NOTICE: This is the author's version of a work that was accepted for publication in **Computers in Human Behavior**. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms, may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in **Computers in Human Behavior**, Volume 40, November 2014, Pages 56-63, DOI: http://dx.doi.org/10.1016/j.chb.2014.07.049.

Please cite as:
Chittaro L., Sioni R. Evaluating mobile apps for breathing training: The effectiveness of visualization, *Computers in Human Behavior*, *40*, 2014, pp. 56-63.

Evaluating Mobile Apps for Breathing Training: the Effectiveness of Visualization

Luca Chittaro and Riccardo Sioni

Human-Computer Interaction Lab
Department of Mathematics and Computer Science
University of Udine
Via delle Scienze 206, 33100, Udine, Italy

Abstract

Deep and slow breathing exercises can be an effective adjunct in the treatment of stress, anxiety, post-traumatic stress disorder, chronic pain and depression. Breathing techniques are traditionally learned in courses with trainers and/or with materials such as audio CDs for home practice. Recently, mobile apps have been proposed as novel breathing training tools, but to the best of our knowledge no research has focused on their evaluation so far. In this paper, we study three different designs for breathing training apps. The first employs audio instructions as in traditional training based on audio CDs, while the other two include visualizations of the breathing process, representative of those employed in current breathing training apps. We carry out a thorough analysis, focusing on users' physiological parameters as well as subjective perception. One visualization produces better results both objectively (measured deepness of breath) and subjectively (users' preferences and perceived effectiveness) than the more traditional audio-only design. This indicates that a visualization can contribute to the effectiveness of breathing training apps. We discuss which features could have allowed one visualization (but not the other) to obtain better results than traditional audio-only instructions.

Keywords: training, mobile devices, evaluation**,** breathing, visualization, health

1. Introduction

Deep and slow breathing, as well as more complex breathing exercises such as Yogic breathing, have been found to be an effective adjunct in the treatment of stress, anxiety, post-traumatic stress disorder (PTSD), chronic pain and depression (Brown & Gerbarg, 2005; Bush et al., 2012; Han, Stengen, De Valck, Clément, & Van de Woestijne, 1996), and in the achievement of relaxation (Grossman, Grossman, Schein, Zimlichman, & Gavish, 2001). These benefits can be explained by the fact that deep and slow breathing exercises contrast the effects of fast and shallow chest breathing, which is a common automatic habit, e.g. in patients with anxiety disorders (Hazlett-Stevens, 2008). In general, hyperventilation can lead to physical sensations resembling anxiety (Hazlett-Stevens, 2008) and symptoms typical of panic attacks (Conrad et al., 2007). Furthermore, several medical studies have reported that breathing exercises can also have positive effects on the circulatory system, by helping to lower blood pressure (Grossman et al., 2001; Joseph et al., 2005; Radaelli et al., 2004). Integrating breathing exercises with more general relaxation training can also ameliorate respiratory symptoms in patients with asthma (Holloway & West, 2007).

Breathing exercises are traditionally learned in courses with trainers, e.g., courses for pregnant women in preparation to labor, programs for dealing with medical conditions such as hypertension or stress-related disorders, stress management courses for different professions that encounter stressful situations in the field (e.g., soldiers, first responders, nurses, doctors…). Moreover, deep and slow breathing is useful also for handling stress in personal situations such as school exams, job interviews, public speaking and various kinds of social interactions. Audio CDs are often used to support breathing practice at home. Recently, some authors (Elliott et al., 2004; Liu, Huang, & Wang, 2011; Mitchell, Coyle, O’Connor, Diamond, & Ward, 2010; Schein et al., 2001) have focused on specialized hardware to guide the trainee during breathing exercises. In general, these systems exploit wearable devices like elastic girth sensors to record users’ breathing activity (e.g., Liu et al, 2011; Mitchell et al., 2010), sometimes combining them with other devices, e.g., a portable music player (Elliott et al., 2004; Schein et al., 2001). The goal of these systems is to provide real-time adaptation of the training to users’ physiology. Unfortunately, specialized hardware may be costly and not always available when required, e.g., when stress strikes. Furthermore, the transition to applying the skills learned through exercises in everyday life remains a challenge (Morris & Guilak, 2009).

Smartphones can be a novel opportunity to improve breathing training. They follow their users anywhere, so a breathing training app could be always available to support users at any moment. Moreover, the cost of a smartphone app is typically low, and it is relatively easy to find free apps. In general, smartphones are increasingly seen as a versatile m-health instrument for treatment and training, in medicine as well as psychology (Miller, 2012), and some authors predict that the mobile phone will emerge as the preferred personal coach for the 21st century (Morris & Guilak, 2009). A few mobile apps for breathing training have been proposed in the literature, e.g., in the Mobile Heart Health project (Morris & Guilak, 2009). Moreover, app stores such as Apple’s App Store and Google Play Store are making available a growing number of mobile apps for breathing training, developed both by small enterprises (e.g., Universal Breathing: Pranayama (Saagara, 2011)) and organizations such as the US National Center for Telehealth and Technology (T2), a part of the US Military Health System (MHS). In particular, T2 has recently launched Tactical Breather (National Center for Telehealth and Technology, 2011a), a mobile app for repetitive breathing training. The goal of Tactical Breather is to help soldiers in gaining a better control of heart rate, emotions, concentration, and other physiological and psychological responses during stressful situations (National Center for Telehealth and Technology, 2011a).

Unfortunately, to the best of our knowledge, no research study has yet focused on the evaluation of mobile breathing training apps. In this paper, we study three different designs for such apps. The first employs audio instructions as in traditional training based on audio CDs. The other two include also visualizations of the breathing process, which are representative of approaches followed in current breathing training apps. We carry out a thorough analysis of participants' physiological parameters as well as their subjective perception and preferences.

