Mindfulness Meditation and Respiration: Accelerometer-based Respiration Rate and Mindfulness Progress Estimation to Enhance App Engagement and Mindfulness Skills

MOHAMMAD NUR HOSSAIN KHAN, Worcester Polytechnic Institute, USA

DAVID CRESWELL, Carnegie Mellon University, USA and Equa Health Inc., USA

JORDAN ALBERT, Carnegie Mellon University, USA and Equa Health Inc., USA

PATRICK O'CONNELL, Carnegie Mellon University, USA and Equa Health Inc., USA

SHAWN FALLON, Carnegie Mellon University, USA and Equa Health Inc., USA

MATHEW POLOWITZ, Carnegie Mellon University, USA and Equa Health Inc., USA
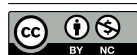
XUHAI "ORSON" XU, Columbia University, USA

BASHIMA ISLAM, Worcester Polytechnic Institute, USA

Mindfulness training is widely recognized for its benefits in reducing depression, anxiety, and loneliness. With the rise of smartphone-based mindfulness apps, digital meditation has become more accessible, but sustaining long-term user engagement remains a challenge. This paper explores whether respiration biosignal feedback and mindfulness skill estimation enhance system usability and skill development. We develop a smartphone's accelerometer-based respiration tracking algorithm, eliminating the need for additional wearables. Unlike existing methods, our approach accurately captures slow breathing patterns typical of mindfulness meditation. Additionally, we introduce the first quantitative framework to estimate mindfulness skills—concentration, sensory clarity, and equanimity—based on accelerometer-derived respiration data. We develop and test our algorithms on 261 mindfulness sessions in both controlled and real-world settings. A user study comparing an experimental group receiving biosignal feedback with a control group using a standard app shows that respiration feedback enhances system usability. Our respiration tracking model achieves a mean absolute error (MAE) of 1.6 breaths per minute, closely aligning with ground truth data, while our mindfulness skill estimation attains F1 scores of 80–84% in tracking skill progression. By integrating respiration tracking and mindfulness estimation into a commercial app, we demonstrate the potential of smartphone sensors to enhance digital mindfulness training.

CCS Concepts: • **Mobile Computing**; • **Ubiquitous Computing**;

Additional Key Words and Phrases: Mindfulness, Concentration, Sensory Clarity, Equanimity, System Usability, User Engagement.

Authors' Contact Information: Mohammad Nur Hossain Khan, Worcester Polytechnic Institute, Worcester, Massachusetts, USA, mkhan@wpi.edu; David creswell, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA and Equa Health Inc., Pittsburgh, Pennsylvania, USA, creswell@andrew.cmu.edu; Jordan Albert, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA and Equa Health Inc., Pittsburgh, Pennsylvania, USA, jordan@equahealth.io; Patrick O'Connell, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA and Equa Health Inc., Pittsburgh, Pennsylvania, USA, pat.oconnell108@gmail.com; Shawn Fallon, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA and Equa Health Inc., Pittsburgh, Pennsylvania, USA, shawn@equahealth.io; Mathew Polowitz, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA and Equa Health Inc., Pittsburgh, Pennsylvania, USA, mathew@equahealth.io; Xuhai "orson" Xu, Columbia University, New York, new York, USA, xx2489@cumc.columbia.edu; Bashima islam, Worcester Polytechnic Institute, Worcester, Massachusetts, USA, bislam@wpi.edu.

1 INTRODUCTION

Mindfulness training, a therapeutic technique that cultivates attention and awareness of the present moment, has gained increasing recognition for its role in improving mental health outcomes [61, 64, 65]. Recent studies highlight its potential to mitigate the rising rates of depression, anxiety, and loneliness among young adults in the United States, a population experiencing an unprecedented surge in mental health challenges [71, 99, 100]. Epidemiological data reveal that the prevalence of mental illness in this age group is nearly doubling every five to seven years. Compared to standard care and active control groups [63, 65, 66, 74, 88], mindfulness training has shown greater efficacy, largely due to higher acceptance and engagement within this demographic [25]. Clinical trials consistently demonstrate that mindfulness not only reduces stress, loneliness [27, 28, 54], cortisol levels, and blood pressure reactivity to stress but also significantly enhances daily well-being [60, 64, 65].

With smartphones now an integral part of daily life for young adults, smartphone-based mindfulness training offers a highly effective, accessible, and practical solution, evidenced by the widespread adoption of mindfulness apps [81, 85] like Headspace and Calm, which have collectively been downloaded over 180 million times. Randomized controlled trials (RCTs) have demonstrated that these digital interventions can significantly improve both psychological and physical well-being [60, 64, 65]. However, despite their promise, sustained user engagement remains a critical challenge [14, 75]. For instance, a recent review found that mindfulness meditation apps achieve higher daily usage (median 21.7 minutes) and retention rates compared to other mental health apps [14]. Yet, even the retention rate for these apps remains low at just 4.7%, underscoring a persistent issue with user engagement [14]. This highlights the urgent need for more engagement mindfulness applications, as research consistently shows that greater user engagement leads to improved clinical outcomes [24, 28, 70].

Several decades of research highlight the potential of biosignal feedback to enhance user engagement and outcomes on mobile health platforms [40, 43, 51]. Notable examples include EEG-measured brainwave feedback in the RelaxWorld app [51], EEG and vibration-based haptic feedback in the AttentiveU platform [50], and respiration biofeedback in gamified platforms [33, 91]. In the context of mindfulness meditation, respiration tracking emerges as a particularly promising and predictive signal [8, 52]. Although the relationship between mindfulness practice and respiration patterns remains underexplored, preliminary studies suggest distinct alterations in breathing during mindfulness sessions, even when users are not consciously controlling their breath [2, 106]. This presents an opportunity to provide users with insights into their mindfulness progression and skill development. Despite the growing popularity of commercial mindfulness apps such as Headspace and Calm, which are integrated into wearable devices like Oura rings, Fitbits, and Whoop sensors, few products incorporate real-time biosignal feedback into their training. While respiration-based biofeedback has been explored in gamified platforms [91], it remains absent in both commercial and non-commercial mindfulness meditation apps aimed at improving system usability and user engagement. Existing apps—such as Muse (EEG-based brain signals), Flowly (heart rate and HRV via external sensors), Inner Balance (HRV using HeartMath sensor), and Core Meditation Trainer (ECG)—rely on external biosignal acquisition devices, limiting scalability. Notably, none of these systems incorporate respiration biofeedback, despite its central role in mindfulness practices. Respiration is not only more intuitive and directly connected to meditative techniques like breath awareness, but it can also be captured using minimally obtrusive sensors, making it a practical and scalable feedback modality. Respiration also plays a pivotal role in the physiological manifestation of mindfulness; specifically, slow breathing rates (4–9 bpm) are known to
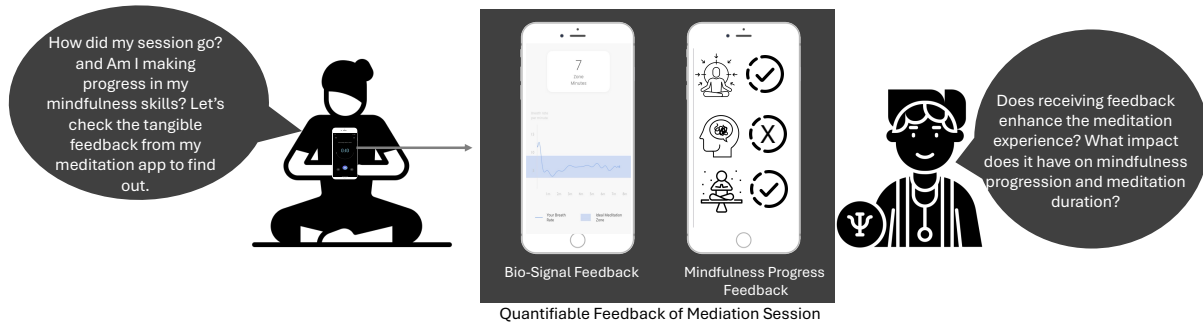
Fig. 1. We integrate a feedback module into a commercially available mindfulness meditation app, Equa, featuring a real-time respiration rate chart, respiration statistics, and mindfulness progress feedback. This is achieved by developing a respiration rate tracking algorithm and a mindfulness skill change estimation algorithm. To assess the impact of this feedback, we conducted a user study evaluating user satisfaction, mindfulness progression, and engagement.

optimize a parasympathetically dominated restful state via vagal stimulation [35, 86], and recent research shows that breathing rhythms modulate brain activity related to emotion regulation and attention [97].

To achieve this goal, this paper addresses three novel research questions: (1) Can respiration biosignals be effectively tracked during digital mindfulness training using the smartphone's built-in sensors? (2) Can these sensor signals predict improvements in mindfulness skills? (3) Does providing respiration biosignal feedback after each session enhance sustained engagement with digital mindfulness training? To answer these questions, we track and visualize respiration biosignal and the change in mindfulness skills as feedback to users into a smartphone application using the phone's on-board accelerometer that delivers a unique mindfulness training curriculum focused on cultivating core skills: concentration— "your ability to focus on what you want to," sensory clarity— "your ability to track and explore your sensory experience in real-time," and equanimity— "your ability to allow sensory experiences to come and go,"—based on previous digital mindfulness prototypes [21, 61, 64, 65, 93]. As illustrated in Figure 1, our system integrates a feedback module that provides real-time respiration tracking, breathing statistics, and mindfulness progress estimation, allowing users to reflect on their meditation sessions.

To effectively track respiration and mindfulness skills in the app, we address two key technical challenges. First, although respiration tracking has been well studied [5, 30, 39, 78, 83, 84, 94, 101], existing methods are inadequate for detecting the slow breathing rates (<10 bpm) characteristic of mindfulness practices. Most respiration rate estimation algorithms are developed and tuned for natural breathing rate detection or for monitoring pulmonary patients [84, 101]. To overcome this, we develop an accelerometer-to-respiration rate algorithm capable of accurately tracking both low (4–9 bpm) and regular (10–30 bpm) respiration rates. This algorithm is validated against clinical-grade Hexoskin lifeshirts, demonstrating reliable performance across a wide range of respiration patterns [44, 49]. While respiration biofeedback and smartphone-based tracking have been explored, our work is the first to use respiration as a biomarker for assessing and enhancing mindfulness skills. Moreover, prior smartphone-based methods (e.g., InstantRR [84, 101]) focus on natural breathing rates (12–20 bpm) and struggle with the slower rates critical to mindfulness practice.

Building on the development of our accurate low respiration tracking, we design a user feedback interface screen that visually represents the change in breaths per minute during each session. It also highlights the total time spent within the optimal breathing zone, which prior research suggests falls between 4 and 9 bpm [15, 58]. Previous studies suggest that this slow breathing zone may optimize a restful parasympathetically-dominated restful state mediated by vagus [35, 86]. In addition to these metrics, participants receive continuous respiratory

signal feedback throughout their sessions. While respiration biosignal feedback has been shown to enhance system usability in other contexts [33, 91], its application within mindfulness meditation apps has not yet been fully explored. This paper is the first work to integrate respiration feedback and breathing statistics into a mindfulness app.

The second technical challenge is the lack of any quantitative method to measure changes in mindfulness skills—such as concentration, sensory clarity, and equanimity [65]—during digital mindfulness training. Current approaches rely on self-report questionnaires [32, 68], which are time-consuming, intrusive, and subjective, often leading to lower user engagement. To overcome this limitation, we develop a deep learning (DL) algorithm that uses accelerometer data to estimate changes in mindfulness skills after each guided session. This DL-based solution eliminates the need for self-reporting, providing a scalable and objective way to measure mindfulness progress while minimizing user burden.

In this study, we collect data from 40 participants, completing 261 mindfulness sessions, to evaluate the impact of respiration biosignal feedback on system usability. We develop a respiration rate tracking algorithm that not only estimates respiration rate in the range of 4-9 that is crucial for mindfulness meditation, but also accurately tracks natural breathing rate (12-20 bpm). Our evaluation of existing methods shows that the algorithms developed using natural breathing rate fail to detect the low breathing rate. Additionally, we propose the first study to quantify the change in mindfulness skills during meditation using accelerometer data based on respiration changes. We train session-level data to develop the deep learning model to detect the mindfulness change. We integrate the respiration chart, respiration statistics, and mindfulness change in the commercial mindfulness app, Equa, to improve our app's usability, user engagement, and mindfulness experience. We conducted a user study spanning upto 21 days using the session-level feedback. Our results show higher system usability, mindfulness skills change, and user engagement.

The contribution of this paper is as follows:

- We develop a respiratory rate detection algorithm that accurately tracks both slow (4–9 BPM) and normal (10–30 BPM) respiration rates, addressing the challenge of monitoring breathing patterns during mindfulness meditation.
- We design a deep learning model to estimate mindfulness skill progression—concentration, sensory clarity, and equanimity—using accelerometer-derived respiration data, making it the first approach to quantitatively track mindfulness skill development using physiological signals in digital mindfulness interventions.
- We integrate real-time respiration tracking, mindfulness skill assessment, and breathing statistics into a mindfulness app, providing users with personalized feedback to enhance their meditation experience.
- We conduct a user study to evaluate the impact of respiration biosignal feedback on system usability, demonstrating its effectiveness in improving engagement and self-awareness.

2 BACKGROUND AND RELATED WORK

2.1 Respiratory Biomarker Estimation

Respiration rate is clinically measured using sensors to detect air pressure near the mouth and nose [7]. Various sensors, e.g., motion [84], audio [3, 53], and camera [11], estimate respiration rate with wearable systems.

Motion sensors, i.e., inertial measurement unit (IMU), are placed on the chest, head, and wrist to sense breath-related body movement and infer the respiratory parameters from it [38, 39, 84, 94]. Unlike smartphones, other wearable motion sensors on the chest are not ubiquitous and scalable. Smartphone IMU-based breathing bio-signal, e.g., breathing phase and breathing rate, estimators have shown promising results for pulmonary patients [4, 20]. Rahman et al. [84] achieves impressive performance using three different algorithms contextually: FFT, peak detection, and Zero cross-detection. However, they only consider short-duration breathing tasks and do not consider low-breathing rate detection.

Some recent works exploited video-based solutions to estimate breathing rates [9, 11, 87]. Bae et al. [11] place a camera at the same height as the participant's face and record the change in motion during inhalation and exhalation. This approach requires adequate lighting conditions and proper placement. Alnaggar et al. [9] also use video to estimate respiration rate with 1.62 MAE, but the pearson correlation coefficient (PCC) is very low. However, these approaches are unsuitable for passively monitoring mindfulness meditation without imposing additional costs or distracting users from the mindfulness training. Besides, video-based solutions fail to measure low breathing rates from 4-9 bpm.

