Navigating the Challenges of Wearables: A Case Study of Research-Grade Devices

Jemma L König
School of Computing and Mathematical Sciences
University of Waikato
Hamilton, New Zealand
jemma.konig@waikato.ac.nz

Ben J Bidois
Health, Sport and Human Performance
University of Waikato
Hamilton, New Zealand
bb140@students.waikato.ac.nz

Ryan Manchester
School of Computing and Mathematical Sciences
University of Waikato
Hamilton, New Zealand
rm534@students.waikato.ac.nz

Mitchell A. Head
Te Kotahi Research Institute
University of Waikato
Hamilton, Waikato, New Zealand
mhead@waikato.ac.nz

Mahonri W Owen
Te Kura Mata Ao
Te whare Wānanaga o Waikato
Hamilton, New Zealand
mahonri.owen@waikato.ac.nz

Merel Cornelie Janna Hoskens
Te Huataki Waiora School of Health
University of Waikato
Hamilton, New Zealand
merel.hoskens@waikato.ac.nz

Anany Dwivedi
Artificial Intelligence Institute
University of Waikato
Hamilton, New Zealand
anany.dwivedi@waikato.ac.nz

Luke R.J. August
University of Waikato
Hamilton, New Zealand
la265@students.waikato.ac.nz

Abstract

The use of wearable devices in HCI research is becoming more common. However, it is also widely accepted that there are challenges associated with the use of wearable devices. Researchers have investigated generalized challenges, and more recently, challenges associated with low-cost commercial devices. Yet, little work has been published on the challenges of working with higher-cost research-grade devices. In this paper, we use a case study to explore the challenges associated with research-grade devices, and compare these challenges with those that have been identified by other researchers. The case study involves the use of two research-grade devices to record a series of physiological data-points. The case study identified one new challenge that is associated with research-grade wearable devices (setup complexity (researcher adoption)) and highlighted similarities between generalized challenges, low-cost device challenges, and research-grade challenges. The findings from this work can be used to further investigate challenges associated with wearable devices.

CCS Concepts

• **Human-centered computing → Empirical studies in HCI**; **User studies**; **Ubiquitous and mobile devices**; **Empirical studies in ubiquitous and mobile computing**.

OZCHI '25, Sydney, Australia

ACM ISBN 979-8-4007-2016-1/25/11
https://doi.org/10.1145/3764687.3769942

Keywords

Challenges of wearable devices, participatory study, wearable devices, physiological devices, research-grade devices

ACM Reference Format:
Jemma L König, Ben J Bidois, Ryan Manchester, Mitchell A. Head, Mahonri W Owen, Merel Cornelie Janna Hoskens, Anany Dwivedi, and Luke R.J. August. 2025. Navigating the Challenges of Wearables: A Case Study of Research-Grade Devices. In *37th Australian Conference on Human-Computer Interaction (HCI) (OZCHI '25), November 29–December 03, 2025, Sydney, Australia.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3764687.3769942

1 Introduction

The integration of wearable devices in Human-Computer Interaction (HCI) research has become increasingly prevalent, with applications ranging from Brain-Computer Interfaces [7] and stress recognition [10] to physiological feedback in Virtual Reality [11]. Despite the growing use of wearable devices for measuring physiological data, numerous challenges persist. A comprehensive set of general challenges associated with wearable devices, including issues with data acquisition, processing, transmission, and concerns over security and privacy, have been outlined by [15]. Additionally, [9] highlighted challenges specific to low-cost, commercially available wearable devices, such as data accessibility, interoperability, and fitness-for-purpose. (Discussed in detail in Section 2.1).

While much attention has been given to the generalized challenges, and more recently, to low-cost commercial devices, there has been limited research on the challenges associated with higher-cost, research-grade wearable devices. These devices, often considered more reliable and accurate, are frequently adopted in studies to overcome some of the limitations of commercial alternatives. However, the unique challenges posed by these research-grade devices remain underexplored.

In this paper, we focus on investigating the challenges associated with research-grade wearable devices. We utilize two research-grade devices, the Shimmer Consensys Bundle Development Kit[1] and the Muse EEG Headband[2], which were selected for their comprehensive capabilities in physiological data collection. The Shimmer Consensys Bundle Development Kit is an all-in-one solution designed for research, offering a range of sensors, hardware, and software [25]. The Muse EEG Headband, a 4-channel EEG sensor [13, 14], measures brain activity, particularly in the frontal and temporal lobes.

In this work, we employed these research-grade devices to measure a variety of physiological data points—including brain activity, heart rate, heart rate variability, electrodermal activity, muscle activity, and movement—from 18 participants. The focus of this paper is on the challenges encountered while working with these research-grade wearable devices. By analyzing these challenges, we aim to contribute new insights into the difficulties faced by researchers when using high-end wearable technology.

This paper is structured as follows: we first discuss the challenges identified by previous research and review the general use of wearable devices in physiological data measurement. We then present a case study that involved recording physiological data from 18 participants using research-grade wearable devices. Finally, we discuss the specific challenges associated with research-grade wearable devices, compare our findings with those related to low-cost, commercially available devices, and highlight areas that require further investigation in the context of high-end wearable technology.