The paper is organized as follows. Section 2 briefly reviews the various apps for breathing training, while Section 3 provides additional motivations for our research. In Section 4, we describe the three designs considered in the present study. Then, Section 5 and 6 illustrate in detail the experiment and its results, Section 7 discusses the results, while Section 8 presents conclusions and future work.

2. Related Work

Visualization plays an important role in many kinds of medical applications (Chittaro, 2001) and is increasingly employed in mobile devices to make the information provided by their applications easier to understand (Chittaro, 2006). While some breathing training apps rely on audio-only instructions (as in the traditional approach based on audio CDs), more innovative apps are attempting to enrich audio instructions with interactive visualizations. Some of these apps share a common approach (called *circle-based visualization* in the following): they display a circle (or a sphere) that expands and contracts, which might suggest the expansion and contraction of human lungs during the breathing activity. For example, Morris et al. (Morris and Guilak, 2009; Morris et al. 2010) animate a simple blue circle to encourage deliberate and slower breathing (Morris et al. 2010). To better highlight the different phases of breathing, Tactical Breather (National Center for Telehealth and Technology, 2011a) also changes the color of the circle. During the inhalation phase, the circle is green and grows in size; in the hold phase (during which users must hold their breath), the circle turns yellow and its size remains constant; in the exhalation phase, the circle turns red and shrinks until it reaches its minimum size, which is marked by a light black circle in the center of the screen. During each phase, voice instructions first pronounce the name of the phase ("Inhale", "Hold" or "Exhale"), then count from 2 to 4 to give users an indication of the progress and duration of the phase. The phase and the pronounced number are also displayed as text on the screen. ColorBREATH (3CUBEs, 2012) substitutes the solid color circle with an iridescent bubble. ECNA-Breath (ECNA LAB, 2011) employs a green sphere that expands and contracts, but it allows users to adjust the maximum size of the sphere before starting the exercise to better fit it with their actual breathing depth. Breathing Zone (Breathing Zone, 2011) employs a multicolor geometric shape similar to a lotus flower instead of a circle. Two different sounds, that can be chosen among various tones (e.g., two guitar chords), announce the beginning of the inhalation and exhalation phases.

Some apps follow a different approach (called *wave-based visualization* in the following), employing a wave-like line to show the optimal respiratory cycle. For example, Vital-EQ Respiroguide (Landelijk Centrum Stressmanagement, 2009) shows the breathing pattern as a sine wave that advances over time during the exercise. A yellow circle, locked on the curve and at the center of the screen, moves upwards and downwards following the wave to guide the user through the inhalation and exhalation phases. A wave-like line is also employed in Paced Breathing (IQPuzz, 2013): in this app, the wave represents a single respiratory cycle and does not move. During the exercise, a white circle follows the line to indicate the current position in the respiratory cycle. A tone with increasing or decreasing pitch is employed as an audio hint to indicate inhalation and exhalation respectively. The shape of the wave in terms of the length of breathing phases, but not their amplitude, is customizable by the user.

Two apps, Universal Breathing: Pranayama (Saagara, 2011) and TotalAwake (TotalAwake, 2012), have tried pie charts instead of waves. Pie sectors represent the different breathing phases and during a respiratory cycle, the chart is gradually filled with color to guide users through the different phases. Breathe2Relax (National Center for Telehealth and Technology, 2011b) employs a cylinder shape which fills up and empties out to indicate when to inhale and exhale. De Stress (Designit, 2009) employs instead animated arrows, which move towards or away from the screen to visually represent the inhalation and exhalation phases. The arrows move with decreasing speed during the exercise to help users gradually slow down their breathing frequency.

Finally, three apps resort to visualizations of the human body. Universal Breathing: Pranayama and Paced Breathing, display a realistic 3D model of a human in the so-called "Burmese" posture, animated to represent the expansions and contractions of chest and abdomen during the exercise. A small blue arrow indicates the air coming in and out through the nose of the human. MyCalmBeat (MyBrainSolutions, 2012) introduces a silhouette of a human chest that encloses two 3D lungs. An animation shows when to inhale or exhale: a tone indicates the start of the inhalation phase; the 3D lungs expand and get gradually filled with oxygen, represented by a blue shade. When the lungs are full, a tone indicates the start of the exhalation phase, and the animation is reproduced in reverse order.

3. Motivations and Goals

The lack of studies of the effects of mobile breathing training apps on users does not allow one to make substantiated claims about the effectiveness of the above described approaches. On one side, visualizations may offer easier to understand instructions and help trainees in reaching the optimal breathing pattern. On the other side, they could possibly distract users' attention from respiratory interoception which is important in controlling breath, and be detrimental to reaching slow and deep breathing. Their effects on breathing must thus be studied and contrasted with the traditional audio-only approach. For this reason, our study will consider the two currently most used types of visualization (summarized in the previous section) as well as the more traditional audio-only approach, to test if the inclusion of a visualization can improve the latter.

Moreover, we should consider that mobile breathing visualizations differ in the amount of information conveyed to users. The most limited ones only suggest the current phase of the respiratory cycle, as it is the case of those circle-based visualizations which convey information only through circle contraction and expansion, without giving an explicit indication of how long the user is supposed to keep inhaling or exhaling. Some apps, like Tactical Breather, try to overcome this limitation by providing additional visual and acoustic hints such as counting the duration of each phase, or marking the minimum size of the circle, i.e., the moment when users will have to stop exhaling. However, these suggestions may not be detailed and clear enough for novice users who have just begun practicing deep and slow breathing. From this point of view, some wave-based visualization such as Vital-EQ Respiroguide might be more clear, because they provide users with temporal information not only about the current breathing phase but also about the previous and the following phase, possibly allowing users to gain greater awareness of where they are in the planned respiratory cycle and thus foresee with greater precision when to switch between phases. For this reason, we hypothesize that wave-like visualizations could allow users to modulate more easily their breathing pattern and better match it with the suggested pattern. Our study thus aims also to provide guidance in the choice between circle-based and wave-based approaches.

