FaceBit: Smart Face Masks Platform

ALEXANDER CURTISS*, BLAINE ROTHROCK*, and ABU BAKAR, Northwestern University
NIVEDITA ARORA, Georgia Institute of Technology
JASON HUANG, ZACHARY ENGLHARDT, Northwestern University
AARON-PATRICK EMPEDRADO, CHIXIANG WANG, Northwestern University
SAAD AHMED, Northwestern University
YANG ZHANG, University of California, Los Angeles
NABIL ALSHURAFA and JOSIAH HESTER, Northwestern University

The COVID-19 pandemic has dramatically increased the use of face masks across the world. Aside from physical distancing, they are among the most effective protection for healthcare workers and the general population. Face masks are passive devices, however, and cannot alert the user in case of improper fit or mask degradation. Additionally, face masks are optimally positioned to give unique insight into some personal health metrics. Recognizing this limitation and opportunity, we present FaceBit: an open-source research platform for smart face mask applications. FaceBit's design was informed by needfinding studies with a cohort of health professionals. Small and easily secured into any face mask, FaceBit is accompanied by a mobile application that provides a user interface and facilitates research. It monitors heart rate without skin contact via ballistocardiography, respiration rate via temperature changes, and mask-fit and wear time from pressure signals, all on-device with an energy-efficient runtime system. FaceBit can harvest energy from breathing, motion, or sunlight to supplement its tiny primary cell battery that alone delivers a battery lifetime of 11 days or more. FaceBit empowers the mobile computing community to jumpstart research in smart face mask sensing and inference, and provides a sustainable, convenient form factor for health management, applicable to COVID-19 frontline workers and beyond.

ACM Reference Format:
Alexander Curtiss, Blaine Rothrock, Abu Bakar, Nivedita Arora, Jason Huang, Zachary Englhardt, Aaron-Patrick Empedrado, Chixiang Wang, Saad Ahmed, Yang Zhang, Nabil Alshurafa, and Josiah Hester. 2021. FaceBit: Smart Face Masks Platform. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 4, Article 151 (December 2021), 44 pages. https://doi.org/10.1145/3494991

1 INTRODUCTION

Due to the ongoing COVID-19 pandemic, face masks have become ubiquitous worldwide. With the incredibly infectious B.1.617.2 (Delta) variant predominant, and the world on pace to pass 300 million confirmed cases by early 2022 [16], the U.S. Centers for Disease Control and Prevention (CDC) and the World Health Organization (WHO) recommend nearly universal masking [23] (August 2021), with many governments enforcing mask mandates. Because of the slow rate of global vaccination, a panel of 175 health experts warned of vaccine resistant

*Both authors (A. Curtiss and B. Rothrock) contributed equally to the work.

Authors' addresses: Alexander Curtiss; Blaine Rothrock; Abu Bakar, Northwestern University; Nivedita Arora, Georgia Institute of Technology; Jason Huang, Zachary Englhardt, Northwestern University; Aaron-Patrick Empedrado, Chixiang Wang, Northwestern University; Saad Ahmed, Northwestern University; Yang Zhang, University of California, Los Angeles; Nabil Alshurafa; Josiah Hester, Northwestern University, josiah@northwestern.edu.

2474-9567/2021/12-ART151 $15.00
https://doi.org/10.1145/3494991

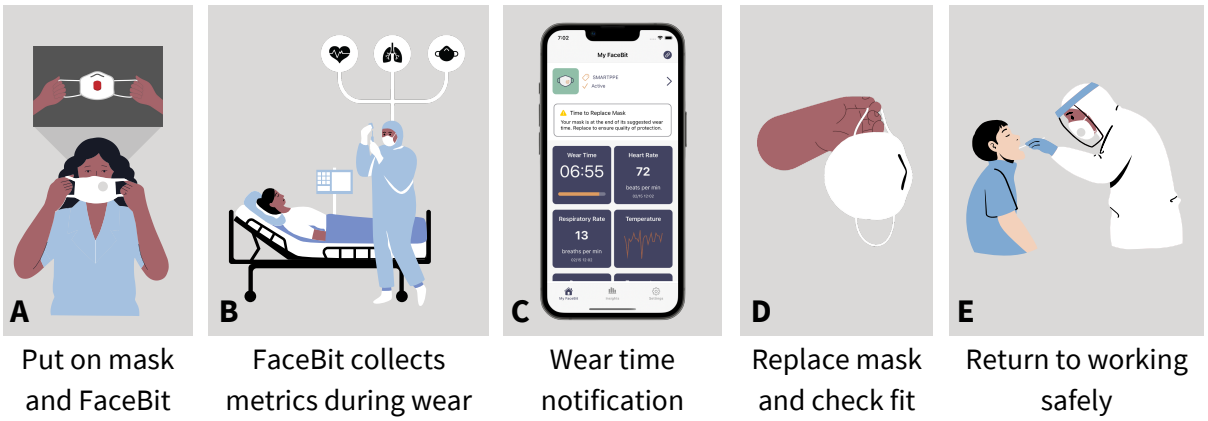


Fig. 1. FaceBit is an energy-efficient research-focused hardware platform that allows face masks to be retrofitted (A) to gather important health and safety metrics on-device like heart rate, respiratory rate, and wear time (B). Notifications are sent to a phone application upon detection of state changes (C). Easy retrofitting when a mask is replaced (D) allows for lower burden usage in difficult work environments (E).

variants possible, potentially lengthening the pandemic [56]. While the pandemic has evolved and public health messaging with it— consistently, beyond a vaccination, masking is one of the best forms of protection.

However, in this moment, the longstanding limits of traditional personal protective equipment (PPE) such as face masks have become apparent. First, face masks must be managed; many types, like N95s, must be professionally fitted, maintained, and eventually replaced or decontaminated [60]. This requires healthcare workers to take time away from caring for their patients [73]. Moreover, traditional PPE is passive protection; problems like an ill-fitting mask that leaves a person exposed will not be caught for hours or days; increasing the risk of infection [26]. Finally, today's face masks miss opportunities for sensing and intelligence that could inform and actively protect wearers.

Instrumentized face masks are exceptionally well placed to provide active protection for frontline health workers and beyond because they are in (necessarily) close proximity to the respiratory system. This system is the target of many infectious diseases (including COVID-19), and also tightly coupled to general health.

Active PPE in the form of "smart" face masks could provide an early warning system when health or environmental metrics indicate danger; such as heart rate, respiratory rate (which is currently difficult to monitor in a wearable setting), or toxic air pollution. A wearer could be informed immediately if a respirator mask was not fitted correctly or if a leak formed during activity.[1] If a mask had been worn too long, the active PPE system could notify the wearer. Finally, if a user could rely on a mask they already must wear to function as a wearable health tracker, this would provide a usability and adherence advantage.

These use cases represent only the start of research questions in the space. With the rapidly changing pandemic world, it is imperative that the mobile computing and health communities have the tools to flexibly perform research that relates to current challenges. Ubiquitous computing researchers need a new platform.

Realizing this vision of smart face masks is challenging. "Smart" PPE that actively protects workers has been proposed, especially for heavyweight applications in high risk environments (for example fire-fighting, construction, mining [53], and electrical work [62]). These systems are high in cost, size, weight, and incur maintenance requirements. Such bulky, technologically complex, and power intensive face masks do not scale to

[1]Workers cannot always detect if a respirator mask is fitting loosely because of numbness of the face, fatigue, and mask wear and tear.

the needs of a global pandemic. Masks must be available and ready to use anytime, they must be either cheap and disposable, or sturdy to repeated use, decontamination, and wear. Masks are constantly replaced in clinical settings, necessitating a mechanism to transfer intelligent sensing capability from one mask to the other, easily and quickly. Masks are small and lightweight, meaning that any extra weight is noticeable and burdensome. Powering the "smarts" (i.e. sensing and inference) would normally require a large battery, yet, a battery sufficiently large enough to last a week or more adds bulk, weight, discomfort, and increases user burden by requiring the user to maintain a state of charge.

1.1 FaceBit: Smart Face Masks Platform

We navigate the previous concerns and tradeoffs and develop **FaceBit**, a low cost, low profile, energy-efficient, energy harvesting, rich sensing platform. FaceBit is also a research platform, and we demonstrate its efficacy and general potential by gathering information about its accuracy and usability. FaceBit is designed to be used in existing masks without modification and is easily temporarily clipped onto N95, and surgical masks with a small magnet.

FaceBit is a modular hardware platform capable of supplementing its small primary cell battery with energy harvested from movement, sunlight, and even breathing. This combination extends the lifetime and reduces battery replacement frequency, but does not sacrifice size and burden. FaceBit collects and analyzes data from a suite of on-device sensors, and wirelessly sends heart rate, respiration rate, mask on/off status and other metrics to a custom phone application which allows for visualization, data analysis, and can also serve as an intervention system. An overview of the platform and concept is shown in Figure 1.

Contributions. We innovate across hardware and software to make smart face masks possible.

(1) After a series of design research studies on energy harvesting capability and need-finding with healthcare workers, we design a hardware and software platform consisting of the FaceBit circuit board and a custom phone application that supports data collection, visualization, and alerts.
(2) We study sensing targets obtainable from a mask in a series of lab experiments and free-living scenarios. We design a low-memory footprint, on-device signal processing pipeline to capture heart rate from subtle head movements, respiration rate from temperature changes in-mask, and explore leak detection via changes in in-mask pressure.
(3) To ensure long battery lifetime and low user-burden, we design an energy-efficient runtime and triggering architecture for *on-device* capture of heart rate, respiratory rate, and mask wear time.
(4) An open-source, open-hardware release of the FaceBit platform and its tools[2]. Marking the first open and end-to-end research platform for smart face masks available to the ubiquitous and mobile computing research community.

2 RELATED WORK

FaceBit sits at the intersection of energy-efficient low power embedded sensing systems and mobile health. It also builds on literature in energy harvesting embedded computing and research focused hardware platforms.

2.1 Energy-efficient Sensing and Computation

Significant prior work has explored how to make mobile computation and sensing less energy intensive, lower user burden, and lower maintenance. FaceBit builds on these works by developing an unobtrusive, low-power runtime system for the collection of health and safety metrics.

[2]The open-source, open-hardware release of the FaceBit platform and its tools can be found at https://facebit.health

Low Power Wearables and IoT. Resource constrained computing has been explored across wireless sensor networks and embedded computing systems. Mate [45] was an early virtual machine to enable lightweight programs with low communication overhead to process data on device. Recently, CapeVM [58] pushed this virtual machine abstraction farther, to enable multiple IoT applications on a single device. TinyOS [46], t-kernel [27], and SOS [28] were all early runtime systems concentrated on low power operation for sensing devices. Mercury [48] was one of the first health based wearables that focused on low power operation, processing features on device before sending them to a basestation. Amulet [32] a smart watch with months long battery life, continued this trend of on-device processing, but focused on enhancing developer tools to allow for low power operation. Unlike Amulet, eBP [10] is a specific purpose device for blood pressure readings, which has a sophisticated on-board signal processing suite, yet suffers from short battery life because of this computational load. FaceBit balances across these approaches by focusing on low power signal processing for health metrics and implementing a runtime system that is health metric aware to save power.

Energy Harvesting. Many devices have adopted energy harvesting techniques to extend lifetimes and reduce maintenance without increasing bulk by adding large batteries. Early work, such as Trio [20], and Prometheus [36], deployed devices with solar panels for long term infrastructure and environmental monitoring. Since then, this area of energy harvesting wireless sensor networks has exploded, as discussed in a recent survey [50]. Of these devices, Permamote [35], is most similar to FaceBit in that it also makes use of a hybrid energy solution, with a primary cell as backup for consistent readings, and energy harvester to extend lifetime. FaceBit extends beyond Permamote in leveraging more predictable energy modalities based on the mask being worn (i.e. breathing, motion). Other work has explored novel energy harvesting scenarios to power interactive devices, like SATURN [7] and Paper Generators [38] use triboelectric materials, and the Battery-free Game Boy [18] and Peppermill [74] which both harvested energy from user actions. Other work like the battery-free cell phone relied on energy harvested from solar panels to make Skype calls [70]. These demonstrations inform FaceBit. It is possible to use FaceBit with a variety of energy harvesters, which we evaluate in the Section 3.3. Like Peppermill and Gameboy, FaceBit can rely on the user to create energy (from breathing and movement).

2.2 Smart Masks and PPE, Wearable Physiological Sensing

There has been much prior work on health-related sensing. Sensing outside conventional medical settings to facilitate low-cost continuous health monitoring has long been investigated. For example, people have deployed sensors in environments to sense personal health [42]. It is also possible to repurpose mobile devices for health sensing [21, 75–77]. Here we review previous wearable systems, which are closer to this work in terms of its wearable form factor.

Wearable bio-sensing. Wearable devices have the advantage of being close to a user's body. This proximity affords better signal quality, sensing coverage, and often improves comfort versus other devices. Recently, we have seen the commercial success of smartwatches and wristbands that feature bio-sensing, such as the PPG-enabled Apple Watch and Fitbit. There have also been brainwave sensors featured on headsets (e.g., Next Mind). In the research domain, clothing has been a promising sensing medium, onto which people have fabricated sensors to monitor blood pressure [11], as well as other physiological signals [39, 54]. The most common sensors on wearables include temperature sensors (e.g., fever detection[24]); cameras (e.g., activity recognition [4, 51, 81]); accelerometers (e.g., building activity profiles [9, 19], and detecting fatigue [47]); and microphones (e.g., cough detection [37] and activity recognition [43, 79]). It is also possible to take a multi-sensor approach, which has been demonstrated in eating detection [80] and blood pressure estimation [12]. For a comprehensive review of literature, we recommend [55, 57, 65]. The sensing front of FaceBit was inspired by these prior systems. However, we confined our sensing to mask form factors, which makes our work more suited for masks and PPE integration.

Face-centric sensing and smart masks. Closer to our platform is prior work that extracts physiological signals from users' faces. There have been commercially available PPG sensors (e.g., Adafruit Pulse Sensor Amped), which makers and hobbyists have used to detect heartbeats off of wearers' ears. Researchers have developed smart glasses with embedded sensors on the nose pads and temples to detect blood pressure [34] and metabolism [64]. It is also possible to instrument sensors into users' mouths for salivary uric acid sensing [41] and other analytes [71]. Finally, we have searched for previous work on smart masks and summarized exemplar ones. Researchers have instrumented masks with pressure sensors to estimate wearing comfort [15]. In response to the COVID-19 outbreak, there has been an on-going effort in swift detection of COVID using on-mask stickers [2]. Lut et al. [49] explored the design space for a smart filter for PPE that monitors its filter quality (e.g., contamination, clog), however, this concept was never built. Hernandez et al. [31] explored heart rate signals at rest from the face captured from smart glasses, however they calculate heart rate using an offline pipeline, not on-device, and used complex signal processing that would be higher power. Hernandez, et al. [31] has more recently obtained relatively high quality results from a head-based wearable by sensing motion using a custom "Motion Level" value derived from an accelerometer. Again, however, this processing was done off-device and utilized both linear and rotational acceleration sensors, leading to higher power usage.

Unlike previous systems, FaceBit was built with practical power constraints considered on day one of its system design. To address these constrains, FaceBit features a low power hardware design with energy optionally partially contributed by various harvesters. Moreover, we adopted a user-centric approach, with our system inspired, developed, and evaluated with medical professionals in the loop. All of these efforts lead to a promising and practical path to future smart masks and PPE.

2.3 Hardware Platforms Enabling the Research Community

FaceBit is inspired by longstanding traditions of the ubiquitous and mobile computing research community to engineer hardware platforms for pushing forward research on important and emerging topics. Wearable smart watch platforms such as Mercury [48] and Amulet [32] were designed by the research community to support high fidelity motion sensing and long term health studies research, respectively. Or to explore specific health problems like eating (NeckSense [80], FitByte [8]) and stress (AutoSense [22]). Wireless smart device research is supported by recent platforms like Hamilton for indoor sensing applications [6], and SignPost [3] for outdoor multi-application smart cities deployments. Devices like SenseCam [33] are used and reprogrammed by researchers to capture different data in free living scenarios. These platforms fill a gap where expert researchers may not have hardware capability, but can (re)use and (re)program the platform to design, test, and deploy novel approaches to signal processing and inference for unique applications and alternative aims.

2.4 Gap in State of the Art

None of the previous works have investigated the intersection of energy harvesting, smart face masks, and on-device signal processing. Indeed, most methods demonstrate health metric capture by either relying on a phone and large battery, or processing data offline. However, these previous techniques in health based sensing are not as feasible for the realities of disposable, interchangeable face masks, which require a special form-factor and particularly high ease of use. No previous work has developed a smart face mask hardware platform, tuned to the unique constraints of on-mask sensing, or released the platform for use by the research community. As the COVID-19 pandemic continues, this is an unmet need.

3 FACEBIT: SYSTEM DESIGN AND IMPLEMENTATION

In the following sections we detail the FaceBit system including the Design Goals (Section 3.1), the need-finding survey of clinicians, design exploration of energy harvesting, the hardware platform and energy harvesting

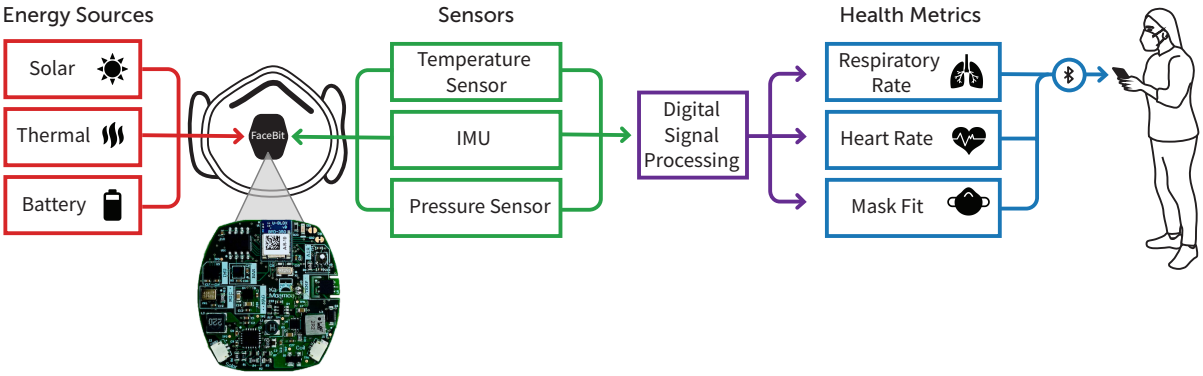


Fig. 2. FaceBit overview. Data from multiple sensors is gathered and then processed on-device for health metrics. This is reported to the phone app. The circuitry is designed to automatically switch between a primary cell and harvested energy.

circuitry (Section 3.4), the runtime system that captures health metrics on-device including heart rate, respiration rate, wear time, and mask fit (Section 4), the phone application (Section 5), and research tools that facilitate community use cases of FaceBit (Section 6).

3.1 Design Goals

We extract design goals for FaceBit, including our exploration of energy harvesting modalities and clinical need-finding (discussed in later sections), our balance across what is feasible to sense in terms of energy and availability, and general requirements for smart face masks gleaned from the literature, CDC guidelines, and Occupational Safety and Health Administration (OSHA) guidelines.

Goal 1: Reliable and Valid Health Sensing. FaceBit must provide reliable monitoring of the wearer's important health signals (e.g., respiration- and heart-rate), and should provide capability to expand to new health signals with expansion sensors and software.

Goal 2: Enables Real-time Interventions. While accurate sensing is critical, having access to wireless data in real-time to make informed decisions or deliver a just-in-time adaptive intervention (JITAI) [68] is a crucial feature for clinicians, workers, patients, and the general population. Therefore, the FaceBit platform must be able to communicate with any mobile device, e.g., the wearer's smartphone. An informed decision must be made to ensure that FaceBit is compatible with all types of mobile devices and provides a reliant and flexible interface.

