TOWARDS THE DESIGN OF A SMARTPHONE-BASED BIOFEEDBACK BREATHING TRAINING: IDENTIFYING DIAPHRAGMATIC BREATHING PATTERNS FROM A SMARTPHONE’S MICROPHONE

Research in Progress

Shih, Chen-Hsuan Iris, Department of Management, Technology and Economics, ETH Zurich, Zurich, Switzerland, ishih@ethz.ch
Kowatsch, Tobias, Institute of Technology Management, University of St. Gallen, St. Gallen, Switzerland, tobias.kowatsch@unisg.ch
Tinschert, Peter, Institute of Technology Management, University of St. Gallen, St. Gallen, Switzerland, peter.tinschert@unisg.ch
Barata, Filipe, Department of Management, Technology and Economics, ETH Zurich, Zurich, Switzerland, fbarata@ethz.ch
Nißen, Marcia Katharina, Department of Management, Technology and Economics, ETH Zurich, Zurich, Switzerland, mnissen@ethz.ch

Abstract

Asthma, diabetes, hypertension, or major depression are non-communicable diseases (NCDs) and impose a major burden on global health. Stress is linked to both the causes and consequences of NCDs and it has been shown that biofeedback-based breathing trainings (BBTs) are effective in coping with stress. Here, diaphragmatic breathing, i.e. deep abdominal breathing, belongs to the most distinguished breathing techniques. However, high costs and low scalability of state-of-the-art BBTs that require expensive medical hardware and health professionals, represent a significant barrier for their widespread adoption. Health information technology has the potential to address this important practical problem. Particularly, it has been shown that a smartphone microphone has the ability to record audio signals from exhalation in a quality that can be compared to professional respiratory devices. As this finding is highly relevant for low-cost and scalable smartphone-based BBTs (SBBT) and – to the best of our knowledge - because it has not been investigated so far, we aim to design and evaluate the efficacy of such a SBBT. As a very first step, we apply design-science research and investigate in this research-in-progress the relationship of diaphragmatic breathing and its acoustic components by just using a smartphone’s microphone. For that purpose, we review related work and develop our hypotheses based on justificatory knowledge from physiology, physics and acoustics. We finally describe a laboratory study that is used to test our hypotheses. We conclude with a brief outlook on future work.

Keywords: Non-communicable disease, NCD, stress, health information system, design science research, diaphragmatic breathing, biofeedback.
1 Introduction

Non-communicable diseases (NCDs) such as heart diseases, asthma, hypertension, diabetes or major depression impose a major burden on global health (Krug, 2016; WHO, 2011; WHO, 2015). That is, a loss of US$ 47 trillion is expected by 2030, which equals approximately 75% of the global gross domestic product in 2010 (Bloom et al., 2011). Stress, which is defined as “the psychological and physiological state that results when the resources of the individual are not sufficient to cope with the demands and pressures of the situation” (Michie, 2002, p. 67), is linked to both the causes and consequences of NCDs, and thus, plays a significant role in the health condition of individuals (Harrison and Cooper, 2011; Kozora et al., 2009). During the last couple of decades, psychophysiological research has shown that relaxation techniques have a positive effect on coping with these demands and pressures, i.e. to reduce stress (e.g. Chittaro and Sioni, 2014a; Pastor et al., 2008). In particular, diaphragmatic breathing, i.e. deep abdominal breathing (vs. shallow chest breathing) where individuals are instructed to inhale with their nose and exhale with their mouth, belongs to the most distinguished relaxation techniques (Chen et al., 2016; Lehrer et al., 2000; Wang et al., 2010). Diaphragmatic breathing does not only lead to a state of relaxation, but is also an effective adjunct in the treatment of anxiety (Busch et al., 2012), eating disorders (McIver et al., 2009), hypertension (Dickinson et al., 2008), depression (Kim and Kim, 2005) and asthma (Chiang et al., 2009).

