DOI: 10.61782/fa.2023.1122

MULTIMODAL CUES FOR GUIDED BREATHING

Georgios Marentakis[1*] **Victoria Grace**[2] **Jonathan Berger** [3]
[1] Østfold University College, Norway
[2] MUVIK Labs, New York, USA
[3] CCRMA, Stanford University, USA

ABSTRACT

Guided breathing has been found useful for stress reduction and relaxation. Different types of cues have been used to guide breathing including visual, audio, haptic, and multimodal combinations. However, the possibility for a constructive or destructive interference due to multimodal cues to guided breathing has not been evaluated. In a controlled experiment, we investigated the effect of congruency between auditory and visual stimuli on guided breathing. Conditions included unimodal audio and visual cues as well as multimodal visual and audio cues whose congruency was manipulated to yield congruent as well as incongruent breathing cues. Participants reported stress levels after exposure using a visual analog scale. Stimuli were effective in reducing subjective stress levels in all conditions. However, there was no evidence for constructive or destructive interference in the subjective responses of participants.

Keywords: *guided breathing, stress relief, multimodal interaction, audiovisual*

1. INTRODUCTION

Breathing biofeedback is an increasingly focal factor in e-health [1]. Breathing biofeedback helps regulate stress levels [2] with significant implications for improved physical and mental health [3]. Numerous commercial applications (e.g., Inner Balance, emWave2, myBreath, Pranayama), games e.g., [4], tangible interfaces e.g., [5], or virtual reality and artistic interventions e.g. [1] use breathing biofeedback. Convenience users, veterans [6], drivers [7], children [8], students [9], athletes [10], and other user groups [2] have benefited.

Breathing biofeedback may be delivered in an open and/or a closed loop. In closed loop applications, users typically adjust their breathing while monitoring relevant psychophysiological parameters. Relevant physiological markers used in breathing biofeedback applications are often related to stress and include respiration rate, heart beat, heart rate variability (HRV), or their derivatives.

In open loop applications, users adjust their breathing based on visual, auditory, tactile, or multimodal guided breathing cues. Several feedback designs have been proposed in the literature but few studies evaluate the efficacy of the feedback cues in guiding breathing. Investigating feedback design is therefore relevant. Audiovisual cues are used increasingly often as they are easy to provide and can lead to a more immersive experience. Here, we focus specifically on evaluating audiovisual feedback for guided breathing. Our research question investigates how congruent or incongruent audiovisual cues compare to unimodal visual or auditory cues when it comes to the stress relief of users engaging in guided breathing.

The article is organized as follows. First, we provide background on breathing biofeedback on which we ground our research question. Subsequently, we present a quantitative evaluation study, the results, and a discussion of the findings.

2. BACKGROUND

Visual breathing feedback can be provided using simple geometric shapes [2], but may also involve animations as done when creating biofeedback games e.g., [8]. Visual breathing biofeedback may also be used to create vir-

*__Corresponding author__: georgios.marentakis@hiof.no.

tual reality scenes but also sonic or visual installations [1]. There are also 'social' biofeedback applications which may involve shared tangibles e.g., [11], avatars [12], or photos [13]. Visual open-loop feedback for guiding respiration may simply be a written instruction to inhale or exhale or an animation, for example, an opening and closing circle e.g., [6] or peripheral screen lighting cues.

Physiological parameters in auditory closed-loop applications may be mapped to the frequency or amplitude of a continuous fixed timbre tone or in the signal-to-noise ratio of a musical signal or the number of channels used to playback sound [14]. Heart rate has been mapped to musical parameters e.g., [15] or controlled algorithmic composition e.g, [16]. Natural sounds have also been mixed with music [17]. Auditory open-loop feedback can be speech instructions about when to inhale or exhale, non-speech sound such as metronome or simple tones [18] that cue to breathing phases, musical sound e.g., [19], or pink or white noise [11]. Breath-like cues have also been used to guide breathing [20].

Haptic biofeedback has also been given e.g., in [5,21] and tactile cues may be given using a transducer array e.g., [22]. Some designs feature multimodal cues [11] for example created a tangible combining visual, auditory and tactile cues. Multimodal cues may also appear in commercial applications e.g., Breathing Zone, but also in guided breathing videos found on online platforms. Audiovisual cues are very common as they are easy to provide in contemporary devices.

2.1 Summary and Research Question

Visual, auditory, and tactile cues have often been used in open-loop breathing biofeedback applications. These typically signal discrete states, such as when using a tone or a voice instruction to signal thee start of the inhale and exhale phases. Although several studies use cues of different modalities, few studies investigate the impact of combinations of cues.