Goal 3: Long Battery Lifetime, Low Bulk/Weight. Charging and battery replacement are burdensome to the wearer. Long battery life enables near continuous monitoring of wearer's health markers and reduces abandonment. When a device can run for multiple days or weeks, new applications are enabled and interesting data and trends can be captured. Large batteries make for large and uncomfortable sensors, and small size and weight is an important design criterion to allow for long term use. To accomplish this goal, techniques for energy-efficient runtime operation, and energy replenishment through energy harvesting must be explored.

Goal 4: Flexible, General Purpose, Research-ready, & Easy to Use. Face masks are ubiquitous passive wearables that, when retrofitted with FaceBit devices, could become important sensing platforms for future health studies in the medical and UbiComp communities. FaceBit should be small enough to fit in the mask without bothering the wearer, should provide capability for general purpose sensing, should be easily reconfigurable in

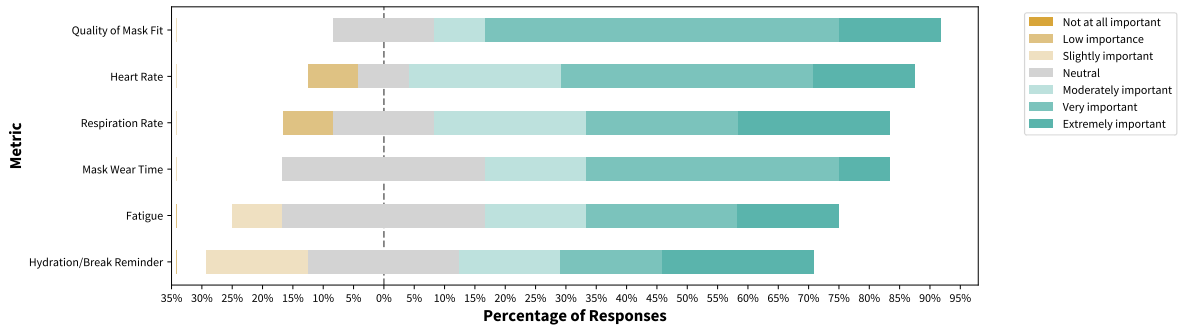


Fig. 3. 1-7 Likert Scale of importance for metrics of focus. The majority of the metrics presented were deemed to be important by healthcare workers (to the right of the 0% mark). The length of each color represents the percentage of people that provided the corresponding response. A full list of metrics can be found in Appendix A.

software and health metric collection, and should have associated tools and infrastructure to enable streamlined data collection and interpretation.

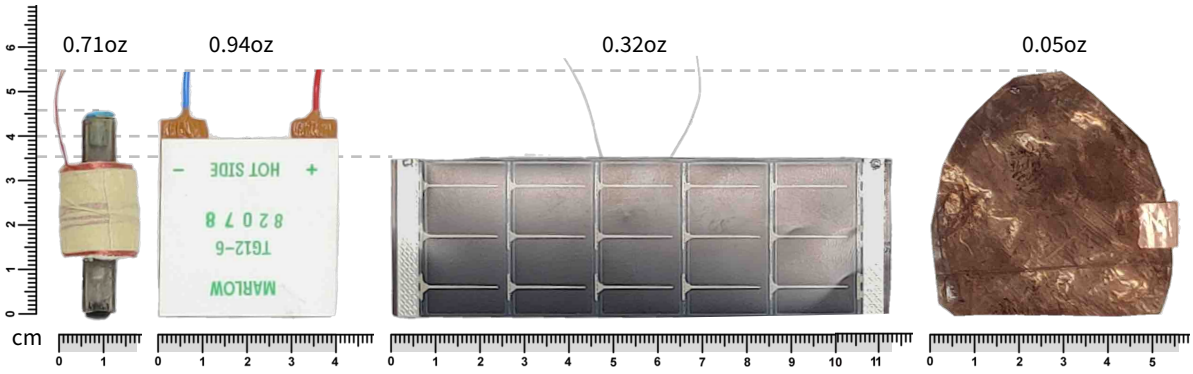
3.2 Design Exploration #1: Clinical Need Finding

FaceBit is meant to be general purpose, however, healthcare professionals have the most risk, and the most constraints on PPE usage (via CDC and OSHA), so we opted to inform the design of FaceBit prioritizing healthcare workers assuming that the general population would also benefit. To understand clinicians' needs for smart face masks and guide our platform design, we conducted a study of 12 healthcare workers (4 medical doctors, 2 nurse practitioners, and 6 medical assistants) at a nearby Pediatrics clinic. We interviewed participants using a semi-structured interview that included a survey with a combination of 20 multiple-choice and free text questions. Additional details of the survey and study are provided in A, we provide the design-relevant details here.

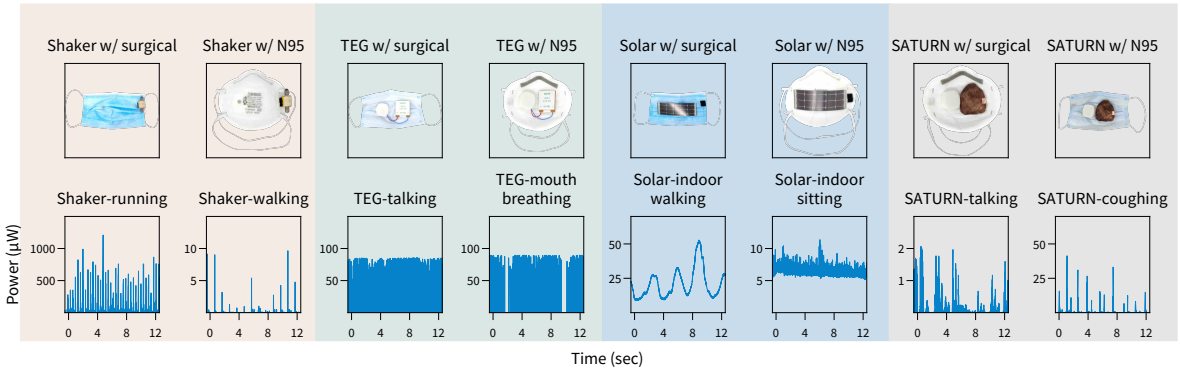
First, we asked a series of baseline questions to assess their typical masking behavior before and during the COVID-19 pandemic. 9 of the 12 (75%) healthcare workers reported never wearing a face mask before the pandemic. All 12 participants reported, while at work, always wearing a face mask (surgical) throughout the COVID-19 pandemic, only removing the mask when eating. At the time of the study, N95 shortages were widespread, which may have played a part in the use of surgical masks.

We asked each healthcare worker an open question about what features they would find helpful from a smart face mask. Common themes from these discussions included statistics about a mask's efficiency, such as airflow or when it is time to dispose of the mask, detection of exposure to COVID-19 or other viral infections, and vital health metrics such as heart rate and respiratory rate. Next, we presented each clinician with a compiled list of 18 potential health metrics drawn from health guidelines and that are attainable by wearable devices based on prior literature. Each metric is considered a candidate for FaceBit based on the platform's current or future envisioned capabilities. These metrics can be categorized into three interest groups: Mask Information and Safety (i.e, is the mask fitted and working), Personal Health (i.e., respiratory system signals and real time physiological monitoring), and Environmental Factors (i.e, air quality).

Clinicians ranked each metric on a Likert response scale. All metrics were considered important in some regard by the clinicians we surveyed. For this study, we chose to focus on metrics that were both of interest to clinicians and attainable by the first-generation implementation of FaceBit considering size, energy, and computational constraints. Figure 3 shows the relative importance of the metrics, discussed in following sections.



(a) Energy harvesters: shaker, TEG, flexible solar panel, and SATURN are shown along with their size and dimensions.



(b) Power harvested from activities with different harvesters along with mask placement.

Fig. 4. Placement of different energy harvesters on surgical and N95 masks is shown (top). Power harvested from activities (bottom). Shaker and SATURN generate energy impulses as the activity is being performed; solar panels give continuous energy as they are always exposed to some light; and thermoelectric generators gives a PWM like signal because of the boost converter used to step-up the very low voltages generated by the small temperature differential between the two plates of the TEG.

3.3 Design Exploration #2: Energy Harvesting On the Face

We investigated energy harvesting opportunities around the face understand the potential for extending the battery lifetime of Facebit. Gadgets on the face are often exposed to ambient light, motion energy from head movements, and air flow which causes thermal and kinetic changes that can be harvested. We considered routine activities that mask wearers perform in daily settings, and mapped these activities to harvesters – some we found off-the-shelf and others built from scratch. We outline our design exploration, discuss tradeoffs of each harvester, and the energy harvesting potential of daily activities. For all experiments we used a Rocketlogger device [66] to record the energy harvested while an experimenter performed these activities with the mask on. The results of the exploration are shown in Figure 4.

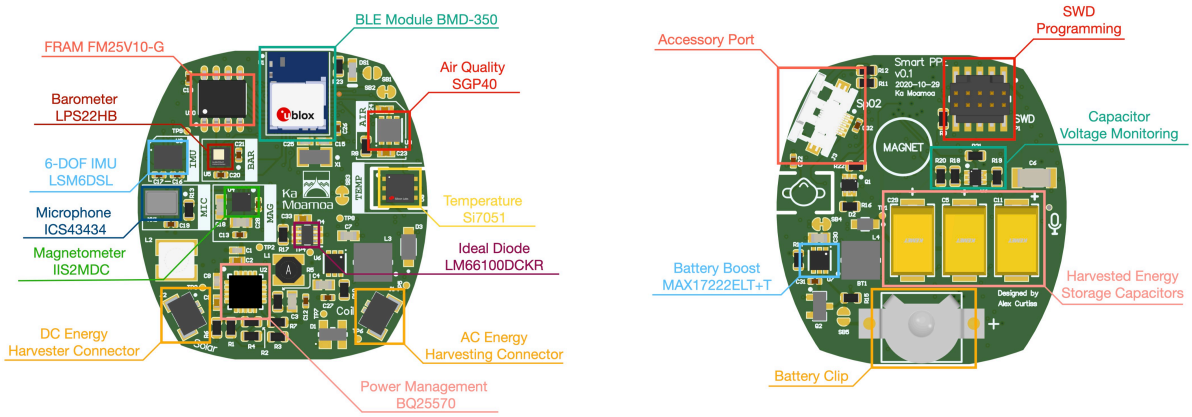
3.3.1 Light. We attach a paper-thick flexible PowerFilm MP3-37 4.49×1.44in 150mW flexible solar panel [67] (Fig. 4a) to the outside of a face mask. A disposable surgical mask is 7×3.875in. We clip the solar panel to surgical and N95 masks using small magnets (Fig. 4b) and perform different activities: walking indoors/outdoors on cloudy and sunny days, sitting indoors without a light source nearby, and sitting indoors near a light source.

3.3.2 Head Movement. To the best of our knowledge, there exists no such harvester that can capture head movement and is small enough to be attached to a surgical mask. We fabricated a lightweight harvester, a "shaker" that generates energy from head movements (Fig. 4a). Our shaker follows the principle of electromagnetic induction. A coil is made by winding a 42-gauge coated copper wire (typically used in motors and generators) on a small hollow plastic cylinder. A cylindrical magnet is placed inside the core with both ends blocked by caps. Another magnet, smaller than the one in the core, is attached to one of the caps, such that the core magnet and side magnet repel one another. When the harvester is shaken by head movement, the core magnet moves inside the coil. This movement changes the magnetic field around the coil, resulting in voltage induction (as per Faraday's Law of Induction).

3.3.3 Thermal. Face masks trap the heat expelled by breathing. A thermoelectric generator (TEG) can convert this heat to useful electrical energy. We attach a 40.13×40.13×3.91mm TEG TG12-6-01S [52] to the inside of a mask with the help of magnets (Fig. 4b). As the wearer breathes, the side of the TEG facing the user becomes hotter than the other side, and this generates a very low voltage that must be boosted by a specialized DC/DC converter to a usable level. We recorded how much energy we can harvest from activities like nasal breathing, mouth breathing, and talking.

3.3.4 Kinetic Energy from Breath. We use a triboelectric nanogenerator (TENG) to convert the kinetic force of breath into electricity [78]. We design a variant of the SATURN [7] platform, a TENG that is flexible and consists of conformable light-weight materials – a film of dielectric (Fluorinated Ethylene Propylene), and a paper with tiny laser cut holes. Both the paper and the dielectric have a thin layer of copper deposited onto them, and are arranged on top of each other such that the dielectric is sandwiched between the two copper layers. Simple fabrication and a configurable shape (Fig. 4a) allows it to be embedded easily into different types of masks using simple magnetic attachments. Fig. 4b shows the TENG device embedded into both N95 and surgical masks.

3.3.5 Summary of Energy Generation. Fig. 4a shows the weight and dimensions of all the harvesters we used for harvesting energy from face masks. Overall, different energy harvesters have different sets of pros and cons, and a user may choose to simply not integrate an energy harvester. As expected, solar panels provide the highest output, even indoors, but might be considered odd looking when placed on a mask. Thermoelectric harvesters provide a short burst of energy when first put on the face, but once temperature on both sides of the TEG equalizes, the energy generation stops. TENG harvesters are the lightest but generate the smallest amount of energy among all options. Despite this low power output, the light weight and flexible form factor still make TENG harvesters an attractive option to be explored in the future. The Shaker outputs significant amounts of energy when its wearer is running (our second best harvester), however, it is heavy and bulky. The shaker harvests more when attached to a surgical mask. We hypothesize this is because surgical masks are not as tightly fitted as N95 masks which amplifies the motion of the magnet. We refer the reader to the appendix for more details on the device fabrication and underlying physical principles of each harvester, and to Table 3 in the appendix which provides total energy harvested for each activity.



(a) Top

(b) Bottom

(c) In-mask

(d) Outer Mask

Fig. 5. The top of the FaceBit board contains the compute, sensing, and most of the energy harvesting circuitry. The application runs on the BMD-350 Module, which incorporates Nordic's NRF52832 (512k Flash/64k RAM). The bottom of the FaceBit board contains the energy storage elements and the programming interface. It also includes an accessory port which allows additional modules to be added in the future. FaceBit is shown placed in a N95 in (c) via a magnet (d).

3.4 Hardware Platform Design

FaceBit is a two part system, as shown in Figure 2. FaceBit comprises a small printed circuit board (PCB) with microcontroller and sensor suite that attaches to any surgical, N95, or cloth face mask via a magnetic clip. Figure 5 shows an annotated render of the device, as well as placement in an N95 mask (with magnetic attachment).

In alignment with our design goals, FaceBit is small enough to fit comfortably in either an N95 or surgical mask (1.2" wide x 1.3" tall). The board features an NRF2832 SoC from Nordic Semiconductor, which is packaged with a chip antenna on the pre-certified BMD-350 BLE Module. With a goal of using this platform to explore a wide range of health and environmental sensing applications, this version includes six sensors. Only three are used in this paper: the LPS22HB barometer, the LSM6DSL 6-DOF Inertial Measurement Unit (IMU), and the Si7051 temperature sensor. In future versions, we expect to reduce the number of sensors and (therefore) reduce the size and cost of the board.

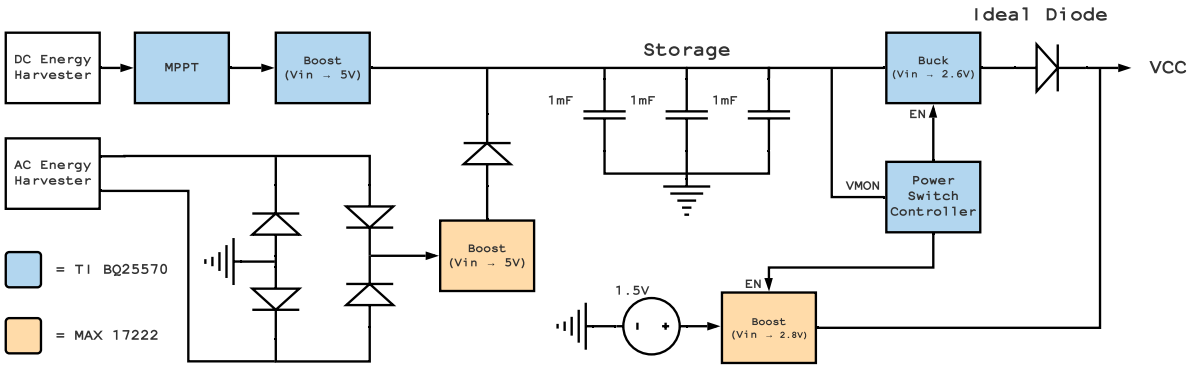


Fig. 6. Simplified hybrid power architecture of the FaceBit platform. The circuit is able to power itself from either a primary cell or energy harvested from AC and DC sources. The system switches intelligently between these sources, favoring harvested energy when a sufficient quantity has been stored.

We adopt a flexible hybrid approach with regards to power (figure 6). We include a battery holder for a small, 105 mWh primary cell, but also include the circuitry and storage to power the board using energy harvested from both DC (e.g. solar) or AC (e.g. shaker) sources. Three tantalum capacitors on the bottom of the board combine to provide 3 mF of storage capacity for harvested energy. These are charged by Texas Instruments' BQ25570 power management IC, which incorporates a boost converter with MPPT, intelligent charging circuitry, and a buck converter to reduce the voltage from the storage elements to a level tolerated by the system. The tantalum capacitors are charged to 5V (80% of their 6.3V rating), and so are able to store a total of 10.4 μWh. We use the "power good" indicator on the BQ25570 IC to implement an intelligent switch that automatically decides which source to draw power from based on the voltage of the capacitors. By default, the system draws its power from the battery (when one is present). When the storage capacitor voltage exceeds a resistor-programmed threshold (3.0V in our application), the BQ25570 begins to draw from them and the battery boost converter is disabled, which protects the battery from reverse current. When the storage capacitors fall below 2.6V, the BQ25570 switches off it's buck converter and system power is once again drawn from the battery. An ideal diode (LM66100DCKR) prevents reverse flow into the energy harvesting circuitry while the battery boost converter is active. This power architecture allows for a relatively long battery life (the energy density of primary cells being greater than that of rechargeable cells), while still maintaining the ability to store and use harvested energy. The device can also operate in a mode where it uses only harvested energy, activated simply by not placing a battery in the connector. This flexible architecture suited the development process, but it is also possible to further simplify and miniaturize the board by using either only a rechargeable cell or only energy storage capacitors.

3.5 Runtime System.

Unlike other work [65, 80] our health metrics are computed in real time on the device itself. This requires careful attention to energy cost so as to save battery. This starts with our energy-efficient runtime which eliminates the need to manually turn FaceBit off or on each time a mask is donned or doffed. An overview is shown in Figure 7, and the algorithm is described in detail in the following section. Once it is established that the mask is being worn, the remaining tasks (heart and respiratory rate sensing, BLE broadcast) are scheduled on a timer. These intervals can be adjusted to fit the application's needs. Figure 7 shows the intervals used in our evaluation. Next, we describe the health metrics and their sensing enabled by FaceBit.

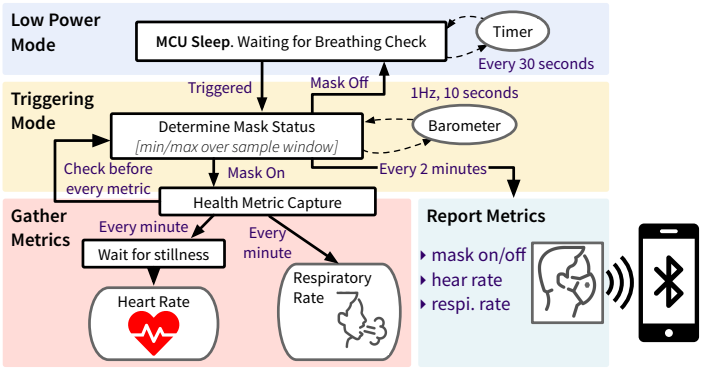


Fig. 7. The runtime system saves energy by using body signals to guide when to turn on sensors, process data and metrics, and report those to the phone. All operation frequencies are configurable.

4 HEALTH METRIC CAPTURE: ENERGY-EFFICIENT RUNTIME AND SIGNAL PROCESSING

We focus on sensing health metrics that are technically feasible for commodity sensors (e.g., detecting COVID-19 micro-droplets is likely not possible), have a definitive ground truth to compare against, and that are of interest to clinicians as identified by our study. These health metrics are wear time, heart rate, respiration rate, and mask fit.