In contrast to breathing trainings that are solely guided by health professionals (e.g. Dickinson, 2008), the efficiency of it can be increased by applying biofeedback that guides individuals based on their own biosignals (Astor et al., 2013; Chen et al., 2016; Wang et al., 2010). In more detail, biofeedback is a technique which trains individuals to adjust or improve their health and performance by controlling certain physiological activities which normally happen involuntarily, such as heart rate, blood pressure, respiratory rate, or skin temperature (Glick, 2010). These physiological measures are usually recorded by medical devices which generate feedback information to the individuals. Moreover, biofeedback enables the training of self-control capabilities to better cope with emotional or stressful situations (Gross and Thompson, 2007). For example, IS research has shown that biofeedback can improve emotional self-regulation capabilities in decision-making processes (Astor et al., 2013). Other studies in the healthcare context have shown a positive effect of biofeedback-based breathing trainings (BBTs) on health outcomes, e.g. to effectively control pre-hypertension (Choi et al., 2016), to reduce blood pressure (Grossman et al., 2001) or to reduce the degree of stress (Chittaro and Sioni, 2014b; Pastor et al., 2008).

Despite the potential of BBTs, adoption of such trainings is limited because of two major shortcomings: (1) high costs and (2) limited scalability. First, BBTs usually require high-cost medical devices (e.g. for electromyography recordings), which are either used for the provision of biofeedback or for analysing a subject’s breathing patterns, for example, by a blood pressure monitor or a respiratory sensor (Viskoper et al., 2003; Zucker et al., 2009). Second, scalability is strongly limited due to daily interventions, which require feasibility on a large scale including the availability of health professionals as well as time and travel effort on the patient’s side.

The use of health information technology (HIT) may address these shortcomings. Innovative HIT has not only the potential to improve outcomes of preventive or therapeutic health interventions, but also to significantly reduce their costs (Agarwal et al., 2010; Anderson and Agarwal, 2011; Fichman et al., 2011; Martin et al., 2010). More precisely, a high potential of HIT has been identified not only for digital health interventions in general (Kraft et al., 2009), but also for biofeedback-enabled self-regulation interventions (Astor et al., 2013). With an empowered patient in mind (Agarwal et al., 2010), digital health interventions made available through mobile applications have increasingly received attention (Free et al., 2013; Kvedar et al., 2016). Even more intriguing and related to BBTs, it has been shown that a smartphone-microphone has the ability to record audio signals from exhalation (in an attempt to validate a spirometry algorithm for people with asthma) in a quality that can be compared to professional respiratory devices (Goel et al., 2016; Larson, 2012). Without the use of additional medical equipment, it might be therefore feasible to acquire an individual’s acoustic fingerprint.
of diaphragmatic breathing and, in turn, to provide biofeedback in a way that guides the individual to perform the BBT correctly. However, and to the best of our knowledge, there does not exist such a BBT and it is therefore an open question how to design a smartphone-based BBT (SBBT) that uses breathing patterns obtained from a smartphone-microphone. Therefore, our research has the following two objectives:

1. To investigate the relationship of diaphragmatic breathing and its acoustic components by just using a smartphone’s microphone.

2. To test the efficacy of a SBBT in relation to a state-of-the-art BBT that utilizes high-cost medical devices for biofeedback and on-site face-to-face instructions by a health professional.

The current research-in-progress addresses the first objective by outlining the theoretical background and design of a laboratory study which aims to identify the acoustic components of diaphragmatic breathing. In terms of design science research (Gregor and Hevner, 2013; Gregor and Jones, 2007; Hevner et al., 2004), this work identifies relevant justificatory knowledge from physiology, physics and acoustics for the design of SBBTs and for the purpose of hypotheses development.

The remainder of this paper is structured as follows. Next, we describe related work with a focus on BBTs and their technical design. Then, we develop and present our hypotheses. Afterwards, we describe a laboratory study which is used to test our hypotheses. We finally conclude with a brief summary and outlook on future work.

2 Related Work

In this section, we first provide a brief overview of existing BBTs and their characteristics. Then, related work is presented with respect to the design and technical implementation of a SBBT.