2 Background and related work

2.1 Challenges in wearable technology

It is widely acknowledged that the use of wearable devices presents challenges for researchers. [15], for example, carried out a state-of-the-art review of wearable devices. In this review, they discuss the historical background of wearables and examine current wearable technologies. As part of this work, [15] identify nine key challenges, as follows.

(1) **Data acquisition and processing** focuses on the quality, quantity, and resolution of the data, both during collection and processing. For instance, various devices may differ in accuracy, employ different sampling rates, and offer varying degrees of filtering and smoothing.
(2) **Data transmission** focuses on data latency and network load. This is especially important when working with devices that rely on a centralized cloud for data management, processing, and storage.
(3) **Security and privacy** focuses on concerns regarding the reliability and trustworthiness of data transmission and the associated cloud providers. Although these concerns are relevant to other areas of computing, they are particularly significant for wearable devices due to the nature of the data typically recorded, such as continuously monitored location-based and physiological data.
(4) **Localization** addresses the challenge of precision when using embedded positioning technology. This is especially problematic with wearable devices used indoors. For example, working with radio signals, Non-Line-of-Sight can lead to position inaccuracies.
(5) **Communication and architecture** focuses on the challenge of developing protocols and architectures that facilitate communication between heterogeneous devices.
(6) **User adoption** addresses the psychological and emotional impacts of wearable device adoption on users, particularly in the medical field where the use of these devices is often not voluntary.
(7) **Hardware constraints** focuses on the resource limitations associated with wearable hardware, such as battery consumption, limited memory space, and constrained processing power.
(8) **Interoperability** focuses on the diverse nature of wearable devices. Beyond the communication challenges between different devices, interoperability issues also involve the integration of wearable components, services, and applications.
(9) **Management and scalability** focuses on managing a network of wearable devices. Currently, there is often a disconnect between devices owned by different users. While a single user may have a connected network of devices, these devices are generally unaware of those owned by other users.

In addition to these challenges, [9] identified three further challenges that tend to occur when working with low-cost commercially available wearable devices [9].

(1) **Data accessibility:** Commercial wearable devices generally do not offer access to raw data. Research-based wearable devices typically provide access to raw data, either for free or through subscription-based software. In contrast, commercial devices are usually not designed for applications that require raw data access and thus do not provide it. This limitation presents a significant challenge when using commercial devices for research purposes.
(2) **Interoperability:** This challenge was identified by both [15], when considering wearable devices in general, and by [9], when considering low-cost commercially available wearable devices. Interoperability focuses on the diverse nature of wearable devices. In the case of [9], the challenges of interoperability were compounded by data acquisition and processing issues, as well as the use of two devices and two mobile applications. More broadly, interoperability challenges can involve the variety of wearable components, communication methods, services, and applications.
(3) **Fit-for-purpose:** [15] did not address the use of wearable devices outside of their intended environments. In contrast, when using low-cost commercial devices for research, one of the biggest challenges [9] encountered was determining whether a device is suitable for its target environment. In this case, the issue arose when a fitness tracker and a stress reduction wearable were used on-site in industry. [9] suggest that similar challenges could occur when using low-cost commercial wearables anywhere outside of their intended purpose, for example in any hazardous or fast-paced physical setting.

[1] https://shimmersensing.com/product/consensys-bundle-development-kit/
[2] https://choosemuse.com/

While [15] highlight a series of common challenges, they generalize these challenges to all wearable devices. To build on this, [9] highlighted challenges involved in adopting physiological-data based research purposes. However, their focus was on low-cost commercial devices. In this paper, we investigate the challenges involved in adopting *research-grade* devices for physiological-data based research purposes. While the challenges outlined by [15] and [9] may apply, we suggest that there are additional challenges that apply directly to research-grade physiological wearable devices.

2.2 Measuring physiological responses

This paper describes a case study that was used to observe challenges with research-grade physiological wearable devices. The purpose of the case study was to measure physiological data during a cognitively intensive task. As such, this section introduces and discusses these data points, and their relationship with cognitive function.[3]

It is becoming more and more common for researchers to use wearable devices to measure physiological responses to cognitive tasks, such as those observed during cognitive workload and cognitive fatigue. This can include Electrocardiogram, Electrodermal Activity, Electromyography, Inertial Measurement Units, and Electroencephalograph. This is because physiological data points can be used to record changes in the Sympathetic Nervous System (SNS). The SNS is part of the autonomic nervous system that becomes more active in response to stress, danger, or exertion, triggering the body's "fight-or-flight" response. Its activity increases in situations like stress, danger, or exertion.

Electrocardiogram (ECG), for example, can be used to identify a stress response to mental fatigue. ECG measures the electrical activity of the heart, and can be used to determine heart rate and heart rate variability [1, 6, 17]. Under stress, the SNS is more active which can cause changes to heart rate and can affect heart rate variability. Our body is continuously trying to be at equilibrium and our bodies sympathetic and parasympathetic nervous systems are constantly changing. When our body is under stress, the SNS activity is elevated. With this elevation comes less variability in heart rate and a trend towards an increase in heart rate [26].

