Can You Hear the Growth of Your Mindfulness Skills?

Self-paced mindfulness training improves people's mindfulness skills, such as concentration, sensory clarity, and equanimity. Respiration dynamics play a vital role in managing mindfulness. Measuring and quantifying these skills are required to deploy a scalable smartphone-based mindfulness meditation for the user. However, the present methods of measuring these mindfulness skills rely on either self-reports or multiple questionnaires-based systems, such as Mindful Attention Awareness Scale (MAAS) and Toronto Mindfulness Scale (TMS). Besides, no previous studies have analytically explained the relationship between respiratory signals and mindfulness skills. This paper first shows that slow-paced breathing can translate the relationship between breathing rate and mindfulness skills. Next, we assess the potential of using acoustic sensors to sense the improvements in mindfulness skills passively. Finally, we study the impact of respiratory signal feedback in enhancing mindfulness skills and propose a respiratory rate estimation model. We develop and evaluate our model on data collected from 20 participants performing 108 meditation sessions using a smartphone application for mindfulness training. Our results show that we can predict improvements in three mindfulness skills – concentration, sensory clarity, and equanimity with an 83%, 86%, and 84% F1-score, respectively, from acoustic signals recorded through in-ear earphones. We estimate breathing rate from acoustic signals with 29-60% less breath per minute error than state-of-the-art solutions. Finally, our user study with 12 participants shows that immediate bio-feedback elevates the mindfulness training experience. This paper is the first to empirically establish the relationship between respiratory dynamics with mindfulness and provide a robust, passive, and accessible solution to monitor the improvement in mindfulness skills from acoustic signals quantitatively.

CCS Concepts: • **Human-centered Computing → Ubiquitous Computing**.

Additional Key Words and Phrases: mindfulness, concentration, sensory clarity, equanimity

ACM Reference Format:
. 2023. Can You Hear the Growth of Your Mindfulness Skills?. In . ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Mindfulness is a therapeutic technique to achieve a calm mental state by focusing one's awareness on the present while acknowledging one's feelings, thoughts, and bodily sensations [1, 3]. Unlike breathing exercises, where the focus is on regulating breathing to match a technique, e.g., box breathing [10], mindfulness meditation focuses on three sensory activities – see, hear, and feel [2, 3]. Studies show that even without conscious regulation of breathing, mindfulness meditation alters respiratory signals in ways unique from other forms of rest or activity [9**?**]. Correlation and intervention studies also show that uniquely altering respiration dynamics improves sleep quality, relaxation, positive emotions, vigor, and alertness and can reduce anxiety, depression, and physiologic markers such as blood pressure [][72-78 grant]. Despite the well-established relationship between respiration and parasympathetic nervous system activation [][77, 82 grant] (as indexed by HRV), the relation between mindfulness and respiration features is yet to be explored. Few works suggested a correlation between slow-paced breathing and mindfulness [][83-86 grant].

Slow-paced breathing is characterized by a low respiration rate of 4.5-6.5 breaths per minute (bpm) []. Slow-paced breathing training can effectively strengthen cardiac health and mental well-being [4, 6, 7]. This training can be self-paced (e.g., during mindfulness meditation) [], externally guided by following acoustic [], seismic [], visual instructions [], or biofeedback []. Measuring slow-paced breathing is crucial as it can translate mindfulness skills analytically. However, measuring slow-paced breathing requires sensing biosignals that are not scalable due to the need for costly trained therapists and additional sensing equipment [].

Thus, smartphone-based solutions are recently gaining popularity in providing a scalable and accessible solution for respiration training and monitoring (both active and passive) [**? ? ?**]. Several sensing domains, including motion, audio, and video, have been used for approximating respiration rates. However, no studies examine the best sensing modality for meditation tasks. Current works on motion-based (with inertial measurement units, IMUs) solutions either require expensive equipment like smart shirts or need to hold smartphones on the chest, making them unsuitable for ubiquitous usage. Video solutions do not perform well in low-light situations or slow-paced breathing. Current audio-based solutions also focus on guided breathing, pulmonary patients, or physical exertion scenarios, e.g., running. However, most of these works focus on regular or fast breathing (10-45 bpm) [] and ignore the slow-paced breathing zone (5-10 bpm). As audio is the most convenient passive monitoring method, we study how acoustic sensing can detect slow-paced breathing rates.

Besides, none of these works focus on the impact of the training, especially its effect on mindfulness skills. The current approach to mindfulness measurement uses self-reports, Mindful Attention Awareness Scale (MAAS), or the Five Factor Mindfulness Questionnaire (FFMQ)) []. These measurements are not feasible to develop a scalable and ubiquitous solution for mindfulness training. Thus, there is a need to automatically predict mindfulness skills.

To understand the efficacy of slow-paced breathing and mindfulness meditation on mindfulness skills, we present MindGain. MindGain is the first solution to quantitatively estimate the changes in mindfulness skills during self-paced meditation using acoustic sensors of commodity earphones and off-the-shelf mindfulness training smartphone applications. We first quantitatively define slow-paced breathing and study its relationship with mindfulness skills. Then, we provide a passive sensing solution using acoustic sensors that can measure changes in mindfulness skills from the sound of breathing collected using in-ear earbuds. Based on previous works [], we focus on three mindfulness skills – concentration, sensory clarity, and equanimity.

We assess the feasibility of using different sensing modalities, e.g., audio and motion (inertial measurement unit or IMU), to estimate the improvement of mindfulness skills by collecting multi-modal data from 20 participants. These participants, in total, performed 102 meditation sessions (each approximately 20 minutes long) using a commercial mindfulness training app. Our data collection further involved, Hexoskin, a commercial respiration rate monitoring smart shirt that provided us with the ground truth respiration rate for analysis. We observe that the acoustic modality outperforms other modalities and achieves 83%, 86%, and 84% F1-scores in predicting change in concentration, sensory clarity, and equanimity, respectively.