Multimodal cues combining several modalities have been shown to be supporting improved performance in a variety of tasks. They may reduce interference with concurrent tasks in other modalities [23] or lead to multimodal facilitation for a single task encoded in several modalities, as in aurally-aided visual search e.g., [24]. Dual-task facilitation has been observed with multimodal warnings provided to participants engaged in a visual task while driving [25,26] and in stock market data monitoring [27]. Dual-task facilitation has also been reported for a flying object detection and classification task performed in parallel with a tracking task e.g. [28].

Multimodal cues are often used in breathing bio-feedback applications as several commercial applications (e.g., Breathing Zone, Breathe+), guided breathing videos, but also researcher papers e.g., [11]. Combining visual and auditory cues in particular is quite common. However, such combinations are often arbitrary and do not consider aspects such as cue congruency in a systematic manner. The design of multimodal cues for guided breathing is therefore a relevant research topic which has not been investigated. As a first step in investigating the design of multimodal cues to guided breathing, our research question is therefore whether audiovisual cue congruency affects the stress relief potential of audiovisual breathing cues in a guided breathing task.

3. METHOD

The research question was evaluated in a controlled study. This was a within-subjects experiment with one independent variable, feedback *modality*, with four levels: visual, auditory, audiovisual congruent and audiovisual incongruent. Both a congruent and an incongruent condition was included in order to check for both constructive or destructive interference due to the varying congruency between the visual and auditory cues. The dependent variable was the perceived stress level of users measured using a Visual Analog Scale (VAS), which is a common and reliable way to assess stress [29]. This was used to calculate stress relief as the difference between the reported stress level before and after exposure to the test trials. It was also used to quantify the effectiveness of the stressor by estimating the difference between the reported stress level before and after the stressor trials. The experiment contained stressor and test trials which were interleaved as seen in Figure 2.

3.1 Stimuli

Figure 1 shows the visual and auditory stimuli used in the test trials. The visual stimulus consisted of concentric circles whose radius increased throughout the inhale phase and decreased throughout the exhale phase. When stimulus modality was visual, a static drone sound without any cues to the breathing cycle was played in the background to provide a similar context as in the other conditions. The auditory stimulus consisted of a bell sound which signaled the beginning of the inhale and exhale breathing cycle. The pitch was higher in the inhale compared to the exhale phase. A drone signal which provided ambience was

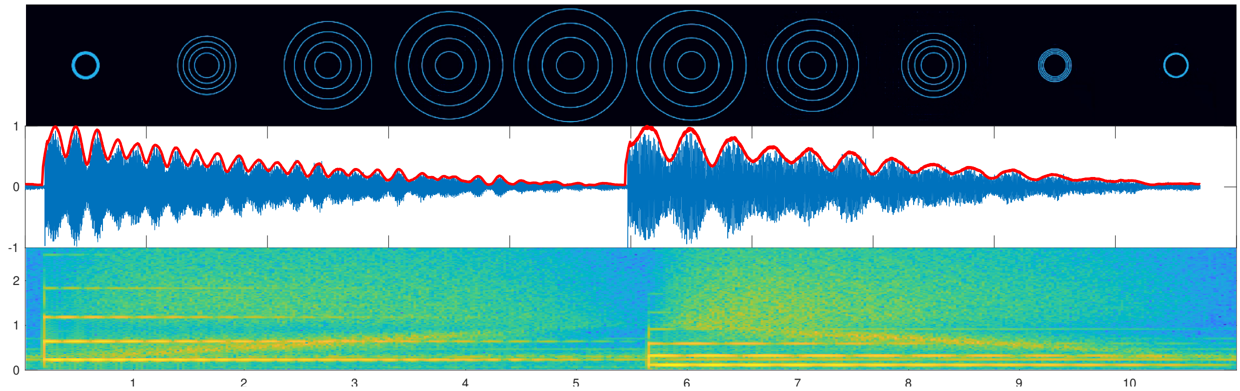

Figure 1: An illustration of the stimuli used in the experiment. The top panel shows instances of the visual stimuli which consisted of oscillating concentric circles of variable radius. The middle panel shows the auditory waveform and the diminishing loudness cues. The lower panel is the sectrogram showing the bell sound and its harmonics and the drone of rising and falling pitch.

I	S1	T1	S2	T2	S3	T3	S4	T4

Figure 2: Setup of test sessions. I is introducion, S1-S4 are stressor trials (1min 15 sec) and T1-T4 are test trials (2 min).

also audible whose pitch swept upwards during the inhaling phase and downwards during the exhale phase. This ambience consists of stacked notes, in harmony with the alternating entrainment sounds (bells), with slow moving LFOs on the amplitude of stacked notes. In the auditory condition, the circle used for visual stimulation was also shown in the screen. However, it remained static and provided no cues to the guided breathing. When modality was audiovisual, stimuli depended on cue congruency. In congruent trials, auditory and visual cues were in sync, pretty much as illustrated in Figure 1. In the incongruent trials, the stimuli started as in the congruent case, however, a delay was gradually introduced to the visual stimulus which led to it increasingly drifting out of sync. The visual stimulus was already delayed compared to the auditory by the end of the first inhale/exhale cycle. In the stressor trials, participants watched a video which instructed them to perform a number of mental arithmetic operations under time pressure while listening to a distressing sound stimulus (a sawtooth wave with a fundamental that swept upwards) and report the result. Mental arithmetic under time pressure is often used to induce stress e.g., [30]. We combined it with a sawtooth wave with a fundamental increasing to uncomfortable levels to accentuate the effect and the time pressure. Test trial duration was 2 minutes and stressor trial duration 1 minute and 15 seconds. The stimulus was prepared for a slow (5 breaths per minute) respiration pace.