4.1 Metric: Mask Wear Time: Are you Breathing?

4.1.1 Signal Existence. In order to compute wear time we must first determine if the mask is on or off. This can be accomplished using the difference in the pressure signal over a short time window that contains at least one respiratory cycle. Our key observation is that the standard deviation of a pressure signal when the mask is on is much higher than when the mask is off, due to the pressure differences caused by respiration. This can be further simplified by only retaining the minimum and maximum pressure values within a time window. This

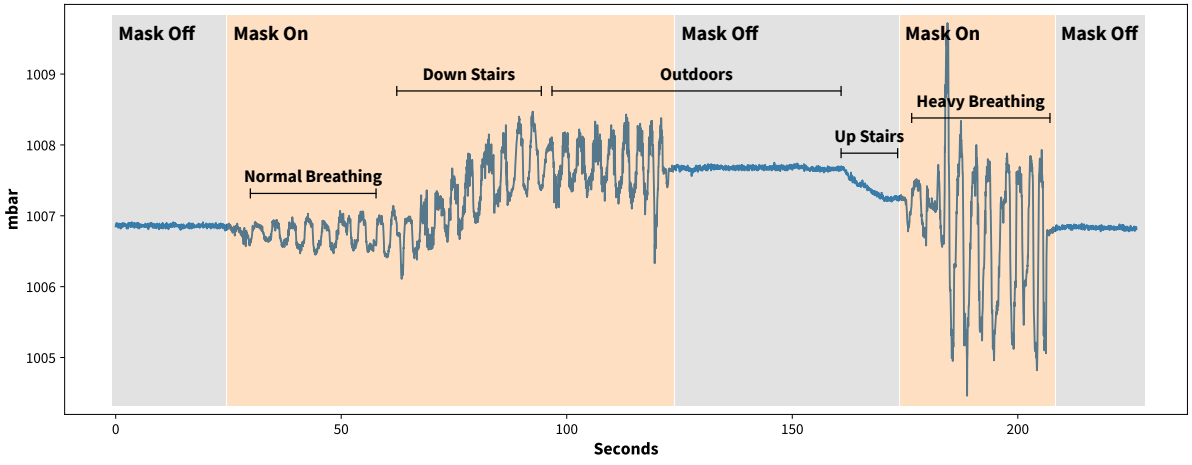
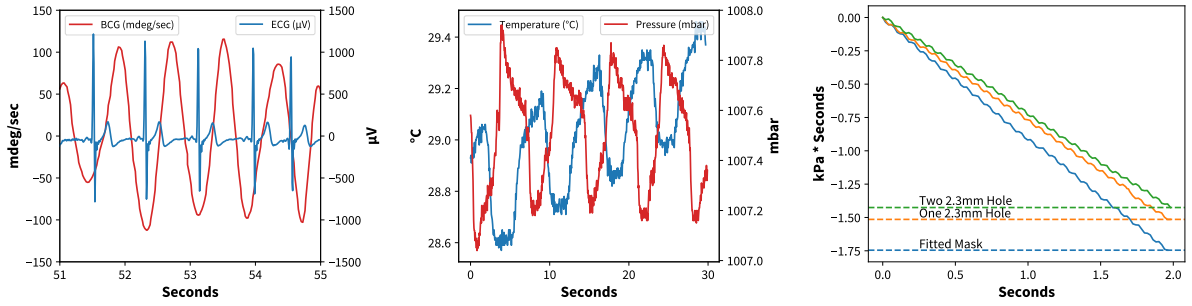


Fig. 8. Example air pressure signal, with mask on/off annotations, collected from FaceBit in an N95 mask in a field test.



(a) Heart Rate (b) Respiratory Rate (c) Mask Fit

Fig. 9. Metric signal existence plots. **(a)** ECG ground truth signal collected from a Polar H10 chest strap compared to a Ballistocardiography (BCG) signal extracted from the FaceBit's gyroscope. **(b)** Respiratory rate signal existence: Raw temperature and pressure signal from the FaceBit sensor board. **(c)** Mask fit exploration. Cumulative integration of the pressure signal sensed inside an N95 mask while under suction from a 12V vacuum pump . Minimum values highlighted represent a measure of mask resistance to airflow.

difference, in combination with a minimum threshold, allows us to clearly see state changes in the mask. An example pressure signal which follows a user through a normal sequence of activities is shown in figure 8.

4.1.2 Signal Processing Pipeline. In idle mode, the device wakes up from sleep intermittently to check if it is being worn. Since this task is performed often, it is important that it is low power. With this in mind, the board uses only the barometer and reads data from it slowly at 1 Hz, sleeping between measurements. It records ten samples over ten seconds, keeping track of the minimum and maximum values. If the difference between these two is greater than 0.1 mbar, it assumes the mask is being worn. Since spurious spikes in pressure can occur when the mask is off, we perform this check before every health measurement task.

These data-points can be aggregated to calculate a running wear time, but in order to determine wear time of any particular face mask, user interaction is required to track disposal or switching. Therefore, wear time is calculated by both the sensor board and via user interaction with the mobile application.

4.2 Metric: Ballistocardiography: Heart Rate from Micro-Movements of the Head.

Heart rate is an important physiological parameter. An elevated resting heart rate is associated with increased incidence of cardiovascular disease in both men and women [17]. Heart rate also rated highly in our healthcare worker survey. Unfortunately, traditional methods of heart rate detection (e.g. ECG, photoplethysmography) require constant, immobile contact with the skin which is unlikely in the context of an easy to don and doff mask. Remote imaging methods such as remote PPG were considered, however it was believed the unstable lighting conditions encountered in everyday life would confound those efforts, and additionally the need for a bright light source or otherwise a sensitive photodetector or camera would raise the cost or energy budget of the platform. As such, we chose to explore another non-contact method that would allow for the gathering of heart rate data at various points throughout the day.

4.2.1 Signal Existence. Ballistocardiography (BCG) is a method of cardiac health assessment whose origins date back to the turn of the 20th century[29], and was significantly refined by Starr, et al. in 1939[69]. The historical methods relied on complicated mechanical setups involving swinging, free-hanging tables or chairs suspended from heavy springs, but they took advantage of the same basic phenomenon we rely on in our FaceBit heart rate

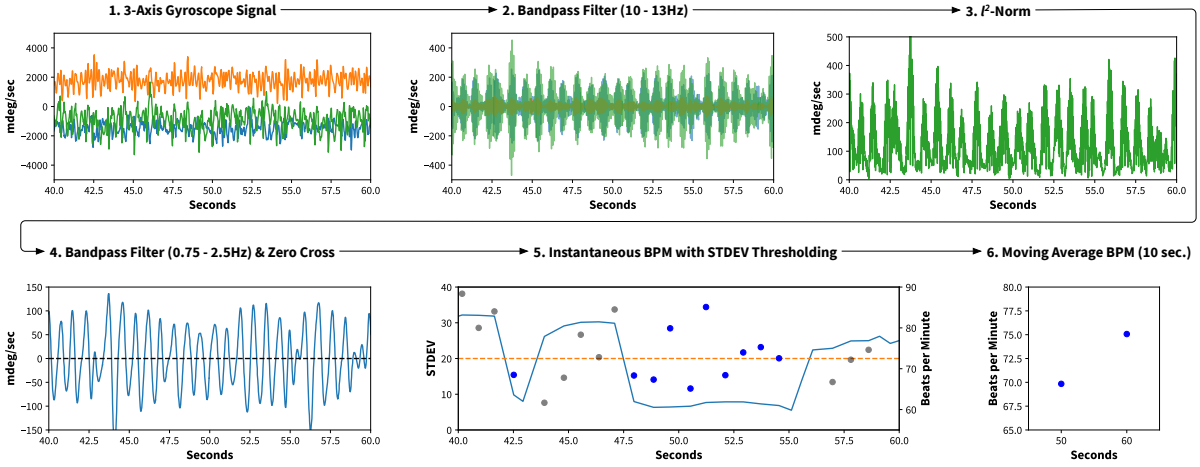


Fig. 10. Heart rate signal processing pipeline. 3-axis gyroscope data is filtered to extract instantaneous heart rate using the time between each zero cross of the signal. A standard deviation threshold is used to determine realistic change in rate from beat-to-beat. Finally, heart rate is calculated as a 10-second average of valid instantaneous rates.

detection algorithm: with each heartbeat, blood is ejected forcefully out of the left ventricle of the heart into the major arteries of the body. This causes a subtle "recoil" throughout the body, which can be detected either by using mechanical contraptions, or by sensitive accelerometers like the LSM6DSL found on the FaceBit circuit board. Each ballistocardiographic pulse is composed of several high-frequency waves, which are hypothesized to arise from interactions between regions of higher and lower pressures in the ascending and descending aorta during each heartbeat [40]. The intervals and amplitudes of these individual waves in the BCG complex are thought to correlate with various cardiovascular health markers [25]. The pre-processing approach adopted by FaceBit exploits this characteristic structure to better locate each heartbeat in the raw 3-axis gyroscopic data, which is fundamentally a noisy signal in a wearable context. Figure 9a shows the recovered BCG/heartbeat signal obtained from a seated participant wearing an N95 mask fitted with the FaceBit, along with an ECG signal recorded at the same time. Each peak in the BCG signal represents a collection of tiny micro-motions of the head caused by the heartbeat represented by the peak in the ECG signal directly preceding it.

4.2.2 Signal Processing Pipeline. Figure 10 shows our signal processing pipeline. Our pre-processing methodology is adapted from Hernandez et al. in the BioGlass paper [30], however we implement the algorithm entirely on the NRF52832 in C++ to run in real-time. The modified BioGlass algorithm is briefly presented here. We begin by collecting 3-axis gyroscope data at 52 Hz. We then apply a fourth-order IIR band-pass filter with cutoff frequencies of 10 and 13 Hz. This passband removes the DC component and low-frequency motion artifacts (e.g. respiration), but retains the high frequency pressure waves associated with the BCG waveform complex. Namely, the I, J, K, and L waves of the typical BCG waveform occur in this frequency range [40]. The signals from each of the three axes are then combined via the euclidean norm (or l^2-norm). Notably, this operation allows the algorithm to function with the FaceBit board placed in any orientation relative to the wearer. Next, the l^2-norm signal is fed through a smoothing, second-order IIR band-pass filter with cutoff frequencies of 0.75 and 2.5 Hz, corresponding to a range of 45-150 beats per minute. This is the recovered BCG-HR signal.

Next we introduce the FaceBeat algorithm, which is tasked with extracting heart rates in real time from this signal, which is easily corrupted by motion artifacts from everyday living like walking or even just looking around.

Moreover, the heavy filtering involved in the pre-processing stage tends to produce signals which resemble valid heart rates when none are present. The FaceBeat algorithm attempts to exclude regions of data that it suspects are due to motion and phantom signals, and use only the regions that correspond with the wearer's true heart rate. In this way, FaceBit is able to detect the user's heart rate during "moments of stillness" throughout the day.

To accomplish this, the algorithm first detects and stores the timestamps of descending zero-crosses. When five such timestamps have accrued, the instantaneous heart rate associated with these timestamps is calculated by 1) measuring the time difference between consecutive elements, yielding $\frac{[seconds]}{[beat]}$ 2) taking the reciprocal of this number to obtain $\frac{[beats]}{[second]}$ and 3) multiplying the reciprocal by 60, thereby arriving at $\frac{[beats]}{[minute]}$. At this point we have four "instantaneous" heart rates associated with the last five suspected heart beats. We rely on the innate predictability of the heart and the inter-beat interval to conclude whether these instantaneous heart rates are genuine or the product of motion artifacts. Specifically, we calculate the standard deviation of the four instantaneous heart rates, and if they fall below some threshold (we used 20 beats/min in our initial evaluation), their average is recorded as a valid heart rate. To save energy in our application, the FaceBeat algorithm is set to run intermittently. Initially, all valid heart rates (>45 BPM and <150 BPM) recorded within a given window were averaged together, and it is this average that was reported to the mobile application.

After an initial in-lab evaluation (section 7), the FaceBeat algorithm was refined. The standard deviation cut-off was lowered from 20 beats/min to 7 beats/min. We also used the amplitude of the BCG waveform as a discriminating factor: it was noted that the amplitude of valid BCG waveforms are typically quite small, less than a maximum of 100 millidegrees/second. This is easily dwarfed by normal body motion, vehicular vibrations, and other everyday phenomena. Therefore, we added a fourth step to filter out suspected beats whose amplitudes exceed a given threshold (150 millidegrees/second in our scenario-based evaluation, section 7.5). Finally, we began reporting any valid instantaneous heart rates rather than an average in order to better compare our results with our ground truth.

4.2.3 *Confounding Scenarios and Response.*

- The rejection of all motion artifacts is complex, particularly those that fall within the 10-13 Hz window used by the first band-pass filter and especially the subset of those which are periodic in the same range as heart beats. In response, we designed the motion-aware approach to only report user heart rate when the motion artifacts are sufficiently small.
- Physiological differences (height, weight, cardiac stroke volume, cardiac arrhythmias) between users and within individual users over time may require adjustments to the thresholds used in the FaceBeat algorithm. A promising remedy is personalization of the algorithm to each user via an initial calibration routine carried out in conjunction with the FaceBit mobile application, as is further discussed in section 8.2.
- Location of the FaceBit circuit board within the mask or differences between masks (e.g. stiffness, size) may alter the raw gyroscopic signals the algorithm uses to detect heartbeats. Taking the euclidean norm of these filtered signals is thought to make the method robust to these effects. We report on experiments to this end in section 7.7.

4.3 Metric: Respiratory Rate from In-Mask Temperature Changes

In an N95 (or equivalent) type mask, normal breathing results in a pressure drop over the filter material that the FaceBit's pressure sensor can easily detect. However, loose-fitting surgical or cloth masks do not generate as large a signal, especially when placement of the FaceBit board is up to the user. Fortunately, breathing also causes the temperature in a mask to fluctuate with a distinct periodicity, and this effect is highly conserved between mask types. Moreover, the thermal mass of the board and surface-mount temperature sensor acts as a sort of low-pass filter, and the insulation properties of still air serves a rudimentary peak-hold function which amplifies

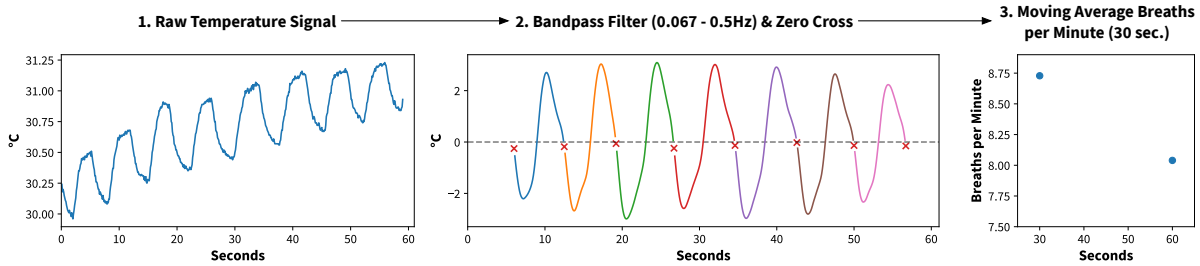


Fig. 11. Respiratory rate detection: tidal breathing and breath detection over 60 seconds using a band-pass IIR filter. The respiratory rate is calculated from the average duration of each breath.

the effect of moving air (and therefore breaths). For all these reasons, we use the on-board temperature sensor as our source for respiratory rate detection.

4.3.1 Signal Existence. Figure 9b shows the signal obtained from FaceBit's temperature sensor and pressure during tidal breathing in an N95 Mask. The temperature signal is shifted relative to the pressure signal due to the thermal mass of the circuit board and temperature sensor, but the frequency is identical.

4.3.2 Signal Processing Pipeline. Our respiratory rate algorithm begins by collecting temperature readings from the Si7051 temperature sensor at 10 Hz. The raw values are sent through a second-order IIR band-pass filter with cutoff frequencies at 0.067 and 0.5 Hz, corresponding to limits of 4 and 30 breaths per minute (BrPM). This filters out the DC component (ambient temperature) as well as trends that might result from donning the mask or entering a room with a different temperature. Next, the timestamps of any descending zero-crosses are buffered. At the end of the sampling window (e.g. 30 seconds), the respiratory rate is calculated from the average time difference between zero-crosses (i.e. length of each breath). Any breaths that correspond to a respiratory rate faster than 30 BrPM or slower than 4 BrPM are excluded from the final calculation, as these are likely to have occurred during speech or other disordered breathing. Figure 11 shows an example raw signal and the output of the algorithm.

This algorithm was updated after our initial evaluation (section 7) to achieve better performance. To measure high respiratory rates that may occur during exertion, we expanded the filter passband to allow frequencies from 0.067 to 1 Hz (4 to 60 BrPM). We also included a temperature threshold that the signal must exceed to register as a breath (0.04 °C) in order to reduce the classification of small oscillations the arise from our filter as valid breaths.

4.3.3 Confounding Scenarios and Response. There are several variables which can potentially affect the performance of our algorithm.

- Sensing respiration rate when the user is talking can affect the output of algorithm. We do not attempt to correct for this error in our current algorithm, although we review potential advanced filtering methods in the discussion.
- Poorly constructed masks often yield a lower signal-to-noise ratio, as the temperature difference between the inside and the outside of these masks is much lower than that seen in high-quality masks. In practice, we found that positioning the FaceBit near the center of the mask mitigates this problem since it puts it directly in the path of each breath. We explore temperature variation within N95 and surgical masks in section 7.7.
- As ambient temperatures approach the user's exhaled breath temperature, signal quality is expected to decrease. Exhaled breath temperature, of course, depends on ambient temperature but reaches a maximum

of approximately 32.5 °C [13]. As such, respiratory rate may not be reliably measured by this method in warm environments. In these conditions, the pressure sensor would likely provide a cleaner signal, suggesting a sensing algorithm that leverages both signals in the future. Since the signal processing pipeline functions equally well for pressure data, it would be relatively easy to have a sensing pipeline that switches its signal source between pressure and temperature.

4.4 Metric: N95 Mask Fit

Mask fit was the most important metric to the clinicians surveyed as part of this study. This is understandable because the difference between a well- and a poorly-fit mask can have significant health-related implications. However, mask fit is a difficult parameter to measure, and currently requires bulky, awkward, and expensive test equipment.

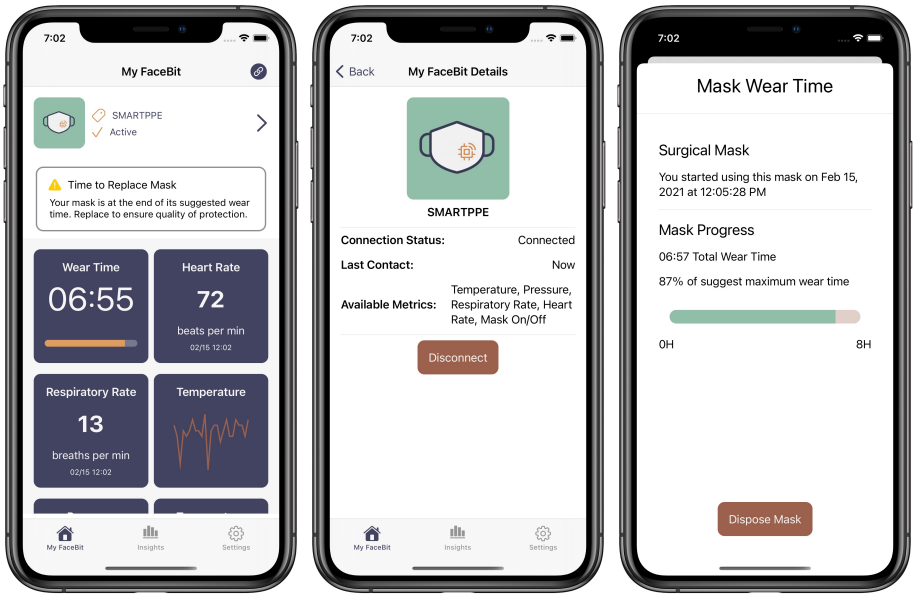
4.4.1 Signal Existence. Initial results from our platform showed a promising correlation between leakage and the integral of the pressure signal while a mask was under suction from an attached vacuum pump (Fig. 9c). A well-fit mask presents a higher resistance to airflow, and so the slope of this integral can serve as a straightforward indicator of mask fit.
However, in practice this would be a cumbersome method to detect mask fit. The vacuum pump is large, loud, heavy, and requires a port that passes through the mask. Instead, in keeping with our aim to achieve a low user burden we seek avoid any mask modification or additional hardware, and rely solely on the user's lungs to pass a constant volume of air through the mask using a special breathing exercise which could one day be facilitated by our mobile application. A mask fit score is calculated using a signal processing pipeline in conjunction with this exercise.