Bio-feedback guided training proves effective for various physiological [] and mental benefits, such as improving self-efficacy []. Breeze [] proposes a smartphone-based bio-feedback training, where feedback is provided as a form of animation. However, no study focuses on the effect of providing respiratory bio-feedback immediately after a meditation session to the users. Additionally, bio-feedback guided training requires active instruction and awareness, which is unsuitable for self-paced mindfulness meditation. Therefore, we develop an audio-to-breathing rate estimator capable of tracking low respiration rate (5-10 bpm) and visualize percentage changes in respiration rate, global average breathing rate, total time spent in the desirable breathing zone, and continuous respiratory signal feedback to eight participants.

We observe that our proposed respiration rate estimator model has 29-60% lower error than state-of-the-art algorithms []. Finally, our study reveals that immediate biosignal feedback elevates the mindfulness training experience.

Through this paper, we make three technical contributions.

- We are the first to define slow-paced breathing in terms of respiration rate and provide a quantitative analysis of how different respiratory features contribute to estimating a change in mindfulness skills during meditation.
- Develop a deep learning model that takes audio signals as input and estimates changes in the three mindfulness metrics – concentration, sensory clarity, and equanimity. We also show how acoustic signals contain more information than the respiratory rate only or even motion data to estimate the changes in mindfulness skills.
- We propose a deep-learning approach to detect respiratory rates for low-breathing rate zones from audio to provide respiration feedback. We further study the effect of providing this feedback immediately on the users.

2 BACKGROUND AND RELATED WORKS

2.1 Mindfulness Detection

Over the past two decades, there has been significant growth in scientific research on mindfulness interventions and the development of mindfulness skills [**?**]. Mindfulness interventions have been shown to reliably increase self-reported mindfulness skills [**?**], and to improve a broad range of health and well-being outcomes in randomized controlled trials [**? ?**]. 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 (UM) 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 mindfulness training programs). Smartphone-based UM training programs in initial placebo-controlled clinical trials seem compelling [**? ? ?**]. However, previous studies utilize trait bases measures of mindfulness skills, such as the trait Mindful Attention Awareness Scale (MAAS) or Five Factor Mindfulness Questionnaire (FFMQ). These kinds of measurements are not adequate for the large-scale implementation of mindfulness meditation. Therefore, some contemporary works measure these mindfulness skills sensitively with a single-item measure of each skill before and after each guided mindfulness training session, which helps us move beyond trait-based measures of mindfulness skills. Thus, in our present work, we define a respiration feature, slow-paced breathing, that can explain the change in mindfulness skills with the training. We also develop an acoustic signal-based classification model to predict these changes for further integration in a ubiquitous smartphone-based meditation.

2.2 Respiratory Biomarker Estimation

Respiration rate or breathing rate is one of the essential biomarkers to estimate physiological signals, airflow, and body movements []. RR is clinically measured using sensors to detect air pressure near the mouth and nose []. Various sensors, e.g., motion [], audio [], and camera [], estimate RR with wearable and mobile 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 []. Such IMU-based breathing bio-signal, e.g., breathing phase and breathing rate, estimators have shown promising results for pulmonary patients [**? ?**]. However, motion-based approaches require additional equipment, e.g., chest bands or active participation of the user, e.g., holding a phone on the chest.

Some recent works exploited video-based solutions to estimate breathing rates []. [] places a camera at the same height as the participant's face and records the change in motion during inhalation and exhalation. This approach requires adequate lighting conditions and proper placement. Besides, such approaches are unsuitable for measuring low breathing rates from 4-10 bpm. Moreover, there need to be more studies on how awareness of self-measurement affects the respiration rate. Therefore, these approaches are unsuitable for passively monitoring mindfulness meditation without imposing additional costs or distracting users from the mindfulness training.

Acoustic solutions for respiratory biosignal estimation are gaining popularity due to the convenience and scalability of the modality []. Recent works on audio-based respiratory signal estimation focus on pulmonary patients [], guided breathing exercises [], or physical exercises []. Breeze [] 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 [] develops an acoustic breathing phase detector that is trained using guidance from the IMU domain. Besides BreathTrack focuses on pulmonary patients and requires multimodal data collection for training the model.

A recent group of works estimated breathing rate instead of the breathing phases as it is easier to get breathing rate labels than breathing phase labels. [] uses a CNN-based model to detect breathing rates from an audio spectrogram. This work is also developed for pulmonary patients and thus does not perform well on low breathing rates (<5 breaths per minute, BPM), indicating slow breathing. [] developed a multi-tasking model to detect breathing 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 focus on estimating low breathing rates, which is crucial for slow-paced breathing mindfulness estimation.

3 BREATHING RATE TO MINDFULNESS SKILLS

Though some previous works have indicated a possible relationship between slow-paced breathing rate and mindfulness, none have shown it analytically. Thus, we first perform a study to define slow-paced breathing in terms of breathing rate and develop a machine learning model to predict changes in mindfulness metrics, i.e., concentration, sensory clarity, and equanimity.

Three features characterize slow-paced breathing -

- **Respiration Rate Variability (RRV)** indicates the decrease in breathing rate during meditation compared to the natural breathing rate of a person.
- **Zone Duration (Z_t)** refers to a person's time spent in the desired breathing zone during meditation. Based on previous studies [], we consider 5–9 breaths per minute (BPM) as the desired breathing zone.
- **Zone Frequency (Z_f)** shows the number of times a person moves into the desired breathing zone.

