaces [35,36]) or enhancing experiences with using biofeedback [37]. However, physiological computing with such platforms often requires a range of technical and computational skills and considerable setup time. Also, to our best knowledge, there is no open-source platform for physiological computing, which integrates and supports both *sensing* (e.g. data acquisition, preprocessing and streaming) and *data analysis* (e.g. signal quality assessment, metrics computation) together. To address these issues, we propose *PhysioKit*, a new open-source physiological computing toolkit for HCI researchers, hobbyists and practitioners. While being cost-effective, *PhysioKit* can facilitate flexible configuration, easy data collection, and support various analyses of physiological data. Our contribution is two-fold:

1. a novel open-source physiological computing toolkit (GitHub repo link) that offers a one-stop physiological computing pipeline spanning from data collection, processing to a wide range of analysis functions including a new machine learning module for the physiological signal quality assessment.
2. a report on validation study results, as well as user reports on the PhysioKit's usability and examples of use cases demonstrating its applicability in diverse applications.

---

1 https://thoughttechnology.com/procomp-infiniti-system-w-biograph-infiniti-software-t7500m/
2 https://www.biopac.com/
3 https://www.empatica.com/
4 https://biofourmis.com/
5 https://www.pluxbiosignals.com/collections/bitalino
6 https://openbci.com/
7 https://www.apple.com/watch/
8 https://www.fitbit.com/
9 https://www.garmin.com
10 https://www.arduino.cc/

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 August 2023 doi:10.20944/preprints202308.0545.v1



We begin this paper with an overview of related work and identify key challenges. We then introduce PhysioKit, describing its hardware and software application layers, and how they are designed to enable high-quality data-acquisition, while considering the flexibility in selecting physiological sensors. We present results from a validation study, outcomes from a usability survey, and an overview of applications in which PhysioKit is used. Finally, we conclude with discussions on the implications and benefits of PhysioKit for different contexts and user communities.

## 2. Related Work

In this section, we discuss prior work on physiological computing and its applications in HCI research and review validation studies on wearable physiological sensors.

### *2.1. Physiological computing and HCI applications*

There are an increasing number of HCI applications that use physiological sensing technology to assess and customize user experience, support health-related practices, and support social and educational contexts. The applications of physiological computing in HCI contexts can also be broadly categorized into two themes based on the role physiological computing plays in the interaction: *interventional* and *passive*.

Interventional physiological computing studies require real-time processing of physiological signals, which is then used to influence interaction with the self, with others, or with technology. Interventional studies have examined how real-time physiological computing can be used in several contexts including health monitoring (e.g., stress [38], diabetes [39]), training healthful practices (e.g., respiration [40]), educating children in anatomy [41], communicating affective states between people during chats [42] or VR gameplay [43], sensing passenger comfort in smart cars [44,45], and personalizing content through adaptive narratives (e.g., interactive storytelling [46], adjustable cultural heritage experiences [47], synchronized content between multiple users [48]).

Passive physiological computing studies, on the other hand, do not require adapting any interaction aspects based on real-time computing of physiological metrics from the acquired signals. Illustrative studies with passive use of physiological computing include the assessment of user workload in different contexts [10,28,49], design and assessment of user experiences [50–53], video game reinforcements [54], game events [55], or objective comparisons with subjective reports [56].

The illustrated HCI studies are not exhaustive and are mentioned to emphasize the two distinct ways in which researchers use physiological computing. This categorization also serves as one of the key design considerations for the development of physiological computing toolkits. For passive studies, the provision of data-acquisition meets the research need, while for interventional studies, real-time computing of physiological metrics is required, which can further be mapped to biofeedback design or to adapt the interaction. While most commercial research-grade and low-cost devices support data-acquisition for passive studies, they are often not designed considering different needs of interventional studies.

### *2.2. Physiological Measurement Devices*

Researchers generally consider consumer devices acceptable to use and prioritize the use of devices based on design and familiarity with the brand, rather than reliability, as well as comfort or ease of use [57]. In Table 1, we highlight validation studies of commercially available physiological computing sensor systems and include studies in which measurements were taken from participants at rest.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 August 2023 doi:10.20944/preprints202308.0545.v1



**Table 1.** Validation studies of commercially available consumer-grade devices as well as low-cost research-grade devices and toolkits with participants at rest.

| Device | Device Category | Cost (USD) | Access to Raw Physiological Signals | Support for Interventional Studies | Support for Multi-user Studies | Validation Results for HR (bpm) | Validation Results for HRV |
|---|---|---|---|---|---|---|---|
| AliveCor KardiaMobile | C | $79.00 | No | No | No | $\kappa = 0.96$ [57] | N.A. |
| Apple Watch 4 | C | $399.00∗ | No | No | No | MAE = 4.4±2.7 [58]; $r = 0.99$ [59] | N.A. |
| Apple Watch 6 | C | $399.99∗ | No | No | No | $\kappa = 0.96$ [57,60] | RMSSD: $\kappa = 0.67$ [60] |
| Fitbit Charge 2 Fitbit | C | $149.95∗ | No | No | No | MAE = 7.3±4.2 [58] | N.A. |
| Fitbit Sense | C | $159.99∗ | No | No | No | $\kappa = 0.88$ [57] | N.A. |
| Garmin Vivosmart 3 | C | $139.99∗ | No | No | No | MAE = 7.0±5.0[58] | N.A. |
| Garmin Fenix 5 | C | $599.00∗ | No | No | No | $r = 0.89$ [59] | N.A. |
| Garmin Forerunner 245 Music | C | $349.00∗ | No | No | No | $\kappa = 0.41$ | RMSSD: $\kappa = 0.24$ [60] |
| Garmin Vivosmart HR+ | C | $219.99∗ | No | No | No | MAE = 2.98, $\kappa = 0.90$ [61] | N.A. |
| Oura Gen 2 | C | $299.00 | No | No | No | $\kappa = 0.85$ | RMSSD: $\kappa = 0.63$ [60] |
| Polar H10 † | C | $89.95 | No | No | No | N.A. | HRV (RR, ms) Spearman $r = 1.00$ [62] |
| Polar Vantage V | C | $499.95 | No | No | No | $\kappa = 0.93$; r = 0.99 [59] | RMSSD: $\kappa = 0.65$ [60] |
| Samsung Galaxy Watch3 | C | $399.99∗ | No | No | No | $\kappa = 0.96$ [57] | N.A. |
| Withings Scanwatch | C | $299.95 | No | No | No | $\kappa = 0.95$ [57] | N.A. |
| Empatica E4 | R | $1690.00∗ | Yes (filtered signals obtained from company servers) | No | No | MAE = 11.3±8.0 [58] | N.A. |
| BITalino (PsychoBIT) | R | $459.78 | Yes | Using third party software | Yes, only co-located | Morphology based validation with ECG signals: $R^2 = 0.83$ [63] | N.A. |
| OpenBCI (EmotiBit) | R | $499.97 | Yes | Using third party software | No | N.A. | N.A. |

∗ Discontinued device †ECG device C.: Consumer Device; R.: Research Grade Device

Most studies reported results as mean absolute error (MAE), interclass correlation (ICC) $\kappa$, or Pearson's correlation coefficient ($r$) for HR (bpm). Where applicable, we also listed results of HRV validation, which were primarily reported as ICC $\kappa$ for RMSSD. To illustrate how affordable, readily available, and useful the devices were, the table also includes the price (USD) and availability of the device, as well as whether raw signal data (e.g., BVP, ECG, EEG signals) is accessible. On average, consumer-grade commercial HR/HRV devices cost $291.77 (SD: $158.77) and had limited access to the mentioned raw signals. Also, by the time validation studies were published, most manufacturers had already discontinued the tested versions of the device. Low-cost research-grade devices range from $459.78 to $1690, and often provide access to raw physiological signals. However, existing low-cost research-grade devices and toolkits are not designed to readily support interventional studies as well as co-located or remote multi-user studies.

Of the physiological sensing devices included in validation studies (e.g., Xiaomi Miband 3, Apple Watch 4, Fitbit Charge 2, Garmin Vivosmart 3) consumer-grade wearables were significantly more accurate HR measurements at rest compared to research-grade wearables (i.e., Empatica E4, Biofotion Everion) when compared to a reference device (MAE: 7.2±5.4 bpm compared to 13.9±7.8 bpm [58]). The Apple Watch showed highest accuracy in several studies (Apple Watch 4 MAE: 4.4±2.7 bpm [58]; Apple Watch 6 ICC: 0.96 [57,60]) compared to other commercial devices. Despite this, we could not find any studies that confirmed it was in good agreement with a gold standard reference device. Research-grade consumer devices showed the highest MAEs (Empatica: 11.3±8.0 bpm, Bioevotion: 16.5±6.4 bpm [58]) in a list of commercially-available weareables.

In terms of results from validation studies, the RR interval signal quality of the Polar H10 HR chest belt was considered in good agreement with the gold-standard (error rate 0.16 during rest) (i.e., medilog AR12plus ECG Holter monitor [62]. Similarly, Garmin 920XT was in good agreement with gold-standard (i.e., ADInstruments Bio Amps) during rest conditions [64]. While validation studies involving physical activity conditions have shown moderate-to-strong agreement for devices, they have consistently revealed a decrease in accuracy as activity levels increase [61,65]. One common explanation for this discrepancy are motion artifacts from movement. Some devices (e.g., Polar H10), have designed elastic chest straps with multiple embedded electrodes to protect against measurement noise. The problem with elastic electrode straps is, however, the reduction in contact between skin and elastic electrode straps [62].

## 3. Problem Statement and Challenges

Based on physiological systems such as those described above, there are several notable challenges to consider when developing physiological sensing toolkit to support different research contexts.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 August 2023 doi:10.20944/preprints202308.0545.v1



### 3.1. Signal Quality Assessment

Environmental and experimental errors can easily contaminate recordings. PPG sensors, for instance, are susceptible to interference from natural and artificial light, and can be highly vulnerable to experimental errors, as movement can alter the way light is transmitted [66]. Users may also experience issues putting on wrist devices [57], which can result in light interference. In addition, a study comparing non-obtrusive physiological sensors for monitoring stress levels in computer users found that signal quality may also be susceptible to different pressures from sensors worn on the wrist while users perform typing tasks [67].

While these challenges cannot be fully addressed, one of the potential ways to circumvent these issues can be to enable flexibility in sensor placement (e.g. PPG sensors can be placed on either ear-lobes or fingers). This flexibility also enhances the accessibility of the sensors to suit users with different physical sensitivities or requirements. Another promising approach would include providing real-time assessment of signal-quality to flag noisy acquisition fragments of signals, which would enable researchers to take appropriate actions during data-acquisition or processing. Widely used methods for signal quality assessment include signal to noise ratio (SNR), template matching, and relative power signal quality index (pSQI) [68], along with recent machine learning based approaches based on SVM classifier [69], LSTM [70], and 1D-CNN [71]. The state-of-the-art (SOTA) performance has been demonstrated by 1D-CNN classifier approach [71] achieving 0.978 accuracy on MIMIC III PPG waveform database. It is noteworthy that stringent signal quality assessment may lower the signal retention, thereby decreasing the usable segments of signal for deriving physiological metrics [72].