Smart watches can be a usable solution, as they don't interrupt the users' natural interaction with the device. Many previous works used IMU from watches [94] to estimate the breathing rate. While SleepMonitor [94] achieved impressive performance, they evaluated their approach during sleep, which does not incorporate the low breathing rates, which is crucial in mindfulness training. Additionally, most works discard noisy signals by employing signal processing or machine learning techniques, which improves accuracy but also discards breathing sessions. For example, WearBreathing [59] achieved impressive accuracy (MAE 1.09 BPM) with only 7−16.5% data retention in a resting position (sitting, standing, and lying) on COPD patients. Such a low retention rate is useless to the user, as they cannot make successful measurements in most attempts. Therefore, we need a more usable solution that estimates the breathing rate with reasonably good accuracy and retention rate. Moreover, many commercial watches do not allow the IMU data from watches with a proper sampling rate, which makes it less desirable for a scalable solution.

Acoustic solutions have also been used for respiratory biosignal estimation [6, 42, 92]. Recent works on audio-based respiratory signal estimation focus on pulmonary patients [42], guided breathing exercises [92], or physical exercises [53]. Breeze [92] detects breathing phases during controlled breathing exercises, consisting of four seconds of inhalation through the nose, two seconds of exhalation through the mouth, and four seconds of pause. BreathTrack [42] develops an acoustic breathing phase detector that is trained using guidance from the IMU collected from the chest. However, BreathTrack focuses on pulmonary patients and aims to differentiate between healthy and clinical participants. The most recent work from Ahmed et al. [6] achieves 1.72-1.94 mean absolute error (MAE) in estimating respiration rate. However, their work collects data from in-ear microphones and IMU, which are publicly inaccessible and, thus, not a scalable solution. Kumar et al. [53] developed a multi-tasking model to detect respiration rate during exercise from audio signals' Mel-Filter Bank energy (MFB). Breathing audio signals during exercise have some advantages over natural breathing in that they are more audible and have more frequency content. Moreover, none of these works support low respiration rate estimation, which is crucial for predicting mindfulness skills.

In this paper, we focus on estimating respiration rates using a smartphone accelerometer placed on the chest, offering a ubiquitous and scalable solution with no additional cost. Our goal is to accurately detect low respiration rates (4-9 bpm) alongside natural breathing rates (10-30 bpm), which are crucial for predicting mindfulness skills.

2.2 Mindfulness Detection

Over the past two decades, there has been significant growth in scientific research on mindfulness interventions and the development of mindfulness skills [26]. Mindfulness interventions have been shown to increase self-reported mindfulness skills reliably [104] and to improve a broad range of health and well-being outcomes in randomized controlled trials [19, 26]. There has been considerable debate on the optimal ways to measure mindfulness skill development in this emerging literature. However, in recent mindfulness intervention studies, one approach is to teach people mindfulness skills such as concentration, sensory clarity, and equanimity in a unified mindfulness curriculum. Then, operational definitions of these measures are described to participants in the mindfulness intervention to measure these skills (i.e., the participants are familiar with them from their
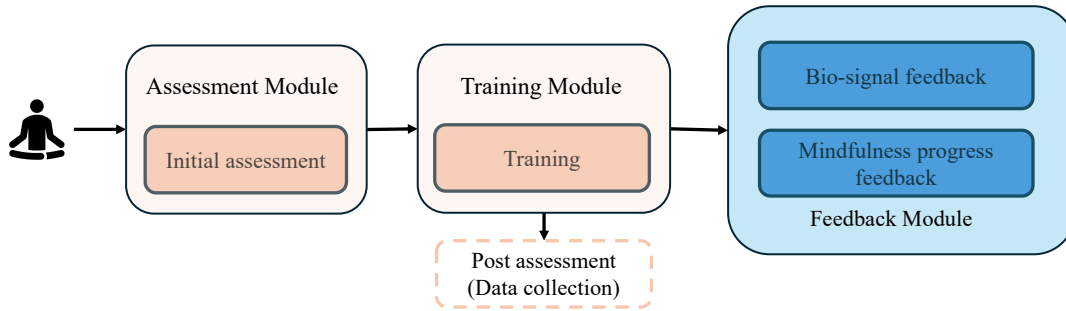
Fig. 2. Overview of the smartphone app modules. Post assessments provide ground truth to develop algorithms for the feedback module. The main contribution of this paper is the development, integration, and study of the feedback module highlighted in blue.

mindfulness training programs). Smartphone-based Unified Mindfulness training programs in initial placebo-controlled clinical trials seem compelling [61, 64, 65]. Previous studies utilize trait bases measures of mindfulness skills, such as the trait Mindful Attention Awareness Scale (MAAS) [68] or Five Facet Mindfulness Questionnaire (FFMQ) [32]. These kinds of measurements are not adequate for the large-scale implementation of mindfulness meditation. In this paper, we measure these mindfulness skills sensitively with a single-item measure of each skill before and after each guided mindfulness training session. Thus, in our present work, we develop an accelerometer signal-based classification model that can predict these changes.

3 OVERVIEW OF THE MINDFULNESS TRAINING SMARTPHONE APP

The mindfulness training app we use in this study is a commercially available app, Equa [41], designed to provide evidence-based instruction in training three core mindfulness skills: concentration, sensory clarity, and equanimity. We integrate guided meditation sessions with biosignal feedback and track mindfulness progress in the app. As illustrated in Figures 2 and 3, the proposed application delivers a structured curriculum tailored to each user's mindfulness journey, offering personalized feedback based on their physiological data and session performance. We integrate the feedback module with the app's existing training and self-assessment modules.

Training Module. This module consists of two submodules: (a) training curriculum and (b) meditation training.

a) *Training Curriculum.* The introductory curriculum consists of 14 interactive and branching lessons. The curriculum has demonstrated an immediate dose-response effect comparable to the Mindfulness-Based Stress Reduction (MBSR) program [45, 67]. The curriculum is evidence-based [26, 61, 62, 64, 65] and interactive, allowing users to navigate through lessons using screen taps. It also encourages the integration of mindfulness into everyday activities to enhance the long-term success of mindfulness training [61, 70]. By leveraging principles from intelligent tutoring systems [10, 102], the smartphone app is the first meditation app to offer this type of in-lesson interactivity, in contrast to traditional digital apps that rely primarily on static, one-way guided meditation audio libraries.

(b) *Meditation Training.* Following the baseline assessment, participants engage in a 10- to 20-minute training session, during which their respiration signals are continuously tracked (see more technical details in section 5).

Self-Assessment Module. The assessment module aims to collect relevant ground truth to develop algorithms for the feedback module. The app begins with a comprehensive pre-training assessment that evaluates users' baseline mindfulness skills and includes self-report ratings on the three core mindfulness skills. After each session, users complete the same self-assessment to measure any immediate changes in their mindfulness skills.

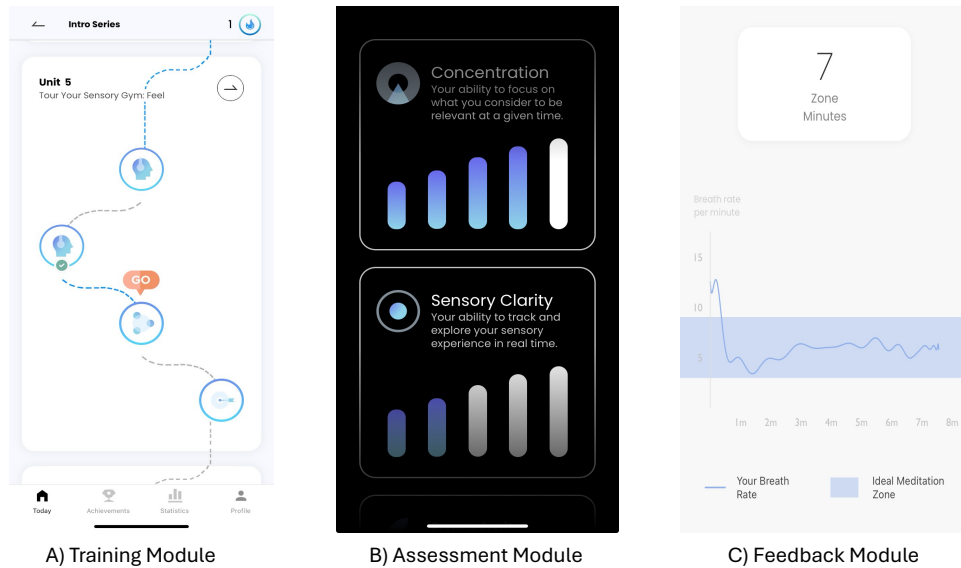A) Training Module   B) Assessment Module   C) Feedback Module

Fig. 3. Screenshot of different modules deployed in Equa. The training module consists of necessary instructions for mindfulness meditation training, and the assessment module captures the baseline mindfulness state before the session. Finally, the feedback module provides the users with biosignal feedback such as a respiration chart, respiration statistics, and estimated mindfulness change.

These self-reported ratings of concentration, sensory clarity, and equanimity serve as ground truth for analyzing the collected respiration data and detecting shifts in mindfulness. While self-reporting is the current way to assess mindfulness skills, and these reports help establish initial benchmarks, a key benefit of mindfulness estimation from accelerometer data is that it minimizes the need for repeated, time-intensive self-assessments after each session—an approach that can reduce user satisfaction in digital mental health apps. By relying on physiological and sensor data to measure mindfulness improvement (see section 6), the app aims to streamline the user experience.

Feedback Module. After each session, participants receive two forms of feedback: respiration biosignal feedback and whether their mindfulness skills improved (or not).

- *Biosignal feedback.* We display a respiration chart from the session, offering users insights into their physiological responses during the guided session. Our respiration rate tracking algorithm (see section 5) is designed to accurately track both slow and natural respiration rates using a smartphone accelerometer while also detecting motion artifacts that may compromise the quality of the data. In such cases, users receive a notification indicating potential data issues, ensuring transparency and trust in the feedback.

- *Mindfulness progresses feedback.* The app estimates the user's improvements (or lack thereof) in mindfulness skills, such as concentration, sensory clarity, and equanimity. Our deep learning model allows us to predict the progression of these mindfulness skills based on the smartphone accelerometer data collected during each session. The detailed methodology and performance evaluation of mindfulness change assessment are discussed in the mindfulness progress estimation and algorithmic evaluation sections (see section 6 and 8.2).

Table 1. Participant statistics in different studies. Study 1 and Study 2 are conducted in a controlled environment setting, whereas Study 3 is conducted in the wild. In addition to testing the developed algorithm, study 3 is used for user study.

Study	Environment	Participants	Sessions	Algorithm Development	Algorithm Testing	User Study
1	Controlled (Lab)	8	88	Yes	Yes	No
2	Controlled (Lab)	10	24	No	No	Yes
3	Natural (home)	22	149	No	Yes	Yes

Table 2. Breakdown of participants' demographics in terms of gender, race, and meditation experience in studies 1, 2, and 3. All studies have a balanced distribution of participants.

Category	Group	Study 1 (%)	Study 2 (%)	Study 3 (%)
Gender	Male	50	40	50
	Female	50	40	50
	Non-binary	-	20	-
Race	Black	12	-	4
	East Asian	12	10	9
	South Asian	38	20	14
	White	25	50	68
	Other	-	10	-
	Mixed	13	-	5
Meditation Experience	Not at all	12	20	4
	A few times	37	40	23
	Periodically	25	20	41
	Frequently	13	10	23
	Everyday	13	10	9

4 DATA COLLECTION AND STUDY PROTOCOL

To develop and evaluate our proposed algorithms, we conduct three data collection studies. The first two are conducted in a controlled lab environment, while the remaining two took place in real-world settings with participants using the app. Figure 7 shows the lab setup and the screen of the app for mindfulness training.

Study design and participants. In total, we collect data from 40 participants, with 18 participating in lab studies and 22 in-the-wild studies. Table 1 shows the participant statistics of all the studies. However, some data samples are not used due to missing ground truth scores or data corruption. The demographic distribution of participants is presented in Table 2. Participants are recruited through community outreach, including online postings, mailing lists, and social media platforms. Eligibility criteria included being fluent in English, being at least 18 years of age, being willing to engage in guided mindfulness meditation sessions, and being willing to wear physiological measuring equipment during meditation training. Initial eligibility screenings were conducted via email, followed by a formal in-person screening at the first session to confirm participation. All participants were first-time users of the app, ensuring consistency in baseline familiarity across study conditions.

    a) *Lab study Protocol (Studies 1 and 2)*: In the lab studies, participants complete a 14-day mindfulness meditation program consisting of 20-minute lessons while their respiration data are collected using a Hexoskin smart shirt and an iPhone X mounted on a chest strap. Participants are randomly assigned into two conditions: a biosignal
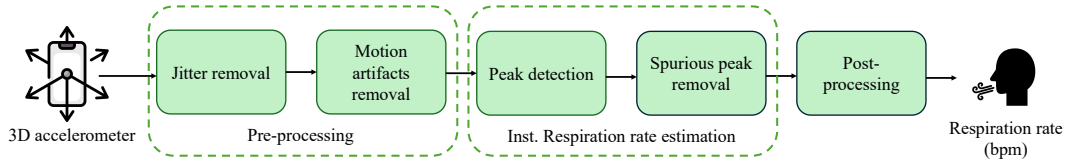
Fig. 4. Overview of respiration rate detection algorithm. We pre-process the accelerometer data by removing jitter and motion artifacts before using the proposed peak detection algorithm. Final respiration rates are shown to the user after additional post-processing.

augmented condition and a control condition. Following the mindfulness lessons, participants in the biosignal augmented condition are exposed to a respiration biosignal feedback chart, which displays the respiration rate trend throughout the lesson. Participants in the control condition are not exposed to this chart. These sessions provided high-fidelity respiration data, which are used as the ground truth for algorithm development. Study 1 included 88 sessions with 8 participants, while Study 2 included 24 sessions with 10 participants. Both studies focused on developing and validating the respiration rate estimation algorithm, testing the accuracy of the mindfulness skill estimation model, and user satisfaction with the app.

b) *In-the-Wild Study Protocol (Study 3)*: In Study 3, conducted remotely, 22 new participants use the mindfulness meditation app in real-world environments. Similar to the lab study, chest straps and smartphones are used for data collection. This study spans over 21 days. Randomization into the two conditions for study 3 is conducted following the same protocol as in Studies 1 and 2. Unlike studies 1 and 2, in this study, participants can finish as many sessions as they want and end any session at any time. After 21 days of study, we use the average number of completed sessions and the average number of minutes in meditation to evaluate user engagement. Participants are instructed to complete sessions in quiet, stable environments to minimize movement artifacts. A total of 149 sessions are collected in this study. This study focused on evaluating user satisfaction, user engagement, and the effectiveness of the mindfulness skill estimation algorithm in natural settings.

Physiological and Psychological Data Measures. Across all studies, chest movements related to breathing are tracked using the smartphone accelerometer. In Studies 1 and 2, physiological data such as respiration rate and motion are also collected using Hexoskin smart shirts, which serve as ground truth for developing the respiration tracking algorithm. Participants also complete self-report assessments of mindfulness skills—concentration, sensory clarity, and equanimity—before and after each session. These self-reported measures provide a baseline for analyzing mindfulness progression and serve as the ground truth for evaluating the deep learning models.