We define slow-paced breathing (B_{SP}) as the convex combination of these three features, which can be represented by Equation 1.

$$B_{SP} = \alpha_1 RRV + \alpha_2 Z_t + \alpha_3 Z_f \tag{1}$$

Here, α_1, α_2, and α_3 are co-efficient, which we determine analytically using principal component analysis (PCA)

Table 1. Statistical feature sets of breathing rate

	Feature Set
F_1	μ_{BR}, σ_{BR}
F_2	μ_{BR}, σ_{BR}, κ_{BR}, γ_{BR}, η
F_3	μ_{BR}, σ_{BR}, κ_{BR}, γ_{BR}, η, t_{min}, t_{max}

Using our definition, we develop classification models with Support Vector Machine (SVM) to differentiate between an increase in mindfulness skills and no unfavorable changes. We achieve 67%, 77%, and 73% F1-score for concentration, sensory clarity, and equanimity, respectively. Figure 2 shows the contribution of each of the three features that characterize slow-paced breathing in determining changes in mindfulness skills

To further validate the definition of slow-paced breathing and assess its relationship with mindfulness, we investigate other commonly used statistical features, including mean breathing rate during meditation (μ_{BR}), the standard deviation of breathing rate during meditation (σ_{BR}), kurtosis (κ_{BR}), skewness (γ_{BR}), entropy across the breathing rates (η) and longest duration of maximum and minimum breathing rates (t_{min} and t_{max}). Table 1 shows the three different feature sets that provide acceptable results. However, Figure 1 shows that different feature sets perform better for different mindfulness metrics. On the contrary, our slow-paced breathing feature (S) provides a better result than all of them for all metrics of mindfulness.

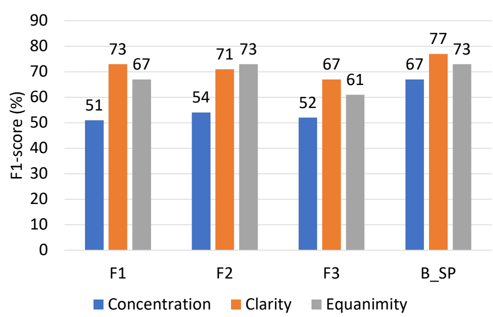


Fig. 1. Performance comparison of different statistical feature sets

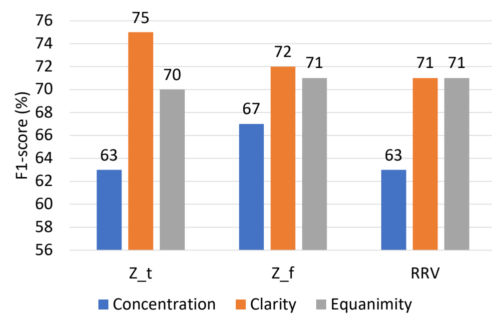


Fig. 2. Performance of individual components of slow-paced breathing

We also develop a deep learning model that learns the underlying features to predict improvement in mindfulness skills and validate slow-paced breathing. We examine the raw breathing rate and the frequency domain information of the breathing rate as input. We observe that the frequency domain information we obtain by taking the Fourier transform of the respiration rate performs better. This is expected as frequency information is similar for the zone frequency, Z_f. Using a recurrent neural network (RNN) with two layers of LSTMs of 64 neurons and two fully connected layers of 64 and 32 neurons, we can predict improvement in mindfulness skills with a 71% F1-score on average. This similar performance with our slow-paced breathing feature indicates that slow-paced breathing is a crucial indicator of mindfulness skill improvement.

4 AUDIO TO IMPROVEMENT OF MINDFULNESS SKILLS ESTIMATION

We propose an audio-based improvement of the mindfulness skills prediction method to provide a passive, non-intrusive, and easy-to-use solution. Given the recent success of deep learning in extracting information from acoustic signals [], we develop a deep neural network that predicts the improvement of mindfulness skills from audio. Our study discovered that estimating the improvement of mindfulness skills from audio outperforms determining them from the breathing bio-signals, as the audio domain captures more information than the breathing bio-signal. This section describes the data pre-processing and architecture of our proposed deep neural network. In the next section, we describe our method to determine breathing rate from acoustic signals for providing biofeedback.

4.1 Data Pre-processing

During every 20 minutes of the mindfulness training session, we collect acoustic signals at an 11.025 KHz sampling rate from the earphone. Acoustic signals collected in the real world comprise different environmental and hardware noises []. Therefore, we first filter the raw acoustic signal using a second-order Butterworth filter [] and a Savitzky-Golay filter []. As breathing sounds reside in the lower frequency zone, these filters eliminate the background noise and unwanted frequency components.

Processing each 20-minute session results in a large raw audio vector input with 13,230,000 parameters, which is infeasible to process. Thus, we segment each session into two-minute chucks and detect its contribution to the mindfulness skills change. Two minutes of data is needed as it is the minimum required mindfulness training time for an effective shift in mindfulness [].

Though features from raw audio extraction with DNN are getting successful, such methods require a large corpus to train []. Inspired by the success of various acoustic features as the input to breathing-related deep neural networks [], we explored three acoustic features – power spectrogram [], Mel filter bank (MFB) energy [], and Mel Frequency Cepstral Coefficient (MFCC) []. We observe through empirical study that MFCC features better represent our required information. Therefore, we calculate the MFCC bands with a window length of 512 and a hop length of 256. For each two-minute-long audio signal sampled at 11.025 KHz, we calculate a 40 x 5550-dimensional feature matrix which we pass as the input to the DNN for mindfulness skills improvement detection.