2.1 Towards biofeedback-based breathing trainings (BBTs)

There are various approaches to guide an individual to effectively perform a diaphragmatic breathing with the overall goal to reduce physiological stress (e.g. bodily indicators such as an increased level of skin conductance or heart rate) and/or psychological (perceived) stress. For example, it has been shown that it is more likely to have a positive effect on stress measures if individuals follow particular duration ratios of breathing (e.g. four-second inhale vs. six-second exhale) or posture such as sitting instead of standing (Kim and Park, 2016; Yamaguti et al., 2012). Because this kind of instruction-only training may be hard to follow without any additional guidance, another group of researchers have developed a stuffed toy that guided individuals through a breathing training by up-and-down movements of the toy’s abdomen (Uratani et al., 2014). Although this approach of guidance is probably more transparent compared to an instruction-only training, it does not measure the actual breathing or state of relaxation and thus, fails to provide personalized biofeedback to guide an individual through a breathing training more efficiently. Until now, only dedicated and expensive medical hardware is required to provide biofeedback in the form of visual or acoustic representations of physiological measures such as skin conductance, heart rate or respiratory rate. For example, elastic sensors attached to a belt are used for recording respiratory activity while dedicated medical software provides guidance on how to breathe correctly (Liu et al., 2010; Mitchell et al., 2010); and there exist also other BBTs that couple sensors with either additional portable devices such as music players or stationary computers to provide biofeedback by visualizations and to increase the effects of a particular treatment (Elliott et al., 2004; Nakao et al., 2000).

Several other findings indicate boundary conditions for performing an effective diaphragmatic breathing which, in turn, can be used to trigger and provide biofeedback: (1) the breathing rate should be below twelve cycles per minute (Kim and Park, 2016; Van Diest et al., 2014); (2) the duration of exhalation should be longer than inhalation following a constant ratio of approximately 2:1 (Chen et al., 2016); and (3) each BBT session should last in between two and five minutes to achieve a certain state of relaxation, for example, to reduce heart rate or blood pressure (Van Diest et al., 2014). Heretofore, however, guidance under these premises has not yet been implemented with respect to a scalable and
low-cost digital health intervention with the smartphone as the intervention delivery platform, i.e. in the form of a SBBT as introduced above the explained in more detail in the next section.

2.2 Towards the design of a smartphone-based BBT (SBBT)

Recently, mobile health interventions have successfully taken advantage of the smartphone-microphone. Applications such as AudioFlow (Natarajan et al., 2016) or SpiroCall (Goel et al., 2016) use the smartphone-microphone to record expiratory sound to perform peak flow measurements for supporting asthma control and to function as a spirometer for measuring the pulmonary function. Furthermore, a smartphone also has the ability to detect an individual’s stress level through voice recordings (Lu et al., 2012). Altogether, these applications suggest that the quality of audio signals recorded by a smartphone-microphone is adequate for the design of digital health interventions. Accordingly, it is worth investigating physiological outcomes such as the respiratory rate, which are traditionally recorded by expensive medical devices, solely based on the smartphone’s hardware in combination with advanced signal processing techniques (Madisetti, 2009) and machine learning (Bishop, 2006). Signal processing techniques can be applied to characterize acoustic patterns and, as a consequence, to extract features of specific audio signals, such as laughing, clapping, singing and the like (Martin, 1999). On the other hand, machine learning can be used to analyze and classify, among others, features from sequential data such as incoming acoustic signals. Thus, with a smartphone-microphone we can acquire acoustic signals and subsequently, with its computational power, we can analyze those signals in quasi real-time to generate biofeedback based on the detected acoustic patterns.

In audio signal processing, the most commonly used representation for an acoustic signal are features extracted from Fast Fourier Transform (FFT) algorithm, which converts signals from the time domain into the frequency domain (Smith, 1997). By applying FFT, it has been shown that acoustic breathing signals characterized by frequency spectra, with a sharp decrease of power at the lower and upper cutoff frequencies at 850 Hz and 1600 Hz (Gaviely et al., 1981). Moreover, peak amplitudes from the original acoustic signal, i.e. the waveform of the sound, in the time domain can also represent important features which separates specific phases of breathing, such as inhalation and exhalation (Yahya and Faezipour, 2014).