Ethical Considerations. The study was approved by the Institutional Review Board (IRB) of Carnegie Mellon University. Before data collection began, informed consent was obtained from all participants. All data were anonymized and stored on encrypted servers to ensure participant privacy. Participants were informed about the nature of the data being collected, their right to withdraw from the study at any time, and the potential use of their anonymized physiological data in future research publications.

5 SLOW-PACED RESPIRATION RATE ESTIMATION USING ACCELEROMETER FOR BIOSIGNAL FEEDBACK MODULE

In the absence of a reliable algorithm for accurately detecting both slow-paced and regular respiration rates, we develop a novel respiration rate estimation method. Our algorithm leverages accelerometer data, which has been widely utilized in regular or high respiration rate estimation [83, 84, 94]. However, accelerometer data is often contaminated by two significant noise sources: jitter and motion artifacts, which obscure the respiratory

signal—especially for slow respiration rates, where the changes in the accelerometer are smaller and closer to the noise level. To address these challenges, we propose a four-step algorithm for respiration rate extraction. The first two steps involve noise reduction, focusing on removing jitter and motion artifacts. The third step evaluates the quality and reliability of the signal post-noise removal. Finally, the fourth step estimates the respiratory rate from a single sensor stream, providing an accurate measurement of both slow and regular breathing patterns. An overview of the algorithm is shown in Figure 4.

Jitter Removal. Jitter refers to high-frequency noise introduced by inertial sensors, often caused by electrical disturbances such as power supply fluctuations, electromagnetic interference, or errors during analog-to-digital conversion. This noise can interfere with the detection of accurate respiratory signals, potentially leading to incorrect peak detection and distorted frequency content. The effect of jitter is particularly significant during slow breathing, where the subtle changes in acceleration are easily masked by high-frequency noise. To address this, we apply a low-pass Butterworth filter [90] with a cutoff frequency of 10 Hz. The Butterworth filter is selected for its superior performance in noise reduction [13, 34, 73].

Motion Artifact Removal. Motion artifacts, commonly observed in IMU data, are caused by sudden changes in the user's posture or movement of the phone. These artifacts often produce large amplitude signals that overshadow respiratory signals, making accurate respiration rate extraction challenging. To mitigate this, we apply a local mean removal technique, where the mean of a short window centered around each value is subtracted from the raw sensor measurements. Local mean removal significantly reduces motion artifacts, enhancing the accuracy of respiration rate estimation. For additional denoising, a 13-point moving average filter (i.e., a window of 1.3s) is subsequently applied.

Reliability Assessment. Although the previous noise removal steps effectively reduce jitter and motion artifacts, continuous user movement or improper phone placement can still lead to inaccurate respiration rate estimations. To address this, we introduce a signal quality and reliability assessment step, which acts as a safeguard for our respiration rate estimation process. This step is crucial because poor signal quality can distort the entire feedback loop. We assess signal quality by calculating the first and third quartiles, along with the interquartile range (IQR), for each 20s window. If any data value at that window exceeds the third quartile with an addition of 0.8 times the IQR value or falls below the first quartile subtracted by 0.8 times the IQR value, we flag that value as compromised. When compromised values account for more than 25% of the total data length, the signal is discarded, and a "signal compromised" message is provided to the user. Additionally, we also detect if the phone is on the chest or staying flat on the table. We observe that if the phone is flat on the surface, we get steady and flat accelerometer data with occasional spikes. To identify if the phone is on a flat surface, we take the average of every 30s segment of data and compare it with the average of the whole file. If the difference between the two averages is more than 0.02, we count that as a faulty segment. If the number of faulty segments is lower than 30% of the total duration of the session, we determine that the data does not have enough variation to be on the chest. We discard the signal with the 'phone is not on chest' message to the user.

Instantaneous Respiration Rate Estimation. After filtering and signal quality assessment, we apply a peak detection algorithm that identifies all local maxima by comparing each data point with its neighboring values [103]. A peak, or local maximum, is defined as any sample whose two immediate neighbors have smaller amplitudes, with each peak corresponding to a breath (inhalation). Our algorithm is optimized to detect respiration rates between 4 and 30 breaths per minute, encompassing both slow and natural respiration rates. To further minimize noise interference, we discard consecutive peaks occurring within 2 seconds and only consider peaks with at least 50% prominence. Peak prominence is defined as the vertical distance between the peak and its lowest contour line and is used to ensure that only significant chest movements associated with breathing are counted. To compute the respiration rate for each minute, we divide 60 by the duration between two consecutive peaks. We then
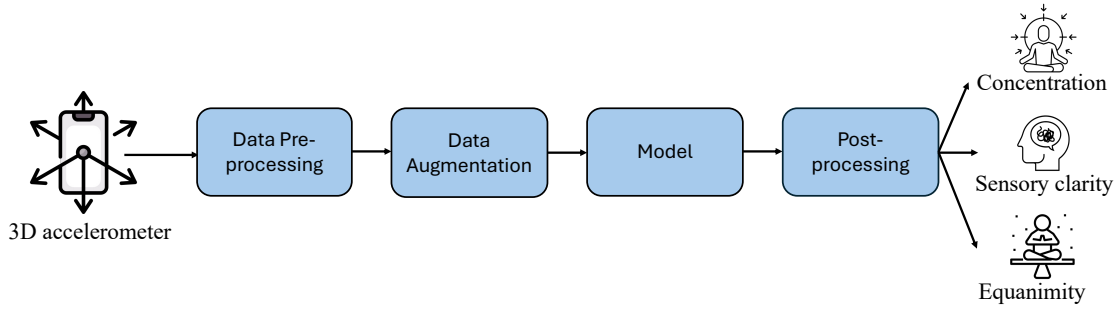
Fig. 5. Overview of the algorithm for estimating mindfulness progress via accelerometer data. The data are passed through preprocessing and augmentation before model inference and finally through a post-processing step to estimate changes in concentration, sensory clarity, and equanimity.

average the respiration rate over seven cycles to calculate the instantaneous rate, ensuring stability, as respiration rates typically do not fluctuate rapidly unless during intentional breathing exercises. This method reduces errors due to inaccurate peak detection and enhances the robustness of respiration rate estimation.

6 ESTIMATING MINDFULNESS PROGRESS VIA ACCELEROMETER FOR MINDFULNESS PROGRESS ASSESSMENT MODULE

We propose a deep neural network (DNN) for estimating changes in mindfulness using accelerometer data from a smartphone. This section outlines our data preparation pipeline, the network architecture, training parameters, and post-processing steps involved in our methodology. An overview of the algorithm is shown in Figure 5.

Data Pre-processing and Augmentation. To process the accelerometer data from mindfulness sessions of varying durations, we segment the data into non-overlapping two-minute intervals. This segmentation is based on prior research showing that even brief periods of mindfulness practice, between 60 and 90 seconds, can significantly influence mindfulness levels [29]. By using two-minute segments, we ensure that the data is both manageable in size and sufficiently detailed to capture meaningful changes in mindfulness. After segmentation, we obtained 287, 263, and 379 positive samples and 374, 398, and 282 negative samples for concentration, sensory clarity, and equanimity, respectively, revealing a substantial class imbalance between the two classes. To handle this imbalance, we augment the dataset by resampling the class that has a lower amount of data. We use raw accelerometer data rather than the filtered or processed data from the previous section, as the deep learning model is capable of denoising the data itself and capturing relevant information [36, 56].

Neural Network Architecture. Figure 6 shows the network architecture of the proposed mindfulness progress estimation model that takes the segmented accelerometer data as input and estimates changes in mindfulness skills—concentration, sensory clarity, and equanimity. The architecture combines a 1D ResNet with a Gated Recurrent Unit (GRU) to effectively handle sequential dependencies in time-series signals like accelerometer data.

We modify ResNet architecture (ResNet-1D) [37] to effectively capture spatial and temporal patterns in the accelerometer data. This ResNet-1D incorporates residual blocks that overcome the vanishing gradient problem and allow the networks to be deeper without loss of performance. Each residual block consists of one-dimensional convolutional layers, batch normalization, and Hardwish activation, which ensures stable and efficient training. The blocks come in two types: identity blocks, where the input is directly passed to the output, and convolutional blocks, which use an additional convolution layer in the shortcut path to adjust the dimensions when needed.
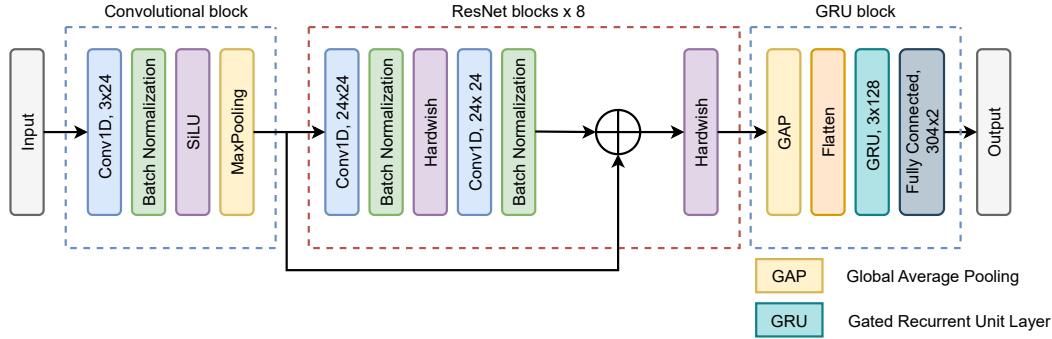
Fig. 6. Network architecture for mindfulness skills estimation. The architecture shows a modified version of ResNet [37] to capture information from accelerometer data with a convolutional block followed by ResNet blocks and a GRU block. GRU block allows the capture of temporal information for the data.

The network begins with an initial one-dimensional convolutional layer, followed by eight residual blocks. To reduce dimensionality while preserving key features, global average pooling is applied, and the architecture concludes with a flattened layer for further processing. Following the ResNet-1D layers, a GRU [22] is integrated to capture the temporal dependencies in the IMU data. GRUs, a variant of recurrent neural networks (RNNs) [72], are well-suited for processing sequential data, making them ideal for identifying patterns related to changes in mindfulness skills over time. We employ one GRU layer with 128 hidden neurons. By retaining relevant information from previous time steps, the GRU helps the network effectively recognize shifts in concentration, sensory clarity, and equanimity throughout the mindfulness sessions.

Training and Model Evaluation. The network is trained using a binary cross-entropy loss function. We use Adam to optimize our model, with a learning rate initially set at 0.0001. We employ a learning rate scheduler to mitigate overfitting by adjusting the learning rate dynamically based on the model's performance and train the model for 50 epochs. The training is conducted using an NVIDIA RTX 3090Ti GPU, and the best model is saved. The final model is evaluated based on its ability to accurately predict changes in mindfulness skills as compared to self-reported measures, which serve as the ground truth.

Post-Processing and Output Interpretation. To ensure robust predictions, a majority voting approach was applied during post-processing, where classification outputs from each two-minute segment were aggregated to determine the overall change in mindfulness skills across the entire session.

7 INTEGRATION OF FEEDBACK MODULE INTO SMARTPHONE APP

To enable real-time biofeedback and mindfulness skill estimation, we integrated our respiration tracking and mindfulness assessment algorithms into the backend of a commercial mindfulness meditation app (name withheld for anonymity). This integration allows the system to collect sensor data during meditation sessions and provide personalized feedback upon completion.

The app records accelerometer and gyroscope data at **100** Hz, storing it in Firebase Storage. Structured session metadata and feedback results are maintained in Firebase Firestore (NoSQL). After each session, Google Cloud Functions are triggered via a monitoring service subscribed to the data storage bucket. These functions execute Python scripts that process the sensor data, estimate respiration rates, and compute mindfulness scores. The

a) Data Collection Setup   b) App training screen
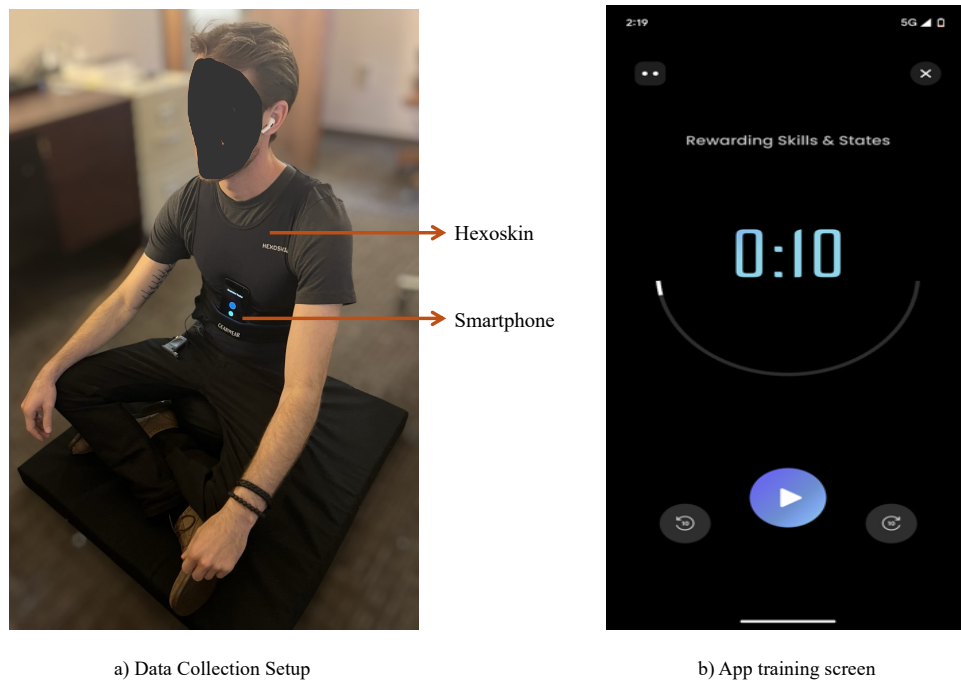
Fig. 7. Data collection setup in the lab and smartphone app training screen. Hexoskin is used to capture the ground truth for developing a respiration rate tracking algorithm. The smartphone captures the necessary accelerometer data utilized to develop our proposed algorithms.

processing typically completes within **0.2** seconds, and results are written back to Firestore via a RESTful API, enabling the app to update feedback visualizations (e.g., trend graphs, score tables) for the user. Feedback is provided post-session rather than in real-time to avoid interrupting the meditation experience and to ensure robust data processing.

The mobile app is cross-platform (iOS and Android), with the frontend dynamically retrieving feedback data from Firestore APIs. The backend architecture scales horizontally using Google Cloud's serverless infrastructure to support concurrent users without significant latency. To maintain data security and privacy, all sensor data is anonymized and transmitted using encrypted channels, following standard company privacy policies. Sessions with insufficient sensor data quality are flagged, and feedback is withheld to maintain accuracy.