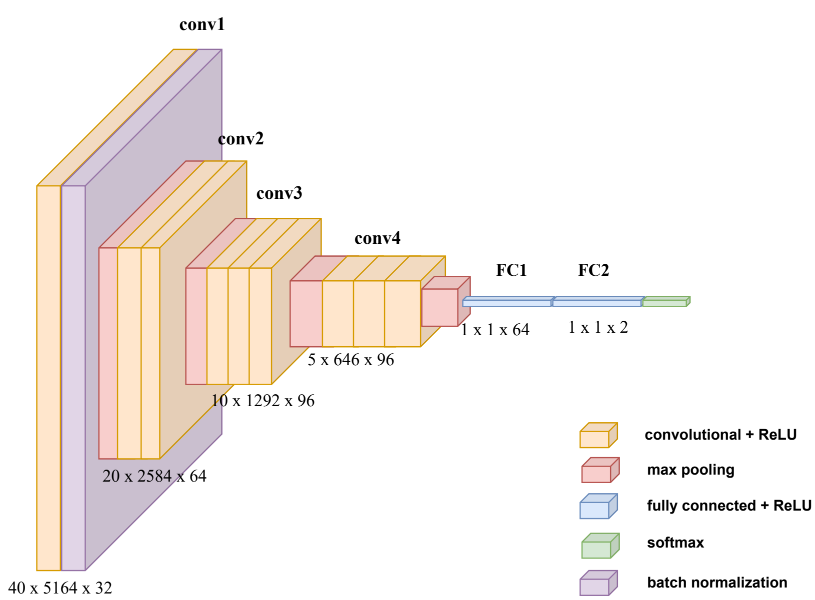


Fig. 3. Network architecture of the acoustic mindfulness skill improvement estimation model.

4.2 Model Architecture

Figure 3 shows the architecture of our proposed deep neural network that predicts improvement in mindfulness skills. The model consists of four convolution layers followed by two fully connected layers. The convolution layers have 32, 64, 96, and 96 kernels, respectively, and the filter size is 3. Each convolution layer is followed by max-pooling to reduce the dimension. We use batch normalization after each layer to mitigate the risk of overfitting. These convolution layers extract the necessary features to differentiate between the improvement and non-improvement of mindfulness skills. The two fully connected layers are responsible for using these features and performing the classification. Each of the fully connected layers has 64 neurons. We use ReLU activation as the non-linear function throughout the network.

We train individual networks for each mindfulness skill: concentration, sensory clarity, and equanimity. We empirically choose the Adam optimizer with a learning rate of 0.001 during training. We use binary cross-entropy as the loss function given by the following equation.

$$\mathcal{L}_{CE} = -\frac{1}{n}\sum_{i=1}^{n} y_i log\big(p(y_i)\big) - (1 - y_i) log\big(1 - p(y_i)\big) \tag{2}$$

Here, n is the number of samples, y_i is the ground truth, and $p(y_i)$ is the predicted probability of improvement in mindfulness skills.

5 RESPIRATION BIOSIGNAL FEEDBACK

To enhance the mindfulness training experience, we aim to provide respiratory bio-feedback after the training by visualizing various respiration signal dynamics. Based on previous studies, we choose four features to track throughout each meditation session – (1) percentage change in respiration, (2) total zone minutes, (3)breathing rate, and (4) global average breathing rate. Here, the total zone minute indicates minutes spent in the ideal breathing zone, which has 6-9 breaths per minute (bpm). To calculate these features, we require to estimate the breathing rate from the audio. Thus, we design a deep neural network that estimates breathing rates from acoustic signals. This section describes the data pre-processing step and the network architecture of our proposed model.

5.1 Data Pre-processing

We use a similar filtering technique on the audio signal provided in section 4.1. While we require 2-minute data to estimate the change in mindfulness skills, we require smaller audio length for estimating breathing rate as a longer signal duration fails to detect high breathing rates. On the other hand, signals smaller than 15 seconds might miss the whole breathing cycle in one segment as breathing during mindfulness training comprises between 4 to 13 bpm compared to natural breathing, which lies between 12–25 bpm. For every 15 seconds audio segment, we calculate MFCC with a window length of 512 and a hop length of 256. Thus, we get a 40x1889 dimensional feature matrix as the breathing rate estimator network's input.

5.2 Network Architecture

As breathing rate is temporal information, our proposed deep neural network (Figure 4) consists of a single long-short-term memory (LSTM) layer and two fully connected layers. The LSTM layer has 32 neurons and aims to extract the temporal relationships among the data. Each fully connected layers also have 32 neurons and ReLU activation for nonlinearity. As LSTM layers are prone to overfitting, we use 50% drop-off, batch normalization, and early stopping to mitigate the overfitting. This network takes 40x1859 dimension MFCC as input and outputs estimated breathing

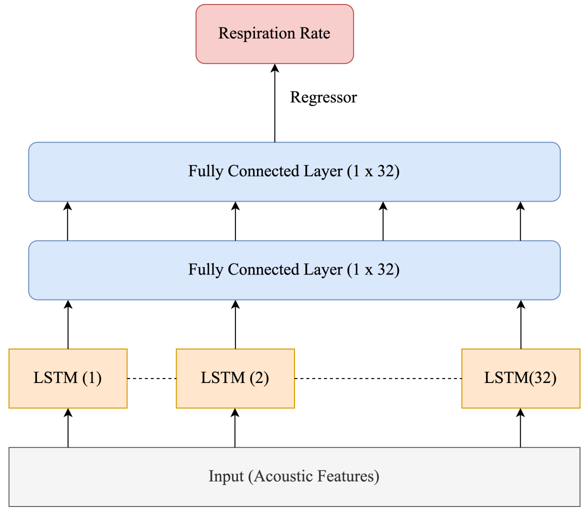


Fig. 4. Network architecture of the breathing rate estimation model.

rate. We use the mean squared error described below as the loss function with the same optimizer and learning rate as Section 4).

$$\mathcal{L}_{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{3}$$

Here, n is the number of samples, y_i is the ground truth respiration rate, and $\hat{y}_i$ is the predicted respiration rate.