### 3.2. Authority over how data is collected, stored, and processed

With consumer wearables, data is first acquired and processed using the manufacturer's proprietary algorithms and then exported directly from the manufacturer's applications and often including remote servers [60]. In addition to limiting control over data collection, manufacturer-controlled data output can also impact how and what data can be analyzed. The transfer and storage of data on proprietary platforms and servers also raises ethical concerns about data privacy.

### 3.3. Validation Studies

The rapid nature of the release of new consumer-grade wearable devices presents additional research challenges. Although it is reasonable to assume that newer models can perform better, it is difficult to conduct and publish validation studies on the same timescale of of new technology releases. Furthermore, despite the continuous development of new devices, the way in which commercial manufacturers maintain proprietary algorithms limits researchers from accessing raw data signals which can contribute towards such validation studies. On the other hand, owing to the access of raw signal data, research community has demonstrated several application use-cases of research-grade devices and toolkits including OpenBCI, Empatica E4 and BITalino and has contributed towards their validation.

### 3.4. Support for Interventional and Multi-User Studies

Despite of the popularity of consumer-grade as well as research-grade devices and toolkits, support for conducting interventional (§ 2.1) and multi-user studies has remained fairly limited. Furthermore, for settings in which multi-user studies are conducted with remotely located users, researchers have either used time-stamping information [73] or have deployed manual approaches [74] for synchronizing the time-series data. These approaches of synchronizing the data-acquisition are not suitable for interventional studies and may result in a varied amount of time-lag between the physiological signals of different users. It is therefore crucial to design a toolkit considering these design objectives and existing research gaps.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 August 2023 doi:10.20944/preprints202308.0545.v1



## 4. The proposed physiological computing toolkit: PhysioKit

Figure 1 details the system architecture of PhysioKit. The toolkit consists of two layers: i) sensor and hardware layer and ii) software application layer. The first layer enables a modular setup of physiological measurements, which offers flexibility in the configuration of physiological sensors depending on research needs. The software application layer includes a data collection module, a real-time streaming module and a data analysis module. Our software application is built with Python and is compatible with Windows, Mac and Linux operating systems.

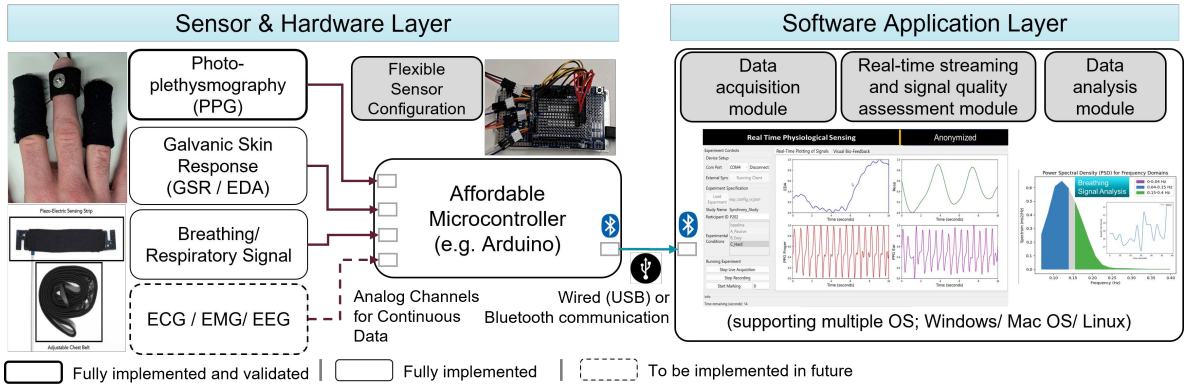


**Figure 1.** PhysioKit is a novel physiological computing toolkit which is open-source and cost-effective. Its sensor and hardware layer provides flexibility in using low-cost physiological sensors, and the software application layer enables data acquisition, visualization, and real-time ML based signal quality assessment, while supporting passive as well as interventional studies for single and multi-user settings.

### *4.1. Hardware layer*

The hardware layer consists of a set of physiological sensors and a firmware module that were developed in a way that addresses the above-mentioned challenges. For example, this layer is designed to facilitate the flexibility of sensor placement, and provide options to configure the toolkit for single and multi-user settings for passive as well as interventional studies.

#### 4.1.1. Physiological Sensors

PhysioKit utilizes a variety of inexpensive physiological sensors that are compatible with Arduino boards and enables smooth integration of these sensors. In this work, we focused on demonstrating the functionality and validation of the toolkit using two PPG sensors[11], which is presented in § 5.1. There, we focus on validating HR and HRV measurements from PPG signal, as it is the most widely used physiological sensing channel. However, the interface and toolkit can be easily extended to other sensing channels, including RSP, EDA, ECG, EMG and EEG. To demonstrate the readiness of extending other sensors, we also integrated a RSP sensor [12] and an EDA sensor [13]. Though the measurements of these alternative sensing channels were not validated in the current work, we discuss future plans to implement these in Figure 1.

To demonstrate an example of alternative ways in which PhysioKit can be designed to accommodate flexible placement of physiological sensors, we designed a printable 3D-CAD model of a mountable PPG sensor wristband, as shown in Figure 2. This wristband provides an alternative way to acquire PPG signals that can be used simultaneously, and in addition to the ear or finger placement.

---

[11] https://pulsesensor.com/
[12] https://www.pluxbiosignals.com/products/respiration-pzt
[13] https://seeeddoc.github.io/Grove-GSR_Sensor/

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 August 2023 doi:10.20944/preprints202308.0545.v1



This ability to simultaneously collect physiological signals from a versatile combination of sensor locations is also useful in lessening motion artifact noise.

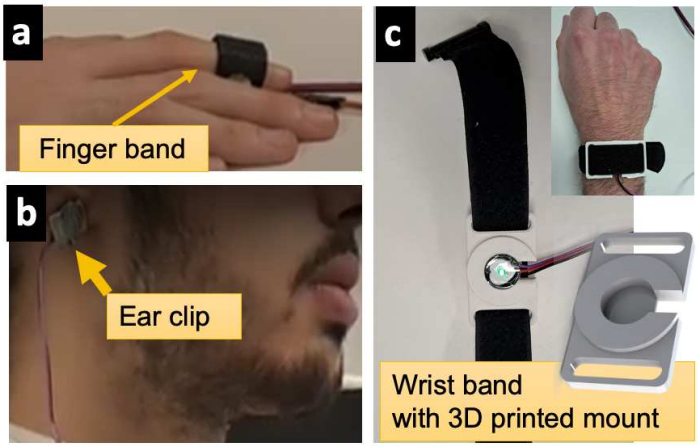


**Figure 2.** Flexible PPG setups available in PhysioKit to enhance accessibility.

4.1.2. Firmware Module

To accompany the versatility of the sensor setup, PhysioKit also includes a set of Arduino programs that are compatible with different combinations of single or multiple physiological sensors. Users can easily configure low-level acquisition parameters, including sampling rate and resolution of analog-to-digital conversion with these Arduino programs. The firmware module transmits the acquired data to the connected PC via wired (USB) or wireless (Bluetooth) communication, which is enabled using a HC-05 Bluetooth Module [14] , as explained in this tutorial [15].

Table 2 lists some of the key parameters of the hardware layer of PhysioKit. The Arduino board governs the hardware specifications, which can be selected according to research needs. For instance, some of the most cost-effective microcontroller boards (e.g. Arduino Uno, Arduino Nano) can support up to six physiological sensors and a sampling rate of up to 512 samples per second, whereas Arduino Due and Arduino Mega can support 12 and 16 channels respectively.

**Table 2.** PhysioKit - Hardware Layer Specifications

| **Parameter** | **Specification** |
|---|---|
| Sampling Rate | 10 - 10000 samples per second |
| ADC | 10/ 12-bit |
| Vref | 3.3V/ 5V |
| Baudrate | 9600 - 2000000 |
| Data transmission mode | USB/ Bluetooth |
| No. of Channels | 1 - Max supported by specific Arduino board |

*4.2. Software application layer*

The software application layer includes the user interface (UI), depicted in Figure 3A. The challenges concerning access to raw, unfiltered physiological signals and flexible configurations for passive and interventional studies involving both co-located and remote multi-user settings were carefully considered when developing the UI. The resulting interface includes features to facilitate data acquisition and signal visualization, as well as bio-feedback visualization. Integrated Jupyter

[14] https://www.deshide.com/Bluetooth-Module.html
[15] https://www.makerguides.com/arduino-and-hc-05-bluetooth-module-complete-tutorial/

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 August 2023 doi:10.20944/preprints202308.0545.v1



notebooks also streamline the process of extracting metrics from the acquired physiological signals. The software module is implemented using Python programming language, as it is a multi-platform language.

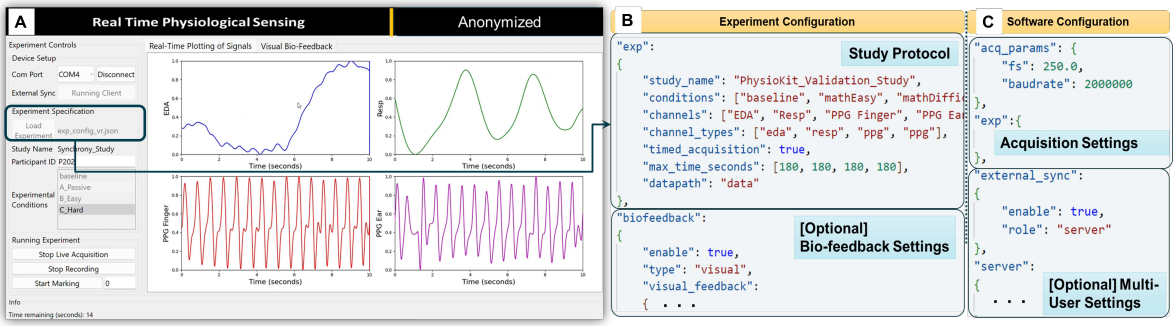


**Figure 3.** A) User-interface for Physiological Computing Toolkit; B) Configuration file for specifying experiment protocol and C) Configuration file for controlling data-acquisition.

#### 4.2.1. User Interface

To optimize acquisition, plotting and user controls, the PhysioKit UI is implemented with a multi-threaded design. Figure 3A shows a PhysioKit interface, which was designed using Qt design tools [16], with controls to configure an experimental study alongside display of real-time signal visualization. Acquired raw data is stored locally and appropriate signal conditioning is applied for plotting each physiological signal, while the quality of the acquired signals is assessed in real-time using a novel 1D-CNN based signal quality assessment module (§ 4.2.2). In addition, the UI enables options for different visual presentations of biofeedback (§ 4.2.4).