4. The Considered Designs

For our study, we developed three versions of a mobile app for the Android platform that differ only in the way instructions are presented. The app guides users in practicing deep and slow circular breathing: a breathing cycle is composed of two phases (inhalation and exhalation) of the same length and the breathing frequency is 6 cycles per minute (0.1 Hz). As in many breathing training apps, the first version of the app (called *Voice-only* in the following) provides instructions through audio only. In particular, we employ the same voice instructions of Tactical Breather, i.e. “inhale”, “2”, “3”, and “4” for the inhalation phase, and “exhale”, “2”, “3”, and “4” for the exhalation phase. The time between any two consecutive instructions is set at 1.25 s.

The second version of the app augments Voice-only with a circle-based visualization. This version of the app (called *Sphere* in the following) was implemented to replicate the visualization used for inhalation and exhalation in Tactical Breather: it employs a green inflating sphere (see Fig. 1) to instruct users to inhale, and a red deflating sphere to instruct them to exhale. As in Tactical Breather, a light black circle is shown at the center of the screen to indicate the minimum size of the sphere and the audio instructions are also visually presented: numbers are shown in the center of the sphere (“1” is simultaneous to the voice instruction “inhale” or “exhale”, while “2”, “3” “4” appear simultaneously to their audio versions); the textual indication about the current breathing phase (“inhale” or “exhale”) is shown at the bottom of the screen. Two screenshots of the visualization are shown in Fig. 1.

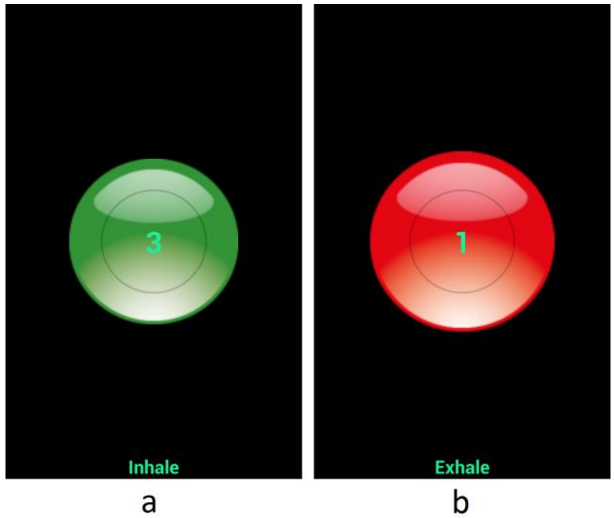

Figure 1. Two screenshots from the Sphere design during the inhalation (a) and exhalation (b) phases, in which the sphere is green and red respectively.

The third version of the app augments Voice-only with a wave-based visualization. This version of the app (called *Wave* in the following) employs a moving green triangle wave (see Fig. 2) to instruct users about when to inhale (the wave is rising) and exhale (the wave is falling), similarly to existing apps like Vital-EQ Respiroguide and Paced Breathing. A vertical translucent white strip, at the center of the screen, is superimposed over the wave to indicate where the user should be in the respiratory cycle at the current instant of time. The previously described audio instructions are also visually presented in the top left corner of the screen. As shown by Fig. 2, during an exhalation (respectively, inhalation) phase, users can see the falling (rising) phase of the triangular wave and also the following rising (falling) phase and the previous rising (falling) phase of the wave.

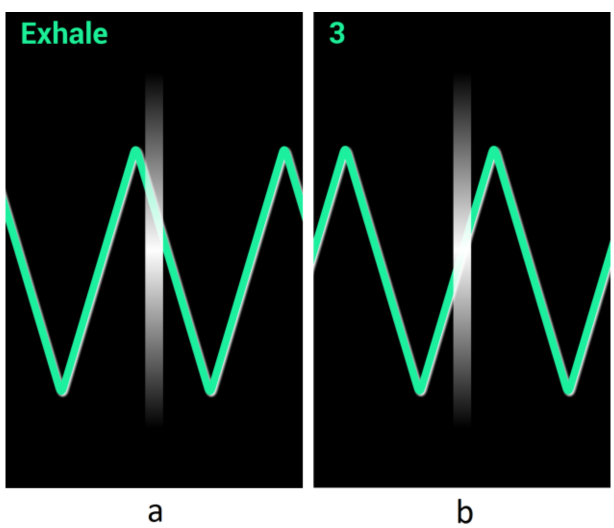

Figure 2. Two screenshots from the Wave design during the exhalation (a) and inhalation (b) phases.

5. Materials and Methods

Our experiment follows a within-subject design with app version (Wave, Sphere, and Voice-only) as the independent variable. We measure possible differences in participants' physiological responses during the exercises, as well as in their subjective perceived effectiveness of the three app designs.

5.1. Materials

The app was run on an HTC Desire equipped with a 3.7'', 480 x 800 pixels display. During the evaluation, the device was placed on a tilted rigid stand placed in a preset position, to maximize visual clarity by avoiding issues like glare and reflections from light sources. Participants' physiological data was recorded and processed on a PC, whose screen could not be seen by participants. To record participants' physiological data, we employed a clinical system (Thought Technology Procomp Infiniti) equipped with 3 sensors: (i) an electrodermal activity (EDA) sensor placed on the intermediate phalanges on the middle and little fingers; (ii) a photopletysmograph (PPG) sensor for measuring blood volume pulse, placed over the distal phalanx of the index finger; (iii) an elastic girth sensor, placed over the participant's navel, to correctly record diaphragmatic respiratory activity.

A demographic questionnaire asked participants about their age and gender, as well as their experience with relaxation and meditation techniques, as suggested in (Feldman, Greeson, & Senville, 2010).