3.2 Participants

A total of 16 healthy adults (μ = 24.9 years, σ=8.67 years, 10 male, 6 female) participated in the study. Subjects did not report any respiration problems. They provided informed consent before the experiment. Participation was voluntary and could be terminated at any time.

3.3 Procedure

Each experiment session started with explaining what will happen in the experiment followed by a demonstration of breathing along the stimuli from the experimenter. Subsequently, interleaved test and stressor trials were presented as illustrated in Figure 2. The order of presentation of the feedback modalities was counterbalanced across participants. After each test or stressor trial participants provided a stress rating on a VAS scale between 0 and 100. Each session lasted about 1 hour. Participant entered stress levels using the VAS also before engaging in the first trial.

3.4 Instructions

At test trials, participants were instructed to 'Try to relax and sync your breath to the stimuli in the video'. After exposure to each video participants were instructed to 'Enter how stressed you feel using the slider below'.

3.5 Setup

The stimuli for the test and stressor trials were prepared and saved in video files. The presentation of the correct video was controlled by a MAX/MSP patch running on a PC. The patch guided participants through the trials. In each trial, participants pressed a button when ready to start. This triggered the display of the relevant video which was either a stressor or test video containing auditory, visual, or audiovisual cues as appropriate. After exposure, participants indicated their perceived stress level using the VAS displayed as a slider. The experiment was performed in a quiet laboratory environment. Sound was played by a Steel Series headset.

3.6 Hypotheses

We hypothesize that (H1) participants reported stress levels will be reduced after exposure to the stimuli in the test trials and (H2) increased due to exposure to the stressor stimuli in the stressor trials.

Given that both visual and auditory cues have been used effectively for guided breathing in the literature, we hypothesize (H3) towards a null effect of modality in reported stress relief after test trials in the unimodal visual and auditory conditions. In the congruent audiovisual condition, both cues guide along the same guided breathing pattern following which is the principal reason for any eventual reported stress relief. This is the very same pattern that is used in the unimodal conditions and should have the same potential for stress relief also when presented using audiovisual cues. We therefore hypothesize (H4) that there will be no increase in reported stress level in the congruent audiovisual condition compared to the unimodal auditory and visual conditions.

On the other hand, incongruency might disturb the ability of users to focus on the guided breathing cues and disturb the breathing cycle. This may in turn limit stress relief. Therefore, we hypothesize (H5) that stress relief would be lower in the incongruent audiovisual condition compared to the three other modality conditions.

4. RESULTS

In Figure 3a, the average absolute stress levels of participants after exposure to the test trials can be seen. Absolute stress levels after exposure are relatively uniform across conditions. A 1-way repeated-measures ANOVA with *modality* as within-subjects factor showed no main effect of *modality* on reported absolute stress level after test trials.

In Figure 3b, we see the difference in stress level before and after exposure to the test stimuli. We see that exposure to the test stimuli resulted in stress relief indicated by a reduction in reported stress levels for all *modalities*. The extent to which the reduction in stress due to exposure to the test stimuli was significantly smaller than 0 was examined using a t-test. The outcome was significant ($p<0.01$) for all conditions in the experiment. On average, some variation in stress reduction as a function of modality appears most notably for the incongruent condition. However, a 1-way ANOVA with *modality* as within-subjects factor showed no main effect of *modality* on the reduction of stress level due to exposure to the test stimuli.

In Figure 3d, we see the difference in reported stress level before and after exposure to the stressor stimulus. We see that overall exposure to stressor trials increased stress ratings. The extent to which the increase in stress due to exposure to the test stimuli was significantly higher than 0 was examined using a t-test. The outcome was significant ($p<0.01$) for all conditions in the experiment. The amount of increase in stress level varies slightly depending on the interval in which the stressor was presented. On average, the increase in stress level is highest the first time participants engage in the mental arithmetic. However, a 1-way ANOVA with *stressor presentation order* as within-subjects factor showed no main effect of *stressor presentation order*. Finally, in Figure 3c, we see that stress level after the stressor trials was similar irrespective of the stressor trial interval. A 1-way ANOVA with *stressor presentation order* as within-subjects factor showed no main effect of *stressor presentation order* on reported stress level. These results indicate that the stressor was effective and functioned in a uniform way throughout the experiment.