6 EXPERIMENTAL SETUP

We have conducted a study to collect breathing data from 20 participants using a pair of in-ear earphones (Apple AirPods), a smartphone (iPhone XXX), and respiration tracking smart shirt (Hexoskin []). The study is approved by the institutional review board (IRB), and all the mindfulness training session is completed in a lab environment. A certified meditation instructor designed and monitored the sessions.

6.1 Systems and Applications

To measure reference respiration dynamics, we use a clinical research grade respiration shirt, Hexoskin Smart Shirts, that has been used in over 160 published studies []. Hexoskin uses two respiratory inductive plethysmography (RIP) sensors to measure respiration rate (bpm) and minute ventilation (L/min). We collected respiratory signals using Hexoskin, which provides a respiration rate every second.

Participants were instructed to wear earbuds before the meditation session started. The audio data were recorded at 44.1 KHz sampling rate stereo, which was downsampled at 11.025 KHz. The smartphone's inertial measurement unit (IMU) consists of an accelerometer and a gyroscope that samples at 60 Hz. The smartphone is placed vertically in the chest-strap pouch to collect the IMU data, with the screen facing away from the body. Hexoskin passive sensing data was stored on the Hexoskin private website. Audio and IMU data were stored on a secure cloud-based app development platform, Firebase [].

We use a consumer mindfulness training app, *Equa* [], that is developed on top the Labbuilt Digital Mindfulness Meditation trainer []. We choose *Equa* as it offers evidence-based training in learning and developing the three core mindfulness skills (concentration, sensory clarity, and equanimity). The *Equa* introductory curriculum offers 14 interactive and branching 20-minute lessons using an evidence-based Unified Mindfulness curriculum. These lessons are designed to develop the three core mindfulness skills and have been proven to have an immediate dose-response effect relative to Mindfulness-Based Stress Reduction (MBSR)[50], often considered the scientific gold standard in mindfulness training programs[59,60]. Finally, using design principles from the science of skill acquisition[61], *Equa* provides mindfulness skill development tracking using self-report, which serves as the ground truth of our collected dataset. In this self-report, users rate their concentration, sensory clarity, and equanimity at the beginning and end of each formal practice lesson.

6.2 Study Description

6.2.1 Participants Recruitment. We recruited participants through advertisements posted in the local community and educational institute campus through mass emails and viva voce. For pre-screening, our eligibility requirements include the following questions - "are you at least 18 years old?"; "are you fluent in English"; "are you willing to participate in a guided meditation?"; "are you willing and able to wear in-ear headphones and a form-fitting smart shirt with embedded sensors to track motion and physiological measures?" After the initial email screening, each participant visited our lab facility, participating in up to 11 30-minute study sessions and completing brief surveys to practice mindfulness training while physiological measures were collected.

6.2.2 Study Protocol or Design. Our study protocol aims to assess respiration patterns, perceived mindfulness skills (concentration, sensory clarity, and equanimity), perceived stress, usability satisfaction, and mobile app rating after exposure to respiration biosignal feedback. Along with collecting respiration bio-signal ground truth (using Hexoskin), regular breathing sounds (using AirPods), and chest movements due to breathing (using IMU of a smartphone), we further collect self-reported psychological questionnaire data before and after mindfulness training using the *Equa* app. Each mindfulness training consisted of a 20-minute practice. Participants who chose to complete two training sessions within the same day could take a 15-minute break between meditation sessions if they chose to do so. For their participation, participants received $8 per session and a $10 bonus for completing all sessions.

Appropriate Institutional Review Board (IRB) approval was obtained before the data collection procedure, which was conducted in a research lab on the university campus.

- Upon arriving at the first scheduled session, participants were formally screened to meet eligibility requirements. Those who met the inclusion criteria were then informed that they would participate in a study focusing on monitoring physiological responses, such as heart rate, during mindfulness meditation by collecting passive sensing data using the Hexoskin smart shirt, in-ear headphones, and smartphone. During instruction, the emphasis was deliberately applied to the physiological measure, heart rate, directing participants' attention away from respiration to prevent participants from consciously focusing on altering their breathing during the training.
- Participants were given a brief overview of mindfulness skills we would assess throughout the study – 1) **concentration** refers to the ability to focus on what you choose to focus on when you want to, 2) **sensory clarity** refers to the ability to track and explore sensory experiences in real-time, and 3) **equanimity** refers to

accepting sensory events, allowing them to come and go without resistance, or attachment to thoughts and feelings.