Even more advanced and related to automated speech recognition applications is the compact frequency spectrum representation known as Mel-Frequency Cepstrum Coefficient (MFCC) (Davis and Mermelstein, 1980; Mermelstein, 1976). This coefficient imitates the human auditory perception and attempts to eliminate speaker dependent characteristics. It is computed in order to identify features of an acoustic signal such that a machine learning classification algorithm can learn to differentiate between a target acoustic pattern (e.g. saying “Iris” to activate a speech recognition system) and some acoustic noise (Muda et al., 2010).

In summary, features derived from the raw audio signal (e.g. peak amplitudes), FFT or MFCC establish a strong foundation for machine learning algorithms to train the mathematical (classification) model. Thus, this model is served as a predictor based on the analysis of relationships between features and sounds. A sleep monitoring application demonstrates the advance by employing MFCC feature extraction on recorded sound during sleep and subsequently utilizes machine learning to detect e.g. snore, cough, turnover and getup (Ren et al., 2015). However, as these advanced computational techniques are not yet used for the design of SBBT, we aim at bringing them into focus with our research of which the current work must be seen as a very first step.

3 Hypotheses Development

In order to provide adequate biofeedback with regard to correct diaphragmatic breathing, the SBBT must be able to basically distinguish (1) inhalation via nose from exhalation via mouth and (2) abdominal breathing from chest breathing (Chen et al., 2016; Lehrer et al., 2000; Wang et al., 2010). We therefore derive in the following hypotheses that link diaphragmatic breathing with justificatory knowledge from physiology, physics and acoustics.
First, an audible breathing signal is caused by airflow traveling through the glottis of an individual and generates vibrations in the tissue near the trachea. By observing breathing sounds over the trachea, expiration was found to be louder than inhalation with respect to frequencies higher than 130Hz (Murphy, 1981). Moreover, it has been studied that during exhalation there is a strong velocity of airflow generated and expelled through mouth (Crawford-Brown, 1997). While it is known that airflow velocity is correlated with sound frequency, higher frequency sound is more likely to happen during exhalation than inhalation. As already outlined above, acoustic features of the breathing sound can be derived from its raw signal (e.g. the peak amplitude), by applying FFT (e.g. frequency spectrum) or MFCC (e.g. the power spectrum). The features calculated by MFCC represent the energy distribution of the sound based on the human auditory perception system (Muda et al., 2010). Because inhalation and exhalation sounds are different, the extracted MFCC features will differ from each other numerically. We therefore hypothesize the first relationship between inhalation and exhalation and their acoustic sound representations as follows.

\[ H1: \text{Higher (lower) peak amplitudes in combination with higher (lower) frequency spectra distributions of acoustic breathing signals are related to exhalation (inhalation) via mouth (nose).} \]

Second, the differences between chest breathing and abdominal breathing are based on the movement of an individual’s central tendon and rib cage (Moore, 2015). According to the passive elastic characteristic of our respiratory system, chest breathing based on rib cage muscles creates rapider motion than abdominal breathing through abdomen muscles (Sharp et al., 1975). This points towards duration differences while performing different types of breathing. Furthermore, during chest breathing only a small volume of the lungs is used to deliver a relatively small amount of oxygen to the blood stream, while abdominal breathing uses the full lung capacity for the maximum of oxygen intake (Quanjer et al., 1993). Thus, there is also a greater air outlet during abdominal breathing compared to chest breathing. It is also known that the amplitude of a breathing sound increases with airflow (Gavriely and Cugell, 1996). Abdominal breathing is therefore more likely to generate higher amplitude sounds than chest breathing. Indeed, recent research has shown the link between an individual’s lung function and exhalation sound recorded through a smartphone’s microphone, in which the duration of the exhalation played a significant role in the determination of the lung volume, as well as the peak amplitude measure (Goel et al., 2016). We therefore formulate our second hypothesis as follows:

\[ H2: \text{Longer (shorter) durations in combination with higher (lower) peak amplitudes of exhalation are related to abdominal (chest) breathing.} \]