Electrodermal Activity (EDA, also known as Galvanic Skin Response or GSR) can also be used to measure a stress response to fatigue, through measurement of the conductivity of the skin [2, 12, 17]. The changes in the skin conductivity occurs due to changes in our sweat glands. When we are under stress our body goes into more of a sympathetic nervous system state. During this time our sweat glands become more active therefore our skin conductivity is increased [12]. Readings can vary depending on temperature, physical activity, and other external factors. However, baseline readings can be taken by asking the participant to rest and/or take a slow deep breath in and a slow breath out. This activates both the sympathetic and parasympathetic nervous systems in a balanced manner [23].

Electromyography (EMG) can be used to measure local muscle fatigue via the measurement of muscle activity [5]. This is because EMG tracks electrical activity within a muscle. This can be measured from any superficial muscle location, e.g. muscles that are touching the skin. However, EMG electrodes can be placed on certain muscles, e.g. in the forearm, to see if there are any functional changes with the presence of cognitive fatigue.

Inertial Measurement Units (IMUs) can be used to measure changes in movement to provide insights into whether participants show indications of more or less activity when they are fatigued. IMUs measure acceleration and rotation [20]. They have many uses outside and within human research, such as orientation sensors for robots and gait analysis with humans [18, 20]. IMUs can be placed anywhere on the body to measure changes in movement. Furthermore, IMU can be used, in conjunction with EMG, to determine whether any changes in muscle activity are due to increased movement, or fatigue.

Finally, Electroencephalograph (EEG) measures brain function, such as fatigue. This is done through tracking the electrical activity in the brain [1, 8, 24]. Different frequency EEG data can be used to measure different cognitive functions. For example, in the frontal cortex, gamma is associated with heightened perception/learning (32-100 Hz), beta is associated with being awake/thinking (13-32 Hz), alpha is associated with being mentally relaxed (8-13 Hz), theta is associated with reduced consciousness (4-8 Hz), and delta is associated with deep sleep (0.5-4Hz).

3 Case study

As touched on previously, we use a case study to observe challenges that apply when working with research-grade physiological wearable devices. The purpose of this study was to measure physiological data points before, during, and after a cognitively intensive task. This, in turn, would allow us to identify any patterns or trends in the data that could be used to identify cognitive fatigue, and identify any challenges that arise.

3.1 Participants

The study was conducted with 18 participants between the ages of 21 and 53. Seven of the participants identified as female and eleven identified as male. Participants were university students who participated in the study as part of an undergraduate course. Ethical consent was applied for and approved before the commencement of the study (HREC(HECS)2023#20). Participants performed the study in a controlled environment to ensure a non-disruptive experience. The study was run in the university's Human-Centered Computing Usability Lab, which is a dedicated space that includes a foyer area and private study room.

3.2 Equipment

For each participant, the physiological measures recorded in this study were ECG, EDA, EMG, IMU, and EEG, discussed in more detail in Section 2.2. Two wearable sensors were used throughout the study: (1) the Shimmer Consensys Bundle Development kit, and (2) the Muse EEG headband.

Shimmer Sensing is a provider of research-grade wearable products and solutions. The Shimmer Consensys Bundle Development kit is an all-in-one kit that contains all of the physiological sensors, hardware, and software provided by Shimmer [25]. Figure 1 shows

[3] Here, it should be noted that the focus of this paper is on the challenges associated with research-grade devices, and not the analysis of the data collected. Nevertheless, we felt it was important to outline the types of data that were being collected.

Figure 1: Shimmer Consensys Bundle Development kit

Figure 2: The Muse EEG Headband

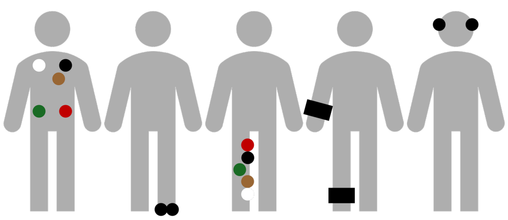

Figure 3: Sensor placement, left to right: ECG, EDA, EMG, IMU, EEG

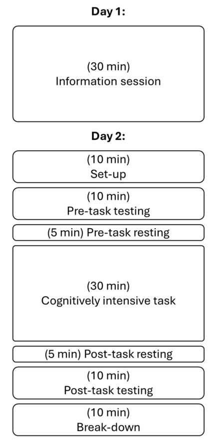

Figure 4: Study protocol

the full Shimmer Consensys Bundle Development kit, which includes an ECG sensor, an EMG sensor, an EDA (GSR) sensor, and an IMU sensor.

The Muse EEG headband is a light weight EEG sensor [13, 14]. It includes 4-channels (TP9, TP10, AF7, and AF8) which measure brain activity in the frontal and temporal lobes. Figure 2 shows the Muse EEG headset that was used in our case study.

Finally, our case study included a cognitively intensive task. This task was performed using the Multiple Attribute Task Battery system, OpenMATB. OpenMATB is a flight simulator with a monitoring task, tracking task, and resource management task [3]. OpenMATB was run on a Dell latitude laptop, using a keyboard and a joystick.