- The participants then provided their height and current weight using the provided scale. This information was collected as the Hexoskin smart shirt software calculates respiration by calibrating individual height and weight.
- At the start of every study session, participants responded to questions related to current heart issues and airway congestion. Before and after every mindfulness training, participants were asked to complete a brief 3-minute self-report psychological questionnaire to assess the impact of the intervention on their mindfulness skills (concentration, sensory clarity, and equanimity.) and perceived stress (PSS). The mindfulness skills are rated on scales ranging from 1 (Low), 4 (Moderate), to 7 (High).
- After completing the self-report psychological questionnaire before the mindfulness training, participants were properly fitted with the correct size Hexoskin Smart Shirt according to their height and weight to wear over their fitted shirt. Participants were then given an athletic chest strap and instructed to place it over the Hexoskin Smart Shirt and around their diaphragm, with the open pouch facing away from the body. A pair of in-ear headphones and a smartphone were then provided to the participants. Participants were instructed to place the smartphone vertically in the chest-strap pouch, with the screen facing away from the body. All participants were informed that all technologies would track and collect passive sensing physiological measures throughout the duration of each mindfulness training.
- All participants sat either on a stabilized chair or a cushion on the ground, with both postures positioned to face a nature-emulated setting throughout the mindfulness training. Due to the passive sensing tracking and the posture options, participants were instructed to refrain from making sudden movements or sounds that would introduce noise into the data or risk de-anonymization.
- Finally, the participants followed the smartphone-based interactive mindfulness meditation without the presence of anyone in the room.
- To obtain the user feedback on immediate exposure to respiration biosignal, we randomly (using a random number generator) assigned one of the two study conditions – (1) Biosignal Augmented: provide respiration biosignal feedback immediately following the mindfulness training or (2) Control: no respiration biosignal feedback was provided immediately following the mindfulness training.
- Succeeding the final session, participants were debriefed and completed the post-study debriefing questionnaire as the final task before completing the study.

6.3 Participants

We have conducted a study for data collection and to complete a user study to evaluate the impact of respiration biosignal feedback and usability of the Equa app. We have collected data from 20 participants. Eight of these participants completed 11 mindfulness sessions, while the other 12 completed two sessions. Every session is 20 minutes long, and we have collected data from 112 meditation sessions. Among these, we use 88 sessions to develop the models described in Sections 4 and 5, we use the last 24 sessions for the user study in Section 7. The demographic of all participants is provided in Figure 6

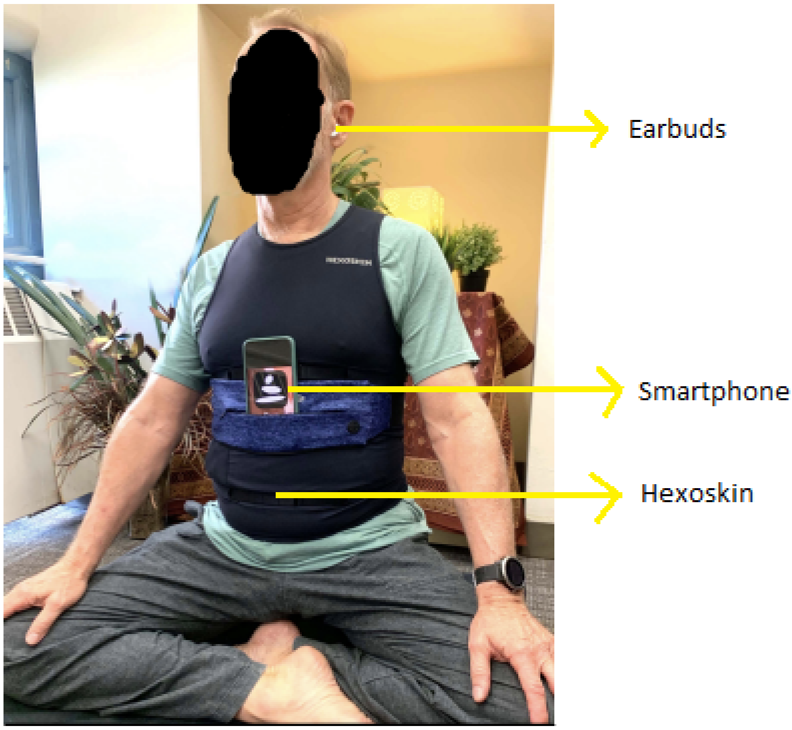
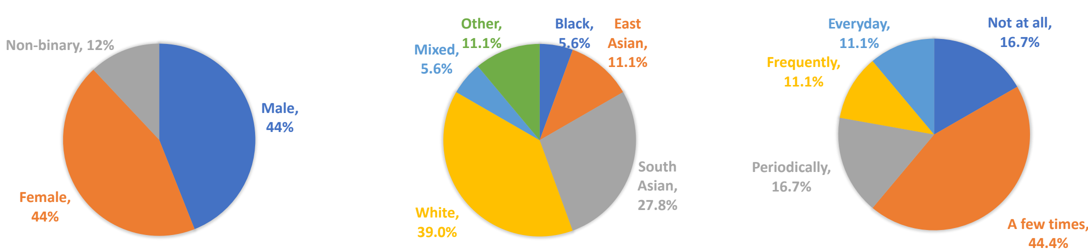


Fig. 5. Data collection setup



(a) Gender (b) Race (c) Meditation Experience

Fig. 6. Demographics of Participants

Mindfulness	Concentration	Sensory Clarity	Equanimity
Increase	43.5% (282)	39.8% (258)	57.4% (372)
Decrease	26.9% (174)	19.4% (126)	15.7% (102)
No Change	29.6% (192)	40.7% (264)	26.9% (174)

Table 2. Data distribution for predicting improvements in mindfulness skills

6.4 Data Size and Data Preparation

Each 20-minute mindfulness training session includes an introduction, homework check, technique introduction, technique practice session, and post-practice check and interaction. We use data that contain technique practice sessions where mindfulness training occurs. Thus, we get 12-15 minutes of data from every session. We segment each of these data into two-minute segments for the mindfulness skills improvement estimation. Audio, IMU, and ground truth breathing signals are segmented similarly. Thus we get 648 data segments; among them, 43.5% have an improvement in concentration, and the other 56.5% corresponds to no change or decrease in concentration. Table 2 shows the overall data distribution. We use the leave-one-session-out validation method to evaluate the classification model and the leave-one-person-out validation method to evaluate the respiration rate estimation model.