#### 4.2.2. Signal Quality Assessment Module for PPG

The signal quality assessment (SQA-Phys) module of PhysioKit extends 1D-CNN approach [71] that demonstrates the use of 1D-CNN network for classifying signal quality of PPG waveforms as "Good" and "Bad". Contrary to classification, SQA-Phys introduces a novel task in which the ML model is trained to infer signal quality metrics for the entire length of the PPG signal segment. For this, we implement encoder-decoder architecture with 1D-CNN layers that generate high temporal resolution signal quality vector. Figure 4 compares the SOTA architecture and the proposed architecture, and highlights how the inferred outcome of SQA-phys differs from the classification task as implemented by existing methods.

---

[16] https://www.qt.io/product/ui-design-tools

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 August 2023 doi:10.20944/preprints202308.0545.v1



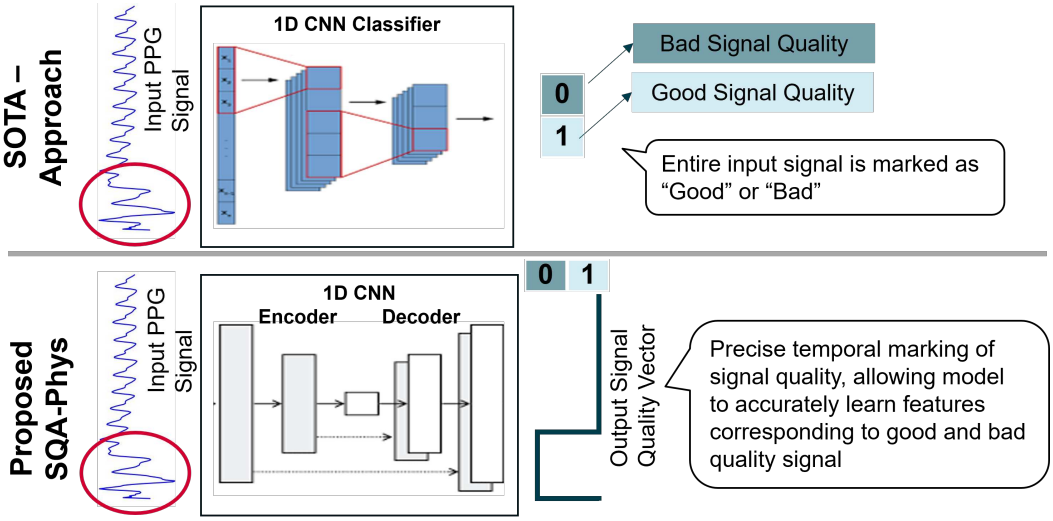


**Figure 4.** SQA-Phys: 1D-CNN based encoder-decoder architecture for high temporal precision signal quality assessment.

To train the model, we used in-house collected data acquired using the Infiniti Procomp and Empatica E4 wristband PPG sensors. This dataset, which includes 170 recordings of PPG signal [5 minutes] from 17 participants, was manually labelled for the signal quality. As the signal quality labels for training were marked for each 0.5 seconds of segment, without overlap, the time-resolution of SQA-phys inference was maintained as the same as the training labels, i.e. 0.5 seconds. However, in theory, the time-resolution could be further increased up to per-sample inference for signal quality.

We used a signal segment of 8 seconds as an input to the model and trained it with the batch size of 256 along with Adam optimizer. Training and validation split was made based on participant IDs, ensuring that the signals of same individual were not represented in both the training and validation sets. Our validation on a novel signal quality assessment task shows 96% classification accuracy with an inference vector having 0.5 second temporal-resolution, whereas the SOTA approach yielded 83% classification accuracy. The SQA-phys is integrated with the PhysioKit to present signal quality assessment with high temporal precision. The same is also stored alongside the data to provide signal quality annotation, and thereby indicating the clean and noisy segments of the signals to researchers. Though SQA-phys is currently implemented for PPG signals, it can be easily applied to other physiological signals.

#### 4.2.3. Configuring the UI for Experimental Studies

Using experiment and software configuration files Figure 3B,C, the UI adapts to the hardware configuration based on the researcher's data collection protocol. The experiment configuration file (Figure 3B) enables users to set up their study protocol by defining the experimental conditions, the acquisition duration for each condition, the required physiological sensors, and the directory path where the acquired data will be saved locally. In contrast to storing data on cloud servers, as most commercial devices do, storing data locally allows researchers to have complete control over the data. In addition, the experiment configuration file allows the number of channels to be selected for real-time plotting of acquired physiological signals. While the maximum number of channels is limited by the number of analog channels on the microcontroller board, a maximum of four channels can be selected for real-time plotting. The acquisition duration for each experimental condition is defined by "max_time_seconds" (see Figure 3B) when "timed_acquisition" is set to "true". However, when "timed_acquisition" field to "false", the UI ignores "max_time_seconds", allowing data acquisition to continue until the user manually stops it.

The software configuration file (Figure 3C) allows users to configure acquisition parameters, such as sampling rate and baudrate for serial data-transfer from the micro-controller to the computer. In

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 August 2023 doi:10.20944/preprints202308.0545.v1



addition, users can acquire physiological signals simultaneously from multiple users from different hardware layers. Here, different computers running the UI would communicate using TCP/IP messaging in order to synchronize the acquisition from the different users. This setup requires each computer to be accessible with an IP address, either on a local intranet or remotely using a virtual private network (VPN). To enable multi-user synchronization, one computer is designated as a TCP server in the software configuration file, while the others are TCP clients and specify the server IP in the software configuration file. In these settings, the live acquisition is initiated on all nodes. The client UIs remain idle until the server triggers the synchronized recording by broadcasting a TCP message. The TCP/IP based messaging can be further extended to the stimulus presentation software (not part of PhysioKit) for synchronized delivery of the study intervention. Figure 5 provides an overview of the multi-user setup.

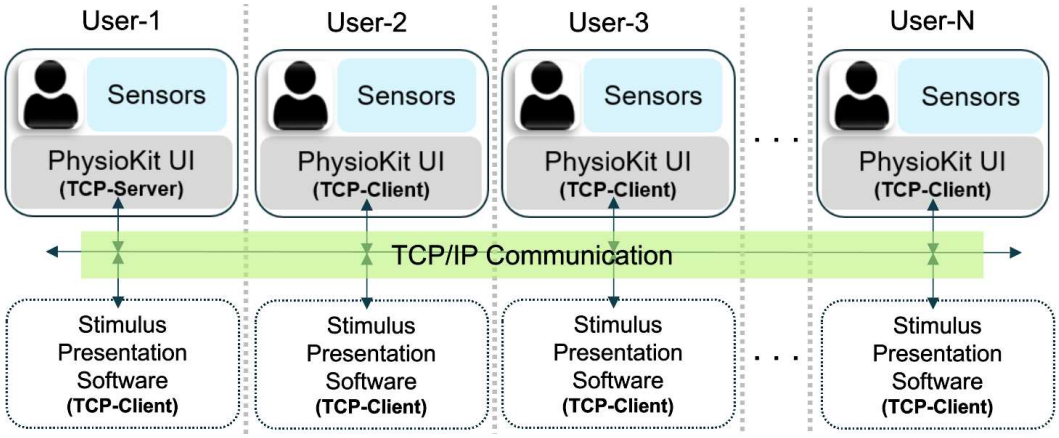


**Figure 5.** Multi-user setup of PhysioKit for synchronized data acquisition and stimulus presentation.

Oftentimes, HCI studies require capturing asynchronous events during experiments for qualitative and quantitative analyses. Based on this requirement, we designed a simple way for users to mark such asynchronous events directly in the PhysioKit UI by activating and deactivating the marking function as needed. This function is associated with an event-code to enable marking for different types of events during data-acquisition. The acquired data, along with the event markings, are stored in a comma-separated value format (CSV) for easy access and further analysis.

#### 4.2.4. Support for Interventional Studies and Biofeedback

PhysioKit supports real-time computing of physiological metrics, which allows adapting interaction for interventional studies. One of the most widely researched interventional study types involves using biofeedback [9,75,76]. To allow researchers to explore this approach, the experimental configuration file offers a way to map selected physiological metrics as dynamic biofeedback visualizations, as well as set window length and step interval to compute these metrics (see Figure 3B). PhysioKit currently provides options for basic biofeedback visualization using geometric shapes that vary in size and colors. However, the mapping implementation can be easily adjusted to include different biofeedback modalities, such as auditory or haptic, according to the study requirements.

#### 4.2.5. Data Analysis

To analyze the data acquired using the PhysioKit, the software layer facilitates using existing physiological signal processing libraries, such as NeuroKit [77] and HeartPy [78]. The data analysis module of the PhysioKit provides Jupyter notebooks as tutorials to process PPG, EDA and RSP signals. These tutorials demonstrate how to load physiological signals from CSV files, along with common steps for signal pre-processing. The tutorials further illustrate how use the libraries mentioned above

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 August 2023 doi:10.20944/preprints202308.0545.v1



to compute physiological metrics, as well as facilitate the analysis of specific events marked during data acquisition.

## 5. Evaluation

### *5.1. Study 1: Technical Validation*

To evaluate the validity of our toolkit, we performed a technical validation study by comparing metrics derived from PPG signals acquired using PhysioKit and a gold-standard reference system (i.e., Procomp Infiniti System [17] by Thought Technology) that has been widely used by researchers over two decades. The Procomp Infiniti system is a medical grade device that consists of PPG, EDA, and breathing belt sensors. It includes an encoder that connects to a computer through Bluetooth connection and is used with an accompanying software (Biograph Infiniti) that includes an interface displaying raw signals. The software also allows raw data to be collected and exported to a CSV formatted file. We set the sampling rate for the Procomp Inifiniti system to 256 samples per second. In order to assess facial movements that could potentially affect signal quality, we also recorded and assessed facial videos at 30 FPS.

#### 5.1.1. Participants and Study Preparation

Physiological signals were collected from 16 participants recruited through an online recruitment platform for research. Everyone reported having no known health conditions, provided informed consent ahead of the study, and was compensated for their time following the study. After being welcomed and briefed, participants were asked to remove any bulky clothing (e.g., winter coats, jackets) and seated comfortably in front of a 65 by 37 inch screen where they were fitted with both Infiniti and Physiokit sensors.

Three physiological signals (PPG, EDA, and RSP) were collected and compared based on the extracted metrics. Respiration belts from both systems were worn just below the diaphragm, one above the other without overlapping. One PhysioKit PPG sensor was fastened to the participant's left ear with a metal clip and the second was strapped around the participant's middle finger of their non-dominant hand with a velcro strap along with the Infiniti PPG. EDA sensors from both PhysioKit and Infiniti systems were placed on the index and ring fingers of the same hand without overlapping.

#### 5.1.2. Data Collection Protocol

The study followed a methodological assessment protocol for the analysis of physiological signals obtained simultaneously with PhysioKit and a gold standard reference. To induce a variety of physiological states, as well as moderate movement, each participant experienced four conditions, including a) a controlled, slow breathing task, b) an easy math task, c) a difficult math task and d) a guided head movement task, as depicted in Figure 6. While a higher agreement between the test device and the reference device could potentially be achieved in the absence of these variances, it could also lead to misleading results.