3.3 Sensor placement

The Shimmer ECG sensor includes five electrodes. One electrode was placed at a point between the heart and each limb (left leg, right leg, left arm, right arm) and one electrode was placed four-intercostal-spaces down on the left side of the sternum [22]. Once the electrodes were placed, they were connected to the Shimmer ECG sensor via colour-coded wires. The white wire connected the right arm electrode, black for the left arm electrode, green for the right leg electrode, red for the left leg electrode, and brown for the electrode at the fourth intercostal space. Figure 3 shows an estimate of the sensor placements. From left to right: ECG, EDA, EMG, IMU, EEG.

The Shimmer EDA sensor includes two electrodes, which must be placed close together. For this study, the electrodes were placed on the participants' inner foot. This allowed participants to have their hands free to complete the simulator task.

The Shimmer EMG sensor includes three electrodes. Two electrodes must be placed at the most substantial point of a single muscle 2 cm apart in parallel with muscle fibers. The third electrode (the reference electrode) is placed at a boney point of the body (such as the knee, ankle, or wrist) [21]. For this study, the electrodes were placed on the right leg's medial gastrocnemius (calf muscle) and right vastus medialis (inner quadricep muscle), with the knee cap as the reference point. Once the electrodes were placed, they were connected to the Shimmer EMG sensor via coloured wires. White and black wires are connected to one electrode for lower frequency readings of the muscle, brown and red wires are connected to the other electrode for higher frequency readings of the muscle, and green is connected to the reference electrode point [21].

In contrast to the other sensors, IMU does not use electrodes. Instead, it is a small sensor that is placed directly onto the participant's body via an elastic strap. For this study, we placed one onto the participants' right wrist and one above their right ankle. The right wrist was used to see if there were any changes in movement patterns while participants were undertaking the cognitive task (discussed in section 3.4), while above the right ankle was chosen to qualify if any changes to the EMG readings in the gastrocnemius and vastus medialis are caused by movement or fatigue.

Finally, the Muse headset was used to record EEG data. The Muse headset was placed so that the electrodes sit approximately one fingers width above the eyebrows, and with the reference electrodes tucked behind the ears.

3.4 Protocol

The protocol for the study included eight phases, as shown in Figure 4: (1) information session, (2) study set-up, (3) pre-task testing, (4) pre-task resting, (5) cognitively intensive task, (6) post-task resting, (7) post-task testing, and (8) study break-down.

First, participants were asked to come in at least 24 hours before the main study session for an information session. This session included an explanation of the study motivation and protocol, the request for participants to restrict their caffeine and alcohol consumption for at least 12 hours prior to the main session, and a period of time for the participants to ask questions.

Second, at the start of the main session, participants were fitted with the ECG, EDA, EMG, IMU, and EEG sensors. The skin was cleaned with an alcohol wipe prior to adhering the electrodes.

Third, following the placement of the sensors, the study moved into the "pre-task testing" phase. In this phase, participants were asked to complete questionnaires based on how they felt at that moment in terms of cognitive fatigue, sleepiness, motivation, and Choice Reaction Time.

Fourth, following the completion of the pre-task testing, participants were asked to rest for five minutes. During the resting phase, participants were seated in a comfortable position with their eyes closed. During this time, all sensors were running and collecting baseline resting data.

Fifth, the participants underwent the "cognitively intensive task" phase. This phase involved undergoing 30 continuous minutes of OpenMATB. During this time, all sensors were running and collecting workload data.

Sixth, following the completion of the cognitively intensive task phase, participants were asked to rest for another five minutes. During the resting phase, participants were again seated in a comfortable position with their eyes closed. During this time, all sensors were running and collecting additional resting data.

Seventh, following the completion of the post-task resting phase, the study moved into the "post-task testing" phase. In this phase, participants were again asked to complete questionnaires based on how they felt at that moment in terms of cognitive fatigue, sleepiness, motivation, and Choice Reaction Time.

Eighth, once the post-task testing was complete, all sensor recordings were stopped and the sensors were removed from the participants.

4 Results

This study resulted in the collection of data (ECG, EDA, EMG, IMU, and EEG) during (a) pre-task resting, (b) a cognitively intensive task, and (c) post-task resting. While the analysis of this data for fatigue classification is still underway, this section outlines the results of the data collection.

Table 1 shows the data points that were collected. As shown in the table, we were unable to collect the full set of data for any of the participants. ECG and IMU data was collected for all participants. However, EDA, EMG, and EEG data was only collected for a subset of participants.

EDA data collection was attempted for all 18 participants. However, the extraction of the data was only successful for ten participants (P3, P4, P5, P10, P11, P14, P15, P16, P17, and P18). The EDA sensor is part of the Shimmer Consensys Bundle Development kit. Each of the Shimmer sensors are Bluetooth enabled. This allows the data transfer from the sensors themselves, to the Consensys software, to be done wirelessly. However, during our case study, we encountered significant challenges enabling the Bluetooth data transfer for the EDA sensor. In some cases, the Bluetooth was enabled and the data was successfully transferred wirelessly. In other cases, the Bluetooth connection would fail, in which case we were required to manually extract the data from the EDA sensor. In addition to this, there were eight cases where we could neither transfer data via Bluetooth, nor manually extract the data from the EDA sensor. In these cases, no EDA data was collected for these participants (P1, P2, P6, P7, P8, P9, P12, P13). Here, it should be noted that the data transfer failure was unable to be determined until after each user participated in the study. It was not until the data was attempted to be extracted from the device manually at the end of a recording session that it was determined that the data was unavailable.