The architecture and integration workflow are modular to support integration with other mindfulness or biofeedback applications. Backend algorithms and deployment pipelines are maintained through version control systems, supporting continuous deployment and future updates. The design also allows for easy integration of additional sensors or algorithms in future iterations.

8 RESULTS

In this section, we present the technical evaluation of our proposed algorithms. To assess the effectiveness of our proposed algorithms[1], we conduct a comprehensive evaluation using the data collected from our studies. Specifically, we measure the accuracy of the respiration rate estimation algorithm and evaluate the mindfulness skill change detection algorithm's ability to identify shifts in mindfulness skills using accelerometer data, demonstrating their effectiveness using data from both controlled laboratory settings and real-world environments.

8.1 Respiration Rate Estimation Algorithm

We assess the performance of our respiration rate estimation algorithm using data collected from study 1. We first compare the performance of our proposed algorithm against the state-of-the-art (SOTA) respiration detection algorithms using IMU. Then, we dive deeper into the performance analysis of our algorithm by analyzing it against the ground truth itself.

Evaluation Metrics. We employ two widely recognized metrics, Mean Absolute Error (MAE) and Pearson Correlation Coefficient (PCC), to evaluate the performance of our algorithm.

- MAE is a common metric used in respiration rate estimation studies [59, 82, 84]. It measures the average absolute value difference between the estimated respiration rates and the ground truth values. MAE is defined as:

$$MAE = \sum_{i=1}^{n} \frac{|BR_{gt_i} - BR_{est_i}|}{n} \qquad (1)$$

  Here, n is the number of samples, BR_{est} is the estimated respiration rate, and BR_{gt} is the ground truth respiration rate. A lower MAE indicates higher accuracy in estimation. In this study, we calculate the MAE by comparing the algorithm's estimated respiration rates to the ground truth data obtained from the Hexoskin smart shirts.

- PCC measures the linear correlation between BR_{gt} and BR_{est} [89]. PCC is an important metric for understanding the algorithm's performance to identify respiration rates for long sessions. PCC provides insight into how well the predictions follow the breathing patterns. PCC is defined as:

$$PCC = \frac{cov(BR_{gt}, BR_{est})}{std(BR_{gt})std(BR_{est})} \qquad (2)$$

  Here, std and cov refer to standard deviation and covariance, respectively. The PCC value ranges between -1 and $+1$ (1) A PCC closer to 1 indicates a strong positive linear relationship, (2) A PCC near 0 suggests no linear correlation, and (3) A PCC closer to -1 indicates a strong negative linear relationship. A higher PCC (closer to 1) signifies that the estimated respiration rates closely follow the patterns of the ground truth, demonstrating the algorithm's effectiveness in tracking respiration trends over time.

Comparison against SOTA IMU To Respiration Rate Detection Algorithms. We evaluate the performance of our respiration rate estimation algorithm against three state-of-the-art accelerometer-based methods: (1) BioWatch [39], (2) SleepMonitor [94], and (3) InstantRR [84]. Both BioWatch and SleepMonitor employ Fast Fourier Transform (FFT)-based algorithms on IMU data collected from smartwatches. InstantRR utilizes a combination of FFT-based algorithms, peak-detection algorithms, and Zero-Crossing Rate (ZCR) algorithms for different body positions, each with specific data processing techniques. As illustrated in Figure 8, our method outperforms BioWatch, SleepMonitor, and all three algorithmic variations of InstantRR. Specifically, our algorithm achieves 13-73% higher PCC and reduces MAE by 40-70% across various respiration rates. This significant improvement is attributed to our algorithm's robust performance in accurately estimating both slow and regular respiration rates.

---

[1]Codes and dataset can be found at https://bashlab.github.io/meditite_project/
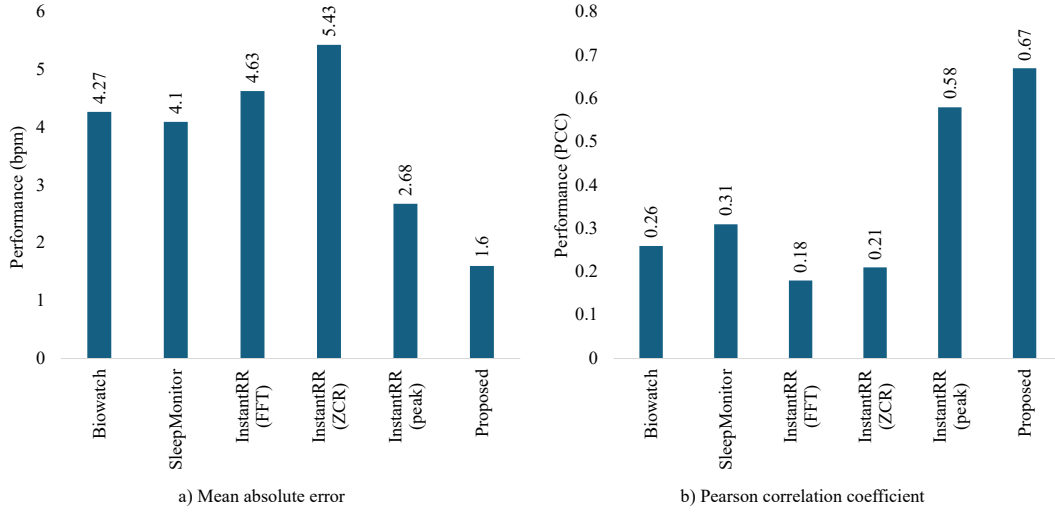
Fig. 8. Performance comparison between the proposed algorithm and baseline algorithms. The proposed algorithm demonstrates less MAE and high PCC compared to existing SOTA methods.
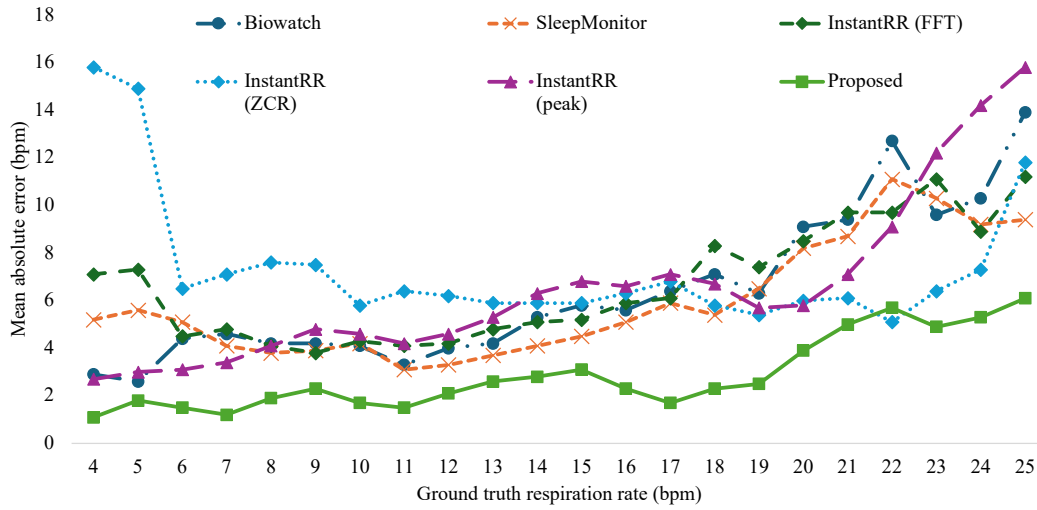

Fig. 9. The mean absolute error for different respiration rates shows that the proposed algorithm performs consistently over a wide range of respiration rates, including low respiration rate zones with a lower MAE compared to the SOTA methods.

Figure 9 illustrates the average performance of all algorithms across different respiration rates observed in study 1, aggregated over all participants. The results demonstrate that our algorithm consistently maintains a low MAE across the entire range of respiration rates, effectively detecting both slow and natural breathing patterns. Notably, our algorithm achieves less than a 2 BPM error up to 19 BPM. In contrast, the compared methods perform adequately within the 8–15 BPM range, but exhibit decreased accuracy at both low respiration
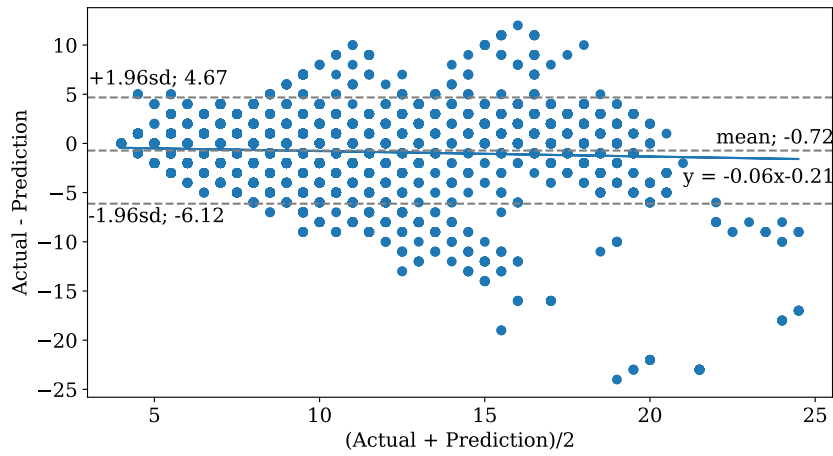
Fig. 10. Bland-Altman plot for respiration rate estimation using the proposed algorithm shows that most of the data are within the limits of agreement, indicating the effectiveness of our proposed algorithm.

rates—which are prevalent during mindfulness practices—and very high respiration rates. Their reliance on FFT and ZCR techniques, which have insufficient frequency resolution and increased susceptibility to noise, makes them less effective for slow-paced respiration. In contrast, our algorithm effectively captures slow breathing patterns through tailored noise reduction techniques and adaptive peak detection mechanisms. By addressing the challenges of low-amplitude signals and minimizing the impact of motion artifacts, our method provides more reliable respiration rate estimations in contexts where slow breathing is prevalent, such as mindfulness training.

Agreement Analysis Using Bland-Altman Plot. To further assess the agreement between our estimated respiration rates and the ground truth, we employ a Bland-Altman plot [18]. The Bland-Altman plot is a method used to analyze the agreement between two quantitative measurement techniques by plotting the difference between the methods against their mean. Figure 10 presents the Bland-Altman plot for our data. We plot a random sample of data points for clarity; however, the entire dataset is used to calculate mean differences and limits of agreement. Most of our algorithm's estimations are within the limits of agreement. Combining these findings with the high PCC reported earlier, we conclude that there is substantial agreement between our estimated respiration rates and the ground truth data from Hexoskin. This validates the accuracy and reliability of our algorithm in estimating respiration rates across various breathing patterns.

Ablation Study and Parameter Tuning. This section examines the impact of various parameters on the algorithm design, providing rational explanations for our design decisions.
Effect of Local Mean Removal. Figure 11 illustrates the rationale and effectiveness of applying local mean removal compared to global mean removal for motion artifact suppression. Global mean removal assumes stationarity over the entire signal, which is often invalid during respiration monitoring, especially in the presence of motion artifacts or shifts in baseline. In contrast, local mean removal dynamically adapts to temporal fluctuations, allowing for better isolation of the respiration component from transient noise and motion disturbances. Local mean removal significantly reduces respiration rate estimation error, validating its role in enhancing signal quality for reliable biosignal tracking in real-world scenarios.
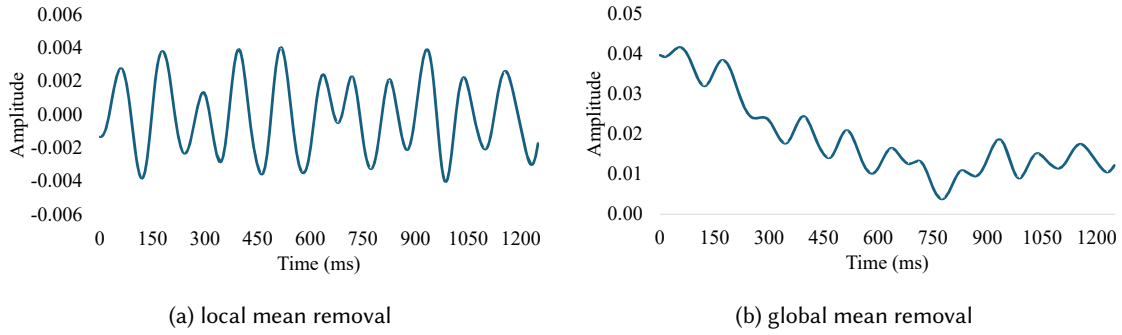
(a) local mean removal

(b) global mean removal

Fig. 11. Effect of applying a) local mean removal, b) global mean removal techniques on data preprocessing. Using the local mean removal can remove the motion artifacts more efficiently than using the global mean removal technique.

Effect of Prominence. The prominence parameter plays a crucial role in peak detection for accurate respiration rate estimation. Figure 12a illustrates how varying prominence values affect algorithm performance. Setting the prominence too high suppresses smaller, yet valid, respiration peaks, leading to missed detections and degraded performance. Conversely, a low prominence threshold allows excessively small-amplitude peaks, often caused by noise or motion artifacts, resulting in false detections. Through systematic fine-tuning, we identified an optimal prominence value of 50%, which balances sensitivity and specificity, achieving the best respiration rate estimation accuracy in our dataset.
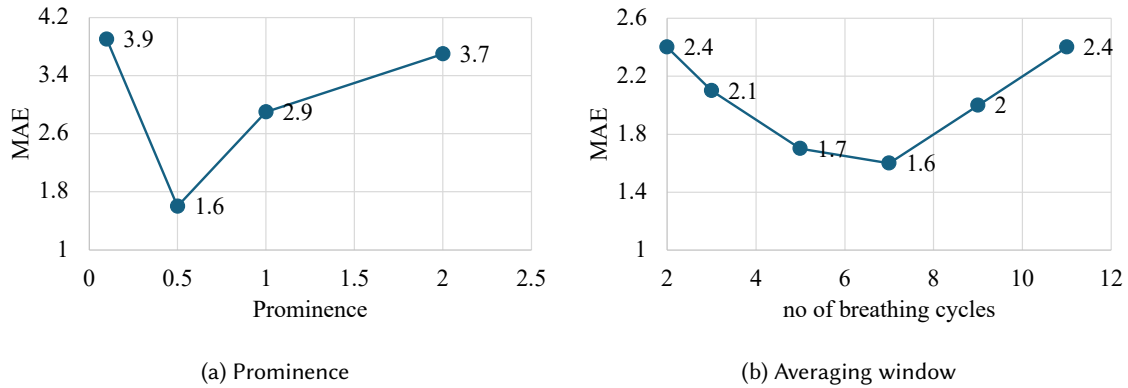

(a) Prominence

(b) Averaging window

Fig. 12. Effect of a) Prominence and b) Averaging window (breathing cycles) on the performance of the respiration rate algorithm allows us to select the optimized parameters to estimate the respiration rate with high accuracy.