---

[17] https://thoughttechnology.com/procomp-infiniti-system-w-biograph-infiniti-software-t7500m/

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 August 2023 doi:10.20944/preprints202308.0545.v1



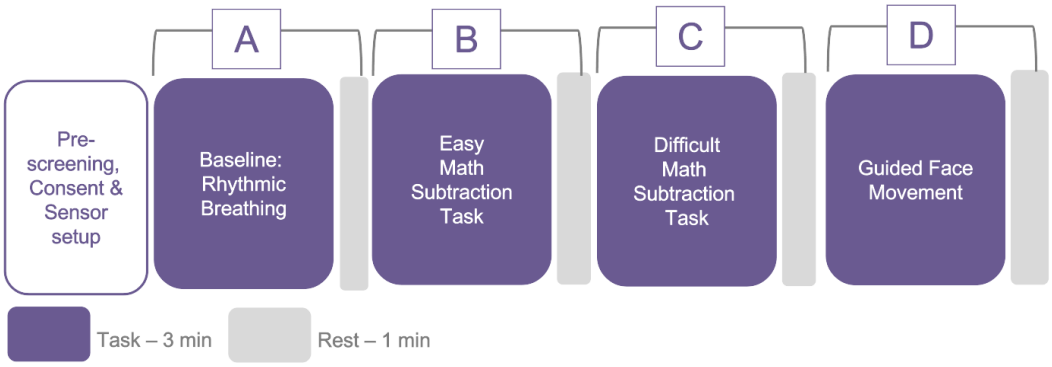


**Figure 6.** Data-collection Protocol for Technical Evaluation of PhysioKit.

Cognitively challenging math tasks with varying degrees of difficulty levels were chosen, as these have been reported to alter the physiological responses [79,80]. Furthermore, as wearable sensors are less reliable under significant motion conditions [81], we added an experimental condition that involved guided head movement. The PPG sensor on the ear remained relatively stable under all conditions, except during head movement, which provided us with the opportunity to investigate the impact of movement on signal quality at different sensor sites under varying motion conditions. Each condition lasted for 3 minutes, with 1 minute of rest after each condition. To randomize the sequence of conditions, we inter-changed "A" with "D" and "B" with "C". The experimental protocol was approved by the Ethics Committee of the university (anonymised for review).

#### 5.1.3. Data Analysis

Of the 16 participants, one was excluded from the analysis due to the incorrect fit of a PPG sensor. With 15 participants and four different experimental sessions, 60 pairs of PPG signals were used for the evaluation. Uniform pre-processing was applied to PPG signals acquired from PhysioKit and the reference device. First, we applied windowing with a window size of 30 seconds and a step interval of 10 seconds to calculate HR. As the HRV could also be extracted from the PPG signal, we refer to it as pulse-rate variability (PRV) [28]. For extracting PRV metrics, the window size was set to 120 seconds, with a step-interval of 10 seconds. In this work, our goal was to validate PRV metrics with the reference device, rather than validate PRV with HRV metrics, which are typically derived from the ECG signal.

We observed high variance in signal quality of the PPG signals acquired using the reference device (see Figure 8). In order to prevent extracting reference metrics from noisy segments, Signal Quality Index (pSQI) was computed for each windowed segment from raw signals, as described in [28]. A threshold of 40% pSQI was used to discard the noisy segments only from the PPG signals acquired using the reference device. No segments acquired using PhysioKit were discarded based on pSQI. PPG signals were then applied with bandpass filtering [0.7Hz - 4.0Hz] and further processed using the NeuroKit library [77] to derive HR and HRV metrics. Among different HRV metrics, we selected pNN50, which provides a proportion of the successive heartbeat intervals exceeding 50ms.

#### 5.1.4. Results

Bland-Altman scatter-plots were used to compare the performance of PhysioKit with the reference device and to assess the agreement between the two devices (see Figure 7) [82,83]. These plots provide a combined comparison for all experimental conditions for HR and HRV metrics. Since PPG signals were acquired from two different sites - finger and ear - we present the comparison for these two sensor sites separately.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 August 2023 doi:10.20944/preprints202308.0545.v1



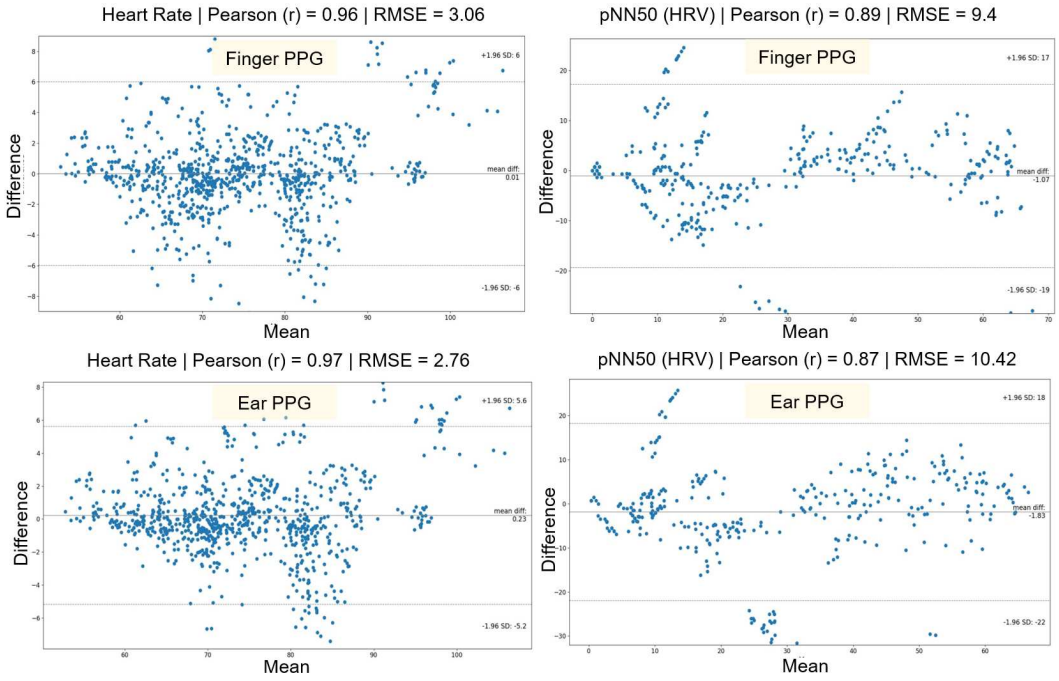


**Figure 7.** Bland-Altman Scatter Plots to Compare Heart-Rate (beats per minute) and Pulse-Rate Variability (pNN50) with the reference device.

Table 3 reports a detailed evaluation across each experimental conditions for both finger and ear sites. Both HR and PRV metrics from PhysioKit show a high correlation, as measured with Pearson correlation coefficient (r), and lower difference, as measured with root mean squared error (RMSE), mean absolute error (MAE) and standard deviation of error (SD). There is a higher correlation for HR (bpm) between PhysioKit and the reference device (ear: $r = 0.97$, finger: $r = 0.97$) than for PRV (pNN50) (ear: $r = 0.87$, finger: $r = 0.89$), which aligns with results from earlier studies listed in Table 1.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 August 2023 doi:10.20944/preprints202308.0545.v1



**Table 3.** Evaluation of PhysioKit: Comparison of Heart-Rate (beats per minute) and Pulse-Rate Variability (pNN50) with the reference device. Evaluation metrics include root mean-squared error (RMSE), mean absolute error (MAE), standard deviation of error (SD) and Pearson correlation coefficient (r), which are mentioned separately for each experimental condition as well as for all conditions combined.

| Metrics | PPG Site | Experimental Condition | RMSE | MAE | SD | Pearson (r) |
|---|---|---|---|---|---|---|
| **Heart Rate (beats per minute)** | **Finger** | Baseline | 3.63 | 2.43 | 3.38 | 0.96 |
| | | Math - Easy | 2.08 | 1.44 | 2.02 | 0.98 |
| | | Math - Difficult | 2.41 | 1.75 | 2.41 | 0.98 |
| | | Face Movement | 2.17 | 1.49 | 2.09 | 0.98 |
| | | All sessions combined | 2.65 | 1.78 | 2.65 | 0.97 |
| | **Ear** | Baseline | 3.71 | 2.46 | 3.41 | 0.95 |
| | | Math - Easy | 1.0 | 0.73 | 0.92 | 1.0 |
| | | Math - Difficult | 2.08 | 1.46 | 2.07 | 0.98 |
| | | Face Movement | 2.11 | 1.40 | 2.05 | 0.98 |
| | | All sessions combined | 2.45 | 1.53 | 2.43 | 0.97 |
| **Pulse Rate Variability (pNN50)** | **Finger** | Baseline | 4.31 | 3.46 | 4.28 | 0.98 |
| | | Math - Easy | 11.33 | 6.54 | 10.6 | 0.91 |
| | | Math - Difficult | 12.11 | 9.53 | 12.09 | 0.75 |
| | | Face Movement | 8.26 | 5.66 | 8.25 | 0.92 |
| | | All sessions combined | 9.4 | 6.26 | 9.34 | 0.89 |
| | **Ear** | Baseline | 4.36 | 3.78 | 4.21 | 0.99 |
| | | Math - Easy | 9.85 | 5.75 | 9.27 | 0.91 |
| | | Math - Difficult | 15.26 | 11.13 | 15.12 | 0.64 |
| | | Face Movement | 9.38 | 7.27 | 9.32 | 0.88 |
| | | All sessions combined | 10.42 | 7.02 | 10.26 | 0.87 |

For PRV, however, the agreement with the reference device during the difficult math task (ear: $r$ = 0.64, finger: $r$ = 0.75) and the face movement task (ear: $r$ = 0.88, finger: $r$ = 0.92) is lower than in the baseline. We can attribute this decrease in performance of PRV metrics for cognitively challenging and movement conditions to motion artifacts. Earlier studies involving commercial wearable PPG devices have also reported similar decreases in PRV accuracy [84].