Table 1: Physiological data recorded during the study. Check marks and crosses represent data recorded and not recorded, respectively

Participants	ECG	EDA	EMG	IMU	EEG
P1	✓	✗	✓	✓	✗
P2	✓	✗	✓	✓	✗
P3	✓	✓	✗	✓	✗
P4	✓	✓	✗	✓	✗
P5	✓	✓	✗	✓	✗
P6	✓	✗	✓	✓	✗
P7	✓	✗	✓	✓	✗
P8	✓	✗	✓	✓	✗
P9	✓	✗	✓	✓	✗
P10	✓	✓	✗	✓	✗
P11	✓	✓	✗	✓	✗
P12	✓	✗	✓	✓	✗
P13	✓	✗	✗	✓	✗
P14	✓	✓	✗	✓	✓
P15	✓	✓	✗	✓	✓
P16	✓	✓	✗	✓	✓
P17	✓	✓	✗	✓	✓
P18	✓	✓	✗	✓	✓

EMG data collection was attempted for all 18 participants. However, the extraction of the data was only successful for seven participants (P1, P2, P6, P7, P8, P9, P12). Like EDA, the EMG sensor is part of the Shimmer Consensys Bundle Development kit. The Shimmer Consensys Bundle Development kit comes with a piece of supporting software called ConsensysPRO. The ConsensysPRO software works alongside the Shimmer sensors, and the Bluetooth data transfer, to synchronize data transfer and aggregate the data streams of all sensors (ECG, EDA, EMG, and IMU). However, like EDA, we encountered significant challenges with the data collection for the EMG sensor. Again, it should be noted that the data transfer failure was unable to be determined until after each user participated in the study. It was not until the data was attempted to be extracted from the device manually at the end of a recording session that it was determined that the data was unavailable. As a result of this, we were unable to collect data from both the EMG sensor and the EDA sensor simultaneously. As can be seen in Table 1, there were no cases where both EMG and EDA data was recorded. Instead, each participant includes either EMG data or EDA data (with the exception of P13, where neither data type was recorded). This is discussed further in Section 5.2

EEG data collection was attempted for all 18 participants. However, the extraction of the data was only successful for five participants (P14, P15, P16, P17, and P18). As discussed in Section 3.2, EEG data was recorded using the Muse EEG Headband. In the case of EEG, all 18 participants were equipped with the Muse EEG Headband, and EEG recording was achieved successfully for all participants throughout the study. However, part way through the study it was determined that the official Muse mobile application does not allow users to export raw EEG data. As such, a third party application was used for EEG recording for the latter portion of the study. As a result, the first 11 participants' EEG data was unable to be extracted. Here, it should be noted that a small pilot study was run previously [16] in order to test the Muse EEG Headband and the data collection process. The third party mobile application was used during the pilot study, and data collection and extraction was undertaken successfully. The return to the official Muse application, and therefore the failure to export participant data, was a result of researcher miscommunication / user error. While the failure to export the data was due to user error on our behalf, we still feel it is important to mention here, as it supports one of the challenges highlighted for other wearable devices – that of the accessibility of raw data when working with low-cost commercial devices. This is discussed further in Section 5.2.

As can be seen in Table 1, the challenges encountered when recording ECG, EDA, EMG, IMU, and EEG data resulted in none of the participants' datasets holding a full set of sensor readings. Five participants' datasets contained four sensor types (ECG, IMU, EEG, and either EDA or EMG); twelve participants' datasets contained three sensor types (ECG, IMU, and either EDA or EMG); and one participant's datasets contained two sensor types (ECG and IMU). These results, and their link to existing challenges with wearable devices, is discussed further in Section 5.

5 Discussion

This paper investigates the challenges of working with research-grade wearable devices through a case study on measuring physiological data for cognitive fatigue research. We have outlined our case study, which uses two research-grade devices (the Shimmer Consensys Bundle Development kit and the Muse EEG Headband) to record ECG, EDA, EMG, IMU, and EEG data.

5.1 A comparison with generalized challenges

As discussed in Section 2.1, it is widely acknowledged that the use of wearable devices presents challenges for researchers. [15] identified a set of nine challenges that they generalize to all wearable devices (outlined in full in Section 2.1). Of these challenges, there are three that we encountered when working with research-grade wearable devices.