To measure participants' perceived effectiveness of each app version, we employed a questionnaire with the following 6 items: (i) The app facilitates relaxation; (ii) The app is pleasant to use; (iii) It is easy to follow the app instructions; (iv) The app effectively teaches how to breath; (v) The app is effective in reducing stress; (vi) The app is effective in increasing attention to breath. Each item was rated on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). The items employed in the questionnaire are based on the structure of existing items and keywords used in usability questionnaires such as the System Usability Scale (Brooke, 1996) (item (iii)), and in different user experience questionnaires focused on flow and engagement such as EGameFlow (Fu, Su, & Yu, 2009) (items (iii), (iv), (v), and (vi)), Player Enjoyment Scale (Chu Yew Yee, Been-Lirn Duh, & Quek, 2010) (item (ii)), and Game Engagement Questionnaire (Brockmyer et al., 2009) (item (i)).

Participants' ratings of the first and fifth items were averaged to assess the effectiveness of each app version in relaxing the participant, because those are the only items which mention effects on well-being (relaxation and stress reduction).

Finally, we asked participants to rank the three app versions, on the basis of which they would prefer to use (1 = best, 3 = worst). Ties were allowed.

5.2. Participants

The considered sample involved 68 participants (43 M, 25 F), recruited among graduate and undergraduate students of our university and people from other occupations. Participants have been recruited by asking for volunteers among students, who then recruited additional volunteers among their acquaintances. Participants did not receive any compensation for participation in the study. Age ranged from 17 to 60 (M = 24.73, SD = 5.57). All participants had little or no experience with breathing techniques and were thus representative of users who could be supported by a training app. More specifically, the majority of participants (55) was completely unfamiliar with exercises that involve breath control (relaxation or meditation), while 13 participants had some familiarity with such exercises but did not practice them.

5.3. Procedure

Participants were informed that the goal of the study was to evaluate three different versions of a mobile app that would guide them in practicing breathing exercises. All participants consented to participate in the experiment and to have physiological signals recorded. They were also clearly informed that all the experimental data were going to be collected and analyzed anonymously for research purposes. After they filled the demographic questionnaire, the girth sensor was applied accurately over their navel to measure diaphragmatic breathing. Then, they were asked to sit, and the PPG and EDA sensors were applied.

Before each condition, participants were asked to relax (the request was deliberately generic and did not mention breathing) for about a minute, to record the baseline for the physiological signals, i.e., the signal values that can be observed when participants are in a resting state. This was done because, when analyzing physiological data in user studies, baseline values have to be subtracted from the data recorded during the experimental conditions, to separate the physiological responses to experimental conditions from the intrinsic physiological differences among participants (Andreassi, 2007). In this way, baseline subtraction allows us to evaluate the physiological responses due to using the breathing training apps, which are likely to require effort to follow and practice the instructions in users who are not already trained. During baseline recording, participants watched a video with nature sounds and images. After each baseline recording, we introduced them to the app version they were going to try next. We clearly explained that, during the exercise, they would have to breathe with the abdomen, trying to follow the instructions given by the app as close as possible. We suggested participants to breathe as deeply as they felt comfortable with. Then, participants practiced breathing training with the given app for 3 minutes. To minimize order effects, as recommended in (Cohen, 2001) when dealing with three experimental conditions, a complete counterbalancing scheme was followed to present the treatments in all possible orders. In this way, no condition should attain an unfair advantage by virtue of its ordinal position (Cohen, 2001).

After trying all three designs, participants filled the perceived effectiveness questionnaires, and then gave the final ranking. They were then debriefed and thanked for their participation.

5.4. Measures

For each participant, we collected the following measures:

- *Skin conductance level* (SCL): the mean value of the SCL measured during each breathing exercise, after the subtraction of the mean baseline value. SCL is commonly employed in the literature as a physiological parameter that measures arousal (Andreassi, 2007; Boucsein, 2006);
- *Heart rate* (HR): the mean value of HR measured during each breathing exercise, after the subtraction of the mean baseline value. Increases in HR are generally related to emotional activation and are used in the literature as a correlate of arousal (e.g., Mandryk & Atkins, 2007). As reported by Boucsein (2006), compared to EDA, HR is well suited as an indicator for the higher arousal range and for pronounced arousal processes;
- *Power of the recommended frequency band*: the signal power in the 0.09–0.11 Hz band of the

respiratory signal power spectrum, measured during each breathing exercise. This measure represents the intensity of the respiratory component near the breathing frequency encouraged by the app (0.1 Hz, i.e., 6 cycles per minute). Higher values for this measure indicate deeper breaths at the encouraged frequency (which is the goal of the deep and slow breathing exercises) when comparing breathing patterns with similar waveforms;

- *Respiratory signal-to-noise ratio* (SNR): the ratio between (i) the power of the recommended breathing frequency band and, (ii) the power in the entire spectrum excluding the band of the recommended frequency and the 0–0.05 Hz band to remove low-frequency fluctuations and the direct current (DC) offset. It describes the ability of participants to correctly follow the instructions received by the app: the more the participant struggles in following the instructions or ignores the instructions, the more the power of the components outside the recommended breathing frequency band, especially high-frequency components, increases. For example, a participant who is unable to inhale further before the inhalation phase is complete might perform a brief exhalation to keep inhaling until the end of the phase. Breathing patterns with similar waveforms will produce similar SNR values;
- *Perceived effectiveness*: the mean score from the questionnaire collecting participants' subjective perception of the effectiveness of each app design. Answers to questions were averaged to form a reliable scale, Cronbach's alpha: Wave = 0.84, Sphere = 0.84, Voice-only = 0.84;
- *Perceived relaxation effectiveness*: the mean score from the first and fifth items of the questionnaire. Answers to questions were averaged to form a reliable scale, Cronbach's alpha: Wave = 0.79, Sphere = 0.82, Voice-only = 0.80;
- *Subjective preference*: rank assigned by participants to each design.