In Figure 8A, we compare the SQI for PPG signals from both PhysioKit and the reference device under different experimental conditions. PhysioKit demonstrates consistent signal quality across different experimental conditions, while the reference device shows higher variance across each experimental condition. For PhysioKit, it can also be observed that the signal quality of the PPG sensor placed on the ear is higher than that of the finger for all conditions except face movements. Figure 8B compares the magnitude of facial movements across different experimental conditions, which is computed as a standard deviation of change in inter-frame rotation angle. A Wilcoxon Signed-Rank test found that motion increased significantly during math task and face-movement conditions from the baseline condition.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 August 2023 doi:10.20944/preprints202308.0545.v1



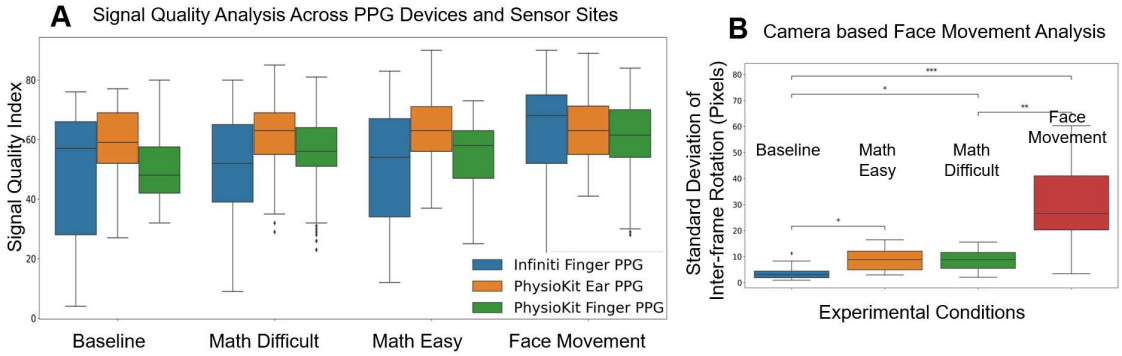


**Figure 8.** Signal Quality Index comparing PPG signal quality from both PhysioKit and the reference device under different experimental conditions (**A**) Analysis for facial movement, computed as standard-deviation of inter-frame rotation of detected facial frame from recorded video is presented in (**B**).

## 6. Study 2: Usability

PhysioKit was distributed to 10 micro-project teams (N = 44 members) to use for their own research purposes for four to six weeks. To gain insights on the usability of PhysioKit, we designed a questionnaire in Qualtrics that included three parts. The first section focused on obtaining demographic data and information on participants' prior experience with physiological sensing and toolkits. For the second part of the survey, we designed a modified version of the Usefulness, Satisfaction and Ease of Use questionnaire (USE; [85,86]) to gather insights on usability. The final section asked optional open-ended questions regarding participant's favorite aspects, and suggestions for improvements and novel features.

One person representing each of the 10 micro-project teams participated in completing the questionnaire (5 female, 5 male; 18 to 34 years old), which took an average of 22.5 minutes to complete. Participation was voluntary, and no identifying information was collected. All participants either had completed a postgraduate degree in computer science or were currently in the process of completing one. 7 participants used consumer-grade wearables (e.g. Apple Watch, FitBit Sense 2, Garmin, Withings Steel HR, Google Watch) either daily (N = 5) or occasionally (N = 2), while 3 had never used one before. In terms of previous experience with open source microcontroller kits, 6 participants had used them for a previous project or class (i.e. Arduino, Raspberry Pi, micro:bit Texas Instruments device), while 4 participants had no previous experience with microcontrollers.

### *6.1. Results*

#### 6.1.1. Usefulness and Ease of Use

Everyone who completed the survey acknowledged that PhysioKit was useful for their projects, with most strongly agreeing (N=7). All participants also agreed that it gave them full control over their project activities ("Strongly agree": N=5) and facilitated easy completion of their project tasks ("Strongly agree": N=4). Several participants with limited computing experience (N=3) considered using other physiological systems (e.g. Empatica, Apple Watch, fnv.reduce), but ultimately chose PhysioKit because it gave them full control over data and signals processing and setup.

Nearly all participants found that the provision of raw data was the most useful aspect of PhysioKit (N=9), while almost all (N=8) found that the ability to extract features, control data acquisition, and access supported data analysis was important. Also, having the flexibility to adapt PhysioKit to different experimental protocols was highly valued (N = 7). Participants found Physiokit easy to use ($M$ = 4.30, $SD$ = 0.675) and quick to set up ($M$ = 4.00, $SD$ = 0.943). They also appreciated that it enabled flexible configurations ($M$ = 4.40, $SD$ = 0.699).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 August 2023 doi:10.20944/preprints202308.0545.v1



#### 6.1.2. Learning Process

Most participants learned to use PhysioKit quickly ($M$ = 3.90, $SD$ = 0.568) and with different sensor configurations for a diverse range of study designs ($M$ = 3.90, $SD$ = 0.568). Once they learned how to use the toolkit, everyone found it easy to remember how to use ($M$ = 4.20, $SD$ = 0.422), regardless of computing their prior experience.

#### 6.1.3. Satisfaction

All participants were satisfied with the way PhysioKit worked ($M$ = 4.20, $SD$ = 0.422) and would recommend it to colleagues ($M$ = 4.10, $SD$ = 0.316). Many also found it essential for the completion of their projects (M = 4.60, SD = 0.699) and most would prefer to use it over other physiological systems for future projects ($M$ = 3.50, $SD$ = 0.527).

#### 6.1.4. Open-ended questions

When participants were asked what they appreciated about PhysioKit, one person with limited programming experience responded: *"It's easy to understand and user-friendly for people without a coding foundation"* (P3). People with high computing proficiency also found PhysioKit quick to setup, well-organized, simple and flexible to use. Lastly, participants left comments encouraging the further distribution of PhysioKit: *"Promote it, make it accessible to more people, [help them] understand the difference between using this product and using physiological sensors directly."* (P3).

### *6.2. Use Cases*

To illustrate how PhysioKit can be applied in practice, we summarize several examples of projects developed using the toolkit. Group members from each of these projects received a 1 hour hands-on tutoring and spent four to six weeks developing a computing system using PhysioKit.

**Table 4.** Project teams specific application cases of PhysioKit.

| **Application** | **Application Type †** | **No. of members** | **Project duration** |
|---|---|---|---|
| Using physiological reactions as emotional responses to music | Interventional | 6 | 4 weeks |
| Emotion recognition during watching of videos | Interventional | 6 | 4 weeks |
| Using artistic biofeedback to encourage mindfulness | Interventional | 1 | 6 weeks |
| Using acute stress response to determine game difficulty | Interventional | 7 | 4 weeks |
| Generating a dataset for an affective music recommendation system | Passive | 7 | 4 weeks |
| Adapting an endless runner game to player stress levels | Interventional | 8<br>7 | 4 weeks |
| Influencing presentation experience with social biofeedback | Interventional | 1 | 6 weeks |
| Mapping Stress in Virtual Reality | Passive | 1 | 6 weeks |
| Assessing synchronous heartbeats during a virtual reality game | Passive | 2 | 6 weeks |

†see § 2.1

#### 6.2.1. PhysioKit for Interventional Applications

Several groups explored using PhysioKit for developing affect recognition systems. For instance, some groups developed adaptive games that adjusted the level of gameplay difficulty by assessing acute stress using a combination of HR, HRV (e.g. pulse rate, pulse amplitude, interbeat interval), skin

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 August 2023 doi:10.20944/preprints202308.0545.v1



conductance and breathing rate. Another group examined how physiological responses from PPG and EDA could be mapped to a valence-arousal model to assess reactions to music. Other projects also examined the effects of biofeedback and social biofeedback visualizations of HR to overcome stressful scenarios (e.g. oral presentations) and promote mindfulness [9].

#### 6.2.2. PhysioKit for Passive Applications

Project groups took advantage of PhysioKit's diverse hardware and software functions to develop passive applications. For instance, a project that mapped stress during a virtual task assessed from an ear PPG showed that signals from this location were less subject to motion artifacts. Team members from a different group used PhysioKit's event-marking function to generate a dataset for an affective music recommender system. Lastly, another research project team used the multi-user client-server function to examine similarities in physiological responses during remote virtual reality experience.

In each of these examples, PhysioKit facilitates the exploration of interactive and passive data collection for a variety of contexts.

## 7. Discussion

In this section, we discuss our main findings and show how PhysioKit relates to and builds upon existing research.

### *7.1. Evaluating the Validation of PhysioKit*

The validation study in § 5.1 highlights very good agreement between PhysioKit and the gold standard for HR. The PRV metrics also show good agreement during the baseline condition, while there is acceptable agreement during experimental conditions involving significant movement. This difference between performance related to HR and PRV can be explained by the findings of a recent study that assessed the validity and reliability of PPG derived HRV metrics, and found that PPG sensors are less reliable for HRV measurements [87] . However, it is worth mentioning that the same study found PPG sensors to be accurate for measuring HR. It is also noteworthy that the overall performance of PhysioKit shows better agreement with the gold standard compared to the performance of existing PPG-based commercial devices mentioned in Table 1.

PhysioKit is a fully open source toolkit for physiological computing that offers cost-effective and flexible solutions to configure different physiological sensors according to research needs. With the overall cost of the PhysioKit hardware, which includes an Arduino Uno and one PPG sensor, being less than $50, this toolkit well below the least expensive of the commercial sensors mentioned in Table 1. Also, the cost of the collection of sensors used in this work (i.e., 2 PPG sensors, one EDA sensor and a one RSP sensor) amounts to less than $200, which is still less than the average cost of commercially available sensors in Table 1.

Regarding flexible configuration, the validation study results show similar performance of the finger PPG sensor and ear PPG sensor, suggesting that alternative sensor sites for PPG can be explored to achieve different research objectives and support accessibility. The PhysioKit UI can also be easily adapted to different study protocols, including both those requiring single-user and multi-user setups. Furthermore, access to raw data not only gives HCI researchers ownership and control over the data, but also enables computational research to develop robust algorithms for handling real-world environments. Taken together, these aspects all contribute to enhancing the usability of the toolkit for researchers.

### *7.2. Assessing the Usability of PhysioKit*

The usability survey results (see § 6) highlight the usefulness, learning experience and favorable aspects of PhysioKit. Participants agreed that PhysioKit was essential for their projects because it provided them with control and flexibility over tasks and allowed them to accomplish what they wanted to do. While access to raw data was valued as the most important aspect of the toolkit, they

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 August 2023 doi:10.20944/preprints202308.0545.v1



also appreciated having control over data acquisition, feature extraction and data analysis. Participants felt PhysioKit was easy to use and seemed enthusiastic about using PhysioKit for future projects as well as sharing it with colleagues and friends because of its open-source features.

### *7.3. Applicability of PhysioKit*

Projects with both interventional and passive application types utilized PhysioKit, demonstrating its versatility in supporting data acquisition in different settings. For instance, one project used PhysioKit to collect simultaneous signals from multiple users during a remote scenario (passive), while another explored the biofeedback function (interventional). In both application scenarios, researchers benefit from robust machine-learning based signal quality assessment, thereby increasing the reliability of computed physiological metrics.

### *7.4. Limitations and Future Work*

In this work, we implemented PPG, EDA and RSP sensors. However, future work could explore integrating other contact-based physiological sensors commonly used in HCI research (e.g., EMG, ECG, and EEG) [88,89]. The validation study [90–92] and signal quality assessment module (SQA-Phys) mentioned in this work currently only apply to data from PPG sensors. However, these can also be extended to other contact-based physiological sensors, such as the ones previously mentioned.

Lastly, our future implementation plan considers implementing a processing pipeline for contact-less sensing methods, such as RGB camera-based remote PPG [93,94], and thermal infrared sensing pipeline, including optimal quantization [27] and semantic segmentation [95] for extraction of breathing [27] and blood volume pulse signals [67,96,97]. We also aim to further enhance the accessibility of the sensor interface (hardware) of PhysioKit for its use in real-world scenarios.