Effect of Averaging Window (Breathing Cycles). Figure 12b illustrates how the number of breathing cycles used for averaging impacts respiration rate estimation performance. Using fewer cycles enables more immediate, responsive estimates, but increases susceptibility to errors from false peak detections. In contrast, averaging over a larger number of cycles smooths out these fluctuations, improving accuracy but introducing latency in the feedback. To balance responsiveness and stability, we empirically determined that averaging over seven breathing cycles provides optimal performance with minimal delay, ensuring reliable respiration rate feedback without compromising timeliness.

Effect of Multiplier on Data Reliability Assessment. To assess the reliability of sensor data, we compute the inter-quartile range (IQR) within each window and apply a multiplier to define the acceptance threshold. Specifically, we use a multiplier of 0.8, selected empirically to balance data retention and algorithm performance.

Figure 13 illustrates how varying the multiplier affects the retention rate of data windows. A smaller multiplier imposes a stricter reliability threshold, effectively filtering out noisy or unreliable data. However, this significantly reduces the retention rate, which is impractical for real-world applications as it limits the amount of usable data, potentially degrading system usability and user engagement.

Conversely, a larger multiplier relaxes the threshold, increasing the retention rate but allowing more noisy data to pass through. This can lead to a drop in the performance of both the respiration rate estimation and mindfulness skill assessment algorithms, negatively impacting user satisfaction due to less reliable feedback.

Through empirical evaluation, we identified 0.8 as the optimal multiplier, ensuring a reasonable retention rate without compromising the accuracy of feedback provided to users.

Parameter Sensitivity Analysis in Flat Surface Detection. Flat surface detection acts as an additional data reliability checkpoint, ensuring that sensor data is processed only when the phone is appropriately placed on the user's chest, not on unintended surfaces. This mechanism improves system usability by preventing unreliable data from influencing feedback. To develop and fine-tune the flat surface detection algorithm, we augmented our dataset with 20 sessions collected while the phone was placed on various flat surfaces (e.g., table, floor, book). We used 15 of these sessions along with 15 valid meditation sessions from Study 1 for training, and tested on the remaining 5 flat-surface sessions and 15 additional meditation sessions.

Figure 14, and 15 present the sensitivity analysis of the detection threshold and window size parameters, focusing on precision performance. Precision is prioritized to minimize false positives, ensuring that valid meditation sessions are not mistakenly classified as flat surface instances, which would otherwise result in data loss and reduced user satisfaction. Through this analysis, we identified an optimal balance between detection sensitivity and data retention, ensuring reliable flat surface identification without compromising the inclusion of valid sessions.
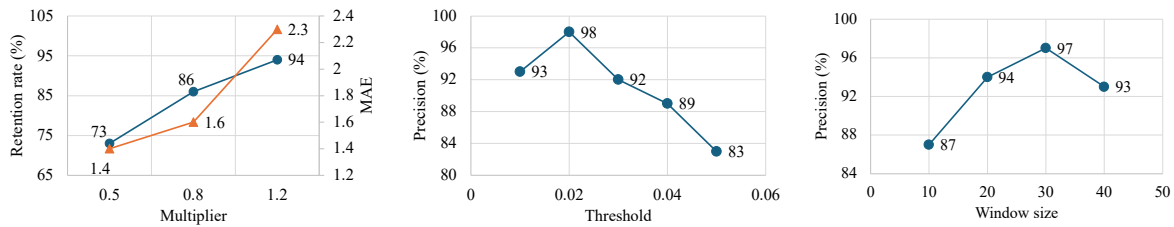

Fig. 13. Effect of multiplier on data reten-tion rate and performance of algorithm.

Fig. 14. Effect of threshold on the perfor-mance of detecting flat surface.

Fig. 15. Effect of window-size on the per-formance of detecting flat surface.

Discussion. Previous works on breathing rate detection mainly target natural or patient breathing, where chest movements are more prominent, making detection easier. Natural breathing rates (12–20 bpm [76]) and pulmonary conditions [84] often produce strong signals, allowing frequent peak detection in accelerometer data. However, focusing solely on high breathing rates risks losing low-frequency inhale/exhale peaks during filtering. As prior algorithms were optimized for natural rates, their performance degrades at low breathing rates. In contrast, our algorithm is specifically designed for low breathing rates, with parameters fine-tuned to capture peaks during both low and natural breathing. Additionally, careful selection of prominence to capture small amplitude peaks and use of an averaging window to post-process the respiration rate ensures high performance
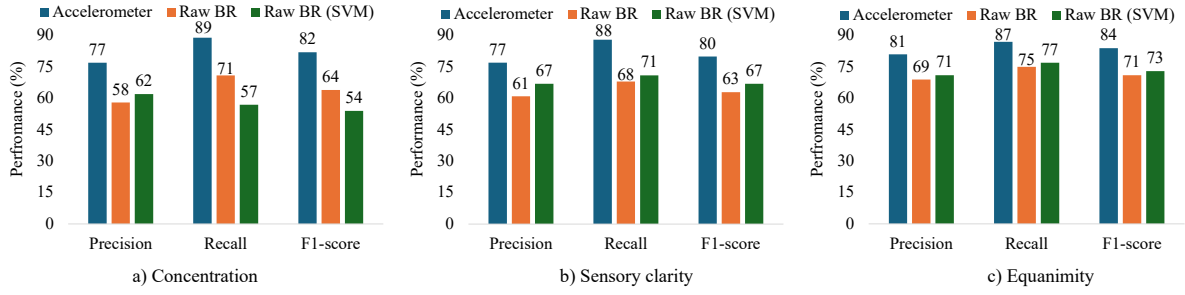
Fig. 16. Performance comparison of smartphone accelerometer with estimated respiration for mindfulness change estimation. Using accelerometer data over raw respiration rate performs better in all three mindfulness change predictions.

in detecting low breathing rates. When a user has a low breathing rate, actual peaks occur after a long duration (15s for 3 bpm). Any instantaneous peaks during this period will risk the accuracy of the algorithms. The use of an averaging window helps us to mitigate this erroneous peak detection and maintain high performance in a low-breathing-rate zone.

Implications and Conclusion. Overall, these results confirm that our algorithm not only maintains low MAE across different respiration rates but also shows strong agreement with established measurement techniques. This underscores its effectiveness in providing accurate respiration rate estimations, which is critical for applications like mindfulness training, where both slow and natural breathing patterns are prevalent.

8.2 Mindfulness Progress Estimation Algorithm

We assess the proposed algorithm's performance by comparing its effectiveness against other network architectures and against other data modalities.

Evaluation Metrics. We evaluate the proposed algorithm using three popular metrics: precision, recall, and F1 score [17, 47, 48, 105].

- Precision measures how many of the predicted positives are truly positive. Precision is defined as:

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

  Here, TP and FP stand for true positive and false positive, respectively.
- Recall measures how many of the actual positives the model correctly identified. Recall is defined as:

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

  Here, FN refers to false negatives.
- F1−score is the harmonic mean of precision and recall, balancing both to give a single performance measure, particularly useful in cases of imbalanced data. F1 score is calculated as:

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{5}$$

Comparison with Ground Truth. Figure 16 illustrates the results of our proposed model against the self-reported ground truth. Our model achieves F1-Scores of 82% for concentration, 80% for sensory clarity, and 84% for equanimity. We observe that the change in equanimity is predicted more accurately by the model than the
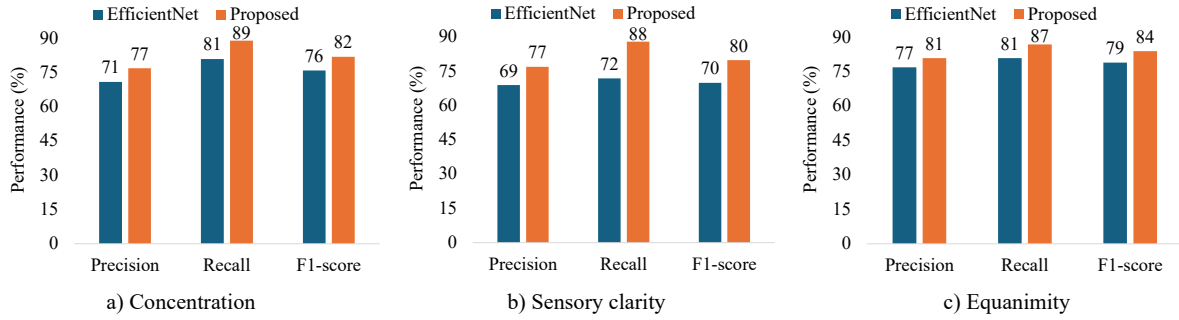
Fig. 17. Performance comparison of various models for estimating change in mindfulness skills. Our proposed model demonstrated superior performance with 80%-84% F1 score, outperforming others and highlighting its effectiveness in capturing mindfulness skill progression.

other two skills. This higher accuracy in predicting equanimity may be attributed to the distinct physiological and behavioral markers associated with this skill during mindfulness practice. Equanimity involves a balanced state of emotional regulation and reduced reactivity to stressors, which may manifest as steadier breathing patterns and minimal physical movement—signals that are effectively captured by accelerometer data. The model's ability to detect these subtle cues underscores the effectiveness of our model to monitor mindfulness skill progression.

Comparison of Data Modalities. We compare our DL model for estimating mindfulness skill improvements using two different data sources: raw accelerometer data from smartphones and true respiration rate data from Hexoskin smart shirts. The inclusion of Hexoskin data serves a crucial purpose—we want to eliminate any potential errors arising from respiration rate estimation using raw accelerometer data. By using accurate respiration measurements from the Hexoskin Smart shirts, we aim to understand the true effectiveness of our algorithm without the confounding effects of estimation inaccuracies. The models are trained on data from study 1, and inference is conducted on an independent set of data from study 3 (see table 1). For each modality, we explore various network architectures and report the best results. We choose to compare our accelerometer-based model with models using true respiration rates because respiration rate is a well-established physiological indicator associated with mindfulness and relaxation states. By including respiration rate as a baseline modality, we aim to assess whether accelerometer data can provide additional predictive power over this traditional measure. Comparing the two modalities allows us to determine the extent to which accelerometer data captures relevant information for mindfulness skill estimation beyond what is offered by respiration rates alone.

For the accelerometer modality, we employ our proposed ResNet-1D architecture combined with a GRU. The raw accelerometer data provides a rich and dense dataset with ample data points per segment, making it well-suited for this DL approach. In contrast, the true respiration rate data from the Hexoskin smart shirts has significantly fewer data points in each segment, which precludes the use of the same architecture. To address this, we experiment with two alternative network configurations for the respiration rate modality: (1) a model consisting of one convolutional layer followed by a GRU layer and a fully connected layer, and (2) a model with two GRU layers followed by two fully connected layers. We find that the model with two GRU layers performed better, and thus, we report the results of this model.

Additionally, we investigate commonly used statistical features of the respiration rate, including the mean respiration rate during meditation, standard deviation, kurtosis, skewness, entropy, and the longest durations at maximum and minimum respiration rates. Using these features extracted from the Hexoskin data, we employ a Support Vector Machine (SVM) to estimate changes in mindfulness skills.

Figure 16 illustrates that the model utilizing accelerometer data achieves a 15.4%–21.9% higher F1 score than the DL model using true respiration rates and a 13.1%–34.1% higher F1 score than the SVM model using statistical features derived from respiration rate data. These results demonstrate the superior performance of the accelerometer modality. The accelerometer data not only captures respiration patterns but also encapsulates subtle bodily movements and postural adjustments that are relevant to mindfulness practices. This richer dataset provides additional contextual information indicative of changes in mindfulness skills, leading to more accurate predictions than models using respiration rates alone. The accelerometer inputs offer a more comprehensive representation, enhancing the model's ability to detect changes in mindfulness skills.

Additionally, the performance of the SVM model with the statistical feature set demonstrates the quantitative relationship between the respiration biosignal and changes in mindfulness skills. However, the SVM model lacks the capability to capture temporal dynamics in the data, and the limited data points per segment due to lower sampling rates constrain the deep learning model using respiration rate data. In contrast, our proposed DL model using accelerometer data effectively captures both spatial and temporal patterns, resulting in better performance than classical machine learning algorithms like SVM and deep learning models constrained by limited data.

Assessment of Network Architecture Choices. We further evaluate the effectiveness of our proposed network architecture by comparing it with EfficientNet, a convolutional neural network (CNN) that has been used in medical image classification and EEG analysis [1, 107]. To train EfficientNet on our dataset, we convert the same two-minute accelerometer segments used in our model into mel-spectrograms [16, 31]. This conversion allows us to represent the time-series accelerometer data in a format suitable for EfficientNet, which is designed for image-like inputs.

As shown in Figure 17, our proposed model outperforms EfficientNet by achieving a 5.9%–12.5% higher F1 score. While the performance of EfficientNet demonstrates that accelerometer data contains sufficient information to identify changes in mindfulness skills, our proposed model achieves better results because it integrates Convolutional Neural Network (CNN) layers with a Gated Recurrent Unit (GRU), effectively capturing both spatial and temporal dependencies in the data.

The superior performance of our model can be attributed to its ability to capture temporal information associated with slow-paced respiration rates, which is crucial for accurately predicting improvements in mindfulness skills. While EfficientNet excels at extracting spatial features from image-like data, it may not effectively model the temporal sequences inherent in accelerometer data related to mindfulness practices. In contrast, our model's architecture is designed to handle time-series data, leading to better overall performance in this context.

Implication and Conclusion. Our approach is grounded in both empirical evidence and theoretical frameworks linking physiological signals to mindfulness components. Our deep learning model utilizes accelerometer data captured during meditation sessions to predict changes in mindfulness skills. This choice is informed by research indicating that respiration patterns, detectable via accelerometry, are closely associated with mindfulness states. Specifically, slow and controlled breathing has been shown to activate the parasympathetic nervous system, promoting relaxation and reducing anxiety. These physiological changes are integral to mindfulness practices [69].

Furthermore, our model captures nuanced patterns in the accelerometer data, including breathing rate, variability, and motion artifacts, which reflect potential distractions or shifts in attention. By analyzing these features, the model differentiates between various aspects of mindfulness: concentration, sensory clarity, and equanimity. This aligns with Shinzen Young's Unified Mindfulness framework [109], which defines mindful awareness as the integration of these three components.

Overall, these results highlight the robustness of our approach to estimating changes in mindfulness skills using readily available smartphone data. By accurately predicting improvements in equanimity, concentration, and sensory clarity, our model demonstrates significant potential for facilitating personalized mindfulness training through feedback. This approach paves the way for scalable, accessible mental health interventions that leverage

everyday technology. By specifically comparing the accelerometer-based model with models using true respiration rates, statistical features, and alternative network architectures like EfficientNet, we establish that accelerometer data provides a more comprehensive and informative input for predicting mindfulness skill improvements. The accelerometer captures nuanced bodily movements and temporal patterns that respiration rate data alone cannot fully represent. Moreover, our tailored DL architecture effectively harnesses this rich data, outperforming architectures like EfficientNet that are not optimized for time-series data. Our findings suggest that utilizing accelerometer data in conjunction with advanced DL architectures can significantly enhance the detection of mindfulness skill progression, offering a valuable tool for personalized mental health support.