However, we take a 15-second segment with a one-second hop for respiration rate detection to achieve detailed frequency information and second-wise estimation. Therefore, for this task, we have collected 28569 audio segments where the respiration rate varies from 4 to 20 bpm.

7 EXPERIMENTAL RESULTS

In this section, we evaluate the performance of the proposed mindfulness skill improvement estimation and breathing rate estimation models using the data collected in section 6.

7.1 Performance of the Acoustic to Mindfulness Skills Improvement Estimation Model

First, we assess our choice of modality for estimating mindfulness skills improvement. We compare the performance of the mindfulness skills estimation from audio, IMU, and the breathing rates measured with Hexoskin. We investigate various models for each modality and report the best result. Figure 7 shows that audio modality achieves an 11%-13% higher F1 Score than Hexoskin and IMU sensors. Audio input works better than the respiration rates from Hexoskins as input because the acoustic signal captures the respiration rate and other contextual information that indicates changes in mindfulness skills. The same is true for the IMU modality. This evaluation also proves that other modalities can be exploited to estimate mindfulness skills from different data modalities. However, it is more convenient for real-time applications to use acoustic signals than IMU sensor data because of the limited data collection ability using smartphones and Hexoskin, which is costly and inconvenient to use in daily life.

Next, we evaluate the design choice of neural network architecture by comparing it with various network architectures. We compare our proposed network described in Section 1 with two other network architectures. The first network consists of a recurrent neural network (RNN). The second network consists of a convolutional recurrent neural network (CRNN). Though RNNs are more suitable for processing temporal data, Figure 8 shows that our network outperforms RNN and CRNN models. CNN model has a 25% higher F1 score than the RNN model and a 23% more F1 score than the CRNN model. The recent success of CNN-based models in different acoustic tasks, e.g., speaker localization [] and speech enhancement [], supports our finding. Therefore, the CNN architecture may successfully learn the features from MFCC to predict improvements in mindfulness skills better than other models.

7.2 Respiration Rate Estimation

We evaluate the performance of the proposed breathing rate estimation model by comparing it with two existing acoustic breathing rate estimation models proposed by Apple [] and Samsung []. Apple's model takes mel-filter bank (MFB) energy from audio as features and passes through a single layer of 32 long-short-term memory (LSTM) units followed by a single layer of an embedding layer. Samsung's proposed model takes the power spectrogram of the audio segments as features and infers through a network with three convolutional layers and one fully connected layer. Figure 9 shows that our proposed network architecture achieves a 29% and 60% less mean absolute breathing per minute error than Apple's and Samsung's models. We observe that MFCC performs better than MFB as acoustic feature. Besides, for temporal information, e.g., breathing rate, recurrent networks, e.g., the LSTM layer of MindGain performs better than only convolution models.

Besides, our data collection method is different from the previous works. To illustrate, in anotesamsung data collection, the participants are asked to breathe at a target breathing rate to develop a balanced dataset. We do not pose any instructions to control breathing as it disrupts the mindfulness training. As a result, our dataset consists of variable breath-per-minute samples. Moreover, anotesamsung annotates every breathing phase individually and uses this information to compute the power spectrogram. On the contrary, to develop a non-intrusive system, we do not have exact knowledge of a single breathing phase. Despite these constraints, MindGain has a 4.9 less mean absolute error (MAE) for breathing rate. This shows the robustness of our model.

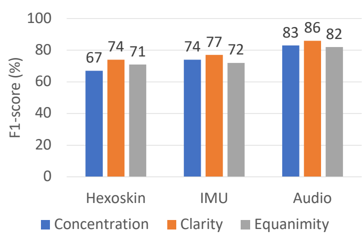


Fig. 7. Performance of different data modality in predicting improvements in mindfulness skills

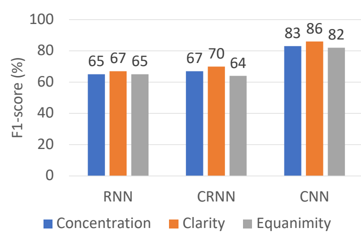


Fig. 8. Performance of different model in predicting improvements in mindfulness skills

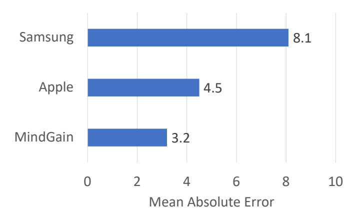


Fig. 9. Comparison of different methods for respiration rate estimation

Apple focuses on breathing rate estimation during physical activities. Therefore the authors proposed a multi-task learning (MTL) method with a combined loss of a different task, which is unsuitable for our task. Moreover, acoustic signals during a workout are more audible with more frequency content compared to breathing during mindfulness training. Both baseline works do not contain a respiration rate lower than ten breaths per minute, whereas mindfulness training often results in a low breathing zone (5-10 bpm). Hence, anoteapples proposed approach is not directly transferable to mindfulness training. Note that we do not compare our proposed network with the breath phase detection models [5, 8] as they require breathing phase annotation, which is unavailable in our scenario.