(1) **Data transmission**. [15] discusses the challenges associated with data transmission. They focus primarily on data latency and network load when working with devices that rely on a centralized cloud for data management, processing, and storage. However, we suggest that this can be extended to other data transmission protocols, such as Bluetooth. As discussed in Section 4, we encountered significant challenges with data transmission via Bluetooth when working with the Shimmer Consensys Bundle Development kit. During our case study, 45% of our EDA data (8 out of 18 participants' data) was lost, largely due to Bluetooth transmission failure. The reliability of data transmission is a crucial element in successful wearable device use. This is applicable to both research-grade wearables, and wearables in general.
(2) **Interoperability**. [15] discussed the diverse nature of wearable devices. They highlighted interoperability challenges associated with communication between different devices, and the integration of wearable components, services, and applications. For our case study, we invested in the Shimmer Consensys Bundle Development kit, in the hopes of avoiding interoperability challenges and the use of heterogeneous devices. The Shimmer kit includes sensors for ECG, EDA, EMG, and IMU, and companion software that integrates with each of the sensors. Nevertheless, one of the main challenges we encountered involved two of the sensors failing to work simultaneously. As discussed in Section 4, there were no cases where both EMG and EDA data were recorded. Instead, each participant includes either EMG data or EDA data (with the exception of P13, where neither data type was recorded).
(3) **Communication and architecture**. [15] discusses the challenge of developing protocols and architectures that facilitate communication between heterogeneous devices. At this point in our research, we do not know the cause of the interoperability issue with EDA and EMG sensors. However, we hypothesize that it may fall under the Communication and Architecture category, with the architecture and communication protocols built into the Shimmer Consensys Bundle Development kit being unable to transmit both sensor data types simultaneously.[4]

5.2 A comparison with commercial-grade challenges

In addition to the challenges identified by [15], [9] identified a further three challenges specific to low-cost, commercially available wearable devices (outlined in full in Section 2.1). Of these challenges, there is one that we encountered when working with research-grade wearable devices.

(1) **Data accessibility:** [9] highlight that commercial wearable devices generally do not offer access to raw data. This tends to be because commercial devices are not designed for applications that require raw data access and thus do not provide it. We experienced this challenge in our case study. As discussed in Section 4, 72% of our EEG data (13 out of 18 participants' data) was lost due to the official Muse mobile

[4]This is an assumption based on our experience. We cannot find documentation detailing the same issue by other researchers.

application not allowing access to raw data. This is not typical for research-grade devices. Research-grade devices are designed and developed for research purposes, and therefore tend to allow access to raw data, either through free or paid services. Here, it should be noted that the Muse EEG headband is marketed as a "research-grade EEG headband" that uses "clinical-grade technology" [14]. However, it is also marketed as a "meditation and sleep headband". It is significantly more light weight, and lower cost than other EEG-based research-grade wearables, such as the Emotiv headset range[5] or BioSemi,[6] both of which provide access to raw data. As such, while the Muse EEG Headband is marketed as research-grade, we suggest that it aligns more closely with lower-cost commercially available devices.

5.3 Considering research-grade challenges

In addition to the challenges identified by [15] and [9], we propose the addition of one challenge that is specific to research-grade wearable devices.

(1) **Setup complexity (researcher adoption):** As illustrated in Section 3.3, the setup of research-grade physiological wearable devices can be complex, requiring both precise sensor placement and a wired connection. For the Shimmer Consensys Bundle Development kit, each sensor includes clinical-grade electrodes and corresponding colour-coded wires that must be placed and connected at precise locations on the body. The ECG electrodes, for example, must be placed at precise points between the heart and each limb (left leg, right leg, left arm, right arm) and four-intercostal-spaces down on the left side of the sternum (see Section 3.3 for full descriptions of sensor placements). The placement of these sensors, and correct connection of the colour-coded wires require significant knowledge of the wearable devices, human anatomy, and the musculoskeletal system. In contrast, the Muse EEG Headband and the low-cost commercially available devices used by [9] involve very little setup. We suggest that the specialized knowledge required for research-grade wearable device setup makes the adoption of research-grade devices more difficult. Our research group includes researchers with knowledge in computing, engineering, health, and sport science. As such, we were able to leverage individual expertise to ensure the correct placement of electrodes, correct setup of the wearable devices, and correct setup of the software and data extraction. Nevertheless, we suggest that the complexity of setup may discourage researchers from adopting research-grade wearable devices. [15] discusses **user adoption** – i.e. the psychological and emotional impacts of wearable device adoption on users. We propose that **researcher adoption** is equally important.

5.4 Generalizability

This paper outlines a case study that allowed us to identify a challenge unique to research grade devices—setup complexity (researcher adoption)—discussed in Section 5.3. While we recognize that our case study only investigated two devices, we argue that the challenge of setup complexity is generalizable to other research-grade devices, particularly those that require precise setup and/or the precise placement of electrodes.

Research-grade EEG devices are one such example of this. The Emotive Flex 2 head cap,[7] for example, is a 32 channel saline EEG head cap. Significant setup is required in order to ensure that data is being recorded correctly for this type of device. Researchers must ensure that the electrodes have been properly lubricated (with saline), that the cap and therefore the electrodes, are sitting in the correct position on the head, and that the electrodes are getting the least amount of impedance from the participant's hair.

EEG is one such example of a research-grade device with complex setup. Another example includes EMG sensors such as the Plux Electromyography (EMG) Sensor[8], which requires precise electrode placement and skin preparation (shaving to avoid impedance). This type of setup (EEG and EMG) takes knowledge and training. Researchers in health, neuroscience, and medical fields may have the expertise, but it can be imagined that researchers in other fields may be discouraged from using such devices based on their complex setup requirements. They may instead, chose to use more commercial grade versions of the same devices. This tends to allow for easier setup, but can limit the accuracy of the readings, and comes with it's own set of challenges (as discussed in Section 2.1).