9 USER STUDY

In this section, we examine how integrating a respiration biosignal feedback chart into our mindfulness training application influences both system usability and mindfulness skill development. Usability was assessed using the System Usability Scale (SUS), while changes in mindfulness skills were evaluated through session-level feedback.

Participants were randomly assigned to one of two groups: (1) the biosignal-augmented group, which received respiration biofeedback after each session, and (2) the control group, which did not receive any feedback post-session.

The SUS scores and mindfulness skill changes were analyzed using data from Studies 2 and 3, which together included 32 participants. User engagement—measured by the number of completed sessions and total meditation minutes—was evaluated only in Study 3, which involved 22 participants who self-regulated their meditation practice. In contrast, participants in Studies 1 and 2 were required to complete all sessions, limiting natural variability in engagement behavior.

Details of each study's design are provided in Table 1, and participant demographics are described in Section 4.

9.1 Number of Participants selection from Power Analysis

We select the number of participants for our user study by conservatively using the effect size of previous studies. There is a wide range of studies that examine the effects of adding gamification or user support features to digital mental health apps on user engagement and performance outcomes [23, 96, 108] . As might be expected, there's a wide range of effect sizes for these outcomes (d's 0.37-1.62). No previous studies to our knowledge have examined the role of respiration feedback to help us more precisely estimate an effect size, but Tsay et al. show that performance is significantly better in their gamified learning courses relative to the nongamified condition (Cohen's d=1.51) and engagement is positively related to performance [98]. From this previous literature, we conservatively estimate an effect size of d=1.32 on user engagement in the respiration biosignal feedback augmented condition relative to the standard no feedback condition. A power analysis using G*Power [46] for this independent samples t-test, assuming power 0.8 and two-tailed p=.05, indicates that a sample size of at least 22 participants is needed. For our user engagement observation, we use a sample size of 22, and in other studies, we use a sample size of 32, combining the participants from the lab.

9.2 System Usability Scale

We assess application usability using the System Usability Scale (SUS) [95] during the post-study debriefing questionnaires of all studies. The SUS prompted participants to evaluate the usability of the mindfulness meditation app with statements such as: "I think that I would like to use the app frequently," "I found the app unnecessarily complex," and "I felt confident using the app." Responses are given on a scale from 1 (strongly disagree) to 5 (strongly agree). A score of 68 is considered average, while a score of 80 corresponds to an "A" rating [57, 79].

We hypothesize that the biosignal augmented condition (participants who are exposed to the respiration biosignal feedback) would yield higher SUS scores in comparison to the control condition. We consider the

(a) Average change in mindfulness skills

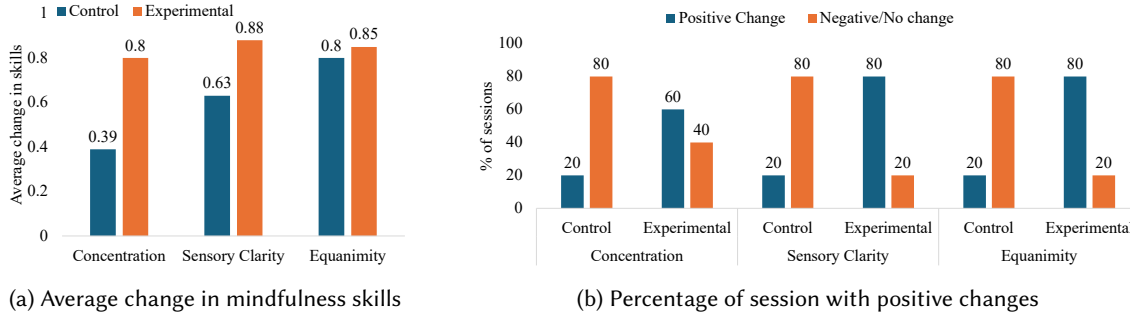(b) Percentage of session with positive changes

Fig. 18. Impact of respiration biosignal feedback shows higher skill change and higher number of sessions with positive changes in the experimental group than the control group.

participants from studies 2 and 3 for the SUS score. In total, scores from 32 participants (16 participants in the biosignal augmented condition and 16 participants in the control condition) are used to prove our hypothesis. An independent samples t-test (N=32) reveals that participants in the biosignal augmented condition report significantly higher SUS scores (mean 84.43, SE 1.80) in comparison to those in the control condition (mean 74.04, SE 3.49; $t = 2.649$; $P = 0.013$; Cohen's $d = 0.83$; 95% CI = [2.37,18.43]).

The SUS analysis demonstrates statistically significant results with a p-value <0.05 and a high effect size of 0.83. It implies that incorporating respiration biosignal feedback into the app enhances user satisfaction and perceived usability. Participants reported higher satisfaction and confidence when using the app with biosignal feedback, suggesting that this feature contributes positively to the overall user experience. Improved usability is crucial for sustained engagement with mindfulness applications, ultimately supporting better mindfulness skill development, which has been shown to improve mental health outcomes.

In-Lab Study Results. Ten participants from study 2 who completed their sessions in the lab are divided into experimental and control conditions. An independent samples t-test using all available data collected in lab environment (study 2) revealed significantly higher user satisfaction in the biosignal augmented condition on the SUS (Mean=85.0, SE=2.11) relative to control condition on the SUS (Mean=78.57, SE=1.99), $t = -2.21$, $p = .04$, Cohen's $d = 0.67$, 95% confidence interval (CI) = [-1.91, - 0.04]. While the sample size is small, the effect size (Cohen's d) indicates that participants may be more satisfied with their training experience when exposed to respiration biosignal feedback.

In-The-Wild Study Results. Similar to the lab study, an independent samples t-test using all available data from study 3 revealed significantly higher user satisfaction in the biosignal augmented condition on the SUS (Mean=84.36, SE=3.83) relative to control condition on the SUS (Mean=67.08, SE=8.52), $t = 2.02$, $p = 0.020$, Cohen's $d = 1.04$, 95% confidence interval (CI) = [-1.31, 6.96]. The result is statistically significant with a p-value < 0.05 and has a high effect size of 1.04 (Cohen's d), which indicates that exposure to respiration biosignal feedback increases participants' satisfaction in the wild.

9.3 Impact on Mindfulness Skill Change

We evaluate the impact of the respiration biosignal feedback chart on changes in mindfulness skills following meditation. Figure 18a shows the average change in mindfulness skills for both the control and experimental conditions. Participants who receive respiration biosignal feedback exhibit up to 2 times greater improvement in mindfulness skills compared to the control group. In addition, Figure 18b indicates that the experimental group experiences more sessions with a positive change in mindfulness skills. These findings suggest that respiration

a) Average number of completed sessions

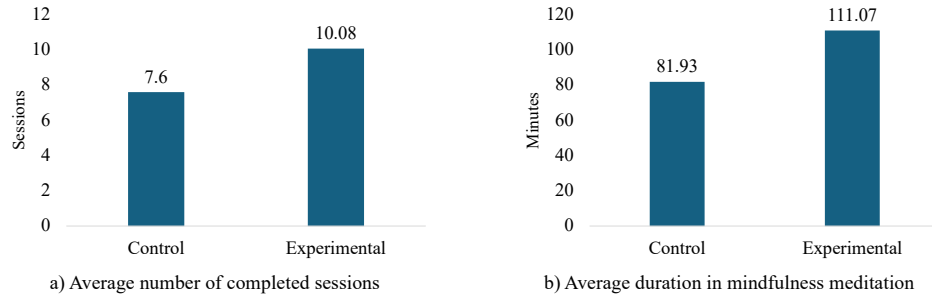b) Average duration in mindfulness meditation

Fig. 19. Participants in the biosignal-augmented condition exhibited higher engagement with the app, completing more sessions on average and spending more time in meditation.

biosignal feedback not only enhances app usability but also promotes ongoing improvements in mindfulness skills as sessions progress.

Additionally, we conduct a t-test on the average change in mindfulness skills and a chi-square test on the number of positive changes in mindfulness skills due to its categorical nature. The average change in concentration and sensory clarity is highly significant with a large effect size, resulting in a p-value of 0.0006 and 0.026, respectively. However, the average change in equanimity is not statistically significant, with a p-value of 0.64 and Cohen's d of 0.17. The change in concentration and sensory clarity also has a large effect size with Cohen's d of 1.37 and 0.83, respectively, while the change in equanimity has a low effect size of 0.17.

On the contrary, the chi-square test shows that the number of sessions with positive changes in skills is highly significant for all three skills of concentration, sensory clarity, and equanimity, with a p-value of 1.8×10^{-8}, 7.2×10^{-7}, and 7.2×10^{-7} and a Cohen's h of -0.86, -1.29, -1.29, respectively, showing large effect size.

9.4 User Engagement

In addition to assessing system usability, we also explore user engagement by providing an experimental group of users with respiration biosignal feedback. Because all participants in Studies 1 and 2 were required to visit the lab, we do not track their engagement data, as it would not accurately reflect real-world usage patterns. As mentioned in section 4, participants from study 3 are divided into a biosignal augmented condition and a control group, and they can complete as many sessions as they want in 21 days.

As shown in Figure 19, participants in the in-the-wild study who received biosignal-augmented feedback completed an average of 10.08 meditation sessions, while the control group averaged 7.60 sessions. Those in the biosignal-augmented condition spent more time meditating, averaging 111.07 minutes, while the control group averaged 81.93 minutes. These trends highlight the potential of respiration biosignal feedback to enhance engagement and encourage longer meditation sessions. A t-test (N=22) proves the results are approaching significance, with a p-value of 0.11. The effect size is medium to large with a Cohen's d of 0.71. A larger longitudinal study with more participants will reinforce our hypothesis.

10 DISCUSSION, LIMITATION, AND FUTURE WORKS

This section highlights several key limitations and areas for improvement in our study, including using a custom mindfulness scale and binary classification for skill detection. We also address challenges involving sample size, device dependence, and environmental artifacts, all of which underscore the need for further investigation and methodological refinements.

10.1 Scale for Mindfulness Skills

One limitation of our study is the use of a custom 1–7 scale for participants to self-report their current levels of concentration, sensory clarity, and equanimity. While this scale provided a concise method for capturing mindfulness skills, it differs from standard validated scales commonly used in mindfulness research. Due to the necessary brevity of our self-report questionnaires and considering that previous studies [61, 64] indicate increases in self-reported mindfulness measures even in control conditions not exposed to mindfulness training [77], we opt for this concise scale. In future work, we plan to incorporate more widely used and validated scales, such as the Mindful Attention Awareness Scale (MAAS) [68] and the Toronto Mindfulness Scale (TMS) [55], to enhance the validity and comparability of our findings.

10.2 Ground Truth and Practicality of Mindfulness Assessment

In this study, we use self-reported scores as the ground truth for training our mindfulness skills estimation model. While self-report measures are subject to potential biases such as mood variability or subjective interpretation, they remain the most practical and widely accepted method for evaluating mindfulness experiences in large-scale, real-world studies.

More objective physiological assessments—such as EEG or fMRI—can offer higher precision in controlled laboratory settings. However, these modalities are often cost-prohibitive, lack scalability, and may disrupt the natural meditation experience. As such, they are not well-suited for mobile, in-the-wild mindfulness interventions. Our approach aims to balance ecological validity and practical deployment by leveraging self-reports alongside non-intrusive, smartphone-based sensing methods.

We acknowledge that self-reported scores alone may not fully capture the complexity of mindfulness. To address this, future work will explore the integration of additional physiological markers, such as heart rate variability (HRV) and galvanic skin response (GSR), to enable multi-modal validation. These extensions will help enhance the robustness and generalizability of our mindfulness assessment framework in naturalistic settings.

10.3 Granularity of Mindfulness Skill Prediction

Our proposed model currently focuses on binary classification— predicting whether mindfulness skills are improving or not during training. This binary approach, while useful, may not capture the full spectrum of changes in mindfulness skills. In future work, we aim to extend the model to a multi-class classification framework that can quantify varying degrees of change, such as large positive change, low positive change, no change, and negative change in concentration, sensory clarity, and equanimity. By providing more granular predictions, we can offer users a deeper understanding of their mindfulness progression, potentially enhancing the personalization and effectiveness of the training.

10.4 Limited Scope of Respiratory Biosignals

While our proposed model for estimating respiration rate achieves an MAE of 1.6 BPM, outperforming state-of-the-art models, it focuses solely on respiration rate estimation. Future research will aim to develop a more robust model to further improve the accuracy of respiration rate estimation, particularly under varying conditions. Moreover, we plan to expand our focus to estimate additional respiratory biomarkers, such as minute ventilation, inhalation and exhalation durations, and respiratory variability. These biomarkers are vital for comprehensive monitoring of user health and well-being [12, 80], and their inclusion could enhance the biofeedback provided to users during mindfulness training.

10.5 Exploring Alternative Modalities for Respiration and Mindfulness Estimation

While this study leverages smartphone accelerometer data to estimate respiration rate and track changes in mindfulness skills during meditation, other physiological sensing modalities—such as EEG, audio, and galvanic skin response (GSR)—were not explored. Recent work has demonstrated the feasibility of estimating respiration from in-ear audio signals [6, 42, 92]. However, capturing respiration using external microphones presents additional challenges, particularly in mindfulness contexts where breathing is slow, soft, and often inaudible. This makes external audio sensing less reliable for detecting low breathing rates in real-world scenarios.

While in-ear microphones may have the capability to capture subtle breathing signals, microphone data access from commercial earbuds (e.g., Samsung Galaxy Buds, Apple AirPods) is restricted. Moreover, although these devices also include embedded IMUs, their sensor data are not publicly accessible—limiting their applicability for research and large-scale deployment. Similarly, EEG and GSR, though capable of providing rich physiological insights, require external sensors and more complex setups, reducing their practicality for everyday use.

In future work, we aim to explore the potential of external microphone audio to estimate both respiration rate and mindfulness skills, while carefully evaluating its limitations under low-breathing-rate conditions. These efforts will help determine whether complementary sensing modalities can enhance estimation accuracy without compromising usability or scalability.

10.6 Long-Term Engagement and Future Directions

Our findings indicate that integrating the respiration feedback module enhances user engagement, as evidenced by an increase in both the number of meditation sessions and the total minutes spent meditating. These results suggest that providing real-time respiration biofeedback and mindfulness skill estimation may positively influence user commitment to mindfulness practices. While this trend is promising and statistically significant, the current study was conducted over a 21-day period and represents an initial step toward understanding engagement dynamics.

To evaluate the sustainability of this effect, a longitudinal study is necessary. In future work, we plan to conduct an extended 6–12 month deployment with a larger and more diverse participant pool. This will allow us to assess long-term user engagement and determine whether the initial increase in usage is maintained over time. Additionally, future studies will incorporate more controlled experimental designs to better isolate the impact of biosignal feedback from other confounding variables. These efforts will help strengthen the evidence for the role of physiological feedback in enhancing digital mindfulness interventions and inform the design of scalable, personalized mindfulness tools.