4.4.2 Breathing Exercise and Signal Processing Pipeline. First, the user is asked to exhale as much air as they can from their lungs (reaching their "residual volume"), and then hold their breath for five seconds. This breath-hold allows the inside of the mask to equilibrate with ambient pressure, which is recorded and used as a baseline. The user is then asked to inhale as deeply as possible at a normal rate until their lungs are filled as much as possible, and then hold their breath for another 5 seconds to obtain another atmospheric pressure reading. This exercise (breathing in as much air as possible after breathing out as much air as possible) is also known as an inspiratory vital capacity test. Vital capacity is a static lung volume, and changes only slowly over time with age (though it can be affected in the short term by posture or disease conditions).

Pressure inside the mask is sampled at 50 Hz throughout the fit test. Atmospheric pressure (obtained at the beginning and end) is subtracted from the pressure readings to obtain gauge pressure. The area under this curve [Pa*s] is proportional to the total volume of air inhaled, with mask resistance $[kg^{-1}meter^{4}second]$ as the proportionality constant. Because we do not know the total volume of air inhaled, which is necessary to compute a value for mask resistance, we ask the user to calibrate the device by pressing the mask against their face to ensure a good seal. Note that this calibration needs only to be performed once per mask type per user. Future mask fit scores are calculated as a percentage of this initial calibration test.

4.4.3 Confounding Scenarios and Response. The size of the leak and the consistency of inspiratory vital capacity are the two variables that affect this algorithm.

A natural question is: what size leak do we need to detect to be useful? Rengasamy et al. [59] studied the effect of leak size on particle penetration in N95 masks, and found that a mask with two leaks (3mm diameter each) allowed a total inward leakage of 10.67% ± 4.6% (ratio of particle concentration inside the mask to outside the mask). This, compared to 0.31% ± 0.4% for a respirator with no leaks. Four leaks (3mm diameter) increased the TIL to 30.54% ± 4.2%. Meanwhile, OSHA requires a fit factor greater than 100 (less than 1% TIL) during quantitative



(a) Home Screen (b) FaceBit Device Details (c) Mask Wear Screen

Fig. 12. FaceBit Companion Application on iOS. **(a)**: The Facebit Phone Application homepage displaying general details, current wear time, respiration rate, heart-rate, and temperature data from the FaceBit Sensor Board. Also shown is a wear time warning indication suggesting a mask change. **(b)**: The mask wear time interface for tracking mask disposal. **(c)**: FaceBit Sensor Board detail screen outline current connection status

fit testing procedures, and so we can piece together that we should strive to detect leaks smaller than 3mm (x2). It isn't clear from the available literature, however, what the absolute minimum allowable leak size is.

Inspiratory vital capacity can change with posture, however this can be mitigated by guiding the user through the exercise. More troublingly, it may be difficult for some people with chronic lung diseases to perform. Other potential methods are discussed in section 8. We are unaware of any reports detailing the consistency of inspiratory vital capacity for healthy individuals on a day-to-day basis, but some combination of this consistency and the magnitude of the effect observable in the signal will inform the efficacy of this method.

5 FACEBIT PHONE APPLICATION

The FaceBit Companion Application is a phone and desktop application that serves two purposes. First, it is the research tool that allows interaction with the FaceBit sensor board. All evaluations utilize the FaceBit Companion Application for data collection. Second, the application serves as a proof-of-concept user interface for a Smart PPE platform.

The FaceBit Companion Application is developed for iOS and macOS using the Swift Programming Language. The user interface is written with the SwiftUI framework and utilizes Mac Catalyst, which supports iOS applications deployed on macOS. The app communicates with the FaceBit board via Bluetooth Low Energy (BLE) using a custom GATT profile. The application handles both high frequency time-series data stream for debugging, as well as low frequency computed metrics from the sensor board for actual deployment. Data is stored in a local SQLite database.

We designed three sections of the app to demonstrate consumer interaction with FaceBit and two additional sections to assist in research. The user interface is compartmentalized such that user interaction does not interfere with data collection or vice-versa. Figure 12 depicts the application's home screen, sensor device details, and mask wear-time interface, the latter of which allows for tracking of the replacement of a mask via user logging and wear time detection. The homepage is a dashboard for at-a-glance information about the state of FaceBit. Notifications can be displayed or pushed to the user, and goals set for the desired amount of wear time. Additionally, this time-tracking interface can assist other reminders such as hydration or break reminders, notifying the user the mask has been worn for continuous lengths of time. The vision is that a user will be able to customize the homescreen with desired widgets to receive timely information on the signals of interest. Additionally, developers will be able to easily add new widgets. Based on feedback from healthcare workers (see section 3), we developed widgets for respiratory rate, heart rate, and wear time. We also included widgets for raw temperature and pressure readings for debugging purposes during development.

6 FACEBIT ENABLED RESEARCH

The FaceBit platform provides capabilities that support research in the ubiquitous and mobile computing research community as well as basic clinical health research. We detail three use cases we support directly that are demonstrated in this paper. We release the project as open-source and open hardware to support these directions and more, facilitating the research community's efforts to build and validate smart PPE. To support the evaluation and exploration of the FaceBit platform, the research team developed a set of tools and a data-pipeline methodology to streamline research, including an event recording interface for the application, a data explorer on the phone application, and a streamlined evaluation system integrated with gitlab. These tools, documentation, source code, and hardware files are hosted at facebit.health. We detail the use cases below separated by skillset.

Use Case #1: Embedded Systems UbiComp Researcher. FaceBit is essentially a body-worn health tracker, in the vein of FitBit, but extensible and with long battery life. FaceBit allows an embedded systems researcher to easily add aditional sensors using the expansion port. These could support applications like blood pressure or air quality monitoring. With the Facebit firmware written using mbed (C/C++), a popular embedded programming platform, this researcher could easily add these capabilities on top of the core FaceBit firmware. They could then deploy the entire platform for testing and evaluation, gathering data (air quality, for example) via the phone app. In this case, novel research questions around on-device and energy efficient activity monitoring, or novel sensing signals in the mask would be explored.

Use Case #2: Signals and Inference UbiComp Researcher. FaceBit allows UbiComp researchers without hardware expertise to explore novel activity and behavior based research from the new form factor of the face mask. FaceBit can be used as a data collection mechanism to train recognition models, hosted on the phone or used offline. For example, exploring fatigue prediction from the raw signals available in the mask would be an interesting research topic in this scope: it is sufficiently challenging, does not require hardware expertise, would benefit from the continuous monitoring capability, and could be done using the FaceBit platform. In this case researchers would gather raw data, saved to the phone or cloud, and then use traditional approaches like PyTorch and TensorFlow to train a fatigue prediction model. The trained model could be deployed with FaceBit as an intervention for participants. Each step of this process, from data collection, model training, and intervention deployment, would be facilitated by FaceBit hardware, reducing the barrier to entry for mask based studies.

Use Case #3: Clinical Researcher. We expect clinical researchers to use FaceBit as a data collection platform to support clinical trials and validation for various populations. Participants in a trial or pilot study could wear the FaceBit device and download the FaceBit application. The physiological signals like respiration and heart rate are generally useful baselines for many types of trials and interventions; for example understanding the effects of

blood pressure medication, dietary interventions, stress level change based on meditation exercises, etc. Being able to easily gather respiration and heart rates without a bulky device, merely retrofitting a mask already likely worn, simplifies the execution of a trial. We anticipate that a clinical researcher would collect this physiological data as part of the study, and then analyse it for statistical correlation with the interventions they deployed in the trial.

7 EVALUATION

We conduct an extensive evaluation covering many aspects of the FaceBit platform. The goal of our evaluation is to measure the capabilities of FaceBit, test the robustness and performance in the face of a diversity of situations, identify the places where improvement is needed, and understand the impacts of various factors on performance, battery lifetime, and user burden.

We conduct an experimental campaign to answer the following questions in our evaluation.

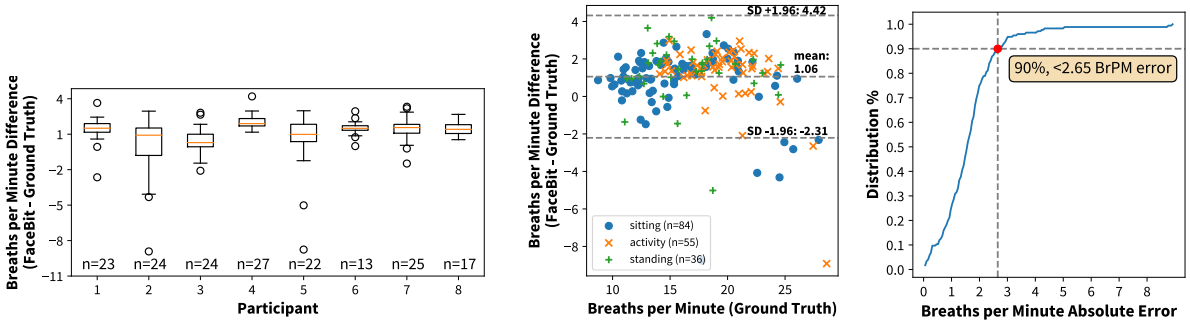
(1) How accurate are key FaceBit metrics in varied and confounding in-lab scenarios? Respiratory Rate (Section 7.2), Heart rate (Section 7.3, and Mask Fit (Section 7.4)
(2) How does FaceBit perform in noisy free-living scenarios: such as riding in the car, on a train, or walking down a busy street? (Section 7.5)
(3) What is the impact of device location and mask type on bio-physiological signal quality? (Section 7.7)
(4) What is the impact of human/physiological differences on signal quality? (Section 7.6)
(5) What is the the energy consumption and battery lifetime of FaceBit? (Section 7.8)
(6) What are the barriers and facilitators for user burden and comfort? Conducted via a user burden survey and interview with healthcare workers. (Section 7.9)

7.1 In-Lab Health Metric Evaluation Methodology

We conducted an in-lab study to understand the accuracy of the heart and respiratory rate metrics, calculated on-device. We then evaluated calculating mask fit based on leak detection in a mask. This section details the methodology, then the results follow for each metric. The evaluations of respiratory rate and heart rate were completed in unison, with FaceBit wirelessly transmitting both metrics to the FaceBit phone application during a set of activities performed in sequence by each participant. The evaluation included nine graduate and undergraduate participants (4 female), however, one participant's data was lost due to a ground-truth device error, therefore we report on eight participants.

Materials. All participants wore an N95 respirator (Demetech, Cup-Style) with an attached FaceBit device running the application firmware. Participants also wore a Polar heart rate monitor chest strap (H10), and a NeuLog respiration belt with a USB data logger (NUL-236, USB-200). FaceBit data, Polar heart rate data, and NeuLog data were collected using the FaceBit phone application (on macOS), a custom iOS application using Polar's BLE SDK, and NeuLog's experimentation application, respectively. Due to a documented low-power-mode related bug in the mbed BLE stack [1], FaceBit broadcast data to the companion application every 2-minutes and then reset.

Participants and In-lab Activities. Participants were evaluated for a total of 25 minutes, excluding setup and transition times. To obtain initial understanding of the reliability of the signal, we designed a three-phase structured study involving sitting (ten minutes), activity (two rounds of one minute of squats followed by four minutes of sitting), and standing (five minutes). While seated and standing, participants watched a documentary playing on a laptop screen. Evaluation with six participants took place outside on a University campus and three inside a lab space. One participant recorded inside was excluded due to an error in the NeuLog respiratory sensor



(a) Box-plot (b) Bland-Altman (c) CDF

Fig. 13. Respiratory Rate Evaluation Results for 8 participants (186 samples). **(a)** box-plot of each participant in our evaluation highlighting number of samples and outliers. **(b)** Bland-Altman plot comparing FaceBit's error to ground truth, showing a tendency to report slightly above the ground truth and to be less accurate with higher ground truth reports. **(c)** Cumulative distribution function over the evaluation samples.

data capture (the signal recording stopped mid-session). The evaluation followed a strict protocol for consistency and sanitation purposes, including following COVID-19 guidelines.

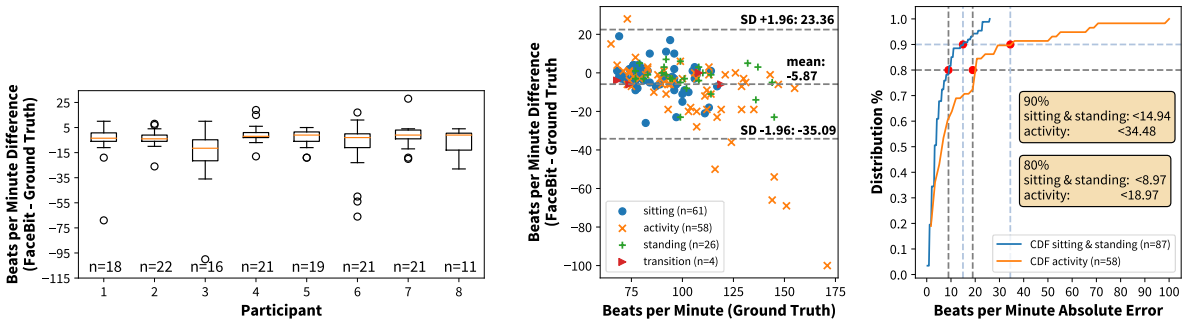
7.2 Respiration Rate

The Neulog respiration belt was fit around a participant's mid-section and recorded an arbitrary analog signal corresponding to the pressure of the belt's air-bladder. Data from the respiration belt was excluded from evaluation during air-squats and transitions since the belt's air-bladder compressed when moving from a standing to sitting position, resulting in inaccurate estimates. FaceBit attempted to measure respiration rate once per minute with a 30-second window of data.

The respiration belt collected data at 5Hz. Post-experimentation, ground truth respiration rate was calculated from this signal by first passing the signal through a Savitzky-Golay filter for smoothing. Peaks were then identified using an empirically set prominence value of 50 [Arb Units], a window length of 5-seconds and a minimum peak-peak distance of 1-second. An instantaneous breaths-per-minute (BrPM) rate was calculated at each peak of the respiration belt signal, defined by the length of time in seconds since the previous peak. To match FaceBit's data sampling window, we define the ground truth respiration rate as an average of instantaneous estimates over the previous 30-seconds at the time of a FaceBit measurement.

Results for the 8 participants (186 samples) are shown in Figure 13. On average, FaceBit reported a respiratory rate once every 55 seconds. The mean difference, highlighted by (13b), is 1.06 BrPM, demonstrating that on average FaceBit's algorithm reports slightly higher respiration rates than ground truth. FaceBit reported 90% of recorded samples within +/- 2.65 BrPM of ground truth. Note that manual calculations of respiratory rate fall within 1 bpm error by definition, counting the number of chest raises over one minute.

Discussion of Results: After analyzing the results, we found a few explanations that could account for the error we see compared to our ground truth. First, our algorithm for calculating ground truth is not perfect and does not account for anomalies that produce peaks in the signal that are not breaths, such as throat clearing, coughing, or talking. After exploring our outliers in detail, we discovered small peaks in the ground truth signal indicating some irregularity in breathing or perhaps a shift in weight on the air-bladder; this would cause the respiratory rate to increase briefly. Second, we were limited to a 5 Hz signal frequency by the NeuLog software, which could



(a) Boxplot (b) Bland-Altman (c) CDF

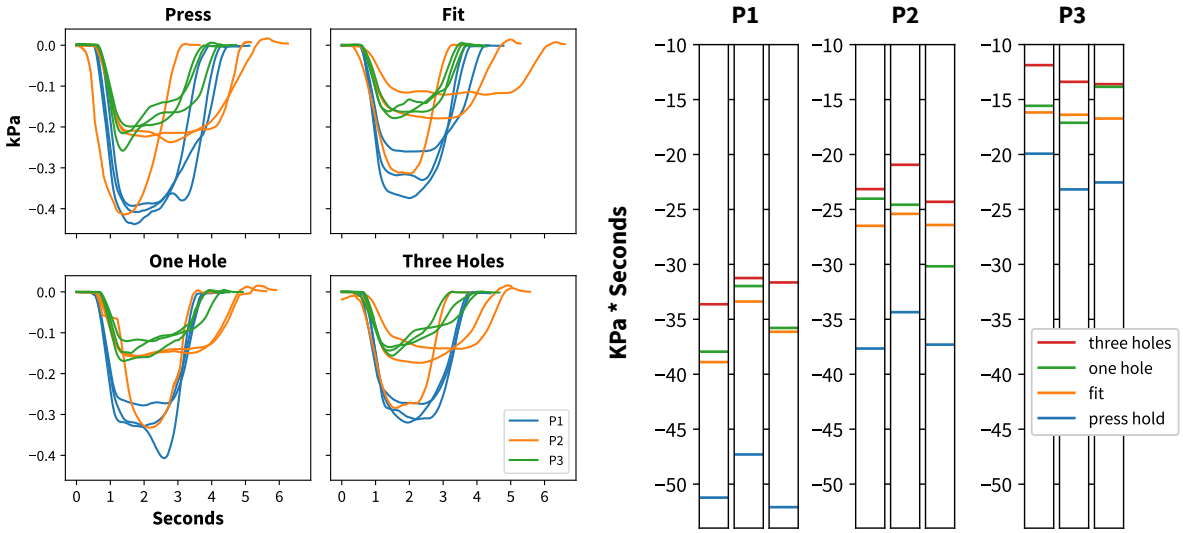
Fig. 14. Heart Rate Evaluation Results for 8 participants (149 Samples). **(a)** box-plot of each participant in our evaluation, highlighting samples per participant and outliers. FaceBit requires low motion levels to report a reading. **(b)** Bland-Altman plot showing FaceBit reporting better accuracy in the lower heart rate ranges. Data is split into activities to show issues with detecting during periods of high motion. **(c)** Probability and cumulative distribution function over the evaluation samples. Reported metrics during transition (n=4) are excluded.

cause shifted peaks, especially in higher frequency breathing. Additionally, after evaluation, we discovered a firmware bug which caused some respiratory rate calculations to overflow on the sensor board. Most of these values were easily identified in post-processing, as they fell below our minimum valid respiratory rate (4 BrPM). The outliers in the -8 to -9 range we suspect is due to this bug since it occurred when ground truth reported a high respiratory rate after a set of squats. However, we cannot determine with complete confidence, so it is left as reported by the FaceBit sensor board.

7.3 Heart Rate

A Polar H10 chest strap was fitted around each participant at the sternum and adjusted to the recommended tightness. Two streams of data were collected from the Polar device using a custom mobile application developed for this evaluation: live heart rate (as calculated by Polar), and a live ECG trace. The Polar data streams were timestamped in order to sync them with the FaceBit data. Polar live heart rate was reported on average every 1 second, and ECG was recorded at a rate of 130.14 Hz. During the evaluation, FaceBit attempted to measure heart rate once every minute with a sampling window of ten seconds. FaceBit will only report heart rate if the algorithm is confident in the reading section 4.2.2, otherwise it will broadcast a null result, indicated by a value of 1. FaceBit reported a null value for 14 of 163 total sampling periods. Participants #3 and #8 accounted for the majority of these null reports.