6. Results

To evaluate the factorial validity of the questionnaire, we performed an exploratory factor analysis, as suggested in (Coolican, 2009), on the three questionnaire data sets (one for each condition). . Post-hoc indicators of data and sampling adequacy (Kaiser-Meyer-Olkin, KMO, and Bartlett's Test of Sphericity) (Table 1) show that, for all data sets, KMO was greater than 0.60, and Bartlett's test of sphericity was significant ($p < 0.05$). The data sets were thus suitable for factor analysis (Pallant, 2011).

Condition	KMO	Bartlett's Test of Sphericity p<
Wave	0.73	0.001
Sphere	0.74	0.001
Voice-only	0.81	0.001

Table 1. KMO and p values of Bartlett's Test of Sphericity observed on questionnaire data.

For all three data sets, Kaiser's criterion determined the presence of only one component, explaining a total of 55.26%, 55.95%, and 55.96% of the variance respectively in Wave, Sphere and Voice-only. The factor analysis shows that the loadings of all six items are above 0.40 (Table 2), which indicates that the items load quite strongly. The observed results confirm that the six items can be employed, as in the present study, to measure the perceived effectiveness construct.

Condition	Item (i)	Item (ii)	Item (iii)	Item (iv)	Item (v)	Item (vi)
Wave	0.77	0.71	0.64	0.85	0.79	0.67
Sphere	0.82	0.80	0.64	0.79	0.76	0.66
Voice-only	0.85	0.83	0.66	0.71	0.83	0.56

Table 2. Loadings of Questionnaire items in the three conditions.

We performed a multivariate analysis of variance (MANOVA) of the physiological data and questionnaire results. As recommended by (Cohen, 2001), to reduce the risk of type I errors, we adopted a conservative alpha level, setting it at 0.0083 (Bonferroni correction) for subsequent univariate tests and pairwise comparisons. This value is obtained by dividing the desired alpha level for the MANOVA (0.05) by 6, which is the number of dependent variables included in the MANOVA. A separate non-parametric analysis of variance (Friedman test) was performed on the subjective preference data. Physiological data for the respiratory activity recorded from 5 participants had to be removed due to a malfunctioning girth sensor during a day of the recording sessions. HR data from 9 participants had to be removed because of a large number of artifacts in the BVP signal due to hand movements during the recording session.

The analysis revealed a main effect of visualization (Wilks' $\lambda = 0.68$, $F(2, 106) = 3.52$, $p < 0.001$, $\eta_p^2 = 0.17$). Univariate tests with Greenhouse-Geisser correction showed significant differences in power of the recommended frequency band ($p < 0.0083$, $\eta_p^2 = 0.10$), perceived effectiveness ($p < 0.001$, $\eta_p^2 = 0.17$), and perceived relaxation effectiveness ($p < 0.0083$, $\eta_p^2 = 0.09$) (Fig. 3a, 3c, 3d). Differences in Respiratory SNR, SCL, and HR did not reach significance.

Pairwise comparisons (with Bonferroni correction) for power of the recommended frequency band revealed a significant difference between the Wave (M = 10.78, SD = 10.46) and the voice-only (M = 7.98, SD = 8.18) conditions ($p < 0.0083$). Similarly, pairwise comparisons among perceived effectiveness scores revealed a significant difference between the Wave (M = 3.85, SD = 0.69) and the Voice-only (M = 3.18, SD = 0.91) conditions ($p < 0.001$), and pairwise comparisons among perceived relaxation effectiveness scores revealed a significant difference between the Wave (M = 3.56, SD = 0.88) and the Voice-only (M = 2.96, SD = 1.22) conditions ($p < 0.0083$).

Finally, Friedman test revealed a significant difference in subjective preference among the three conditions ($\chi^2 = 14.84$, $p < 0.001$) (Fig. 4), and a pairwise comparison with Dunn's test (with Bonferroni correction) revealed a significant difference between the Wave (M = 1.59, SD = 0.74) and the Voice-only (M = 2.24, SD = 0.83) conditions ($p < 0.001$).