4 Method

In order to test our hypotheses we will conduct a laboratory study in which breathing sound will be recorded. The study protocol has already been approved by the institutional review board of the first author. A secondary explorative objective of this study is to increase the accuracy to detect diaphragmatic breathing in real life conditions which contains acoustic noise. That is, we plan to record several additional control sounds such as throat clearing, coughing, laughing, talking or background noise (note that breathing exercises should be conducted per se in a quite environment). Being aware of the acoustic components of these control sounds will probably increase the feasibility to perform SBBTs in practical settings beyond a controlled laboratory study. However, due to the focus of the current paper, i.e. to test the two hypotheses related to diaphragmatic breathing, evaluation of the practical feasibility of our envisioned SBBT under the condition of various control sounds and background noise will be future work. Next, we describe the sampling of the subjects. Then, we outline the study procedure, the data acquisition instruments and how we plan to analyze the collected data with respect to our hypotheses.

4.1 Subjects acquisition and compensation

The present study is a preliminary research focusing on the feasibility and practicability of our basic assumptions and hypotheses in a HIT context. Therefore, we will recruit a relatively small and ho-
mogenous sample, which will be extended in future research depending on the results of this very first study. The sample will consist of a minimum of 40 bachelor and master students, balanced with regard to gender. The subjects will be invited to participate with the help of web-based announcements (e.g. via relevant bulletin boards or Facebook groups) and flyers that are distributed around the University’s campus. Each subject will receive a financial compensation of 15 US dollars for participation.

4.2 Procedure and data acquisition

The schematic procedure of our study is illustrated in Figure 1. Each session will approximately take one hour. In the beginning, subjects will be welcomed by the instructor and briefed regarding the objective and procedure of the study. To participate in the study, subjects will be required to provide their informed consent and to explicitly review several exclusion criteria as described in the following. Subjects in a laboratory setting are usually unlikely to perform a relaxation training like diaphragmatic breathing in the same way they would apply those trainings in their everyday life. That is, relaxation trainings are intended as a coping strategy in stressful situations (Cohen and Williamson, 1979). In order to simulate the natural setting of these trainings, we will induce stress in our study. Consistently, the external validity of the recorded breathing sounds is expected to be higher when stress is induced. Students suffering from cardiovascular diseases (e.g. tachycardia), severe respiratory diseases (e.g. chronic obstructive pulmonary disease) or other diseases prone to be triggered by stress (e.g. post-traumatic stress disorder) will be therefore excluded from the study. These exclusion criteria will be enforced by the instructor.

![Study procedure diagram](image-url)

**Figure 1. Study procedure**

After this welcome phase, the subject will be guided by the instructor to the laboratory with the pre-arranged study set-up. The subject will be asked to sit down in front of a desktop computer (Apple iMac 5K 27”), which will run a pre-configured survey program (LimeSurvey). During the study, a subject’s physiological data will be continuously tracked and all audio signals will be recorded. In terms of physiological data, subjects will be connected to the MindMedia NeXus 10 medical device and the Biotrace software, which measures skin conductance and heart rate as a physiological proxy of stress (for the manipulation check) and the respiratory rate as a physiological control variable for the degree of abdominal breathing. Audio signals will be recorded using the Audacity software and six different microphones to account for various recording hardware: A Rhode studio microphone (as the “gold standard”), a built-in microphone of an Android tablet (Google Nexus 7 2013) and Android smartphone (HTC M8), an iPhone 5 built-in microphone and two plug-in earphones for one Android...
device (Samsung Galaxy s5) and the iPhone 5. Once the study preparation is completed and the subject’s physiological data and acoustic signals are set to be recorded, the subject will be asked to start LimeSurvey, which guides the user through the actual procedure and automatically sets corresponding markers in the physiological and audio recording software.