Every heart rate reported by FaceBit is an average of any "valid" heart rates recorded over the last ten seconds (see Section 4.2.2). Therefore, we compare each FaceBit heart rate with an average of all recorded Polar-calculated heart rates over the same ten seconds. Results for our 8 participants are shown in Figure 14. The mean error over the course of the evaluation was -5.87 beats per minutes (BPM), and the standard deviation of the error was 14.91 BPM. In general, the heart rates reported from FaceBit during our "sitting" and "standing" test conditions were more accurate than during our "activity" condition, with 90% of reported heart rates being less than 14.95 BPM off from our ground truth rate, and 80% being less than 8.97 BPM off. The largest errors were seen during the activity condition (Figure 14b), where 90% of our results fell within 34.48 BPM of ground truth, and 80% fell within 18.97 BPM. There is a clear trend of increased error magnitude at greater heart rates.



(a) Inspiratory Vital Capacity Test on FaceBit

(b) Integral Values of Vital Capacity Test on FaceBit

Fig. 15. Results from initial mask fit exploratory study capture using FaceBit. **(a)** Signal from inspiratory vital capacity test (IVCT) recorded on-board FaceBit showing filtered pressure signals for each participant (n=3) grouped by fit. **(b)** Integral values from initial evaluation for participant (3) and each experiment (3) compared to each leak configuration.

Discussion of Results: A significant percentage of FaceBit reported heart rates are within 5 BPM of ground truth (68% of measurements in the seated and standing conditions, 43% of measurements in the activity condition).

With that said, our evaluation pointed to a need for enhanced rejection of motion artifacts. There are a number of significant outliers in the sitting and standing conditions (20 BPM different from ground truth), and more significant outliers in the activity condition (>50 BPM). Lower accuracy at rates approaching and above 150 BPM are to be expected due to our second bandpass filter whose cutoff point lies in this range. There is also likely more motion from heavy and rapid breathing that often accompanies these faster heart rates.

An interesting finding is that our algorithm, on average, underestimated the heart rate. Even after removing outliers greater than 1.96 standard deviations from the mean, the mean difference (FaceBit - ground truth HR) was -3.71 BPM. An observation from testing during the development of the algorithm was that it detected heartbeats more successfully during exhalation than inhalation. Moreover, respiration is known to affect hemodynamics: inhalation decreases arterial pressure and increases heart rate through a series of physiological changes [14]. This drop in arterial pressure during inhalation may reduce the amplitude of the BCG waveform enough to prevent detection in some cases, which in turn may bias the detection of heart beats during exhalation when arterial pressure increases and (importantly) heart rate slows.

7.4 Leak Detection

Evaluation Methodology To explore the feasibility of leak detection on the FaceBit platform, we conducted two exploratory studies using raw pressure data from FaceBit and an external differential pressure sensor module. Both studies used an N95 mask fitted with three plastic bulkhead fittings with an internal diameter of 2.3 mm (4.2 mm^2), to which we attached a short length of silicone tubing with a removable end-cap to simulate a small mask

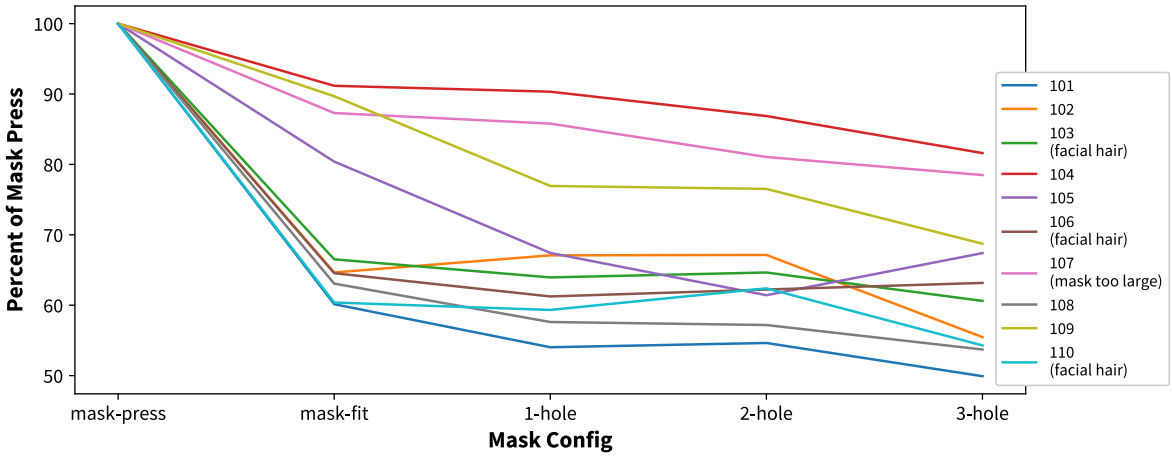


Fig. 16. Secondary mask fit exploratory study using an external differential pressure sensor. The integral value of five mask configurations were measured while performing a inspiratory vital capacity test (IVCT). Shown is the percent of change over the mask-press configuration which allows us to compare individual capacities side-by-side. The trend shows the integral value generally decreases as more leaks are introduced. Physiological differences are noted in the legend.

leak. During each experiment, participants performed a self-guided inspiratory vital capacity test (IVCT) where they exhale their lung capacity, hold their breath for 5-seconds, inhale their total lung capacity, then again hold their breath for 5-seconds. We hypothesize that we can use this relatively constant volume of air to quantify mask leakage (see section 4.4.2).

Exploratory Study Our first study was exploratory ($n = 3$) and was conducted using FaceBit. We instructed participants to perform IVCTs under three different mask-fit configurations. The configurations were **self-mask fit**: Donned using CDC guided user-seal check, **mask press**: cupping one's hand and pressing the mask against the face to form a stable seal, **one hole open**, and **three holes open**). We conducted three trials per participant, randomizing the order of mask configurations in each trial.

Results The pressure signal from FaceBit was collected wirelessly and analyzed in a Python notebook. We computed the integral of the IVCT pressure signal, which is proportional to total mask resistance. These values are presented in 15b. All trials showed a larger IVCT integral during the mask-press configuration than the other configurations. Seven of nine trials measured a larger IVCT integral with the self-fit mask than a mask with a single hole. The IVCT integral was smallest in the three-hole condition in all trials.

Follow-up Study Taking these results as promising, we conducted a larger study ($n = 10$) with a similar setup, except that pressure was measured with an external, differential pressure sensor connected via a length of tubing to the N95 mask. A differential sensor has the advantage of automatically subtracting atmospheric pressure from the measurement. Participants performed the same IVCT over five mask configurations (mask fit, mask press, one, two, and three holes). We conducted three trials per participant, randomizing the order of mask configurations in each trial. Between each trial, the participant doffed and re-donned the mask. Before the experiment, all participants were presented with a CDC video on donning and doffing an N95 face mask and performing a user seal check.

Results. Taking advantage of the increased participant count, we performed a statistical analysis of our results. First, we combined each participant's three trials into one by taking the mean of the IVCT integral of each mask configuration. Then, we performed a repeated-measures analysis of variance (ANOVA) test across the five mask configurations. It was found that the data violated the sphericity assumption, and so we applied the Greenhouse-Geisser correction. This test yielded a corrected p-value of $6.57e-5$, showing a significant difference between IVCT integrals between configurations. We therefore continued with a post-hoc paired T-test for each pair of mask configurations, which showed a significant difference between both the mask-press and mask-fit configurations with all other mask configurations. However, the difference between configurations with one, two, and three holes were not enough to show significance compared with each other. Our results are visualized in 16 and further detailed in Appendix C table 5) and figure 24.

Discussion of Results: The statistical analysis of our second trial suggests this is a valid method for detecting mask leakage. Still, with our sample size we could not determine the minimum leak size we can confidently detect, and additionally we require more testing to show leak detection feasibility with FaceBit alone. Nonetheless, these results are promising in that they show that pressure sensing used in conjunction with an easily-performed lung capacity exercise may be able to quantify mask-fit in real-time.

7.5 Scenario Based Free-living Study

To further test the robustness and performance of FaceBit, we conducted a small scale study where three participants wore N95 masks retrofitted with a FaceBit, and performed targeted, noisy, daily life activities. By placing participants in noisy situations we see concentrated usage of the device and find interesting confounding situations that can spark future work.

We made four changes for this experiment based on lessons learned during the in-lab study. First, we used a new respiration rate ground truth sensor (a Vernier Respiration Belt), as the NeuLog was not capable of maintaining accurate respiration rate measurements while mobile.

Second, instead of relying on the calculated heart rate measurement from the Polar as ground truth, we instead calculated heart rate from the raw EKG data of the Polar using the HeartPy Python library [72]. In our previous experiments, we found that the Polar calculated heart rates were potentially phase-shifted and filtering out heart rate values that we were measuring, behaving much like a moving average filter.

Third, building on the lessons learned in our in-lab evaluation, we modified the runtime BCG algorithm to more aggressively reject time windows where significant motion was occurring (changes detailed in 4.2.2). This would necessarily reduce the number of samples we might gather, since we might reject windows where it was possible to calculate a accurate heart rate, but we contend this tradeoff is worthwhile since the FaceBit could be worn for twelve hours or more in a day in a clinic, or while working at a desk due to indoor masking requirements, providing ample time to gather a large amount of heart rate measurements. The sampling window was also increased: in these scenarios, FaceBit attempts to measure heart rate once per minute with a sampling window of 30 seconds.

Finally, we also modified the respiratory rate algorithm to allow measurement of high respiratory rates (in the range of 30-60 BrPM) and better reject oscillations that may result from our filter (detailed in 4.3.2.

Otherwise, the setup for these experiments is similar to our first evaluation: a participant wears an N95 instrumented with Facebit, and several ground truth devices. The Vernier Respiration Belt is strapped around the chest and measures the forces associated with chest expansion during respiration, and a Polar H10 chest strap is worn under the clothes to capture EKG data from which we calculate heart rate. Once a participant is instrumented, a study coordinator accompanies the participant (but does not interfere in any way beyond observation and data collection) on the scenarios they conduct, and uses a phone to record the ground truth values, as well as capture GPS and speed data using the Strava app.

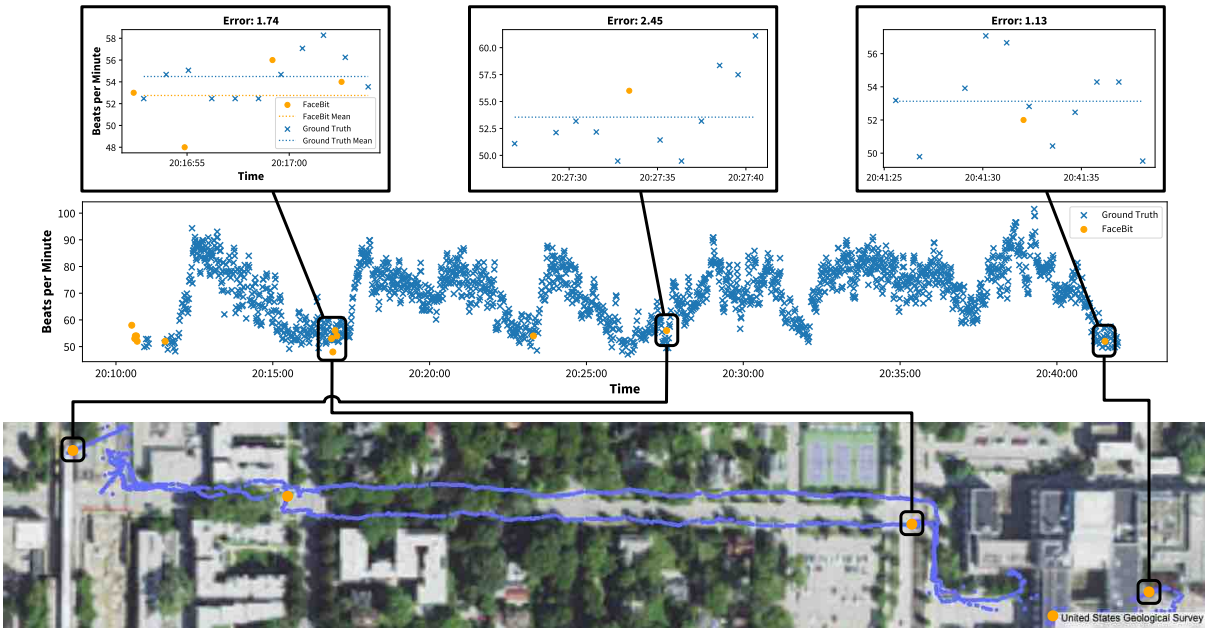


Fig. 17. The free-living scenario "walk to lunch". FaceBit is able to capture accurate heart rate at times of low-motion despite noisy conditions, and environmental factors like heat and humidity. **(middle)** shows annotations where heart rate measurements were gathered by FaceBit, plotted over the ground truth heart rate where each marker is the instantaneous heart rate associated with each beat. **(top)** Zoomed-in data-sections to highlight error. **(bottom)** GPS plot of the walking path with indicators of FaceBit heart rate recordings.

We conducted four different scenarios, ranging in length from 13 to 27 minutes. The scenarios were captured during a high heat warning (exceeding 90 degrees Fahrenheit/32 degrees Celsius), and high humidity (from 50-70%), further testing the robustness of the temperature based respiratory algorithm. The scenarios are outlined in brief:

(1) **Walk through town to pick up lunch.** Including multiple stops on busy cross streets, waiting in line to pick up food, entry/exit from air conditioning to hot outside, and sitting at table before eating (shown in Figure 17).
(2) **Ride on train.** Including waiting at the train stop, entry and exit, and the train ride itself with multiple stops.
(3) **Ride in car through busy town.** Including turns, stop lights and signs, and quick accelerations/decelerations.
(4) **Sit at desk and work.** We capture this scenarios as a baseline, where the participant just works and types at a desk normally.

Each scenario was performed once. Different participants were used for the "walk to lunch" and "work at desk" scenarios, while the same participant was used for the "ride on train" and "ride in car" scenarios.

Discussion of Results: Results from the "walk to lunch" scenario are distilled in figures 17 and 18. Similar graphs for the other scenarios are shown in appendix C, figure 25 The figures show that the respiration and heart rate data reported from FaceBit agree well with our ground truth data. We report the root mean square error (RMSE) for both metrics in each scenario in Table 1.

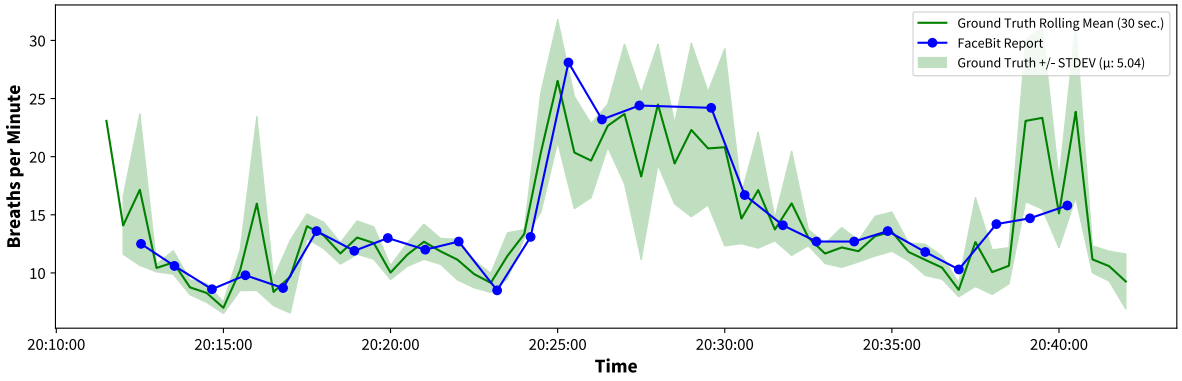


Fig. 18. FaceBit reported respiration rate plotted over the calculated rolling average of instantaneous respiration rates from the ground truth respiration belt. The data shown is from the "walking to lunch" scenario.

Table 1. Root mean square error for FaceBit predicted heart rate and respiration rate compared to ground truth for all in-wild scenarios.

Scenario	HR (RMSE)	RR (RMSE)	# HR Samp.	# RR Samp.	Avg. HR FB	Avg. RR FB	Duration
At computer	2.52	1.48	49	18	70.22	11.49	00:27:04
Lunch	1.55	2.97	14	26	53.50	14.29	00:27:43
Driving	1.40	3.28	13	19	64.15	18.93	00:20:08
Train	2.9	3.47	2	13	71.50	21.62	00:13:47
Overall	2.09	2.80	78	76	64.84	16.58	01:28:42

Overall, we find 2.09 beats per minute RMSE for FaceBit calculated heart rate, and a 2.80 breaths per minute RMSE for respiration, all in noisy scenarios. These results are very promising, demonstrating that FaceBit can be worn and used effectively in diverse, noisy situations (from heat, movement, and other diverse conditions that happen in an urban environment). We are particularly encouraged that our improvements to the FaceBeat algorithm resulted in no obvious heart rate outliers, even in these challenging conditions.

That said, we also see places for improvement, mainly, in balancing the need to throw out high motion periods for accuracy, and the need to gather heart rates when highly active. Indeed, on average FaceBit reported almost two heart rates per minute in the "at computer" condition, but only once every six minutes in the train condition. Figure 17 shows that heart rate data is usually gathered when heart rate is lowest, corresponding to periods of standing still (i.e. at crosswalks, at deli-counter) and at rest (i.e. sitting at table).

7.6 Physiological Differences and Their Effects on Signals

In this section we explore how physiological differences amongst humans may confound or make it harder to preserve signal fidelity for health metrics. We discuss the data we gathered from the participants in our studies, and observations we made. While making a definitive statement about generalizability against human differences is impossible for any wearable device because humans are so diverse, we can draw some interesting findings from our existing data and intuition on FaceBit's function.

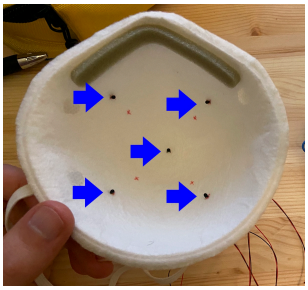


Fig. 19. N95 mask fit with thermistors to test temperature variation within the mask.

FaceBit is intended as a general purpose platform, and the particular metrics chosen as a demonstration of the platform (i.e., heart rate and respiration rate) are gathered from robust signals that originate from common physiological aspects of humans.

Respiration rate, especially, is presumed to be robust to physiological difference as it is gathered from temperature changes in the mask caused by breath. The respiration action in the mask causes replacement of warm air with cool air brought in from outside the mask, and this signal appears to be universal in all our tests, even in participants for whom the mask was too large (n=1), who had facial hair that interfered with mask seal (n=3), and for larger persons over six foot (n=1).

Differences in the cardiovascular system that arise from height, weight, health, athleticism, and biological sex may alter the dominant frequencies, strength, and overall pacing of the BCG signal. In practice we found our algorithm was able to gather data from every participant in both of our evaluations, albeit with a higher degree of accuracy after our algorithmic changes. We believe tailoring the algorithm to an individual by adjusting filter and threshold parameters may allow for more heart rate measurements and increased accuracy, and discuss a general strategy in section 8.2.

Our leak detection methodology depends on a consistent lung capacity, but it does not rely on the magnitude of the lung capacity itself. Lung capacity varies by individual, and that is why we score our IVCTs as a percentage of the press-fit condition, rather than the absolute difference.

Our studies had a small but heterogeneous participant group across gender, size, and facial features. The results suggest the device is able to gather signals from a diverse population.

7.7 Effect of Mask Type and Device Location on Signal Quality

We explore the effect of mask type (N95, Surgical) and FaceBit location in the mask on the signal fidelity of the respiration and heart rates. For the experiment, we isolate each variable for both metrics of interest. We collected data from a single male user with no facial hair, and in the 18-35 age range.

For respiration rate, since we use the change in temperature to calculate respiration, we seek to understand the variation of temperature within masks during normal breathing. Our aim is to establish whether device placement will have an effect on the respiration rate calculations. We instrument a surgical and an N95 mask with five high precision thermistors placed in a star shape at the center and extreme ends of the mask where a FaceBit could feasibly fit without making contact with the face. The placements are shown in a N95 mask in Figure 19, with surgical mask placements similar. We then breathe into the mask regularly for 1-2 minutes, and record the simultaneous output of the thermistors, which give a continuous temperature reading.