**Author Contributions:** "Conceptualization, Y.C. and J.J.; methodology, J.J., K.W. and Y.C.; programming, J.J.; study validation, J.J. and K.W.; artefacts validation, J.J. and Y.C.; investigation, Y.C.; data collection and analysis, J.J., K.W. and Y.C; manuscript preparation, J.J., K.W. and Y.C; visualization, J.J.; overall supervision, Y.C.; project administration, Y.C.; funding acquisition, Y.C. All authors have read and agreed to the published version of the manuscript."

**Funding:** UCL CS Physiological Computing Studentship.

**Institutional Review Board Statement:** The study protocol is approved by the University College London Interaction Centre ethics committee (ID Number: UCLIC/1920/006/Staff/Cho).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** Data is unavailable due to privacy or ethical restrictions.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| 1D-CNN | 1-dimensional convolutional neural network |
| ADC | Analog-to-digital converter |
| API | Application programming interface |
| BPM | Beats per minute |
| BVP | Blood volume pulse |
| CAD | Computer-aided design |
| ECG | Electrocardiogram |
| EDA | Electrodermal activity |
| EMG | Electromyography |
| GSR | Galvanic skin response |

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 August 2023 doi:10.20944/preprints202308.0545.v1



| | |
|---|---|
| HCI | Human computer interface |
| HR | Heart rate |
| HRV | Heart rate variability |
| LSTM | Long short-term memory |
| PPG | Photoplethysmography |
| PRV | Pulse rate variability |
| RGB | Color images with red, green an blue frames |
| RSP | Respiration or breathing |
| SOTA | State-of-the-art |
| SQI | Signal quality index |
| SVM | Support vector machine |
| TCP/IP | Transmission control protocol/ Internet Protocol |
| UI | User interface |

## References


1. Elvitigala, D.S.; Somarathna, R.; Yan, Y.; Mohammadi, G.; Quigley, A. Towards Using Involuntary Body Gestures for Measuring the User Engagement in VR Gaming. The Adjunct Publication of the 35th Annual ACM Symposium on User Interface Software and Technology; ACM: Bend OR USA, 2022; pp. 1–3. doi:10.1145/3526114.3558691.
2. Chigira, H.; Maeda, A.; Kobayashi, M. Area-Based Photo-Plethysmographic Sensing Method for the Surfaces of Handheld Devices. Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology; Association for Computing Machinery: New York, NY, USA, 2011; UIST '11, pp. 499–508. doi:10.1145/2047196.2047262.
3. Norooz, L.; Mauriello, M.L.; Jorgensen, A.; McNally, B.; Froehlich, J.E. BodyVis: A New Approach to Body Learning Through Wearable Sensing and Visualization. Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems; Association for Computing Machinery: New York, NY, USA, 2015; CHI '15, pp. 1025–1034. doi:10.1145/2702123.2702299.
4. Almusawi, H.A.; Durugbo, C.M.; Bugawa, A.M. Wearable Technology in Education: A Systematic Review. *IEEE Transactions on Learning Technologies* **2021**, *14*, 540–554. doi:10.1109/TLT.2021.3107459.
5. Kang, S.; Norooz, L.; Oguamanam, V.; Plane, A.C.; Clegg, T.L.; Froehlich, J.E. SharedPhys: Live Physiological Sensing, Whole-Body Interaction, and Large-Screen Visualizations to Support Shared Inquiry Experiences. Proceedings of the The 15th International Conference on Interaction Design and Children; Association for Computing Machinery: New York, NY, USA, 2016; IDC '16, pp. 275–287. doi:10.1145/2930674.2930710.
6. Bustos-López, M.; Cruz-Ramírez, N.; Guerra-Hernández, A.; Sánchez-Morales, L.N.; Cruz-Ramos, N.A.; Alor-Hernández, G. Wearables for Engagement Detection in Learning Environments: A Review. *Biosensors* **2022**, *12*, 509. doi:10.3390/bios12070509.
7. Malasinghe, L.P.; Ramzan, N.; Dahal, K. Remote Patient Monitoring: A Comprehensive Study. *Journal of Ambient Intelligence and Humanized Computing* **2019**, *10*, 57–76. doi:10.1007/s12652-017-0598-x.
8. Chanel, G.; Mühl, C. Connecting Brains and Bodies: Applying Physiological Computing to Support Social Interaction. *Interacting with Computers* **2015**, *27*, 534–550. doi:10.1093/iwc/iwv013.
9. Moge, C.; Wang, K.; Cho, Y. Shared User Interfaces of Physiological Data: Systematic Review of Social Biofeedback Systems and Contexts in HCI. Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems; Association for Computing Machinery: New York, NY, USA, 2022; CHI '22, pp. 1–16. doi:10.1145/3491102.3517495.
10. Cho, Y. Rethinking Eye-blink: Assessing Task Difficulty through Physiological Representation of Spontaneous Blinking. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems; Association for Computing Machinery: New York, NY, USA, 2021; CHI '21, pp. 1–12. doi:10.1145/3411764.3445577.
11. Xiao, X.; Pham, P.; Wang, J. AttentiveLearner: Adaptive Mobile MOOC Learning via Implicit Cognitive States Inference. Proceedings of the 2015 ACM on International Conference on Multimodal Interaction; ACM: Seattle Washington USA, 2015; pp. 373–374. doi:10.1145/2818346.2823297.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 August 2023 doi:10.20944/preprints202308.0545.v1




12. DiSalvo, B.; Bandaru, D.; Wang, Q.; Li, H.; Plötz, T. Reading the Room: Automated, Momentary Assessment of Student Engagement in the Classroom: Are We There Yet? *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* **2022**, *6*, 112:1–112:26. doi:10.1145/3550328.
13. Huynh, S.; Kim, S.; Ko, J.; Balan, R.K.; Lee, Y. EngageMon: Multi-Modal Engagement Sensing for Mobile Games. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* **2018**, *2*, 13:1–13:27. doi:10.1145/3191745.
14. Qin, C.Y.; Choi, J.H.; Constantinides, M.; Aiello, L.M.; Quercia, D. Having a Heart Time? A Wearable-based Biofeedback System. 22nd International Conference on Human-Computer Interaction with Mobile Devices and Services; Association for Computing Machinery: New York, NY, USA, 2020; MobileHCI '20, pp. 1–4. doi:10.1145/3406324.3410539.
15. Howell, N.; Niemeyer, G.; Ryokai, K. Life-Affirming Biosensing in Public: Sounding Heartbeats on a Red Bench. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems; Association for Computing Machinery: New York, NY, USA, 2019; CHI '19, pp. 1–16. doi:10.1145/3290605.3300910.
16. Curran, M.T.; Gordon, J.R.; Lin, L.; Sridhar, P.K.; Chuang, J. Understanding Digitally-Mediated Empathy: An Exploration of Visual, Narrative, and Biosensory Informational Cues. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems; ACM: Glasgow Scotland Uk, 2019; pp. 1–13. doi:10.1145/3290605.3300844.
17. Moraveji, N.; Olson, B.; Nguyen, T.; Saadat, M.; Khalighi, Y.; Pea, R.; Heer, J. Peripheral Paced Respiration: Influencing User Physiology during Information Work. Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology; ACM: Santa Barbara California USA, 2011; pp. 423–428. doi:10.1145/2047196.2047250.
18. Wongsuphasawat, K.; Gamburg, A.; Moraveji, N. You Can't Force Calm: Designing and Evaluating Respiratory Regulating Interfaces for Calming Technology. Adjunct Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology; ACM: Cambridge Massachusetts USA, 2012; pp. 69–70. doi:10.1145/2380296.2380326.
19. Frey, J.; Grabli, M.; Slyper, R.; Cauchard, J.R. Breeze: Sharing Biofeedback through Wearable Technologies. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems; Association for Computing Machinery: New York, NY, USA, 2018; CHI '18, pp. 1–12. doi:10.1145/3173574.3174219.
20. Fortin, P.E.; Sulmont, E.; Cooperstock, J.R. SweatSponse: Closing the Loop on Notification Delivery Using Skin Conductance Responses. Adjunct Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology; ACM: Berlin Germany, 2018; pp. 7–9. doi:10.1145/3266037.3266084.
21. Yabutani, M.; Mashiba, Y.; Kawamura, H.; Harada, S.; Zempo, K. Sharing Heartbeat: Toward Conducting Heartrate and Speech Rhythm through Tactile Presentation of Pseudo-heartbeats. Adjunct Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology; Association for Computing Machinery: New York, NY, USA, 2022; UIST '22 Adjunct, pp. 1–4. doi:10.1145/3526114.3558704.
22. Liu, F. Expressive Biosignals: Authentic Social Cues for Social Connection. Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems; Association for Computing Machinery: New York, NY, USA, 2019; CHI EA '19, pp. 1–5. doi:10.1145/3290607.3299081.
23. Dey, A.; Piumsomboon, T.; Lee, Y.; Billinghurst, M. Effects of Sharing Physiological States of Players in a Collaborative Virtual Reality Gameplay. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems; Association for Computing Machinery: New York, NY, USA, 2017; CHI '17, pp. 4045–4056. doi:10.1145/3025453.3026028.
24. Ayobi, A.; Sonne, T.; Marshall, P.; Cox, A.L. Flexible and Mindful Self-Tracking: Design Implications from Paper Bullet Journals. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems; Association for Computing Machinery: New York, NY, USA, 2018; CHI '18, pp. 1–14. doi:10.1145/3173574.3173602.
25. Koydemir, H.C.; Ozcan, A. Wearable and Implantable Sensors for Biomedical Applications. *Annual Review of Analytical Chemistry* **2018**, *11*, 127–146. doi:10.1146/annurev-anchem-061417-125956.
26. Peake, J.M.; Kerr, G.; Sullivan, J.P. A Critical Review of Consumer Wearables, Mobile Applications, and Equipment for Providing Biofeedback, Monitoring Stress, and Sleep in Physically Active Populations. *Frontiers in Physiology* **2018**, *9*.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 August 2023 doi:10.20944/preprints202308.0545.v1