To our knowledge, little research has been conducted to investigate the adoption rates of research-grade devices based on their complexity and setup (discussed more in Section 5.5). Instead, researchers tend to focus on commercial devices, or devices that are simple enough for users to setup on their own.

[19] for example, examined over 100 commercial wearable devices, with a focus on security, energy consumption, and computing power. They categorized these wearables as accessories (smart watches, eye-wear, jewelry, etc.), e-textiles (shirts, pants, socks, gloves, etc.), and e-Patches (tattoos, patches). While they discuss research-based prototypes of accessories, e-textiles, and e-patches, they do not touch on research-grade wearables that involve complex setup or the precise placement of electrodes (e.g. EEG caps). In fact, they state that they "review only those devices that can be easily worn by the user ... " [19, p. 2574].

Similarly, [4] focused on studies where participants were given commercial devices and asked to use them for a period of time. Within their study, they highlighted user error as one of the challenges associated with wearable devices. This includes (1) the device not being synced correctly, (2) poor calibration of the device, (3) quality of skin contact, and (4) misplacement of the device on the body. While their focus was on commercial devices that are setup by users, we argue that the same errors could be perpetuated by researchers, particularly when working with research-grade devices that require complex or precise setup.

5.5 Limitations and Future Work

Within this work, we identify two limitations: reproducibility, and generalization.

[5] https://www.emotiv.com/
[6] https://www.biosemi.com
[7] https://www.emotiv.com/products/flex-saline
[8] https://www.pluxbiosignals.com/products/electromyography-emg

First, in Section 5.1 we discussed the challenges of Data transmission (where the Bluetooth connectivity failed for 45% of our EDA data) and Interoperability (where there were no cases where both EMG and EDA data were recorded). We hypothesized that these challenges stem from the architecture and communication protocols built into the Shimmer Consensys Bundle Development kit. However, we were unable to test or validate this hypothesis. Here, it should be noted that while the data was collected for the purpose of one study, it was collected over multiple days, with researchers attempting to trouble shoot this problem on several occasions. Future work could include repeating this study to determine whether the same challenges arise again, and/or whether the issues of data transmission and interoperability are reproducible.

Second, this paper outlined a case study that investigates two research-grade devices: the Shimmer Consensys Bundle Development Kit, and the Muse EEG Headband. While the inclusion of these two systems allowed us to draw comparisons with generalized and commercial-grade challenges, and identify a challenge unique to research grade devices, one limitation of this work is that it does not investigate any additional research-grade devices. Future work could include evaluating a larger subset of devices, including commonly used research-grade wearables such as the Emotive Flex 2 head cap and the Plux Electromyography (EMG) Sensor. In addition to this, it would be worthwhile to investigate the adoption rates of research-grade devices based on their complexity and setup. This could include a survey of researchers undertaking bio-signal and physiological data-based research across a variety of fields, universities, and countries.

6 Conclusion

This paper investigated the challenges associated with research-grade wearable devices. In this paper, we introduced a case study on measuring physiological data for cognitive fatigue research. The case study was used to evaluate two research-grade wearable devices: the Shimmer Consensys Bundle Development kit, and the Muse EEG Headband. The data that was recorded included ECG, EDA, EMG, IMU, and EEG data. Within the case study, 100% of ECG and IMU data was recorded successfully. However, only 55% of EDA, 39% of EMG, and 28% of EEG data was recorded successfully. These results were used to identify links between our experiences and existing challenges with wearable devices (data transmission, interoperability, communication and architecture, and data accessibility), and to identify one new challenge that is associated with research-grade wearable devices (setup complexity (researcher adoption)). For Setup Complexity (researcher adoption), we suggest that the specialized knowledge required to setup research-grade wearable device makes the use of these devices more difficult and may discourage researchers from adopting them. The findings from this work can be applied to further studies using research-grade wearable devices.

References

[1] Gianluca Borghini, Laura Astolfi, Giovanni Vecchiato, Donatella Mattia, and Fabio Babiloni. 2014. Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. *Neuroscience & Biobehavioral Reviews* 44 (2014), 58–75.

[2] Gianluca Borghini, Laura Astolfi, Giovanni Vecchiato, Donatella Mattia, and Fabio Babiloni. 2014. Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. *Neuroscience & Biobehavioral Reviews* 44 (2014), 58–75.

[3] J Cegarra, B Valéry, Eugénie Avril, C Calmettes, and J Navarro. 2020. OpenMATB: A Multi-Attribute Task Battery promoting task customization, software extensibility and experiment replicability. *Behavior research methods* 52 (2020), 1980–1990.

[4] Sylvia Cho, Ipek Ensari, Chunhua Weng, Michael G Kahn, and Karthik Natarajan. 2021. Factors affecting the quality of person-generated wearable device data and associated challenges: rapid systematic review. *JMIR mHealth and uHealth* 9, 3 (2021), e20738.

[5] Mario Cifrek, Vladimir Medved, Stanko Tonković, and Saša Ostojić. 2009. Surface EMG based muscle fatigue evaluation in biomechanics. *Clinical biomechanics* 24, 4 (2009), 327–340.

[6] Jesús Díaz-García, Vicente Javier Clemente-Suárez, Juan Pedro Fuentes-García, and Santos Villafaina. 2023. Combining HIIT Plus Cognitive Task Increased Mental Fatigue but Not Physical Workload in Tennis Players. *Applied Sciences* 13, 12 (2023), 7046.