For both masks, the waveform is nearly exactly the same, with minor shifts in amplitude because of thermistor calibration offsets. After being put through the same bandpass filter as on the FaceBit, the differences shrink but

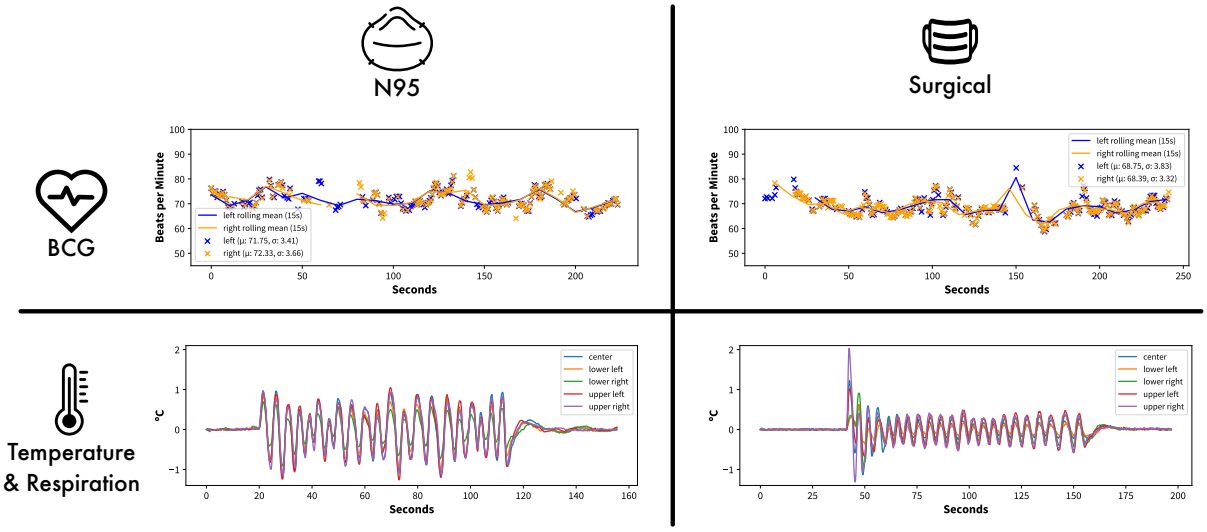


Fig. 20. On top are shown the BCG signals captured for heart rate from two FaceBit devices placed on opposite sides of the mask, reporting signals with a very similar frequency content (just phase shifted). On the bottom are shown the filtered temperature signals at five placements in the mask, captured simultaneously, showing near perfect correlation. The left side shows N95 masks, and the right shows Surgical masks. This demonstrates the robustness of the signals to device location and mask type.

do not entirely disappear. These filtered results are shown in Figure 20. Notably, the temperature fluctuations in the N95 mask are greater than those in the surgical mask (likely due to better trapping of air in the N95). The amplitudes of the breathing signal vary across mask locations, but the signal is clearly distinguishable throughout the mask in both cases.

Temperature change is intuitively a robust mechanism for respiration rate: the pocket of air inside the mask is small, and the volume of each breath is relatively large, therefore the volume of air inside the mask is replaced entirely with each inhalation and exhalation. If the air in the environment is cooler or (less likely) warmer than the exhaled air, these air replacements also change the temperature in the mask. Certain locations may receive a higher flow rate if the air preferentially takes a certain path out of the mask (as may be in the case of leaks in an N95 mask, or in the normal operation of a surgical mask), and therefore see higher temperature variation during breathing. In a surgical mask or small N95 mask, the sensors may be pressed against the skin in which case they receive less airflow, and therefore less temperature variation during breathing.

For heart rate, we conduct a similar experiment where we place two FaceBit devices inside each type of mask at opposite sides. In the N95 mask, the devices are placed as far from each other as possible without being pressed against the skin. In the surgical mask, they are inset 1 inch from the left and right sides. We don the mask and then breathe normally for 3-4 minutes, and record the heart rates reported by both devices via a JLink connection to a desktop computer.

The results are shown in Figure 20. In both mask types we obtained very similar heart rates from the two devices. Occasionally one device would lose the signal and stopped reporting heart rates while the other maintained it, however both devices resumed reporting in short order. We show a moving average trend placed over both scatter-plots, indicating a high degree of agreement between both sensors.

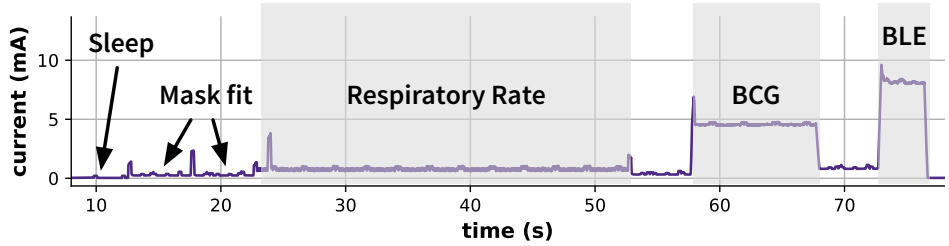


Fig. 21. Application power trace over time while calculating health metrics. Spikes are from current inrush due to hand resetting between tests with our test infrastructure.

This result makes sense because the ballistic force of the heart is carried through the N95 and surgical masks via their attachment to the face. Since the structures are tight to the face in both cases, and therefore rotate with the head, the rotational forces are seen throughout the masks. The BCG algorithm (and namely the combination of the gyroscopic signals across the three axes via the euclidean norm) is responsible for making the signal processing robust to device orientation.

From these experiments, we can see that the fidelity of either signal is not very sensitive to device location. This confirms the results of our two studies (in-lab and scenario based), where the devices were not carefully placed prior to data collection. In our in-lab study, the participants were asked to place the FaceBit device themselves "in the center of the mask". We did not modify or change the placement of the FaceBit, so eight users we captured data on for heart rate made 8 different placements. Similarly for our scenario-based study, the devices were not carefully placed in any specific location.

7.8 Benchmarking Power and Performance of FaceBit

The nRF52832 is capable of a deep-sleep current consumption of 1.9 μA, but various circuitry elements (e.g. pull-down resistors) bring the sleep current of this version of the FaceBit platform up to approximately 32 μA. See Table 2 for a summary of the power and energy requirements of the various states that comprise the FaceBit application, and Figure 21 for a power trace of the various tasks the application is capable of. For the purposes of arriving at a battery life estimate, we assume 8 hrs/day of active on-face measurement, a moderate measurement duty cycle of one respiratory rate and heart rate every five minutes, and one BLE broadcast every five minutes. Mask On/Off status is updated every two minutes. Our coin cell has a nominal capacity of 105 mWh (378 mJ), but we revise this to an estimated usable capacity of 75.6 µWh (272 mJ)[3]. In this scenario, on battery power alone our device is expected to last for approximately eleven days.

7.9 User-Burden Analysis

The FaceBit device is designed for use in many environments, but with special attention paid to clinical environments where health care workers are required to wear protective face masks throughout the workday. As part of the clinician study introduced in 3.2 a portion of the survey was devoted to the FaceBit prototype's comfort, serving as an initial user-burden analysis to drive an ongoing iterative design progress. We provided the clinicians with instructions to wear the FaceBit device with a 3D printed enclosure, and a stand-in SATURN harvester component. They then were instructed to perform a few simple exercises including adjusting the mask, moving

[3]We arrive at this number by taking into account our converter efficiency of ≈90%, and by applying an engineering factor of 80% to account for non-ideal conditions like temperature swings, self-leakage, and unused capacity below our cold-start limit

Table 2. Power and energy consumption of the FaceBit application runtime.

	Average Power [μW]	**Duration [s]**	**Total Energy [mJ]**
Idle/Sleep	48	-	-
Mask On/Off Detection	678	10	6.8
Respiratory Rate Measurement	1300.5	30	38.9
BCG Measurement	7320	10	73.1
BLE Broadcast (Typical)	12360	3.7	45.4

the head from side to side and up and down, talking, and deep breathing. Participants had two magnets to attach, one for the FaceBit device, and one for the harvester below the nose.

Results. All participants succeeded in attaching the device independently without the aid of the study coordinator. It took on average 97 seconds to follow the instructions and attach the FaceBit device to their mask. We then asked the participants to give feedback on the experience, including rating the comfort of wearing a face mask with and without FaceBit attached. The average reported difference of 11 participants on comfort before and after attaching FaceBit was -3.18 on a scale of 1 to 10. One response was removed due to an error in response, where they recorded an increase in comfort which was not reflected in their open-text response.

As the survey's secondary goal, we report on feedback from users to explore future design strategies. Three participants (25%) reported considering wearing FaceBit daily in its current state. Two participants reported no difference in face mask's comfort after attaching FaceBit. Additionally, 7 participants forgot to remove their FaceBit, and continuously wore it during our end-of-session discussion. We consider these initial findings to be promising for the development of a future consumer-ready product.

Discussion of Results: From user feedback, we derive the following themes to guide future design iterations of the FaceBit platform:

- **Material:** 3 clinicians, each, complained about the material of the stand-in harvester and mentioned discomfort in the material of the 3D printed case. The casing was somewhat abrasive due to inconsistency in the 3D printer filament, even after sanding. Two participants specifically mentioned skin irritation concerns where the device contacts the face. One participant suggested a cloth material, similar to that of the mask, for the sensor board's casing. We expected this feedback in part due to the 3D printed casing and in future design iterations a more permanent and comfortable material should be considered.
- **Size:** 5 participants reported noticing the size or felt discomfort with the weight of the sensor board and casing. We are confident that future productions of the sensor board (with fewer sensing modalities) can be more compact as we develop more refined features and confirm which sensing modalities are necessary.
- **Attachment:** 1 participant mentioned in our end-of-survey discussion that FaceBit adds to the already burdensome task of taking a mask on and off. Another pointed out the difficulty in adjusting the position of the components. Any future FaceBit designs should be easy to attach, flexible to fit on multiple masks, adapt to fit an individual's comfort, and be easy to store.

8 DISCUSSION AND FUTURE WORK

This work explores a vision of smart, low-burden, sustainable personal protective equipment. The experimental results of this paper suggest promising future research directions that will enhance the reliability, ease-of-use, and capabilities of face masks. We discuss our results and detail some of these future opportunities below.

8.1 Limitations

Respiratory Rate: Our respiratory rate evaluation did not attempt to measure the effect of confounding scenarios such as speech, laughter, or other such patterns of breathing. Moreover, our algorithm is not expected to compensate for these events. We suggest more advanced signal processing approaches below that may be able to account for these cases.

Ground truth: Our ground truth devices limited the sorts of activities under which our algorithms could be studied. The respiratory belt used in the in-lab evaluation was tethered to a computer via USB), and body motion introduced noise into the signal in both environments. We argue that due to its proximity to the signal, Facebit should enjoy a clearer picture of respiratory activity than almost any other (non-clinical) method of respiratory rate detection.

Generalizability: The focus of the paper is not on full and comprehensive evaluation of the signals captured in the wild. We show promise of generalizable predictive signals via scenario based evaluation in noisy situations, an in-lab study on a diverse cohort, and diverse testing with mask types and device placement. Further exploration is needed to tune the algorithms so they can more cleanly decide when to take a reading.

Energy Harvesting: Not all harvesting modalities are equal: a small solar panel (smaller than we tested) mounted on the outside of the mask would be more than sufficient to power all operation in outdoor environments. However, indoor mask wearing is more common for health applications, while outdoor applications around air quality and pollution avoidance would benefit from the solar panel power. We believe the breath-based harvesters, like the TENG, are the most promising modality despite the TENGs low power output in our tests. The form factor, reliance on an energy source that is periodic and always present, and the small size and easy application on top of the FaceBit, make for a compelling future direction. Finally, using multiple harvesters could result in more comprehensive coverage. For example, TEG or TENG harvesters can complement solar harvesters when light is absent. Overall, the modular energy harvesting input can support any number of harvesters, and multiple are quite feasible for daily wear. Further exploration of harvesting as well as power-failure resilient operation we leave to future work.

8.2 Future Directions

Optimizing Signal Pipeline. Between our in-lab and scenario-based evaluations we demonstrated an ability to tune the FaceBeat algorithm to reduce the number of falsely detected heart rates. Based on the relatively low reporting frequency during our active scenarios, however, it seems that this may have come at the cost of increasing our false negative rate. In other words, it is likely that the algorithm is more conservative than necessary and it did not report some heart rates that were valid. We believe fine-tuning the algorithm's parameters for the general population could be accomplished with a larger study. As previously mentioned, however, we suspect that superior results could be achieved by personalizing the algorithm to each individual. Using the FaceBit mobile application and a phone's camera, heart rate ground truth can be gathered via PPG. This signal would allow us to determine the normal range of standard deviations of the users's inter-beat interval. Two-way communication with FaceBit could help to set the filter passbands that best isolate the individual's BCG pulse complex. Finally, using the application to instruct the user to perform certain activities (e.g. sitting at rest, standing) would allow us to estimate the proper BCG amplitude thresholds for an individual.

Our respiratory rate algorithm can be confounded by activities such as speech or coughing. Our algorithm is already capable of slicing the signal into individual breaths. By analyzing these breaths individually (e.g. calculating length of inhale and exhale, or ratio of the inhale/exhale volume) we may be able to accurately classify

different types of respiratory events. Fusing the data we can gather from the temperature, pressure, and audio sensors (e.g. to detect speech) is also promising.

Mask Fit: Promising Pathways. One method to improve breathing exercise consistency is to incorporate an interface into the companion application that guides a user through a mask fit test similar to the test performed in our evaluation. A guided exercise could assist in making the test more repeatable, especially with real-time feedback from FaceBit's sensors.

A challenge inherent to our simple magnetic clip-on design is calculating the relative pressure difference between the inside and the outside of the mask. Our in-mask leak detection algorithm relies on the user to hold their breath to allow the pressure inside the mask to equilibrate with atmospheric pressure. With a different design that penetrates through the filter material, we could use a differential pressure sensor rather than an absolute pressure sensor, which would simplify both the signal processing pipeline and the breathing exercise. Naturally, it has the downside of requiring a modification to the mask.

Another perhaps promising option is to instead focus analysis on the *step response* of the pressure signal. This could be accomplished by, for instance, asking the wearer to shut their mouth in the middle of an inhale and monitoring the relaxation back to ambient pressure. Similar to an electrical system, it is expected that mask compliance may act as a small capacitive element in parallel with the resistance of the filter material and (importantly) any leaks, thus forming a "low-pass filter" for pressure changes. Therefore, there is expected to a be a time constant associated with any step input to the system, whose length would vary in relation to the total leakage of the mask. An advantage of this approach is that the time constant does not depend on the initial level of pressure in the mask, and so the consistency of the user is taken out of the equation. We tested this approach with our platform, but found that the 75 Hz maximum frequency of our pressure sensor was not fast enough to capture the relaxation curve. Mask design plays a role here: more compliant filter materials will result in a slower (easier to capture) step response.

Lung Health. The FaceBit platform has the potential to be a powerful respiratory health tool. Masks that seal and present any resistance to airflow amplify the pressure differences created by the lungs, diaphragm, and other breathing muscles (much like a shunt resistor does for current monitoring applications). While many find masks to be mildly uncomfortable, it's undeniable that a platform like FaceBit is preferable to spirometer-like devices for anything other than limited, sporadic monitoring. Indeed, the relative comfort of the FaceBit platform and the now widespread acceptability of face masks in everyday settings could allow researchers or interested individuals to investigate in great detail breathing patterns and tidal volume throughout the day, the frequency of coughs and other respiratory symptoms, and even subtle differences in inspiration and expiration that could indicate and monitor restrictive or obstructive lung disorders. The accessory port allows for the addition of other sensors that could be used, for instance, to monitor the gaseous composition of exhaled air (an area of active research [5, 61]).

Combining Health Metrics. The device's capability to monitor both heart rate and respiratory rate is powerful, since these two systems are intricately linked and the ways in which one rate rises and falls in relation to the other may prove to be an interesting physiological parameter. The "Pulse-Respiration Quotient" (HR/RR) is a metric that has been difficult to monitor in everyday settings, but has been shown in limited studies to change with circadian rhythm, age, and in disease [63].

Battery-free Operation. We chose to use a commodity Bluetooth module for ease of community engagement and support. Switching to a low power module (i.e. an Ambiq Apollo3, with 10x lower power draw) would allow for completely battery-free operation and unlimited lifetime, compared to our current 11 day life.

9 CONCLUSIONS

It seems increasingly likely that worldwide face mask usage is not going away. Even before the COVID-19 pandemic, health care workers regularly used PPE, and many countries' populations are regular users of face masks or other forms of personal protective equipment to deal with air quality or prevent viral infections. With tens of billions of face masks manufactured a month, poor air quality and respiratory infections on the rise, we have an unprecedented need for intelligent and active PPE. This paper presents the FaceBit platform as both a novel wearable health device and a new platform enabling a community of diverse researchers in the emerging space of smart face masks. FaceBit balances a large design space, offering rich sensing with a small form factor and long battery lifetime. FaceBit attempts to make the vision of smart PPE more sustainable by taking advantage of energy harvested from the environment and incorporating energy-efficient operation. With innovations in low power design, signal processing for health metrics, energy harvesting, and form factors, the platform offers a foundation for future health research and a jumping off point for new design of smart face masks. While the current COVID-19 crisis is the primary motivation for this paper, the contributions to building smart personal protective equipment go beyond the pandemic, and provide new ways to think about how we protect vulnerable populations. Due to the low user burden of the platform, we believe FaceBit will catalyze mobile health sensing despite the pandemic and assist with long term, high adherence studies within ubiquitous and mobile computing, health and human behavior, and novel on-body sensing. To support these aims we have made the software, hardware, and documentation freely available to the research community at facebit.health.

ACKNOWLEDGMENTS

The authors would like to thank the anonymous associate editor for helpful guiding comments during the revision phase, Thomas Cohen and Alex Cindric for illustrations and visual design for figures, and Andrea Maioli for graciously helping us test FaceBit on many occasions. This research is based upon work supported by the National Science Foundation under award number CNS-2032408, as well as CNS-1850496, CNS-2038853, CNS-2030251, and CNS-2107400. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

[1] [n.d.]. BLE power consumption using CORDIO stack too high · Issue #10669 · ARMmbed/mbed-os. https://github.com/ARMmbed/mbed-os/issues/10669

[2] [n.d.]. Making masks smarter and safer against COVID-19. https://jacobsschool.ucsd.edu/news/release/3206. Accessed: 2021-02-13.

[3] Joshua Adkins, Branden Ghena, Neal Jackson, Pat Pannuto, Samuel Rohrer, Bradford Campbell, and Prabal Dutta. 2018. The signpost platform for city-scale sensing. In *2018 17th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*. IEEE, 188–199.

[4] Rawan Alharbi, Mariam Tolba, Lucia C Petito, Josiah Hester, and Nabil Alshurafa. 2019. To mask or not to mask? balancing privacy with visual confirmation utility in activity-oriented wearable cameras. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 3, 3 (2019), 1–29.

[5] Reef Einoch Amor, Morad K Nakhleh, Orna Barash, and Hossam Haick. 2019. Breath analysis of cancer in the present and the future. *European Respiratory Review* 28, 152 (2019).

[6] Michael P Andersen, Hyung-Sin Kim, and David E Culler. 2017. Hamilton: a cost-effective, low power networked sensor for indoor environment monitoring. In *Proceedings of the 4th ACM International Conference on Systems for Energy-Efficient Built Environments*. 1–2.

[7] Nivedita Arora, Steven L Zhang, Fereshteh Shahmiri, Diego Osorio, Yi-Cheng Wang, Mohit Gupta, Zhengjun Wang, Thad Starner, Zhong Lin Wang, and Gregory D Abowd. 2018. SATURN: A thin and flexible self-powered microphone leveraging triboelectric nanogenerator. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 2 (2018), 1–28.

[8] Abdelkareem Bedri, Diana Li, Rushil Khurana, Kunal Bhuwalka, and Mayank Goel. 2020. Fitbyte: Automatic diet monitoring in unconstrained situations using multimodal sensing on eyeglasses. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–12.