10.7 Sample Size and Participant Diversity

A limitation of our study is the relatively small sample size, which includes only 261 sessions. While our findings are promising, the limited number of participants may affect the generalizability of the results. Future studies should include a larger sample to validate the findings across different populations, ages, and backgrounds.

10.8 Device and Sensor Limitations

Our respiration rate estimation algorithm and mindfulness skill detection model are developed using data collected from specific devices, namely the iPhone X and Hexoskin smart shirts. The reliance on these specific devices may limit the generalizability of our methods to other smartphones or wearable devices with different sensor characteristics. Future work should explore the applicability of our algorithms across a wider range of devices and platforms to enhance their usability and accessibility.

10.9 Environmental and Movement Artifacts

Although our algorithm includes steps for motion artifact removal, continuous user movement or improper device placement can still impact the accuracy of respiration rate estimation and mindfulness skill detection. In real-world settings, uncontrolled environmental factors and varying levels of user compliance may introduce additional noise and artifacts. Future research should focus on enhancing the robustness of the algorithms to handle such variability, possibly through adaptive filtering techniques or machine learning models that can account for diverse real-world conditions.

11 CONCLUSION

This study demonstrates the significant potential of integrating respiration biosignal feedback into smartphone-based mindfulness training apps. Our respiration rate estimation algorithm outperformed existing methods in both accuracy and reliability, particularly excelling at detecting slow respiration rates typical during mindfulness meditation. By maintaining a low MAE of 1.6 for various respiration rates, our algorithm ensures effective biofeedback, which is crucial for mindfulness practices. Additionally, our deep learning model for mindfulness skill estimation using accelerometer data achieved robust performance, with F1 scores of 82% for concentration, 80% for sensory clarity, and 84% for equanimity. Our model surpassed other models that are based solely on respiration rate, highlighting the value of accelerometer data in capturing subtle physiological and behavioral cues associated with mindfulness skills. Participants who receive biosignal feedback reported higher satisfaction with using the app. These findings suggest that incorporating biosignal feedback, combined with accelerometer-based mindfulness estimation, can significantly enhance the effectiveness of digital mindfulness interventions. By harnessing ubiquitous technology and sophisticated data analysis, we can create effective, scalable solutions to support mindfulness practice and mental health in diverse populations.

Acknowledgments

References

[1] Deevyankar Agarwal, Manuel Álvaro Berbís, Antonio Luna, Vivian Lipari, Julien Brito Ballester, and Isabel de la Torre-Díez. 2023. Automated medical diagnosis of Alzheimer´s disease using an efficient net convolutional neural network. *Journal of Medical Systems* 47, 1 (2023), 57.

[2] Asieh Ahani, Helane Wahbeh, Hooman Nezamfar, Meghan Miller, Deniz Erdogmus, and Barry Oken. 2014. Quantitative change of EEG and respiration signals during mindfulness meditation. *Journal of neuroengineering and rehabilitation* 11, 1 (2014), 1–11.

[3] Mohsin Y Ahmed, Tousif Ahmed, Md Mahbubur Rahman, Zihan Wang, Jilong Kuang, and Alex Gao. 2022. Deep Audio Spectral Processing for Respiration Rate Estimation from Smart Commodity Earbuds. In *2022 IEEE-EMBS International Conference on Wearable and Implantable Body Sensor Networks (BSN)*. IEEE, 1–4.

[4] Mohsin Y Ahmed, Md Mahbubur Rahman, Viswam Nathan, Ebrahim Nemati, Korosh Vatanparvar, and Jilong Kuang. 2019. mlung: Privacy-preserving naturally windowed lung activity detection for pulmonary patients. In *2019 IEEE 16th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*. IEEE, 1–4.

[5] Tousif Ahmed, Md Mahbubur Rahman, Mohsin Yusuf Ahmed, Ebrahim Nemati, Minh Dinh, Nathan Folkman, Jilong Kuang, and Alex Gao. 2021. Rrmonitor: A resource-aware end-to-end system for continuous monitoring of respiration rate using earbuds. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2463–2467.

[6] Tousif Ahmed, Md Mahbubur Rahman, Ebrahim Nemati, Mohsin Yusuf Ahmed, Jilong Kuang, and Alex Jun Gao. 2023. Remote breathing rate tracking in stationary position using the motion and acoustic sensors of earables. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–22.

[7] Farah Q AL-Khalidi, Reza Saatchi, Derek Burke, H Elphick, and Stephen Tan. 2011. Respiration rate monitoring methods: A review. *Pediatric pulmonology* 46, 6 (2011), 523–529.

[8] Micah Allen, Somogy Varga, and Detlef H Heck. 2023. Respiratory rhythms of the predictive mind. *Psychological Review* 130, 4 (2023), 1066.

[9] Mona Alnaggar, Ali I Siam, Mohamed Handosa, Tamer Medhat, and MZ Rashad. 2023. Video-based real-time monitoring for heart rate and respiration rate. *Expert Systems with Applications* 225 (2023), 120135.

[10] John R Anderson, C Franklin Boyle, and Brian J Reiser. 1985. Intelligent tutoring systems. *Science* 228, 4698 (1985), 456–462.

[11] Sean Bae, Silviu Borac, Yunus Emre, Jonathan Wang, Jiang Wu, Mehr Kashyap, Si-Hyuck Kang, Liwen Chen, Melissa Moran, Julie Cannon, et al. 2022. Prospective validation of smartphone-based heart rate and respiratory rate measurement algorithms. *Communications medicine* 2, 1 (2022), 40.

[12] Rummana Bari, Md Mahbubur Rahman, Nazir Saleheen, Megan Battles Parsons, Eugene H Buder, and Santosh Kumar. 2020. Automated detection of stressful conversations using wearable physiological and inertial sensors. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 4, 4 (2020), 1–23.

[13] Subhadeep Basu and Samiul Mamud. 2020. Comparative study on the effect of order and cut off frequency of Butterworth low pass filter for removal of noise in ECG signal. In *2020 IEEE 1st International Conference for Convergence in Engineering (ICCE)*. IEEE, 156–160.

[14] Amit Baumel, Frederick Muench, Stav Edan, and John M Kane. 2019. Objective user engagement with mental health apps: systematic search and panel-based usage analysis. *Journal of medical Internet research* 21, 9 (2019), e14567.

[15] Luciano Bernardi, Alessandra Gabutti, Cesare Porta, and Lucia Spicuzza. 2001. Slow breathing reduces chemoreflex response to hypoxia and hypercapnia, and increases baroreflex sensitivity. *Journal of hypertension* 19, 12 (2001), 2221–2229.

[16] Subrata Biswas, Mohammad Nur Hossain Khan, Alex Colwell, Jack Adiletta, and Bashima Islam. 2023. LOCUS: LOcalization with Channel Uncertainty and Sporadic Energy. *arXiv preprint arXiv:2302.09409* (2023).

[17] Subrata Biswas, Mohammad Nur Hossain Khan, and Bashima Islam. 2025. QUADS: QUAntized Distillation Framework for Efficient Speech Language Understanding. *arXiv preprint arXiv:2505.14723* (2025).

[18] J Martin Bland and DouglasG Altman. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *The lancet* 327, 8476 (1986), 307–310.

[19] Kirk Warren Brown, Richard M Ryan, and J David Creswell. 2007. Addressing fundamental questions about mindfulness. *Psychological Inquiry* 18, 4 (2007), 272–281.

[20] Soujanya Chatterjee, Md Mahbubur Rahman, Tousif Ahmed, Nazir Saleheen, Ebrahim Nemati, Viswam Nathan, Korosh Vatanparvar, and Jilong Kuang. 2020. Assessing severity of pulmonary obstruction from respiration phase-based wheeze-sensing using mobile sensors. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.

[21] Xinli Chi, Ai Bo, Tingting Liu, Peichao Zhang, and Iris Chi. 2018. Effects of mindfulness-based stress reduction on depression in adolescents and young adults: A systematic review and meta-analysis. *Frontiers in psychology* 9 (2018), 361432.

[22] Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, and Yoshua Bengio. 2014. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555* (2014).

[23] Robert Ciuchita, Jonas Heller, Sarah Köcher, Sören Köcher, Thomas Leclercq, Karim Sidaoui, and Susan Stead. 2023. It is really not a game: An integrative review of gamification for service research. *Journal of Service Research* 26, 1 (2023), 3–20.

[24] Christopher E Cox, John A Gallis, Maren K Olsen, Laura S Porter, Tina Gremore, Jeffrey M Greeson, Cynthia Morris, Marc Moss, and Catherine L Hough. 2024. Mobile Mindfulness Intervention for Psychological Distress Among Intensive Care Unit Survivors: A Randomized Clinical Trial. *JAMA internal medicine* (2024).

[25] Holger Cramer, Helen Hall, Matthew Leach, Jane Frawley, Yan Zhang, Brenda Leung, Jon Adams, and Romy Lauche. 2016. Prevalence, patterns, and predictors of meditation use among US adults: A nationally representative survey. *Scientific reports* 6, 1 (2016), 36760.

[26] J David Creswell. 2017. Mindfulness interventions. *Annual review of psychology* 68 (2017), 491–516.

[27] J David Creswell and Emily K Lindsay. 2014. How does mindfulness training affect health? A mindfulness stress buffering account. *Current directions in psychological science* 23, 6 (2014), 401–407.

[28] J David Creswell, Emily K Lindsay, Daniella K Villalba, and Brian Chin. 2019. Mindfulness training and physical health: mechanisms and outcomes. *Psychosomatic medicine* 81, 3 (2019), 224–232.

[29] Gene Crumley and Howard Schutz. 2011. Short-duration mindfulness training with adult learners. *Adult Learning* 22, 2 (2011), 37–42.

[30] Ruixuan Dai, Chenyang Lu, Michael Avidan, and Thomas Kannampallil. 2021. Respwatch: Robust measurement of respiratory rate on smartwatches with photoplethysmography. In *Proceedings of the international conference on internet-of-things design and implementation*. 208–220.

[31] Steven Davis and Paul Mermelstein. 1980. Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE transactions on acoustics, speech, and signal processing* 28, 4 (1980), 357–366.

[32] Esther I de Bruin, Maurice Topper, Jan GAM Muskens, Susan M Bögels, and Jan H Kamphuis. 2012. Psychometric properties of the Five Facets Mindfulness Questionnaire (FFMQ) in a meditating and a non-meditating sample. *Assessment* 19, 2 (2012), 187–197.

[33] Jérémy Frey, May Grabli, Ronit Slyper, and Jessica R Cauchard. 2018. Breeze: Sharing biofeedback through wearable technologies. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–12.

[34] Kaustubh Manik Gaikwad and Mahesh Shrikant Chavan. 2014. Removal of high frequency noise from ECG signal using digital IIR butterworth filter. In *2014 IEEE Global Conference on Wireless Computing & Networking (GCWCN)*. IEEE, 121–124.

[35] Roderik JS Gerritsen and Guido PH Band. 2018. Breath of life: the respiratory vagal stimulation model of contemplative activity. *Frontiers in human neuroscience* 12 (2018), 397.

[36] Harish Haresamudram, Irfan Essa, and Thomas Plötz. 2021. Contrastive predictive coding for human activity recognition. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 2 (2021), 1–26.

[37] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 770–778.

[38] Javier Hernandez, Yin Li, James M Rehg, and Rosalind W Picard. 2015. Cardiac and respiratory parameter estimation using head-mounted motion-sensitive sensors. *EAI Endorsed Transactions on Pervasive Health and Technology* 1, 1 (2015), e2–e2.

[39] Javier Hernandez, Daniel McDuff, and Rosalind W Picard. 2015. Biowatch: estimation of heart and breathing rates from wrist motions. In *2015 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)*. IEEE, 169–176.

[40] Xiwei Huang, Dandan Xu, Jin Chen, Jixuan Liu, Yangbo Li, Jing Song, Xing Ma, and Jinhong Guo. 2018. Smartphone-based analytical biosensors. *Analyst* 143, 22 (2018), 5339–5351.

[41] Equan Health Inc. 2025. https://equahealth.io/. Accessed: 2025-07-22.

[42] Bashima Islam, Md Mahbubur Rahman, Tousif Ahmed, Mohsin Yusuf Ahmed, Md Mehedi Hasan, Viswam Nathan, Korosh Vatanparvar, Ebrahim Nemati, Jilong Kuang, and Jun Alex Gao. 2021. BreathTrack: Detecting Regular Breathing Phases from Unannotated Acoustic Data Captured by a Smartphone. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 3, Article 124 (sep 2021), 22 pages. https://doi.org/10.1145/3478123

[43] Simo Järvelä, Benjamin Cowley, Mikko Salminen, Giulio Jacucci, Juho Hamari, and Niklas Ravaja. 2021. Augmented virtual reality meditation: Shared dyadic biofeedback increases social presence via respiratory synchrony. *ACM Transactions on Social Computing* 4, 2 (2021), 1–19.

[44] Shehan Jayasekera, Edward Hensel, and Risa Robinson. 2021. Feasibility of using the Hexoskin smart garment for natural environment observation of respiration topography. *International journal of environmental research and public health* 18, 13 (2021), 7012.

[45] Jon Kabat-Zinn. 1982. An outpatient program in behavioral medicine for chronic pain patients based on the practice of mindfulness meditation: Theoretical considerations and preliminary results. *General hospital psychiatry* 4, 1 (1982), 33–47.

[46] Hyun Kang. 2021. Sample size determination and power analysis using the G* Power software. *Journal of educational evaluation for health professions* 18 (2021).

[47] Mohammad Nur Hossain Khan, Jialu Li, Nancy L McElwain, Mark Hasegawa-Johnson, and Bashima Islam. 2024. Sound tagging in infant-centric home soundscapes. In *2024 IEEE/ACM Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*. IEEE, 142–146.

[48] Mohammad Nur Hossain Khan, Nancy L McElwain, Mark Hasegawa-Johnson, and Bashima Islam. 2024. InfantMotion2Vec: Unlabeled Data-Driven Infant Pose Estimation Using a Single Chest IMU. In *2024 IEEE 20th International Conference on Body Sensor Networks (BSN)*. IEEE, 1–4.

[49] Hamzeh Khundaqji, Wayne Hing, James Furness, Mike Climstein, et al. 2020. Smart shirts for monitoring physiological parameters: scoping review. *JMIR mHealth and uHealth* 8, 5 (2020), e18092.

[50] Nataliya Kosmyna and Pattie Maes. 2019. AttentivU: an EEG-based closed-loop biofeedback system for real-time monitoring and improvement of engagement for personalized learning. *Sensors* 19, 23 (2019), 5200.

[51] Ilkka Kosunen, Mikko Salminen, Simo Järvelä, Antti Ruonala, Niklas Ravaja, and Giulio Jacucci. 2016. RelaWorld: neuroadaptive and immersive virtual reality meditation system. In *Proceedings of the 21st international conference on intelligent user interfaces*. 208–217.