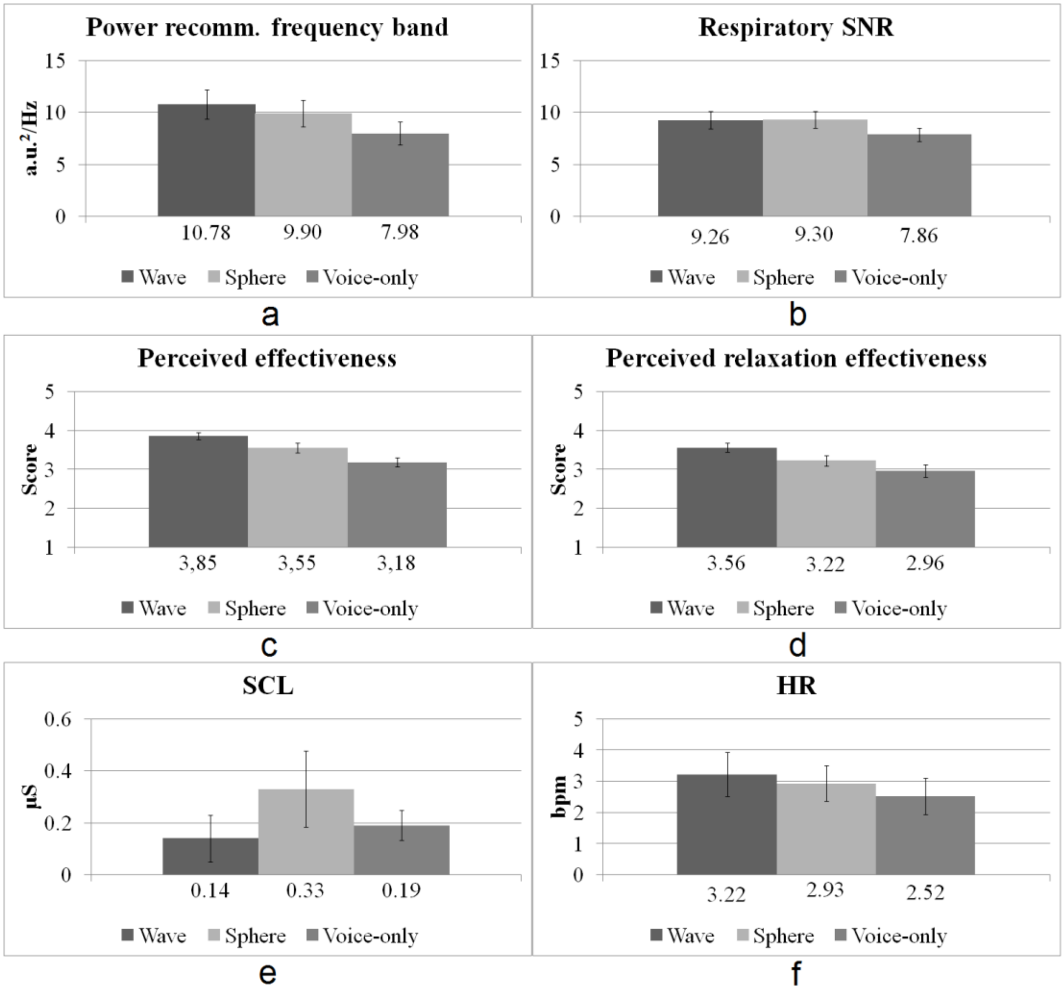

Figure 3. Mean values of power of the recommended frequency band (a), respiratory SNR (b), perceived effectiveness (c), perceived relaxation effectiveness (d), SCL (e), and HR (f). Error bars indicate standard error of the mean. a.u. indicates the arbitrary units provided by the elastic girth sensor to measure its tension.

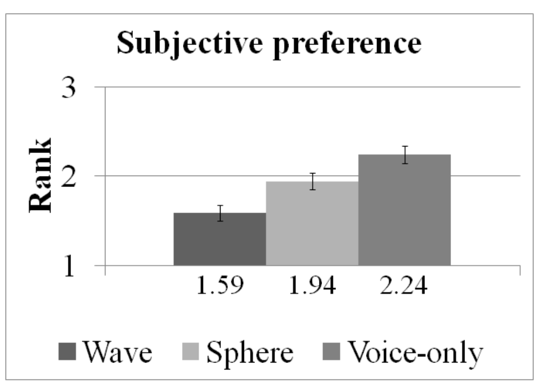

Figure 4. Mean values of subjective preference. A lower value indicates higher preference. Error bars indicate standard error of the mean.

7. Discussion

The waveform of the breathing signal was similar among conditions as indicated by the lack of significant differences in mean respiratory SNR (as well as visual monitoring of physiological recordings). This suggests that the (visual and auditory) instructions provided by the three evaluated designs were good enough to allow participants to quickly adapt their breathing frequency to the suggested one for the duration of the exercise. The lack of significant differences in mean respiratory SNR suggests that the difference in power of the recommended frequency band values between Wave and Voice-only is not due to possibly different waveforms. Therefore, higher values in mean power of the recommended frequency band with Wave are originated by significantly deeper breaths at the encouraged frequency with respect to Voice-only. This is probably due to the fact that Wave allows users to keep better track over time of the ideal breathing pattern they have to follow during the exercise. The visualization provided by Wave could have allowed participants to assess more easily the duration of the breathing phases, thus helping them to maximize the quantity of air to inhale or exhale in each respiratory cycle with respect to Voice-only. On the contrary, receiving only audio instructions as in Voice-only could have at times resulted in a more shallow breathing in participants to avoid finding themselves unable to further inhale (or exhale) before the end of the inhalation (exhalation) phase.

Although Sphere also produced deeper breathing at the suggested frequency with respect to Audio-only (Fig. 3a), the difference did not reach significance. However, it is interesting to note that the average obtained by Sphere is only slightly smaller than Wave. The presence of the light black circle and the inflating and deflating sphere, which can be associated with human lungs' expanding and contracting, might have partially helped participants in calibrating their inhalation and exhalation.

The same considerations apply to subjective measures: perceived effectiveness (Fig. 3c) and perceived relaxation effectiveness (Fig. 3d) were significantly higher with Wave rather than Voice-only, and participants reported a significantly higher preference for Wave over Voice-only (Fig. 4). Mean values for perceived relaxation effectiveness are very similar to the general perceived effectiveness construct, which has a broader scope that includes clarity of the instructions. From a physiological point of view, the perception of better effectiveness is consistent with the deeper breathing that has been physiologically measured through the power of the recommended frequency band.

In physiological variables, while we observed statistically significant differences in power of the recommended frequency band as discussed above, differences in variables indicating arousal (SCL and HR, see Section 5.4) were very small and not statistically significant. This result is not surprising: while a breathing training app that is perceived as very effective by users may likely facilitate them in lowering their arousal level in the long term, we have to consider that in our experiment participants had no or very little experience of relaxation through deep and slow breathing and used the mobile app for just a few minutes.

Since there were more males than females in the sample, we performed a mixed design two-factors MANOVA (with app version as the within-subject variable and gender as the between-subject variable) to assess if participants' gender could be a factor in the observed results. The analysis found neither a main effect of gender nor an interaction with gender.