27. Cho, Y.; Julier, S.J.; Marquardt, N.; Bianchi-Berthouze, N. Robust Tracking of Respiratory Rate in High-Dynamic Range Scenes Using Mobile Thermal Imaging. *Biomedical Optics Express* **2017**, *8*, 4480–4503. doi:10.1364/BOE.8.004480.
28. Cho, Y.; Julier, S.J.; Bianchi-Berthouze, N. Instant Stress: Detection of Perceived Mental Stress Through Smartphone Photoplethysmography and Thermal Imaging. *JMIR Mental Health* **2019**, *6*, e10140. doi:10.2196/10140.
29. Peake, J.M.; Kerr, G.; Sullivan, J.P. A Critical Review of Consumer Wearables, Mobile Applications, and Equipment for Providing Biofeedback, Monitoring Stress, and Sleep in Physically Active Populations. *Frontiers in Physiology* **2018**, *9*.
30. Ometov, A.; Shubina, V.; Klus, L.; Skibińska, J.; Saafi, S.; Pascacio, P.; Flueratoru, L.; Gaibor, D.Q.; Chukhno, N.; Chukhno, O.; Ali, A.; Channa, A.; Svertoka, E.; Qaim, W.B.; Casanova-Marqués, R.; Holcer, S.; Torres-Sospedra, J.; Casteleyn, S.; Ruggeri, G.; Araniti, G.; Burget, R.; Hosek, J.; Lohan, E.S. A Survey on Wearable Technology: History, State-of-the-Art and Current Challenges. *Computer Networks* **2021**, *193*, 108074. doi:10.1016/j.comnet.2021.108074.
31. Baron, K.G.; Duffecy, J.; Berendsen, M.A.; Cheung Mason, I.; Lattie, E.G.; Manalo, N.C. Feeling Validated yet? A Scoping Review of the Use of Consumer-Targeted Wearable and Mobile Technology to Measure and Improve Sleep. *Sleep Medicine Reviews* **2018**, *40*, 151–159. doi:10.1016/j.smrv.2017.12.002.
32. Burd, B.; Barker, L.; Divitini, M.; Perez, F.A.F.; Russell, I.; Siever, B.; Tudor, L. Courses, Content, and Tools for Internet of Things in Computer Science Education. Proceedings of the 2017 ITiCSE Conference on Working Group Reports; Association for Computing Machinery: New York, NY, USA, 2018; ITiCSE-WGR '17, pp. 125–139. doi:10.1145/3174781.3174788.
33. Arias, O.; Wurm, J.; Hoang, K.; Jin, Y. Privacy and Security in Internet of Things and Wearable Devices. *IEEE Transactions on Multi-Scale Computing Systems* **2015**, *1*, 99–109.
34. Sato, M.; Puri, R.S.; Olwal, A.; Chandra, D.; Poupyrev, I.; Raskar, R. Zensei: Augmenting Objects with Effortless User Recognition Capabilities through Bioimpedance Sensing. Adjunct Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology; ACM: Daegu Kyungpook Republic of Korea, 2015; pp. 41–42. doi:10.1145/2815585.2817786.
35. Shibata, T.; Peck, E.M.; Afergan, D.; Hincks, S.W.; Yuksel, B.F.; Jacob, R.J. Building Implicit Interfaces for Wearable Computers with Physiological Inputs: Zero Shutter Camera and Phylter. Proceedings of the Adjunct Publication of the 27th Annual ACM Symposium on User Interface Software and Technology - UIST'14 Adjunct; ACM Press: Honolulu, Hawaii, USA, 2014; pp. 89–90. doi:10.1145/2658779.2658790.
36. Yamamura, H.; Baldauf, H.; Kunze, K. Pleasant Locomotion – Towards Reducing Cybersickness Using fNIRS during Walking Events in VR. Adjunct Publication of the 33rd Annual ACM Symposium on User Interface Software and Technology; ACM: Virtual Event USA, 2020; pp. 56–58. doi:10.1145/3379350.3416184.
37. Pai, Y.S. Physiological Signal-Driven Virtual Reality in Social Spaces. Proceedings of the 29th Annual Symposium on User Interface Software and Technology; ACM: Tokyo Japan, 2016; pp. 25–28. doi:10.1145/2984751.2984787.
38. Lee, J.H.; Gamper, H.; Tashev, I.; Dong, S.; Ma, S.; Remaley, J.; Holbery, J.D.; Yoon, S.H. Stress Monitoring Using Multimodal Bio-sensing Headset. Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems; ACM: Honolulu HI USA, 2020; pp. 1–7. doi:10.1145/3334480.3382891.
39. Maritsch, M.; Föll, S.; Lehmann, V.; Bérubé, C.; Kraus, M.; Feuerriegel, S.; Kowatsch, T.; Züger, T.; Stettler, C.; Fleisch, E.; Wortmann, F. Towards Wearable-based Hypoglycemia Detection and Warning in Diabetes. Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems; ACM: Honolulu HI USA, 2020; pp. 1–8. doi:10.1145/3334480.3382808.
40. Liang, R.H.; Yu, B.; Xue, M.; Hu, J.; Feijs, L.M.G. BioFidget: Biofeedback for Respiration Training Using an Augmented Fidget Spinner. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems; ACM: Montreal QC Canada, 2018; pp. 1–12. doi:10.1145/3173574.3174187.
41. Clegg, T.; Norooz, L.; Kang, S.; Byrne, V.; Katzen, M.; Velez, R.; Plane, A.; Oguamanam, V.; Outing, T.; Yip, J.; Bonsignore, E.; Froehlich, J. Live Physiological Sensing and Visualization Ecosystems: An Activity Theory Analysis. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems; ACM: Denver Colorado USA, 2017; pp. 2029–2041. doi:10.1145/3025453.3025987.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 August 2023 doi:10.20944/preprints202308.0545.v1




42. Wang, H.; Prendinger, H.; Igarashi, T. Communicating Emotions in Online Chat Using Physiological Sensors and Animated Text. CHI '04 Extended Abstracts on Human Factors in Computing Systems; ACM: Vienna Austria, 2004; pp. 1171–1174. doi:10.1145/985921.986016.
43. Dey, A.; Piumsomboon, T.; Lee, Y.; Billinghurst, M. Effects of Sharing Physiological States of Players in a Collaborative Virtual Reality Gameplay. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems; ACM: Denver Colorado USA, 2017; pp. 4045–4056. doi:10.1145/3025453.3026028.
44. Olugbade, T.; Cho, Y.; Morgan, Z.; El Ghani, M.A.; Bianchi-Berthouze, N. Toward Intelligent Car Comfort Sensing: New Dataset and Analysis of Annotated Physiological Metrics. 2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII), 2021, pp. 1–8. doi:10.1109/ACII52823.2021.9597393.
45. Dillen, N.; Ilievski, M.; Law, E.; Nacke, L.E.; Czarnecki, K.; Schneider, O. Keep Calm and Ride Along: Passenger Comfort and Anxiety as Physiological Responses to Autonomous Driving Styles. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems; ACM: Honolulu HI USA, 2020; pp. 1–13. doi:10.1145/3313831.3376247.
46. Frey, J.; Ostrin, G.; Grabli, M.; Cauchard, J.R. Physiologically Driven Storytelling: Concept and Software Tool. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems; ACM: Honolulu HI USA, 2020; pp. 1–13. doi:10.1145/3313831.3376643.
47. Karran, A.J.; Fairclough, S.H.; Gilleade, K. Towards an Adaptive Cultural Heritage Experience Using Physiological Computing. CHI '13 Extended Abstracts on Human Factors in Computing Systems; ACM: Paris France, 2013; pp. 1683–1688. doi:10.1145/2468356.2468657.
48. Robinson, R.B.; Reid, E.; Depping, A.E.; Mandryk, R.; Fey, J.C.; Isbister, K. 'In the Same Boat',: A Game of Mirroring Emotions for Enhancing Social Play. Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems; ACM: Glasgow Scotland Uk, 2019; pp. 1–4. doi:10.1145/3290607.3313268.
49. Schneegass, S.; Pfleging, B.; Broy, N.; Heinrich, F.; Schmidt, A. A Data Set of Real World Driving to Assess Driver Workload. Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications; Association for Computing Machinery: New York, NY, USA, 2013; AutomotiveUI '13, pp. 150–157. doi:10.1145/2516540.2516561.
50. Fairclough, S.H.; Karran, A.J.; Gilleade, K. Classification Accuracy from the Perspective of the User: Real-Time Interaction with Physiological Computing. Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems; ACM: Seoul Republic of Korea, 2015; pp. 3029–3038. doi:10.1145/2702123.2702454.
51. Oh, S.Y.; Kim, S. Does Social Endorsement Influence Physiological Arousal? Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems; ACM: San Jose California USA, 2016; pp. 2900–2905. doi:10.1145/2851581.2892437.
52. Sato, Y.; Ueoka, R. Investigating Haptic Perception of and Physiological Responses to Air Vortex Rings on a User's Cheek. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems; ACM: Denver Colorado USA, 2017; pp. 3083–3094. doi:10.1145/3025453.3025501.
53. Frey, J.; Daniel, M.; Castet, J.; Hachet, M.; Lotte, F. Framework for Electroencephalography-based Evaluation of User Experience. Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems; Association for Computing Machinery: New York, NY, USA, 2016; CHI '16, pp. 2283–2294. doi:10.1145/2858036.2858525.
54. Duarte, L.; Carriço, L. The Cake Can Be a Lie: Placebos as Persuasive Videogame Elements. CHI '13 Extended Abstracts on Human Factors in Computing Systems; ACM: Paris France, 2013; pp. 1113–1118. doi:10.1145/2468356.2468555.
55. Mandryk, R.L. Objectively Evaluating Entertainment Technology. CHI '04 Extended Abstracts on Human Factors in Computing Systems; ACM: Vienna Austria, 2004; pp. 1057–1058. doi:10.1145/985921.985977.
56. Nukarinen, T.; Istance, H.O.; Rantala, J.; Mäkelä, J.; Korpela, K.; Ronkainen, K.; Surakka, V.; Raisamo, R. Physiological and Psychological Restoration in Matched Real and Virtual Natural Environments. Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems; ACM: Honolulu HI USA, 2020; pp. 1–8. doi:10.1145/3334480.3382956.
57. Mannhart, D.; Lischer, M.; Knecht, S.; du Fay de Lavallaz, J.; Strebel, I.; Serban, T.; Vögeli, D.; Schaer, B.; Osswald, S.; Mueller, C.; Kühne, M.; Sticherling, C.; Badertscher, P. Clinical Validation of 5 Direct-to-Consumer Wearable Smart Devices to Detect Atrial Fibrillation: BASEL Wearable Study. *JACC: Clinical Electrophysiology* **2023**, *9*, 232–242. doi:10.1016/j.jacep.2022.09.011.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 August 2023 doi:10.20944/preprints202308.0545.v1