The actual study procedure can be divided into three main blocks, with the first two blocks being almost identical. The first two blocks start with an assessment of the subject’s initial psychological (perceived) stress (PS1 resp. PS3) which will complement the two physiological measures of stress (Moody and Galletta, 2015; Riedl, 2013; Riedl et al., 2013; Tams et al., 2014). We will measure the stress level using the self-assessment manikin (Bradley and Lang, 1994) and the visual analogue scale, which assesses acute stress through one 11-point Likert-scale item (Lesage et al., 2012). After the initial perceived stress measurement, stress will be induced through a math task (TASK1 resp. TASK2). The math task is based on (Wang et al., 2007) and was confirmed to efficiently induce stress in less than one minute by forcing subjects to perform consecutive arithmetic tasks under time pressure. Subsequent manipulation checks will assess whether physiological and psychological stress was induced successfully (PS2 resp. PS4). In the next step, each subject will receive breathing instructions, in block one for diaphragmatic breathing and in block two for chest breathing. The subject will be prompted to perform the corresponding breathing training for three minutes. To account for confounding time and order effects, we will systematically change between subjects whether diaphragmatic breathing or chest breathing will be performed first (see Figure 1). Upon completion, perceived stress will be measured again (PS3 resp. PS5).

In the third and final block, we address the second, more exploratory goal of this study. That is, we will record acoustic fingerprints of control sounds like intentional breath nose-in/out, breath mouth-in/out, breath mouth-in and nose-out, throat clearing, coughing, induced laughter and speech.

The study ends with a debriefing. The cables for physiological measurement will be detached and the subjects will have the opportunity to ask questions. Financial compensation for the participation will be provided and the actual purpose of the study will be revealed.

4.3 Data analysis

The data analysis consists of two parts: The first part contains the manipulation checks for both the stress induction and the effects of breathing exercise on physiological and perceived stress. Applying paired t-tests on the corresponding pre- and post-measurements of perceived stress, we will analyze whether the math tasks (TASK1, TASK2) increased perceived stress (H0: PS2-PS1 ≤ 0 resp. H0: PS4-PS3 ≤ 0) and whether the breathing exercises (DB) decreased perceived stress (H0: PS3-PS2 ≥ 0 resp. H0: PS4-PS3 ≥ 0). The same procedure will be adopted for the two physiological stress measures.

In the second part, we will test our hypotheses by modeling the acquired sound data using two logistic regressions. In contrast to the first part of the data analysis, when the data format changes from between-subjects level to the level of individual acoustic signals (e.g. inhaling sounds). We will examine whether we can successfully classify inhaling via nose vs. exhaling via mouth (H1) and chest vs. abdominal breathing (H2) using the features peak amplitude, FFT and MFCC as predictors. For the classification to be considered successful, the algorithm needs to have significantly better classification results than classification by chance (H0: P(Y=1|X=1) ≤ 0.5 for both H1 and H2). For hypotheses testing, it is important to consider the hierarchical structure of the sound data, as the individual acoustic signals (e.g. inhalation/exhalation) are nested within-subjects. Thus, a multi-level framework will be applied. The study observations, which will be coded by two independent reviewers, will serve as the ground truth in the algorithm development by machine learning. Half of the sound data will be used for algorithm development, the other half for hypotheses testing.

5 Summary and Future Work

In summary, this paper outlines the first step towards a scalable, low-cost and biofeedback-based breathing training with the smartphone as the intervention delivery platform. In our future work, we will evaluate the efficacy of the SBBT, which we will design based on the results of the present study.
Acknowledgements

We would like to thank the anonymous reviewer of the paper, Elgar Fleisch, Gabriella Chiesa, Niklas Elser, Dagmar l’Allemand, Dirk Büchter, Katrin Heldt, Björn Brogle, Dominique Durrer, Nathalie Farpour-Lambert, Pauline Gindrat and Wolfgang Maass for their valuable support and feedback. The work is part-funded by Swiss National Science Foundation (http://p3.snf.ch/project-159289 PathMate2) and CSS Insurance.

References