The contribution of the obtained results is twofold. First, they show that augmenting audio-only breathing training apps with an interactive visualization to help trainees in gaining greater awareness of the breathing process can increase user's effectiveness in obtaining deep and slow breathing. Without the results provided by our evaluation, one could not exclude, as noted in the introduction, that the addition of a visualization might distract users' attention from the respiratory interoception required to carry out breathing exercises, and thus be detrimental to the training. Second, our study sheds light on differences between the two currently most used types of visualizations in breathing training apps. All statistically significant differences from Audio-only were obtained with the Wave visualization, which performed consistently better. Even extending consideration to the measures for which significance was not reached, Sphere and Audio-only never show better values than Wave. It should also be noted that the advantages observed with the Wave design can be easily obtained by existing and future mobile breathing training apps by augmenting them with the Wave visualization described in this study, which should be a straightforward and relatively inexpensive addition.

By looking critically at the results, we must anyway remember that the Wave and Sphere visualizations augment Audio-only with additional information. In particular, Wave always shows two complete breathing cycles and the vertical white strip (see Figure 2), helping users to keep track of the ideal breathing pattern, the current breathing phase (inhaling or exhaling), and the exact point of the phase they have reached at the moment. In Sphere, the light black circle at the center of the screen provides users with a persistent hint about the duration of the exhalation phase, and the color of the sphere indicates at any time the current breathing phase. To obtain the same information in Voice-only, users have to mentally keep track of the current breathing phase as well as the rhythm of the voice instructions provided. A limitation of the study could thus concern the fact that we did not evaluate if more complex audio could be included in Voice-only to provide users with exactly the same information of the Wave and Sphere visualizations, but keeping an approach that frees the eyes from the task of visual tracking.

8. Conclusions

The present study provides experimental support to the recent proposal of using mobile apps for breathing training. We analyzed three major designs which are employed in current breathing training apps: a more traditional one based only on audio instructions, and two novel ones which augment audio instructions with different kinds of visualizations. The design based on a Wave visualization was able to produce better results compared to the Voice-only design, both in objective terms (measured deep and slow respiration) and subjectively (users found it more effective from an instruction as well as a relaxation point of view). To the best of our knowledge, our study is the first to focus on the evaluation of mobile apps for breathing training and to shed light on the positive role that visualization can play in such applications.

Thanks to the availability of mobile devices anytime and anywhere, trainees can have a convenient access to effective and affordable breathing training apps that could be integrated more easily in their daily routine with respect to desktop applications or traditional approaches (following courses given by instructors and using audio CDs). As we have seen, a wave-based mobile visualization also allows to obtain better results with respect to audio-only instructions. The obtained results could be useful also to interventions (e.g., for hypertension or stress-related disorders) which adopt phones as an m-health instrument for patient training.

To evaluate the long-term effects of different apps, the next phase in our research will concern the design of a longitudinal study to gain broader insights about how the applications are used by trainees in their daily life, the possible differences in facilitating the transition from app-supported respiration to independent application of the breathing techniques in everyday life, and the possible differences in the level of wellbeing obtained, e.g., users' ability to keep their stress level under control. We plan to focus on specific user categories, in particular nurses and doctors in the context of workplace stress reduction courses. Furthermore, our intention is to evaluate in more detail the different roles played by the (i) visualization employed to convey the breathing instructions to users, and (ii) the instructions themselves. To this purpose, we plan to carry out a study in which we try to augment the traditional audio-only approach with additional information provided through sound instead of visualization, comparing the obtained eyes-free interface with the Wave-based visual approach. Finally, future studies will also explore further the design space of visualizations, by considering improvements to the two types of visualizations evaluated in the current study as well as extending our attention to other visualizations we described in Section 2. In particular, we are interested to assess if animations that resort to human body visualization might further improve the results obtained in the current study.

References

3CUBEs (2012). *ColorBREATH*. https://itunes.apple.com/us/app/colorbreath/id522637619?mt=8 (last accessed on 11 June 2013).

Andreassi, J. L. (2007). *Psychophysiology – Human Behavior and Physiological Response*. Mahwah, NJ, USA: Lawrence Erlbaum Associates.

Boucsein, W. (2006). *Electrodermal Activity*. New York, NY, USA: Springer.

Breathing zone (2011). *Breathing Zone.* http://www.breathing-zone.com/ (last accessed on 11 June 2013).

Brockmyer, J. H., Fox, C. M., Curtiss, K. A., McBroom, E., Burkhart, K. M., & Pidruzny, J. N. (2009). The development of the Game Engagement Questionnaire: a measure of engagement in video game-playing. *Journal of Experimental Social Psychology 45*(4), 624-634.

Brooke, J. (1996). SUS – A quick and dirty usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmester & I. L. McClelland (Eds.), *Usability Evaluation in Industry*, (189-194) London, UK: Taylor & Francis.

Brown, R. P., & Gerbarg, P. L. (2005). Sudarshan Krya Yogic breathing in the treatment of stress, anxiety and depression. Part II – clinical applications and guidelines. *Journal of Alternative and Complementary Medicine, 11*(4), 711-717.

Bush, V., Magerl, W., Kern, U., Haas, J., Hajak, G., & Eichhammer, P. (2012). The effects of deep and slow breathing on pain perception, autonomic activity, and mood processing – an experimental study. *Pain Medicine 13*(2), 215-228.

Chittaro, L. (2001). Information visualization and its application to medicine. *Artificial Intelligence in Medicine, 22*(2), 81-88.

Chittaro, L. (2006). Visualizing information on mobile devices. *IEEE Computer, 39*(3), 40-45.

Cohen, B. H. (2001). *Explaining Psychological Statistics* (2nd ed.). New York, NY, USA: John Wiley & Sons.

Chu Yew Yee, S. L., Been-Lirn Duh, H., & Quek, F. (2010). Investigating narrative in mobile games for seniors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2010)*. New York, NY, USA: ACM.

Coolican, H. (2009). *Research Methods and Statistics in Psychology* (5th ed.). London, UK: Hodder Education.