58. Bent, B.; Goldstein, B.A.; Kibbe, W.A.; Dunn, J.P. Investigating Sources of Inaccuracy in Wearable Optical Heart Rate Sensors. *npj Digital Medicine* **2020**, *3*, 1–9. doi:10.1038/s41746-020-0226-6.
59. Düking, P.; Giessing, L.; Frenkel, M.O.; Koehler, K.; Holmberg, H.C.; Sperlich, B. Wrist-Worn Wearables for Monitoring Heart Rate and Energy Expenditure While Sitting or Performing Light-to-Vigorous Physical Activity: Validation Study. *JMIR mHealth and uHealth* **2020**, *8*, e16716. doi:10.2196/16716.
60. Miller, D.J.; Sargent, C.; Roach, G.D. A Validation of Six Wearable Devices for Estimating Sleep, Heart Rate and Heart Rate Variability in Healthy Adults. *Sensors* **2022**, *22*, 6317. doi:10.3390/s22166317.
61. Chow, H.W.; Yang, C.C. Accuracy of Optical Heart Rate Sensing Technology in Wearable Fitness Trackers for Young and Older Adults: Validation and Comparison Study. *JMIR mHealth and uHealth* **2020**, *8*, e14707. doi:10.2196/14707.
62. Gilgen-Ammann, R.; Schweizer, T.; Wyss, T. RR Interval Signal Quality of a Heart Rate Monitor and an ECG Holter at Rest and during Exercise. *European Journal of Applied Physiology* **2019**, *119*, 1525–1532. doi:10.1007/s00421-019-04142-5.
63. Batista, D.; Plácido da Silva, H.; Fred, A.; Moreira, C.; Reis, M.; Ferreira, H.A. Benchmarking of the BITalino Biomedical Toolkit against an Established Gold Standard. *Healthcare Technology Letters* **2019**, *6*, 32–36. doi:10.1049/htl.2018.5037.
64. Cassirame, J.; Vanhaesebrouck, R.; Chevrolat, S.; Mourot, L. Accuracy of the Garmin 920 XT HRM to Perform HRV Analysis. *Australasian Physical & Engineering Sciences in Medicine* **2017**, *40*, 831–839. doi:10.1007/s13246-017-0593-8.
65. Müller, A.M.; Wang, N.X.; Yao, J.; Tan, C.S.; Low, I.C.C.; Lim, N.; Tan, J.; Tan, A.; Müller-Riemenschneider, F. Heart Rate Measures From Wrist-Worn Activity Trackers in a Laboratory and Free-Living Setting: Validation Study. *JMIR mHealth and uHealth* **2019**, *7*, e14120. doi:10.2196/14120.
66. Sweeney, K.T.; Ward, T.E.; McLoone, S.F. Artifact Removal in Physiological Signals—Practices and Possibilities. *IEEE Transactions on Information Technology in Biomedicine* **2012**, *16*, 488–500. doi:10.1109/TITB.2012.2188536.
67. Akbar, F.; Mark, G.; Pavlidis, I.; Gutierrez-Osuna, R. An Empirical Study Comparing Unobtrusive Physiological Sensors for Stress Detection in Computer Work. *Sensors* **2019**, *19*, 3766. doi:10.3390/s19173766.
68. Elgendi, M. Optimal Signal Quality Index for Photoplethysmogram Signals. *Bioengineering* **2016**, *3*, 21. doi:10.3390/bioengineering3040021.
69. Pereira, T.; Gadhoumi, K.; Ma, M.; Liu, X.; Xiao, R.; Colorado, R.A.; Keenan, K.J.; Meisel, K.; Hu, X. A Supervised Approach to Robust Photoplethysmography Quality Assessment. *IEEE Journal of Biomedical and Health Informatics* **2020**, *24*, 649–657. doi:10.1109/JBHI.2019.2909065.
70. Gao, H.; Wu, X.; Shi, C.; Gao, Q.; Geng, J. A LSTM-Based Realtime Signal Quality Assessment for Photoplethysmogram and Remote Photoplethysmogram. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 3831–3840.
71. Shin, H. Deep Convolutional Neural Network-Based Signal Quality Assessment for Photoplethysmogram. *Computers in Biology and Medicine* **2022**, *145*, 105430. doi:10.1016/j.compbiomed.2022.105430.
72. Ahmed, T.; Rahman, M.M.; Nemati, E.; Ahmed, M.Y.; Kuang, J.; Gao, A.J. Remote Breathing Rate Tracking in Stationary Position Using the Motion and Acoustic Sensors of Earables. Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems; Association for Computing Machinery: New York, NY, USA, 2023; CHI '23, pp. 1–22. doi:10.1145/3544548.3581265.
73. Gashi, S.; Di Lascio, E.; Santini, S. Using Unobtrusive Wearable Sensors to Measure the Physiological Synchrony Between Presenters and Audience Members. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* **2019**, *3*, 13:1–13:19. doi:10.1145/3314400.
74. Milstein, N.; Gordon, I. Validating Measures of Electrodermal Activity and Heart Rate Variability Derived From the Empatica E4 Utilized in Research Settings That Involve Interactive Dyadic States. *Frontiers in Behavioral Neuroscience* **2020**, *14*.
75. Lehrer, P.M.; Gevirtz, R. Heart Rate Variability Biofeedback: How and Why Does It Work? *Frontiers in Psychology* **2014**, *5*, 756. doi:10.3389/fpsyg.2014.00756.
76. Giggins, O.M.; Persson, U.M.; Caulfield, B. Biofeedback in Rehabilitation. *Journal of NeuroEngineering and Rehabilitation* **2013**, *10*, 60. doi:10.1186/1743-0003-10-60.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 August 2023 doi:10.20944/preprints202308.0545.v1




77. Makowski, D.; Pham, T.; Lau, Z.J.; Brammer, J.C.; Lespinasse, F.; Pham, H.; Schölzel, C.; Chen, S.H.A. NeuroKit2: A Python Toolbox for Neurophysiological Signal Processing. *Behavior Research Methods* **2021**, *53*, 1689–1696. doi:10.3758/s13428-020-01516-y.
78. van Gent, P.; Farah, H.; Nes, N.; Arem, B. Heart Rate Analysis for Human Factors: Development and Validation of an Open Source Toolkit for Noisy Naturalistic Heart Rate Data. 2018.
79. Tonacci, A.; Billeci, L.; Burrai, E.; Sansone, F.; Conte, R. Comparative Evaluation of the Autonomic Response to Cognitive and Sensory Stimulations through Wearable Sensors. *Sensors (Basel, Switzerland)* **2019**, *19*, 4661. doi:10.3390/s19214661.
80. Birkett, M.A. The Trier Social Stress Test Protocol for Inducing Psychological Stress. *Journal of Visualized Experiments : JoVE* **2011**, p. 3238. doi:10.3791/3238.
81. Johnson, K.T.; Narain, J.; Ferguson, C.; Picard, R.; Maes, P. The ECHOS Platform to Enhance Communication for Nonverbal Children with Autism: A Case Study. Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems; Association for Computing Machinery: New York, NY, USA, 2020; CHI EA '20, pp. 1–8. doi:10.1145/3334480.3375206.
82. Altman, D.G.; Bland, J.M. Measurement in Medicine: The Analysis of Method Comparison Studies. *Journal of the Royal Statistical Society. Series D (The Statistician)* **1983**, *32*, 307–317, [2987937]. doi:10.2307/2987937.
83. Giavarina, D. Understanding Bland Altman Analysis. *Biochemia Medica* **2015**, *25*, 141–151. doi:10.11613/BM.2015.015.
84. Georgiou, K.; Larentzakis, A.V.; Khamis, N.N.; Alsuhaibani, G.I.; Alaska, Y.A.; Giallafos, E.J. Can Wearable Devices Accurately Measure Heart Rate Variability? A Systematic Review. *Folia Medica* **2018**, *60*, 7–20. doi:10.2478/folmed-2018-0012.
85. Lund, A. Measuring Usability with the USE Questionnaire. *Usability and User Experience Newsletter of the STC Usability SIG* **2001**, *8*.
86. Gao, M.; Kortum, P.; Oswald, F. Psychometric Evaluation of the USE (Usefulness, Satisfaction, and Ease of Use) Questionnaire for Reliability and Validity. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* **2018**, *62*, 1414–1418. doi:10.1177/1541931218621322.
87. Speer, K.E.; Semple, S.; Naumovski, N.; McKune, A.J. Measuring Heart Rate Variability Using Commercially Available Devices in Healthy Children: A Validity and Reliability Study. *European Journal of Investigation in Health, Psychology and Education* **2020**, *10*, 390–404. doi:10.3390/ejihpe10010029.
88. Cowley, B.; Filetti, M.; Lukander, K.; Torniainen, J.; Henelius, A.; Ahonen, L.; Barral, O.; Kosunen, I.; Valtonen, T.; Huotilainen, M.; Ravaja, N.; Jacucci, G. The Psychophysiology Primer: A Guide to Methods and a Broad Review with a Focus on Human–Computer Interaction. *Foundations and Trends® in Human–Computer Interaction* **2016**, *9*, 151–308. doi:10.1561/1100000065.
89. Wang, T.; Zhang, H. Using Wearable Devices for Emotion Recognition in Mobile Human- Computer Interaction: A Review. HCI International 2022 - Late Breaking Papers. Multimodality in Advanced Interaction Environments; Kurosu, M.; Yamamoto, S.; Mori, H.; Schmorrow, D.D.; Fidopiastis, C.M.; Streitz, N.A.; Konomi, S., Eds.; Springer Nature Switzerland: Cham, 2022; Lecture Notes in Computer Science, pp. 205–227. doi:10.1007/978-3-031-17618-0_16.
90. van Lier, H.G.; Pieterse, M.E.; Garde, A.; Postel, M.G.; de Haan, H.A.; Vollenbroek-Hutten, M.M.R.; Schraagen, J.M.; Noordzij, M.L. A Standardized Validity Assessment Protocol for Physiological Signals from Wearable Technology: Methodological Underpinnings and an Application to the E4 Biosensor. *Behavior Research Methods* **2020**, *52*, 607–629. doi:10.3758/s13428-019-01263-9.
91. Shiri, S.; Feintuch, U.; Weiss, N.; Pustilnik, A.; Geffen, T.; Kay, B.; Meiner, Z.; Berger, I. A Virtual Reality System Combined with Biofeedback for Treating Pediatric Chronic Headache—A Pilot Study. *Pain Medicine* **2013**, *14*, 621–627. doi:10.1111/pme.12083.
92. Reali, P.; Tacchino, G.; Rocco, G.; Cerutti, S.; Bianchi, A.M. Heart Rate Variability from Wearables: A Comparative Analysis Among Standard ECG, a Smart Shirt and a Wristband. *Studies in Health Technology and Informatics* **2019**, *261*, 128–133.
93. Wang, C.; Pun, T.; Chanel, G. A Comparative Survey of Methods for Remote Heart Rate Detection From Frontal Face Videos. *Frontiers in Bioengineering and Biotechnology* **2018**, *6*.
94. Yu, Z.; Li, X.; Zhao, G. Facial-Video-Based Physiological Signal Measurement: Recent Advances and Affective Applications. *IEEE Signal Processing Magazine* **2021**, *38*, 50–58. doi:10.1109/MSP.2021.3106285.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 August 2023 doi:10.20944/preprints202308.0545.v1



95. Joshi, J.; Bianchi-Berthouze, N.; Cho, Y. Self-Adversarial Multi-scale Contrastive Learning for Semantic Segmentation of Thermal Facial Images, 2022, [arxiv:cs/2209.10700]. doi:10.48550/arXiv.2209.10700.
96. Manullang, M.C.T.; Lin, Y.H.; Lai, S.J.; Chou, N.K. Implementation of Thermal Camera for Non-Contact Physiological Measurement: A Systematic Review. *Sensors* **2021**, *21*, 7777. doi:10.3390/s21237777.
97. Gault, T.; Farag, A. A Fully Automatic Method to Extract the Heart Rate from Thermal Video. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2013, pp. 336–341.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.