[52] Mikki Kressbach. 2018. Breath work: mediating health through breathing apps and wearable technologies. *New Review of Film and Television Studies* 16, 2 (2018), 184–206.

[53] Agni Kumar, Vikramjit Mitra, Carolyn Oliver, Adeeti Ullal, Matt Biddulph, and Irida Mance. 2021. Estimating respiratory rate from breath audio obtained through wearable microphones. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 7310–7315.

[54] Willem Kuyken, Rachel Hayes, Barbara Barrett, Richard Byng, Tim Dalgleish, David Kessler, Glyn Lewis, Edward Watkins, Claire Brejcha, Jessica Cardy, et al. 2015. Effectiveness and cost-effectiveness of mindfulness-based cognitive therapy compared with maintenance antidepressant treatment in the prevention of depressive relapse or recurrence (PREVENT): a randomised controlled trial. *The Lancet* 386, 9988 (2015), 63–73.

[55] Mark A Lau, Scott R Bishop, Zindel V Segal, Tom Buis, Nicole D Anderson, Linda Carlson, Shauna Shapiro, James Carmody, Susan Abbey, and Gerald Devins. 2006. The Toronto mindfulness scale: Development and validation. *Journal of clinical psychology* 62, 12 (2006), 1445–1467.

[56] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. 2015. Deep learning. *nature* 521, 7553 (2015), 436–444.

[57] Debra Lerner, David A Adler, Hong Chang, Leueen Lapitsky, Maggie Y Hood, Carla Perissinotto, John Reed, Thomas J McLaughlin, Ernst R Berndt, and William H Rogers. 2004. Unemployment, job retention, and productivity loss among employees with depression. *Psychiatric services* 55, 12 (2004), 1371–1378.

[58] Changjun Li, Qinghua Chang, Jia Zhang, and Wenshu Chai. 2018. Effects of slow breathing rate on heart rate variability and arterial baroreflex sensitivity in essential hypertension. *Medicine* 97, 18 (2018).

[59] Daniyal Liaqat, Mohamed Abdalla, Pegah Abed-Esfahani, Moshe Gabel, Tatiana Son, Robert Wu, Andrea Gershon, Frank Rudzicz, and Eyal De Lara. 2019. WearBreathing: Real world respiratory rate monitoring using smartwatches. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 2 (2019), 1–22.

[60] I-Mei Lin. 2018. Effects of a cardiorespiratory synchronization training mobile application on heart rate variability and electroen-cephalography in healthy adults. *International Journal of Psychophysiology* 134 (2018), 168–177.

[61] Emily K Lindsay, Brian Chin, Carol M Greco, Shinzen Young, Kirk W Brown, Aidan GC Wright, Joshua M Smyth, Deanna Burkett, and J David Creswell. 2018. How mindfulness training promotes positive emotions: Dismantling acceptance skills training in two randomized controlled trials. *Journal of personality and social psychology* 115, 6 (2018), 944.

[62] Emily K Lindsay and J David Creswell. 2017. Mechanisms of mindfulness training: Monitor and Acceptance Theory (MAT). *Clinical psychology review* 51 (2017), 48–59.

[63] Emily K Lindsay and John David Creswell. 2019. Mindfulness, acceptance, and emotion regulation: Perspectives from Monitor and Acceptance Theory (MAT). *Current opinion in psychology* 28 (2019), 120–125.

[64] Emily K Lindsay, Shinzen Young, Kirk Warren Brown, Joshua M Smyth, and J David Creswell. 2019. Mindfulness training reduces loneliness and increases social contact in a randomized controlled trial. *Proceedings of the National Academy of Sciences* 116, 9 (2019), 3488–3493.

[65] Emily K Lindsay, Shinzen Young, Joshua M Smyth, Kirk Warren Brown, and J David Creswell. 2018. Acceptance lowers stress reactivity: Dismantling mindfulness training in a randomized controlled trial. *Psychoneuroendocrinology* 87 (2018), 63–73.

[66] Eric B Loucks, William R Nardi, Roee Gutman, Frances B Saadeh, Yu Li, David R Vago, Lauren B Fiske, Jayson J Spas, and Abigail Harrison. 2021. Mindfulness-based college: A stage 1 randomized controlled trial for university student well-being. *Psychosomatic medicine* 83, 6 (2021), 602–614.

[67] David S Ludwig and Jon Kabat-Zinn. 2008. Mindfulness in medicine. *Jama* 300, 11 (2008), 1350–1352.

[68] James MacKillop and Emily J Anderson. 2007. Further psychometric validation of the mindful attention awareness scale (MAAS). *Journal of Psychopathology and Behavioral Assessment* 29 (2007), 289–293.

[69] Peter Malinowski. 2013. Neural mechanisms of attentional control in mindfulness meditation. *Frontiers in Neuroscience* 7 (2013), 8. https://doi.org/10.3389/fnins.2013.00008

[70] Andrew W Manigault, Jeremiah Slutsky, Julianna Raye, and John David Creswell. 2021. Examining practice effects in a randomized controlled trial: Daily life mindfulness practice predicts stress buffering effects of mindfulness meditation training. *Mindfulness* 12, 10 (2021), 2487–2497.

[71] Farhana Mann, Jingyi Wang, Eiluned Pearce, Ruimin Ma, Merle Schlief, Brynmor Lloyd-Evans, Sarah Ikhtabi, and Sonia Johnson. 2022. Loneliness and the onset of new mental health problems in the general population. *Social psychiatry and psychiatric epidemiology* 57, 11 (2022), 2161–2178.

[72] Larry R Medsker, Lakhmi Jain, et al. 2001. Recurrent neural networks. *Design and Applications* 5, 64-67 (2001), 2.

[73] Roger GT Mello, Liliam F Oliveira, and Jurandir Nadal. 2007. Digital Butterworth filter for subtracting noise from low magnitude surface electromyogram. *Computer methods and programs in biomedicine* 87, 1 (2007), 28–35.

[74] Ramin Mojtabai, Mark Olfson, and Beth Han. 2016. National trends in the prevalence and treatment of depression in adolescents and young adults. *Pediatrics* 138, 6 (2016).

[75] Inbal Nahum-Shani and Carolyn Yoon. 2024. Toward the Science of Engagement With Digital Interventions. *Current Directions in Psychological Science* (2024), 09637214241254328.

[76] Andrea Nicolò, Carlo Massaroni, Emiliano Schena, and Massimo Sacchetti. 2020. The importance of respiratory rate monitoring: From healthcare to sport and exercise. *Sensors* 20, 21 (2020), 6396.

[77] Taehwan Park, Maryanne Reilly-Spong, and Cynthia R Gross. 2013. Mindfulness: a systematic review of instruments to measure an emergent patient-reported outcome (PRO). *Quality of Life Research* 22 (2013), 2639–2659.

[78] Thanakij Pechprasarn and Suporn Pongnumkul. 2013. Estimation of respiratory rate from smartphone's acceleration data. In *2013 10th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*. IEEE, 1–5.

[79] Valentina Perciavalle, Marta Blandini, Paola Fecarotta, Andrea Buscemi, Donatella Di Corrado, Luana Bertolo, Fulvia Fichera, and Marinella Coco. 2017. The role of deep breathing on stress. *Neurological Sciences* 38, 3 (2017), 451–458.

[80] Kurt Plarre, Andrew Raij, Syed Monowar Hossain, Amin Ahsan Ali, Motohiro Nakajima, Mustafa Al'Absi, Emre Ertin, Thomas Kamarck, Santosh Kumar, Marcia Scott, et al. 2011. Continuous inference of psychological stress from sensory measurements collected in the natural environment. In *Proceedings of the 10th ACM/IEEE international conference on information processing in sensor networks*. IEEE, 97–108.

[81] Inmaculada Plaza, Marcelo Marcos Piva Demarzo, Paola Herrera-Mercadal, Javier García-Campayo, et al. 2013. Mindfulness-based mobile applications: literature review and analysis of current features. *JMIR mHealth and uHealth* 1, 2 (2013), e2733.

[82] Jannatul Ferdaous Progga, Mohammad Nur Hossain Khan, and Mg Mostofa Amın. 2023. Meteorological Parameters–Soil Temperature Relations in a Sub-Tropical Summer Grassland: Physically-Based and Data-Driven Modeling. *Atatürk Üniversitesi Ziraat Fakültesi Dergisi* 54, 2 (2023), 48–56.

[83] Md Mahbubur Rahman, Tousif Ahmed, Mohsin Yusuf Ahmed, Ebrahim Nemati, Minh Dinh, Nathan Folkman, Md Mehedi Hasan, Jilong Kuang, and Jun Alex Gao. 2021. Towards motion-aware passive resting respiratory rate monitoring using earbuds. In *2021 IEEE 17th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*. IEEE, 1–4.

[84] Md Mahbubur Rahman, Ebrahim Nemati, Viswam Nathan, and Jilong Kuang. 2020. Instantrr: Instantaneous respiratory rate estimation on context-aware mobile devices. In *13th EAI International Conference on Body Area Networks 13*. Springer, 267–281.

[85] Mary AM Rogers, Kelsey Lemmen, Rachel Kramer, Jason Mann, and Vineet Chopra. 2017. Internet-delivered health interventions that work: systematic review of meta-analyses and evaluation of website availability. *Journal of medical Internet research* 19, 3 (2017), e90.

[86] Marc A Russo, Danielle M Santarelli, and Dean O'Rourke. 2017. The physiological effects of slow breathing in the healthy human. *Breathe* 13, 4 (2017), 298–309.

[87] Shourjya Sanyal and Koushik Kumar Nundy. 2018. Algorithms for monitoring heart rate and respiratory rate from the video of a user's face. *IEEE Journal of translational engineering in health and medicine* 6 (2018), 1–11.

[88] Maya C Schumer, Emily K Lindsay, and J David Creswell. 2018. Brief mindfulness training for negative affectivity: A systematic review and meta-analysis. *Journal of consulting and clinical psychology* 86, 7 (2018), 569.

[89] Philip Sedgwick. 2012. Pearson's correlation coefficient. *Bmj* 345 (2012).

[90] Ivan W Selesnick and C Sidney Burrus. 1998. Generalized digital Butterworth filter design. *IEEE Transactions on signal processing* 46, 6 (1998), 1688–1694.

[91] Chen-Hsuan Shih, Naofumi Tomita, Yanick X Lukic, Álvaro Hernández Reguera, Elgar Fleisch, and Tobias Kowatsch. 2019. Breeze: Smartphone-based acoustic real-time detection of breathing phases for a gamified biofeedback breathing training. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 3, 4 (2019), 1–30.

[92] Chen-Hsuan (Iris) Shih, Naofumi Tomita, Yanick X. Lukic, Álvaro Hernández Reguera, Elgar Fleisch, and Tobias Kowatsch. 2020. Breeze: Smartphone-Based Acoustic Real-Time Detection of Breathing Phases for a Gamified Biofeedback Breathing Training. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 4, Article 152 (sep 2020), 30 pages. https://doi.org/10.1145/3369835

[93] Jeremiah Slutsky, Brian Chin, Julianna Raye, and John David Creswell. 2019. Mindfulness training improves employee well-being: A randomized controlled trial. *Journal of Occupational Health Psychology* 24, 1 (2019), 139.

[94] Xiao Sun, Li Qiu, Yibo Wu, Yeming Tang, and Guohong Cao. 2017. Sleepmonitor: Monitoring respiratory rate and body position during sleep using smartwatch. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 1, 3 (2017), 1–22.

[95] Sharon Toker and Michal Biron. 2012. Job burnout and depression: unraveling their temporal relationship and considering the role of physical activity. *Journal of applied psychology* 97, 3 (2012), 699.

[96] John Torous, Jessica Lipschitz, Michelle Ng, and Joseph Firth. 2020. Dropout rates in clinical trials of smartphone apps for depressive symptoms: a systematic review and meta-analysis. *Journal of affective disorders* 263 (2020), 413–419.

[97] Adriano BL Tort, Diego A Laplagne, Andreas Draguhn, and Joaquin Gonzalez. 2025. Global coordination of brain activity by the breathing cycle. *Nature Reviews Neuroscience* (2025), 1–21.

[98] Crystal Han-Huei Tsay, Alexander Kofinas, and Jing Luo. 2018. Enhancing student learning experience with technology-mediated gamification: An empirical study. *Computers & Education* 121 (2018), 1–17.

[99] Jean M Twenge, Jonathan Haidt, Andrew B Blake, Cooper McAllister, Hannah Lemon, and Astrid Le Roy. 2021. Worldwide increases in adolescent loneliness. *Journal of adolescence* 93 (2021), 257–269.

[100] Jean M Twenge, Cooper McAllister, and Thomas E Joiner. 2021. Anxiety and depressive symptoms in US Census Bureau assessments of adults: Trends from 2019 to fall 2020 across demographic groups. *Journal of anxiety disorders* 83 (2021), 102455.

[101] Sophie Valentine, Adam C Cunningham, Benjamin Klasmer, Mohammad Dabbah, Marko Balabanovic, Mert Aral, Dan Vahdat, and David Plans. 2022. Smartphone movement sensors for the remote monitoring of respiratory rates: Technical validation. *Digital Health* 8 (2022), 20552076221089090.

[102] Kurt VanLehn. 2011. The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational psychologist* 46, 4 (2011), 197–221.

[103] Pauli Virtanen, Ralf Gommers, Travis E Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, et al. 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature methods* 17, 3 (2020), 261–272.

[104] Endre Visted, Jon Vøllestad, Morten Birkeland Nielsen, and Geir Høstmark Nielsen. 2015. The impact of group-based mindfulness training on self-reported mindfulness: a systematic review and meta-analysis. *Mindfulness* 6 (2015), 501–522.

[105] Ni Wayan Surya Wardhani, Masithoh Yessi Rochayani, Atiek Iriany, Agus Dwi Sulistyono, and Prayudi Lestantyo. 2019. Cross-validation metrics for evaluating classification performance on imbalanced data. In *2019 international conference on computer, control, informatics*

and its applications (IC3INA). IEEE, 14–18.

[106] Joseph Wielgosz, Brianna S Schuyler, Antoine Lutz, and Richard J Davidson. 2016. Long-term mindfulness training is associated with reliable differences in resting respiration rate. *Scientific reports* 6, 1 (2016), 27533.

[107] Zhiwen Xiong, Huibin Wang, Lili Zhang, Tanghuai Fan, Jie Shen, Yue Zhao, Yang Liu, and Qi Wu. 2021. A study on seizure detection of EEG signals represented in 2D. *Sensors* 21, 15 (2021), 5145.

[108] Siqi Yin, Xianling Cai, Ziyang Wang, Yuning Zhang, Shenghui Luo, and Jingdong Ma. 2022. Impact of gamification elements on user satisfaction in health and fitness applications: A comprehensive approach based on the Kano model. *Computers in Human Behavior* 128 (2022), 107106.

[109] Shinzen Young. 2016. What is mindfulness? A contemplative perspective. *Handbook of mindfulness in education: Integrating theory and research into practice* (2016), 29–45.