8 USER STUDY

This user study aims to understand the effect of displaying the respiration biosignal feedback immediately after mindfulness training on the participants. This respiration biosignal feedback includes four features tracked throughout each mindfulness training of 20 minutes. They are – (1) percentage change in respiration, (2) total zone minutes, (3) breathing rate, and (4) global average breathing rate. Here, the total zone minute indicates minutes spent in the ideal breathing zone, which has 6-9 breaths per minute (bpm). An example of the respiration signal visualization and bio-feedback is shown in figure 10

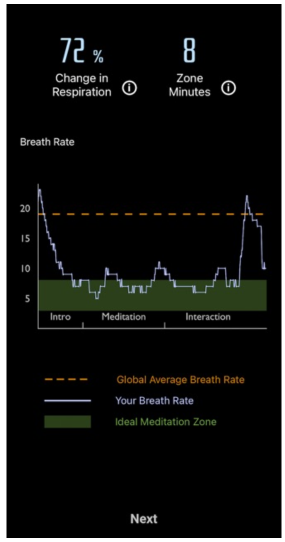


Fig. 10. Equa respiration signal feedback

8.1 Evaluation Scale

We assess the user study using System Usability Scale (SUS) [] and Mobile App Rating System (MARS) []. We assess the SUS during the post-study debriefing questionnaire. The SUS questionnaire assesses the usability of the smartphone-based interactive mindfulness meditation application by asking the participants to answer seven questions, such as "I think that I would like to use [Equa] training frequently," "I found [Equa] unnecessarily complex," and "I felt confident using [Equa]." We shortened the original 10-item scale to 7 items as 3 of the items are not applicable to the participant's interaction with Equa within our study design. Participants answered on a scale ranging from 1 (strongly disagree) to 5 (strongly agree). When scoring this scale, we followed the standard scheme for computing SUS scores but multiplied the sum of the score contributions by 100/28 instead of 2.5 to compensate for the missing questions. A score of 68 is the mean score, and a score of 80 is scored as an "A" [].

MARS is used to assess subjective application quality during the post-study debriefing questionnaire. This 4-item scale asks participants to assess the subjective quality of the smartphone-based interactive mindfulness meditation application (e.g., "Would you recommend this app to people who might benefit from it?", "Would you pay for this app?"). Participants answer on a scale ranging from 1 (e.g., Not at all) to 5 (e.g., Definitely) and sum to create a total MARS score (range 1-20). Higher scores reflect higher subjective quality ratings.

8.2 Analysis

To discover the impact of immediate respiration biosignal feedback on the user, we randomly chose to show the biofeedback (Biosignal Augmented condition) in one of the last two sessions. The session without the biofeedback is called Control Session. N=20 young adults each completed two mindfulness training sessions and then completed the SUS and MARS questionnaire as a measure of user satisfaction after each training session. We performed independent samples t-test with SPSS Statistics 28.0 software [] to compare SUS, and MARS-subjective quality responses between the Biosignal Augmented condition and Control condition.

Measurement	t(18)	p_value	cohen's d	95% CI
SUS	-2.21	0.04	0.67	[-12.54, - 0.32]
MARS	-1.15	0.266	0.36	[-4.53, 1.33]

Table 3. Measurements of independent t-test

We hypothesized that the Biosignal Augmented condition would have a higher SUS score than the Control condition. Table 3 shows the statistics of SUS and MARS of independent samples t-test. The independent samples t-test using all available data (N=20 participants) revealed significantly higher user satisfaction in the Biosignal Augmented condition on the SUS (mean, M=85, SE=2)) relative to the Control condition on the SUS (M=79, SE=2) where t(18) = -2.21, p_value < 0.04, Cohen's d = 0.67, 95% confidence interval (CI) = [-12.54, - 0.32]. This result indicates that users have a more satisfactory training experience when exposed to respiration biosignal feedback immediately following the mindfulness training.

Next, we hypothesized a higher MARS for the Biosignal Augmented condition than the Control condition. The Biosignal Augmented condition and Control condition rated the app favorably (M=3.5, SD=.97, and M=3.9, SD=.57, respectively). However, an Independent samples t-test using all available data (N=20 participants) revealed no difference between the group exposed to the respiration biosignal feedback (M=13, SE=1) and the control condition (M=11, SE=1), where t(18) = -1.15, p < 0.266, Cohen's d = 0.36, 95% confidence interval (CI) = [-4.53, 1.33]. This result indicates that the immediate biosignal feedback did not affect the mobile app's usability but elevated the mindfulness training experience.

REFERENCES

[1] Mohamed Al-Refae, Amr Al-Refae, Melanie Munroe, Nicole A Sardella, and Michel Ferrari. 2021. A self-compassion and mindfulness-based cognitive mobile intervention (Serene) for depression, anxiety, and stress: promoting adaptive emotional regulation and wisdom. *Frontiers in Psychology* 12 (2021), 648087.

[2] Kirk Warren Brown, Daniel Berry, Kristina Eichel, Polina Beloborodova, Hadley Rahrig, and Willoughby B Britton. 2022. Comparing impacts of meditation training in focused attention, open monitoring, and mindfulness-based cognitive therapy on emotion reactivity and regulation: Neural and subjective evidence from a dismantling study. *Psychophysiology* 59, 7 (2022), e14024.

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

[4] V. C. Goessl, J. E. Curtiss, and S. G. Hofmann. 2017. The effect of heart rate variability biofeedback training on stress and anxiety: a meta-analysis. *Psychological Medicine* 47, 15 (2017), 2578–2586. https://doi.org/10.1017/S0033291717001003

[5] 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

[6] Matthew EB Russell, April B Scott, Ian A Boggero, and Charles R Carlson. 2017. Inclusion of a rest period in diaphragmatic breathing increases high frequency heart rate variability: Implications for behavioral therapy. *Psychophysiology* 54, 3 (2017), 358–365.

[7] 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.

[8] 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

[9] 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.

[10] Katherine Ka-Yin Yau and Alice Yuen Loke. 2021. Effects of diaphragmatic deep breathing exercises on prehypertensive or hypertensive adults: A literature review. *Complementary therapies in clinical practice* 43 (2021), 101315.