[9] Giorgio Biagetti, Paolo Crippa, Laura Falaschetti, Simone Orcioni, and Claudio Turchetti. 2018. Human activity monitoring system based on wearable sEMG and accelerometer wireless sensor nodes. *Biomedical engineering online* 17, 1 (2018), 1–18.

[10] Nam Bui, Nhat Pham, Jessica Jacqueline Barnitz, Zhanan Zou, Phuc Nguyen, Hoang Truong, Taeho Kim, Nicholas Farrow, Anh Nguyen, Jianliang Xiao, Robin Deterding, Thang Dinh, and Tam Vu. 2019. EBP: A Wearable System For Frequent and Comfortable Blood Pressure Monitoring From User's Ear. In *The 25th Annual International Conference on Mobile Computing and Networking*. Association for Computing Machinery, New York, NY, USA, Article 53, 17 pages. https://doi.org/10.1145/3300061.3345454

[11] Andrew Carek and Christian Holz. 2018. Naptics: Convenient and continuous blood pressure monitoring during sleep. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 3 (2018), 1–22.

[12] Andrew M Carek, Jordan Conant, Anirudh Joshi, Hyolim Kang, and Omer T Inan. 2017. SeismoWatch: wearable cuffless blood pressure monitoring using pulse transit time. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 1, 3 (2017), 1–16.

[13] Giovanna E Carpagnano, Maria P Foschino-Barbaro, Corrado Crocetta, Donato Lacedonia, Valerio Saliani, Luigi Davide Zoppo, and Peter J Barnes. 2017. Validation of the exhaled breath temperature measure: Reference values in healthy subjects. *Chest* 151, 4 (2017), 855–860.

[14] Robert G. Carroll. 2007. 9 - Integrated Cardiovascular Function. In *Elsevier's Integrated Physiology*, Robert G. Carroll (Ed.). Mosby, Philadelphia, 91–98. https://doi.org/10.1016/B978-0-323-04318-2.50015-7

[15] Sejin Choi, Ryeol Park, Nahmkeon Hur, and Wonjung Kim. 2020. Evaluation of wearing comfort of dust masks. *Plos one* 15, 8 (2020), e0237848.

[16] CNBC. 2021. Global Covid cases on track to add 100 million by early next year without help to poorer nations, WHO says. https://www.cnbc.com/2021/08/11/who-says-the-world-is-on-track-to-surpass-300-million-covid-cases-early-next-year.html. Accessed: 2021-08-01.

[17] Marie Therese Cooney, Erkki Vartiainen, Tinna Laakitainen, Anne Juolevi, Alexandra Dudina, and Ian M Graham. 2010. Elevated resting heart rate is an independent risk factor for cardiovascular disease in healthy men and women. *American heart journal* 159, 4 (2010), 612–619.

[18] Jasper De Winkel, Vito Kortbeek, Josiah Hester, and Przemysław Pawełczak. 2020. Battery-free game boy. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 3 (2020), 1–34.

[19] Scott Duncan, Tom Stewart, Lisa Mackay, Jono Neville, Anantha Narayanan, Caroline Walker, Sarah Berry, and Susan Morton. 2018. Wear-time compliance with a dual-accelerometer system for capturing 24-h behavioural profiles in children and adults. *International journal of environmental research and public health* 15, 7 (2018), 1296.

[20] Prabal Dutta, Jonathan Hui, Jaein Jeong, Sukun Kim, Cory Sharp, Jay Taneja, Gilman Tolle, Kamin Whitehouse, and David Culler. 2006. Trio: enabling sustainable and scalable outdoor wireless sensor network deployments. In *2006 5th International Conference on Information Processing in Sensor Networks*. IEEE, 407–415.

[21] Begum Egilmez, Emirhan Poyraz, Wenting Zhou, Gokhan Memik, Peter Dinda, and Nabil Alshurafa. 2017. UStress: Understanding college student subjective stress using wrist-based passive sensing. In *2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*. IEEE, 673–678.

[22] Emre Ertin, Nathan Stohs, Santosh Kumar, Andrew Raij, Mustafa Al'Absi, and Siddharth Shah. 2011. AutoSense: unobtrusively wearable sensor suite for inferring the onset, causality, and consequences of stress in the field. In *Proceedings of the 9th ACM conference on embedded networked sensor systems*. 274–287.

[23] U.S. Centers for Diseases Control and Prevention. 2021. Interim Public Health Recommendations for Fully Vaccinated People. https://www.cdc.gov/coronavirus/2019-ncov/vaccines/fully-vaccinated-guidance.html. Accessed: 2021-08-01.

[24] MDP Garcia-Souto and P Dabnichki. 2018. Non-invasive and wearable early fever detection system for young children. *Measurement* 116 (2018), 216–229.

[25] Laurent Giovangrandi, Omer T Inan, Richard M Wiard, Mozziyar Etemadi, and Gregory TA Kovacs. 2011. Ballistocardiography—a method worth revisiting. In *2011 annual international conference of the IEEE engineering in medicine and biology society*. IEEE, 4279–4282.

[26] Jan Gralton and Mary-Louise McLaws. 2010. Protecting healthcare workers from pandemic influenza: N95 or surgical masks? *Critical care medicine* 38, 2 (2010), 657–667.

[27] Lin Gu and John A Stankovic. 2006. t-kernel: Providing reliable OS support to wireless sensor networks. In *Proceedings of the 4th international conference on Embedded networked sensor systems*. 1–14.

[28] Chih-Chieh Han, Ram Kumar, Roy Shea, Eddie Kohler, and Mani Srivastava. 2005. A dynamic operating system for sensor nodes. In *Proceedings of the 3rd international conference on Mobile systems, applications, and services*. 163–176.

[29] Yandell Henderson. 1905. The mass-movements of the circulation as shown by a recoil curve. *American Journal of Physiology-Legacy Content* 14, 3 (1905), 287–298.

[30] Javier Hernandez, Yin Li, James M Rehg, and Rosalind W Picard. 2014. Bioglass: Physiological parameter estimation using a head-mounted wearable device. In *2014 4th International Conference on Wireless Mobile Communication and Healthcare-Transforming Healthcare Through Innovations in Mobile and Wireless Technologies (MOBIHEALTH)*. IEEE, 55–58.

[31] Javier Hernandez, Daniel McDuff, Karen Quigley, Pattie Maes, and Rosalind W Picard. 2018. Wearable motion-based heart rate at rest: A workplace evaluation. *IEEE journal of biomedical and health informatics* 23, 5 (2018), 1920–1927.

[32] Josiah Hester, Travis Peters, Tianlong Yun, Ronald Peterson, Joseph Skinner, Bhargav Golla, Kevin Storer, Steven Hearndon, Kevin Freeman, Sarah Lord, Ryan Halter, David Kotz, and Jacob Sorber. 2016. Amulet: An Energy-Efficient, Multi-Application Wearable Platform. In *Proceedings of the 14th ACM Conference on Embedded Network Sensor Systems (SenSys '16)*. ACM, New York, NY, USA, 216–229. https://doi.org/10.1145/2994551.2994554

[33] Steve Hodges, Lyndsay Williams, Emma Berry, Shahram Izadi, James Srinivasan, Alex Butler, Gavin Smyth, Narinder Kapur, and Ken Wood. 2006. SenseCam: A retrospective memory aid. In *International Conference on Ubiquitous Computing*. Springer, 177–193.

[34] Christian Holz and Edward J Wang. 2017. Glabella: Continuously sensing blood pressure behavior using an unobtrusive wearable device. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 1–23.

[35] Neal Jackson, Joshua Adkins, and Prabal Dutta. 2019. Capacity over capacitance for reliable energy harvesting sensors. In *Proceedings of the 18th International Conference on Information Processing in Sensor Networks*. 193–204.

[36] Xiaofan Jiang, Joseph Polastre, and David Culler. 2005. Perpetual environmentally powered sensor networks. In *IPSN 2005. Fourth International Symposium on Information Processing in Sensor Networks, 2005.* IEEE, 463–468.

[37] Prad Kadambi, Abinash Mohanty, Hao Ren, Jaclyn Smith, Kevin McGuinnes, Kimberly Holt, Armin Furtwaengler, Roberto Slepetys, Zheng Yang, Jae-sun Seo, et al. 2018. Towards a wearable cough detector based on neural networks. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2161–2165.

[38] Mustafa Emre Karagozler, Ivan Poupyrev, Gary K. Fedder, and Yuri Suzuki. 2013. Paper Generators: Harvesting Energy from Touching, Rubbing and Sliding. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology (UIST '13)*. Association for Computing Machinery, New York, NY, USA, 23–30. https://doi.org/10.1145/2501988.2502054

[39] Ali Kiaghadi, Seyedeh Zohreh Homayounfar, Jeremy Gummeson, Trisha Andrew, and Deepak Ganesan. 2019. Phyjama: physiological sensing via fiber-enhanced pyjamas. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (2019), 1–29.

[40] Chang-Sei Kim, Stephanie L Ober, M Sean McMurtry, Barry A Finegan, Omer T Inan, Ramakrishna Mukkamala, and Jin-Oh Hahn. 2016. Ballistocardiogram: Mechanism and potential for unobtrusive cardiovascular health monitoring. *Scientific reports* 6, 1 (2016), 1–6.

[41] Jayoung Kim, Somayeh Imani, William R de Araujo, Julian Warchall, Gabriela Valdés-Ramírez, Thiago RLC Paixão, Patrick P Mercier, and Joseph Wang. 2015. Wearable salivary uric acid mouthguard biosensor with integrated wireless electronics. *Biosensors and Bioelectronics* 74 (2015), 1061–1068.

[42] Juha M Kortelainen, Martin O Mendez, Anna Maria Bianchi, Matteo Matteucci, and Sergio Cerutti. 2010. Sleep staging based on signals acquired through bed sensor. *IEEE Transactions on Information Technology in Biomedicine* 14, 3 (2010), 776–785.

[43] Gierad Laput, Karan Ahuja, Mayank Goel, and Chris Harrison. 2018. Ubicoustics: Plug-and-play acoustic activity recognition. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*. 213–224.

[44] Ping-Ing Lee, Ya-Li Hu, Po-Yen Chen, Yhu-Chering Huang, and Po-Ren Hsueh. 2020. Are children less susceptible to COVID-19? *Journal of Microbiology, Immunology, and Infection* 53, 3 (2020), 371.

[45] Phil Levis and David Culler. 2002. Mate: a virtual machine for tiny networked sensors. In *Proceedings of the ACM Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS)*, Vol. 10.

[46] Philip Levis, Samuel Madden, Joseph Polastre, Robert Szewczyk, Kamin Whitehouse, Alec Woo, David Gay, Jason Hill, Matt Welsh, Eric Brewer, et al. 2005. TinyOS: An operating system for sensor networks. In *Ambient intelligence*. Springer, 115–148.

[47] Ping Li, Ramy Meziane, Martin J-D Otis, Hassan Ezzaidi, and Philippe Cardou. 2014. A Smart Safety Helmet using IMU and EEG sensors for worker fatigue detection. In *2014 IEEE International Symposium on Robotic and Sensors Environments (ROSE) Proceedings*. IEEE, 55–60.

[48] Konrad Lorincz, Bor-rong Chen, Geoffrey Werner Challen, Atanu Roy Chowdhury, Shyamal Patel, Paolo Bonato, Matt Welsh, et al. 2009. Mercury: a wearable sensor network platform for high-fidelity motion analysis.. In *SenSys*, Vol. 9. 183–196.

[49] JJM Lut and M Goddijn. 2020. Smart Personal Protective Equipment: Sensing and Control: The future of face masks. (2020).

[50] Dong Ma, Guohao Lan, Mahbub Hassan, Wen Hu, and Sajal K Das. 2019. Sensing, computing, and communications for energy harvesting IoTs: A survey. *IEEE Communications Surveys & Tutorials* 22, 2 (2019), 1222–1250.

[51] Takuya Maekawa, Yutaka Yanagisawa, Yasue Kishino, Katsuhiko Ishiguro, Koji Kamei, Yasushi Sakurai, and Takeshi Okadome. 2010. Object-based activity recognition with heterogeneous sensors on wrist. In *International Conference on Pervasive Computing*. Springer, 246–264.

[52] Inc. Marlow Industries. 2020. Thermoelectric generator TG12-6-01S, 40.13×40.13×3.91mm. https://cdn2.hubspot.net/hubfs/547732/Data_Sheets/TG12-6.pdf. Last accessed: Feb. 11, 2021.

[53] Andrei Nikulin, Dmitry Ikonnikov, and Iliya Dolzhikov. 2019. Smart Personal Protective Equipment in the Coal Mining Industry. *International Journal of Civil Engineering and Technology* 10, 4 (2019), 852–863.

[54] Rita Paradiso, Giannicola Loriga, and Nicola Taccini. 2005. A wearable health care system based on knitted integrated sensors. *IEEE transactions on Information Technology in biomedicine* 9, 3 (2005), 337–344.

[55] Shyamal Patel, Hyung Park, Paolo Bonato, Leighton Chan, and Mary Rodgers. 2012. A review of wearable sensors and systems with application in rehabilitation. *Journal of neuroengineering and rehabilitation* 9, 1 (2012), 1–17.

[56] Washington Post. 2021. 'Act now' on global vaccines to stop more-dangerous variants, experts warn Biden. https://www.washingtonpost.com/health/2021/08/10/health-experts-demand-global-vaccines-pandemic/. Accessed: 2021-08-01.
[57] Tyler R Ray, Jungil Choi, Amay J Bandodkar, Siddharth Krishnan, Philipp Gutruf, Limei Tian, Roozbeh Ghaffari, and John A Rogers. 2019. Bio-integrated wearable systems: A comprehensive review. *Chemical reviews* 119, 8 (2019), 5461–5533.
[58] Niels Reijers and Chi-Sheng Shih. 2018. CapeVM: A Safe and Fast Virtual Machine for Resource-Constrained Internet-of-Things Devices. In *Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems (SenSys '18)*. Association for Computing Machinery, New York, NY, USA, 250–263. https://doi.org/10.1145/3274783.3274842
[59] Samy Rengasamy, Benjamin C Eimer, and Jonathan Szalajda. 2014. A quantitative assessment of the total inward leakage of NaCl aerosol representing submicron-size bioaerosol through N95 filtering facepiece respirators and surgical masks. *Journal of occupational and environmental hygiene* 11, 6 (2014), 388–396.
[60] Occupational Safety and Health Administration. 2019. 1910.134 App A - Fit Testing Procedures (Mandatory). Personal Protective Equipment. Standard Number: 1910.134 App A.
[61] Tarik Saidi, Omar Zaim, Mohammed Moufid, Nezha El Bari, Radu Ionescu, and Benachir Bouchikhi. 2018. Exhaled breath analysis using electronic nose and gas chromatography–mass spectrometry for non-invasive diagnosis of chronic kidney disease, diabetes mellitus and healthy subjects. *Sensors and Actuators B: Chemical* 257 (2018), 178–188.
[62] Sergio Márquez Sánchez, Roberto Casado Vara, Francisco Javier García Criado, Sara Rodríguez González, Javier Prieto Tejedor, and Juan Manuel Corchado. 2019. Smart PPE and CPE Platform for Electric Industry Workforce. In *International Workshop on Soft Computing Models in Industrial and Environmental Applications*. Springer, 422–431.
[63] Felix Scholkmann and Ursula Wolf. 2019. The pulse-respiration quotient: A powerful but untapped parameter for modern studies about human physiology and pathophysiology. *Frontiers in physiology* 10 (2019), 371.
[64] Juliane R Sempionatto, Tatsuo Nakagawa, Adriana Pavinatto, Samantha T Mensah, Somayeh Imani, Patrick Mercier, and Joseph Wang. 2017. Eyeglasses based wireless electrolyte and metabolite sensor platform. *Lab on a Chip* 17, 10 (2017), 1834–1842.
[65] Dhruv R Seshadri, Ryan T Li, James E Voos, James R Rowbottom, Celeste M Alfes, Christian A Zorman, and Colin K Drummond. 2019. Wearable sensors for monitoring the physiological and biochemical profile of the athlete. *NPJ digital medicine* 2, 1 (2019), 1–16.
[66] Lukas Sigrist, Andres Gomez, Roman Lim, Stefan Lippuner, Matthias Leubin, and Lothar Thiele. 2017. Measurement and validation of energy harvesting IoT devices. In *Design, Automation & Test in Europe Conference & Exhibition (DATE), 2017*. IEEE, 1159–1164.
[67] PowerFilm Solar. 2021. MP3-37 3V 150mW 4.49×1.44in Flexible Solar Panel. https://www.powerfilmsolar.com/media/cms/Electronic_Component_Spec_Sheet_Cla_77DEA84523C82.pdf. Last accessed: Feb. 11, 2021.
[68] Donna Spruijt-Metz and Wendy Nilsen. 2014. Dynamic models of behavior for just-in-time adaptive interventions. *IEEE Pervasive Computing* 13, 3 (2014), 13–17.
[69] Isaac Starr, AJ Rawson, HA Schroeder, and NR Joseph. 1939. Studies on the estimation of cardiac ouptut in man, and of abnormalities in cardiac function, from the heart's recoil and the blood's impacts; the ballistocardiogram. *American Journal of Physiology-Legacy Content* 127, 1 (1939), 1–28.
[70] Vamsi Talla, Bryce Kellogg, Shyamnath Gollakota, and Joshua R Smith. 2017. Battery-free cellphone. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 2 (2017), 1–20.
[71] Peter Tseng, Bradley Napier, Logan Garbarini, David L Kaplan, and Fiorenzo G Omenetto. 2018. Functional, RF-trilayer sensors for tooth-mounted, wireless monitoring of the oral cavity and food consumption. *Advanced Materials* 30, 18 (2018), 1703257.
[72] Paul van Gent, Haneen Farah, Nicole Nes, and B. Arem. 2018. Heart Rate Analysis for Human Factors: Development and Validation of an Open Source Toolkit for Noisy Naturalistic Heart Rate Data. In *Proceedings of the 6th HUMANIST Conference*. 173–178.
[73] Emily Vardell. 2012. Chemical hazards emergency medical management (CHEMM). *Medical reference services quarterly* 31, 1 (2012), 73–83.
[74] Nicolas Villar and Steve Hodges. 2010. The Peppermill: A human-powered user interface device. In *Proceedings of the fourth international conference on Tangible, embedded, and embodied interaction*. 29–32.
[75] Edward Jay Wang, William Li, Doug Hawkins, Terry Gernsheimer, Colette Norby-Slycord, and Shwetak N Patel. 2016. HemaApp: noninvasive blood screening of hemoglobin using smartphone cameras. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 593–604.
[76] Edward Jay Wang, Junyi Zhu, Mohit Jain, Tien-Jui Lee, Elliot Saba, Lama Nachman, and Shwetak N Patel. 2018. Seismo: Blood pressure monitoring using built-in smartphone accelerometer and camera. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–9.
[77] Rui Wang, Fanglin Chen, Zhenyu Chen, Tianxing Li, Gabriella Harari, Stefanie Tignor, Xia Zhou, Dror Ben-Zeev, and Andrew T Campbell. 2014. StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing*. 3–14.
[78] Zhong Lin Wang. 2015. Triboelectric nanogenerators as new energy technology and self-powered sensors–Principles, problems and perspectives. *Faraday discussions* 176 (2015), 447–458.

[79] Koji Yatani and Khai N Truong. 2012. Bodyscope: a wearable acoustic sensor for activity recognition. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. 341–350.

[80] Shibo Zhang, Yuqi Zhao, Dzung Tri Nguyen, Runsheng Xu, Sougata Sen, Josiah Hester, and Nabil Alshurafa. 2020. Necksense: A multi-sensor necklace for detecting eating activities in free-living conditions. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 2 (2020), 1–26.

[81] Yun C Zhang, Shibo Zhang, Miao Liu, Elyse Daly, Samuel Battalio, Santosh Kumar, Bonnie Spring, James M Rehg, and Nabil Alshurafa. 2020. SyncWISE: Window Induced Shift Estimation for Synchronization of Video and Accelerometry from Wearable Sensors. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 3 (2020), 1–26.