Conrad, A., Müller, A., Doberenz, S., Kim, S., Meuret, A. E., Wollburg, E., & Roth, W. T. (2007). Psychophysiological effects of breathing instructions for stress management. *Applied Psychophysiology and Biofeedback, 32*(2), 89-98.

Designit (2009). *De Stress.* http://healthmate.eu/destress (last accessed on 11 June 2013).

ECNA LAB (2011). *ECNA-Breath.* https://itunes.apple.com/us/app/ecna-breath/id386514251?mt=8 (last accessed on 11 June 2013).

Elliott, W. J., Izzo, J. L., White, W. B., Rosing, D. R., Snyder, C. S., Alter, A., Gavish, B., Black, H. (2004). Graded blood pressure reduction in hypertensive outpatients associated with use of a device to assist with slow breathing. *Journal of Clinical Hypertension, 6*(10), 553-559.

Feldman, G., Greeson, J., & Senville, J. (2010). Differential effects of mindful breathing, progressive muscle relaxation, and loving-kindness meditation on decentering and negative reactions to repetitive thoughts. *Behaviour Research and Therapy 48*(10), 1002-1011.

Fu, F. L., Su, R. C., & Yu, S. C. (2009). EGameFlow: a scale to measure learners' enjoyment of e-learning games. *Computers & Educations 52*(1), 101-112.

Grossman, E., Grossman, A., Schein, M. H., Zimlichman, R., & Gavish, B. (2001). Breathing-control lowers blood pressure. *Journal of Human Hypertension, 15*(4), 263-269.

Han, J. N., Stegen, K., De Valck, C., Clément, J., & Van de Woestijne, K. P. (1996). Influence of breathing therapy on complaints, anxiety and breathing pattern in patients with hyperventilation syndrome and anxiety disorders. *Journal of Psychosomatic Research, 41*(5), 481-493.

Hazlett-Stevens, H. (2008). *Psychological Approaches to Generalized Anxiety Disorders*. New York, NY, USA: Springer.

Holloway, E. A., & West, R. J. (2007). Integrated breathing and relaxation training (the Papworth method) for adults with asthma in primary care: a randomized controlled trial. *Thorax 62*(12), 1039-1042.

IQPuzz (2013). *Paced Breathing.* https://play.google.com/store/apps/details?id=com.apps.paced.breathing (last accessed on 11 June 2013).

Joseph, C. N., Porta, C., Casucci, G., Casiraghi, N., Maffeis, M., Rossi, M., & Bernardi, L. (2005). Slow breathing improves arterial baroreflex sensitivity and decreases blood pressure in essential hypertension. *Hypertension 46*(4), 714-718.

Kaushik, R. M., Kaushik, R., Mahajan, S. K., & Rajesh, V. (2006). Effects of mental relaxation and slow breathing in essential hypertension. *Complementary Therapies in Medicine, 14*(2), 120-126.

Landelijk Centrum Stressmanagement (2009). Vital-EQ Respiroguide. http://www.heartintelligence.com/11/index.html (last accessed on 11 June 2013).

Liu, G.-Z., Huang, B.-Y., & Wang, L. 2011. A wearable respiratory biofeedback system based on generalized body sensor network. *Telemedicine and e-Health, 17*(5), 348-357.

Lorig, T. S. (2007). The respiratory system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of Psychophysiology* (231-244). New York, NY: Cambridge.

Mandryk, R. L., & Atkins, M. S. (2007). A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. *International Journal of Human-Computer Studies 65*(4), 329-347.

Miller, G. (2012). The smartphone psychology manifesto. *Perspectives on Psychological Science, 7*(3), 221-237.

Mitchell, E., Coyle, S., O'Connor, N. E., Diamond, D., & Ward, T. (2010). Breathing feedback system with wearable textile sensors. *2010 International Conference on Body Sensor Networks (BSN 2010)*, 56-61.

Morris, M. E., & Guilak, F. (2009). Mobile Heart Health: project highlight. *IEEE Pervasive Computing, 8*(2), 57-61.

Morris, M. E., Kathawala, Q., Leen, T. K., Gorenstein, E. E., Guilak, F., Labhard, M., Deleeuw, W. (2010). Mobile therapy: case study evaluations of a cell phone application for emotional self awareness. *Journal of Medical Internet Research, 12*(2), e10.

MyBrainSolutions (2012). *MyCalmBeat*. https://www.mybrainsolutions.com/Pages/TrainingThatsFun/GoMobile.aspx (last accessed on 11 June 2013).

National Center for Telehealth and Technology (2011a). *Tactical Breather*. https://www.t2health.org/apps/tactical-breather (last accessed on 10 June 2013).

National Center for Telehealth and Technology (2011b). *Breathe2Relax*. http://t2health.org/apps/breathe2relax#.UbcZsPm-13s (last accessed on 10 June 2013).

Pallant, J. (2011). *SPSS Survival Manual: A Step by Step Guide to Data Analysis Using SPSS for Windows* (4th ed.). Crows Nest, AU: Allen & Unwin.

Radaelli, A., Raco, R., Perfetti, P., Viola, A., Azzellino, A., Signorini, M. G., & Ferrari, A. U. (2004). Effects of slow, controller breathing on baroreceptor control of heart rate and blood pressure in healthy men. *Journal of Hypertension, 22*(7), 1361-1370.

Saagara (2011). *Universal Breathing: Pranayama*. http://www.saagara.com/apps/pranayama-app (last accessed on 11 June 2013).

Schein, M. H., Gavish, B., Herz, M., Rosner-Kahana, D., Naveh, P., Knishkowy, B., Zlotnikov, E., Ben-Zvi, N., Melmed, R. N. (2001). Treating hypertension with a device that slows and regularizes breathing: a randomized, double-blind controlled study. *Journal of Human Hypertension, 15*(4), 271-278.

TotalAwake (2012). *TotalAwake*. http://totalawake.com/index.php (last accessed on 11 June 2013).