[7] Alexander E Hramov, Vladimir A Maksimenko, and Alexander N Pisarchik. 2021. Physical principles of brain–computer interfaces and their applications for rehabilitation, robotics and control of human brain states. *Physics Reports* 918 (2021), 1–133.

[8] Ivo Käthner, Selina C Wriessnegger, Gernot R Müller-Putz, Andrea Kübler, and Sebastian Halder. 2014. Effects of mental workload and fatigue on the P300, alpha and theta band power during operation of an ERP (P300) brain–computer interface. *Biological psychology* 102 (2014), 118–129.

[9] Jemma L König, Jascha Penaredondo, Emily McCullagh, Judy Bowen, and Annika Hinze. 2023. Let's Make it Accessible: The Challenges Of Working With Low-cost Commercially Available Wearable Devices. In *Proceedings of the 35th Australian Computer-Human Interaction Conference*. Computer-Human Interaction Special Interest Group (CHISIG), 493–503.

[10] Alexandros Liapis, Christos Katsanos, Dimitris Sotiropoulos, Michalis Xenos, and Nikos Karousos. 2015. Stress recognition in human-computer interaction using physiological and self-reported data: a study of gender differences. In *Proceedings of the 19th Panhellenic Conference on Informatics*. ACM, 323–328.

[11] Mohammadhossein Moghim, Robert Stone, Pia Rotshtein, and Neil Cooke. 2015. Adaptive virtual environments: A physiological feedback HCI system concept. In *2015 7th Computer Science and Electronic Engineering Conference (CEEC)*. IEEE, 123–128.

[12] JD Montagu and Edward Michael Coles. 1966. Mechanism and measurement of the galvanic skin response. *Psychological Bulletin* 65, 5 (1966), 261.

[13] Muse. 2023. Muse™ 2 Premium Subscription Bundle. https://choosemuse.com/products/muse-2-premium-subscription-bundle

[14] Muse. 2023. Muse™ EEG-Powered Meditation & Sleep headband. https://choosemuse.com/

[15] Aleksandr Ometov, Viktoriia Shubina, Lucie Klus, Justyna Skibińska, Salwa Saafi, Pavel Pascacio, Laura Flueratoru, Darwin Quezada Gaibor, Nadezhda Chukhno, Olga Chukhno, et al. 2021. A survey on wearable technology: History, state-of-the-art and current challenges. *Computer Networks* 193 (2021), 108074.

[16] Mahonri Owen, Jemma König, Mitchell Head, Merel Hoskens, Katie Sullivan, Jadon Miller, and Te Taka Keegan. 2024. Diversity, Equity, and Inclusion in I&M: Beyond the Technology: Wearables and the Cultural Compass. *IEEE Instrumentation & Measurement Magazine* 27, 9 (2024), 38–43.

[17] Kate O'Keeffe, Simon Hodder, and Alex Lloyd. 2020. A comparison of methods used for inducing mental fatigue in performance research: Individualised, dual-task and short duration cognitive tests are most effective. *Ergonomics* 63, 1 (2020), 1–12.

[18] Gerasimos G Samatas and Theodore P Pachidis. 2022. Inertial measurement units (imus) in mobile robots over the last five years: A review. *Designs* 6, 1 (2022), 17.

[19] Suranga Seneviratne, Yining Hu, Tham Nguyen, Guohao Lan, Sara Khalifa, Kanchana Thilakarathna, Mahbub Hassan, and Aruna Seneviratne. 2017. A survey of wearable devices and challenges. *IEEE Communications Surveys & Tutorials* 19, 4 (2017), 2573–2620.

[20] Shimmer Sensing. 2016. Shimmer3 IMU User Guide. https://shimmersensing.com/wp-content/docs/support/documentation/IMU_User_Guide_rev1.4.pdf

[21] Shimmer Sensing. 2017. Shimmer3 EMG User Guide. https://shimmersensing.com/wp-content/docs/support/documentation/EMG_User_Guide_Rev1.12.pdf

[22] Shimmer Sensing. 2018. Shimmer3 ECG User Guide. https://shimmersensing.com/wp-content/docs/support/documentation/ECG_User_Guide_Rev1.12.pdf

[23] Shimmer Sensing. 2018. Shimmer3 GSR+ User Guide. https://shimmersensing.com/wp-content/docs/support/documentation/GSR_User_Guide_rev1.13.pdf

[24] Mitchell R Smith, Rifai Chai, Hung T Nguyen, Samuele M Marcora, and Aaron J Coutts. 2019. Comparing the effects of three cognitive tasks on indicators of mental fatigue. *The Journal of psychology* 153, 8 (2019), 759–783.

[25] Shimmer Wearable Sensor Technology. 2024. Shimmer Consensys Bundle Development kit. https://shimmersensing.com/product/consensys-bundle-development-kit/

[26] Gaetano Valenza, Luca Citi, Vegard Bruun Wyller, and Riccardo Barbieri. 2018. ECG-derived sympathetic and parasympathetic activity in the healthy: an early

lower-body negative pressure study using adaptive Kalman prediction. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 5628–5631.