A APPENDIX: USER STUDY QUESTIONNAIRE AND RESPONSE

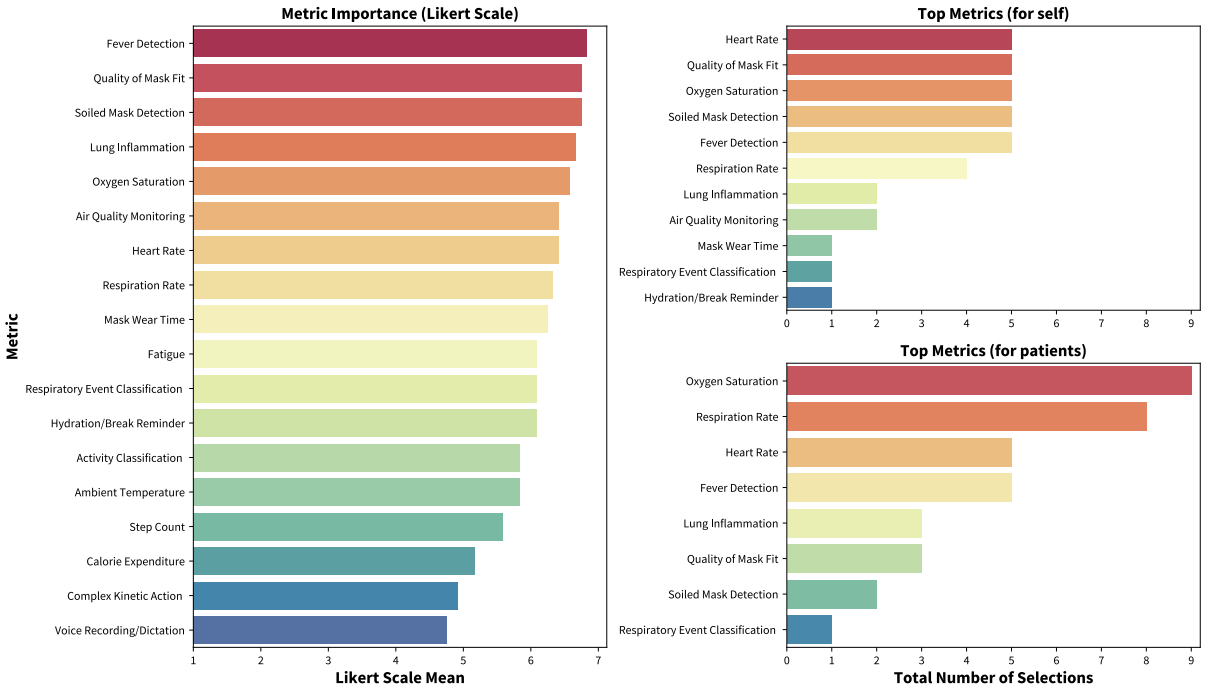


Fig. 22. Right to Left: (1) Ranking metric importance based from 1 (not important at all) to 7 (extremely important) Likert Scale. The bar indicates the mean of all 12 participants with standard deviation indicated by the line on each bar. (2, 3) We asked participants to select 3 metrics that they would like to see for their (personal) use or reporting from patients use. Metrics with zero counts are not shown. The graphic displays total counts across the 12 participants of the survey.

Questionnaire/Interview Procedure We conducted the study in 2-sessions onsite a local office. The Questionnaire consisted of 3 parts with interactions with the interviewers at the beginning, to induce the survey, middle, to give an overview of the FaceBit platform, and end to discuss and conclude. While the participant was working on the questionnaire, the interviewer(s) left the room until a section was completed.

Results of Target Metrics Section After conducting the study in the clinic to determine the most critical measures, we found it challenging to identify clear winners, without identifying the subject of the measure (i.e., the clinician or patient). When asked which metrics were most critical to the patient (scores >=5), responses show the following to be most critical: oxygen saturation, respiration rate, heart-rate, and fever detection. Clinicians on

Table 3. Energy harvested from different activities.

Harvester	Activity	Energy (μJ/min) Surgical Mask	N95 Mask
Shaker	Running	1300	800
	Walking	1.5	2.3
Thermoelectric*	Nasal Breathing	75	194
	Mouth Breathing	288	257
	Talking	415	271
Solar	Walking outdoor–sunny	432431	433191
	Walking outdoor–cloudy	74667	75203
	Sitting indoor–w/o light	422	415
	Sitting indoor–w/ light	6449	63987
	Walking indoor	1015	1008
SATURN	Breathing	0.2	0.1
	Coughing	2.3	4.8
	Talking	3.8	1.5

Note: * Energy harvested is for first minute.

the other hand were also interested in measures of heart-rate and fever detection, however were also interested in knowing their quality of mask fit, oxygen saturation levels and whether the mask was soiled (needed replacement) or not.

Mask Wearing. Before the pandemic, 3 out of 12 (25%) participants noted occasionally wearing a face mask, usually while seeing patients suspected of a viral infection or when the clinician was ill. 4 participants reported wearing N95 masks at some point during the COVID-19 pandemic. However, all participants wore surgical or cloth masks during the study and rarely reported requiring an N95. Overall, we found that all clinicians rely on face masks as their primary personal protective equipment during the COVID-19 pandemic. We should note that given the apparently lower attack rate of SARS-CoV-2 amongst children [44], PPE requirements for a pediatric setting may differ from that of other medical settings, and the fact that the COVID-19 pandemic caused a national shortage of PPE equipment may also have played a part in the limited use of N95 masks in this setting.

Expanded Metrics of Interest.

- **Mask Information and Safety**: The first and primary goal of a face mask is protection. Therefore, metrics related to the protective ability of a mask are sensible.
- **Personal Health**: A face mask's proximity to the respiratory system provides rich, untapped sensing opportunities for personal health monitoring. We can leverage this to provide real-time monitoring of physiological signals. Measuring these values over time can allow us to infer second-order metrics such as fatigue, and when to suggest a break or hydration.
- **Environmental Factors**: It is also possible to use wearable sensors to monitor environmental factors related to health. The proximity to the respiratory system makes face masks an ideal location to study air quality and composition.

Table 4. Theoretical measurements per day and minutes of sleep per day possible with various energy harvesters and activities. Assumes all harvested energy can be stored or utilized as it is harvested. Thermoelectric statistics assume 10 mask dons per day, and that energy is harvested principally for the first minute after donning a mask.

Harvester	Activity	Measurements/Day			Minutes Sleep/Day
		Mask On-Off	BCG	RR	
Shaker	Running	274.84	25.62	48.15	632.60
	Walking	0.32	0.03	0.06	0.73
Thermoelectric	Nasal Breathing	0.11	0.01	0.02	0.25
	Mouth Breathing	0.42	0.04	0.07	0.97
	Talking	0.61	0.06	0.11	1.40
Solar	Walking Outdoor-Sunny	91,423.04	8,520.81	16,015.96	210,428.71
	Walking Outdoor-Cloudy	15,785.84	1,471.27	2,765.44	36,334.31
	Sitting Indoor-w/o Light	89.22	8.32	15.63	205.35
	Sitting Indoor-w/ Light	1,363.42	127.07	238.85	3,138.20
	Solar Indoor Walking Light	214.59	20.00	37.59	493.92
SATURN	Breathing	0.04	0.00	0.01	0.10
	Coughing	0.49	0.05	0.09	1.12
	Talking	0.8	0.07	0.14	1.85

B APPENDIX: ENERGY HARVESTING AND ACTIVITY EFFECT ON BATTERY LIFETIME

Solar. Table 3 presents the energy harvested while performing these activities. We found that the quantity of energy harvested with the surgical and N95 masks is similar. We note that the energy harvested outdoors is orders of magnitude higher than the energy we harvest indoors. However, even at indoor conditions, solar energy harvesting can yield considerable accumulated energy given their continuous power output across a long period of time. The wires that connect the solar panel to the FaceBit are very small in diameter, but we note that any object that passes across the face seal of an N95 mask may cause leaks, and so is discouraged.

Head Movement. According to Faraday's law of electromagnetic induction, the induced voltage depends the number of turn in the coil (N), and the rate of change of the magnetic field (Φ).

$$V = -N\frac{d\Phi}{dt}$$

With harvester weight as one of our main constraints, we limit the number of turns to 5000. We attach the shaker to both surgical and N95 masks using a small Velcro strip (Fig. 4b), as using magnets affects the motion of the core magnet. We again use Rocketlogger to record the energy harvested while an experimenter walked and ran with the mask on. Table 3 shows the energy output from these harvesters. We found that energy harvested from running is significantly higher than from walking. We also found that the shaker harvests more when attached to a surgical mask. We hypothesize this is because surgical masks are not as tightly fitted as N95 masks which amplifies the motion of the magnet. Also, the harvested energy and the impulse frequency is dependent on how fast the person is moving. This can be seen in Fig. 4b as the frequency is higher when running.

Thermal. Again using a Rocketlogger, we record how much energy we can harvest from activities like nasal breathing, mouth breathing, and talking, as all of these generate heat whenever a mask is worn. Table 3 summarizes the energy output from these harvesters. We found that the temperature differential between the two sides of the

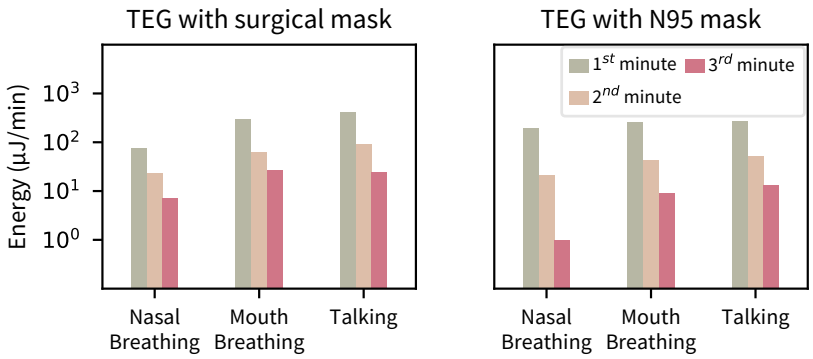


Fig. 23. Energy harvested from thermoelectric generator (TEG) while doing different activities. Energy decreases with time as the temperature differential between two plates of the TEG decreases and reaches steady state after 3 minutes.

Table 5. Secondary leak detection evaluation post hoc paired T-test within mask configurations.

A	B	p-value
1-hole	2 -hole	0.9556
1-hole	3-hole	0.9977
1-hole	mask-fit	0.0149
1-hole	mask-press	0.0005
2-hole	3-hole	0.9847
2-hole	mask-fit	0.0334
2-hole	mask-press	0.0005
3-hole	mask-fit	0.0007
3-hole	mask-press	0.0005
mask-fit	mask-press	0.0005

TEG decreases over time because of the heat conducted internally. This leads to a drop in energy generation over time as can be seen in Fig. 23. Energy generation reaches a steady state after 3 minutes.

Using the results of our energy harvesting and energy consumption measurements, we show the possibility of supplementing and extending our battery life with renewable sources of energy. Table 4 summarizes our findings. Notably, certain energy harvesting strategies and activities vastly outperform others, but all are somewhat capable of supplementing battery life, especially with hardware and software optimizations that reduce the energy consumption of the platform and runtime. Of course, we cannot achieve the results shown in the table unless the platform matches its energy usage to the harvested energy available in our energy storage elements (i.e. energy-adaptive runtime), and this is a promising future direction to pursue.

C APPENDIX: EVALUATION

C.1 Leak Detection

Figure 24 shows the raw signal values and measured integrals for the participants in our secondary mask fit exploratory study. Each signal sample depicts the differential pressure between inside the mask and the ambient pressure of the participants environment while performing an inspiratory vital capacity test (IVCT). Each signal

is filtered using a Savitzky-Golay filter. The relatively stable signal before and after the dip in each signal is the participants breath hold where the pressure inside the mask equalizes to the pressure outside the mask. The dip is the participant inhaling completely after first exhaling their entire lung capacity. In most cases the decrease in integral value is apparent by just visually comparing each plot. The most notable exception is with 107, where the participant reported that they were unable to completely pass the user seal check because the mask was too large. The statistical significance of this study is outlined in section 7.4. Each configuration was completed 3-times and randomized with each trail. All samples were used in analysis.

C.2 Scenario Based Free-Living Study

Included in figure 25 are the remaining FaceBit, ground truth comparisons free-Living studies; "working at computer", "riding in car", and "riding on train". Talking was limited in these scenarios and therefore, respiratory rate is captured at consistent frequency throughout each scenario. The ground truth rolling average of instantaneous respiratory rate (calculated at each breath) and FaceBit's measurements follow a similar trend. Error is reporting in table 1. Heart rate is only captured when the participant is at rest and therefore is much more sparse in scenarios with noise ("riding in car", and "riding on train"). "Working at computer" captures heart rate at a consistent frequency since the user is stationary.

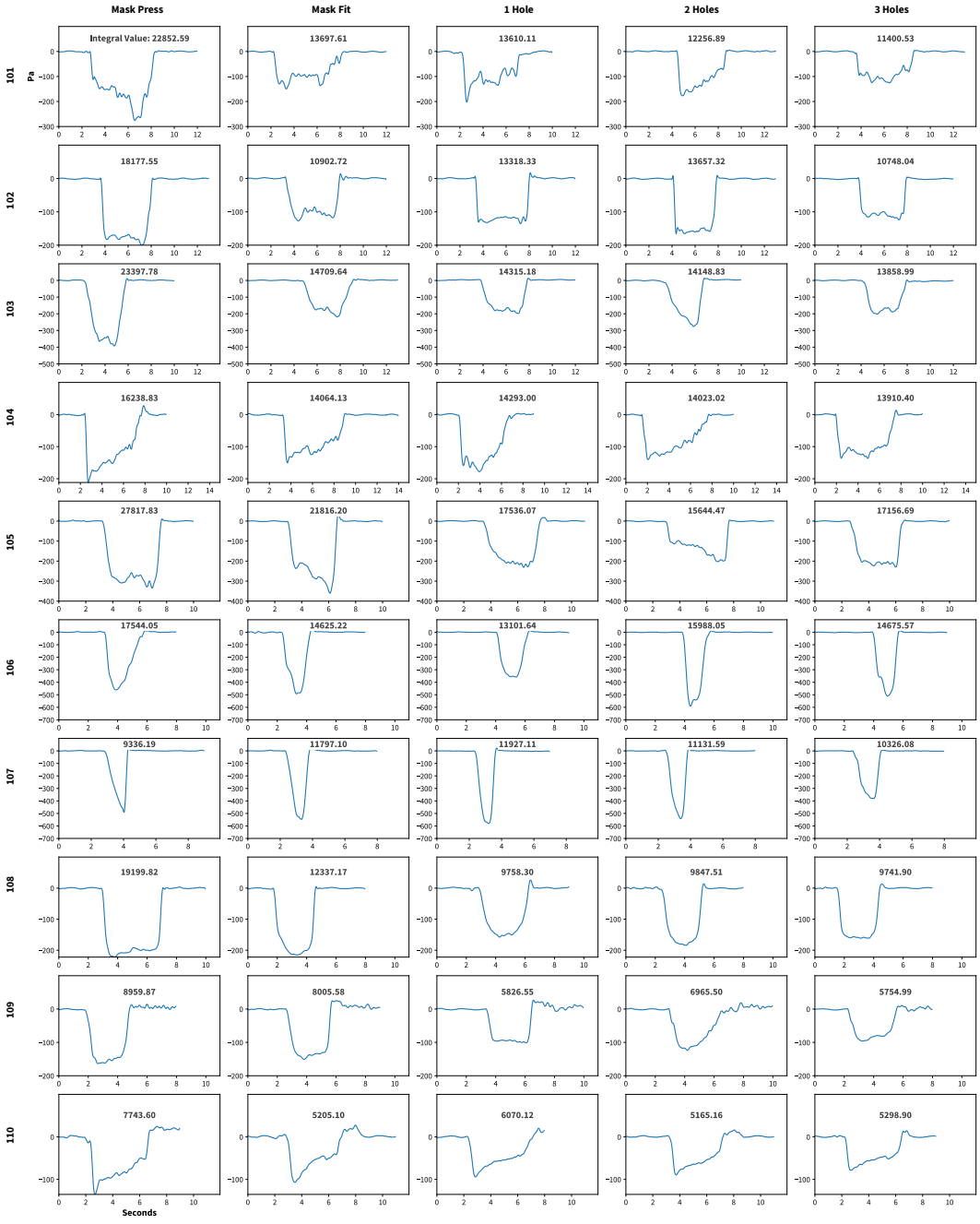


Fig. 24. Cropped and filtered differential pressure signals from the first trial of the secondary mask leak detection experiment for each participant. Absolute value integral is presented with each sample.

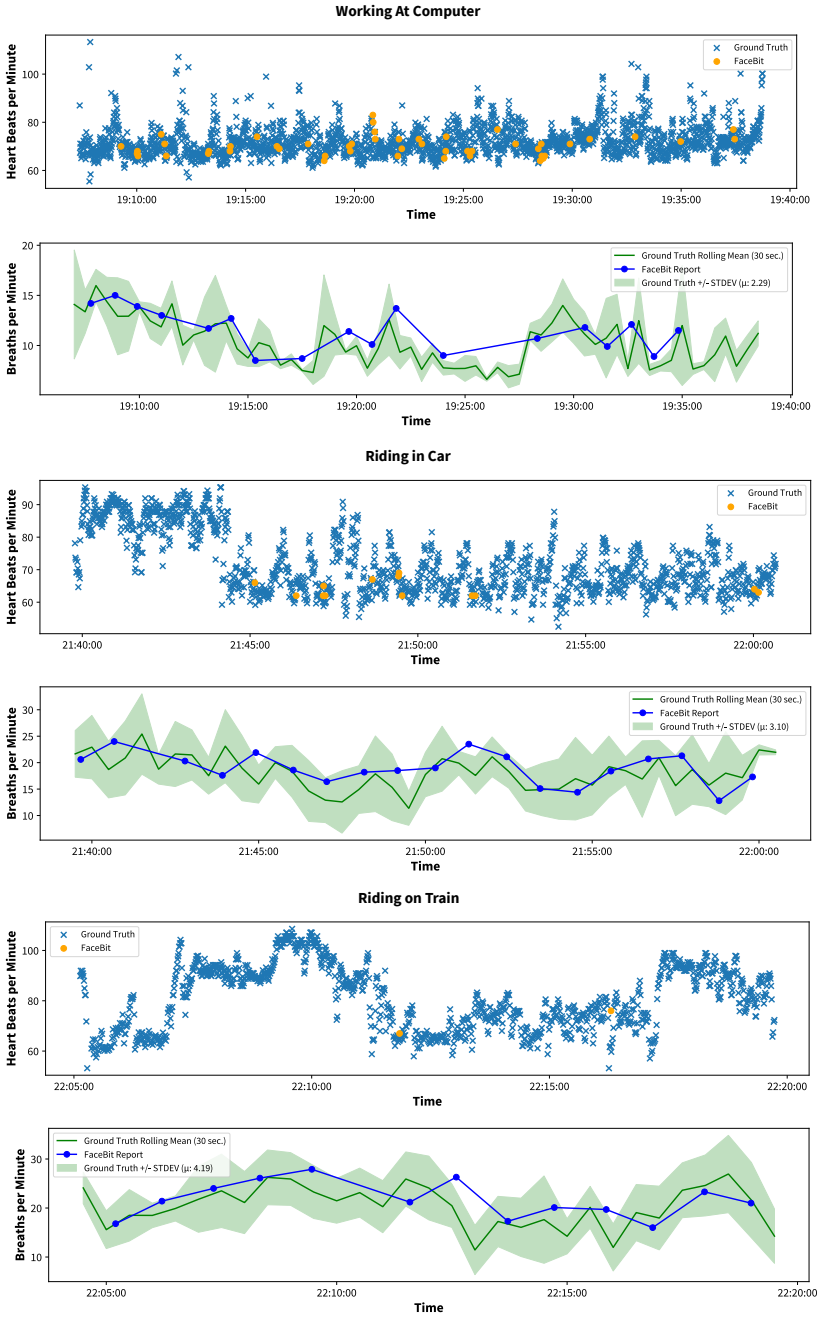


Fig. 25. Heart rate and respiratory rate samples from FaceBit and ground truth for free-living scenarios