5. DISCUSSION

This study was a first step in examining whether introducing multimodal audiovisual cues for guided breathing may affect the reported stress relief due to the performance of a

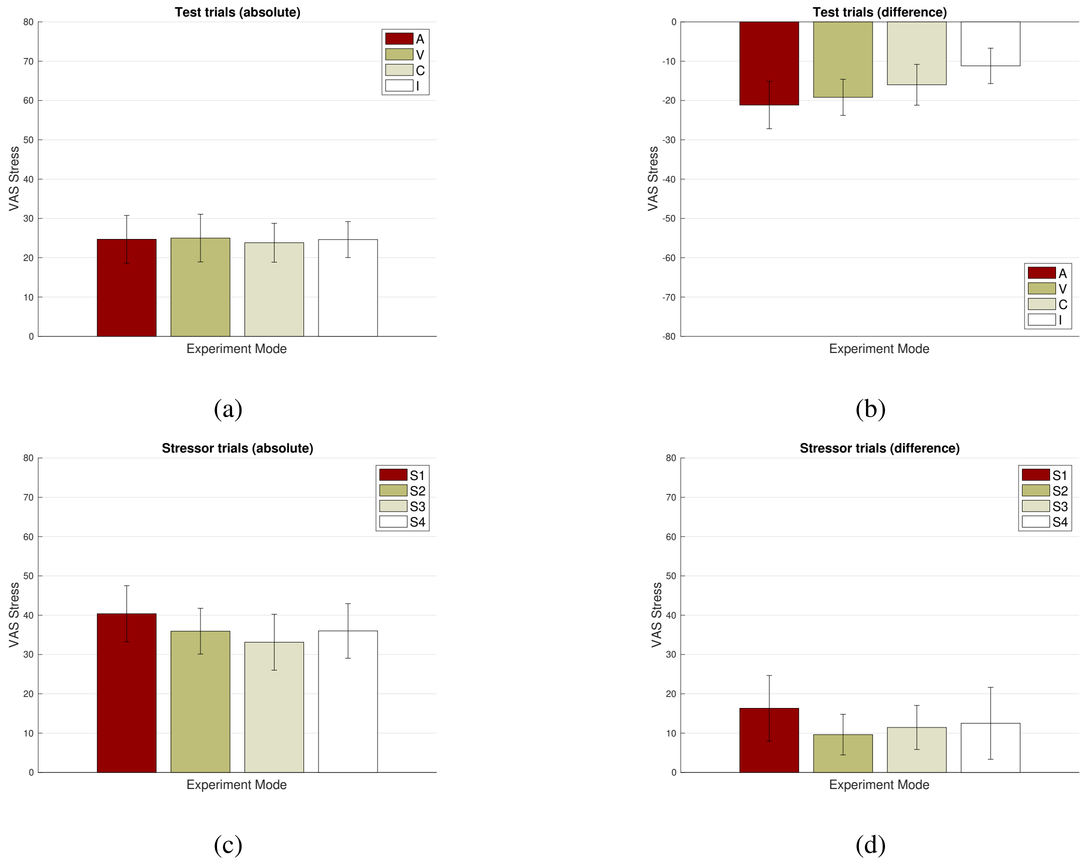

Figure 3: a. Absolute VAS stress level on test trials. b. Difference in VAS stress level after test trials. c. Absolute VAS stress level on the stressor trials. d. Difference in VAS stress level after stressor trials. Brackets indicate standard error.

guided breathing task. Specifically, we evaluated whether congruent cues may lead to increased reported stress relief and incongruent to a reduced reported stress relief relative to unimodal visual and auditory conditions. In a controlled experiment, participants were presented with interleaved stressor and test trials. The latter contained visual, auditory, and audiovisual guided breathing stimuli which were either congruent or drifted out of sync and became incongruent. Participants reported their perceived stress level using a visual analogue scale after each trial was completed.

Reported stress level increased significantly after exposure to the stressor trials. This is in agreement with H2 and testifies to the effectiveness of the stressor stimuli used in the experiment to induce stress. Furthermore, the increase in stress level did not depend on the interval in which the stressor video was presented, which shows that the stressor resulted in a relatively uniform increase in the stress level of participants throughout the experiment.

Furthermore, the results showed that exposure to the guided breathing stimuli in the test trials significantly decreased reported stress level. This is in accordance with hypothesis H1. However, the magnitude of the reduction in reported stress level after the test trials did not depend on the *feedback modality*. While this in agreement with hypotheses H3 and H4, it is against hypothesis H5.

It appears therefore that the rationale behind H1, H3, and H4 generally holds. The visual and the auditory unimodal cues were equally effective in guiding the breathing pattern and reducing stress levels. No additional stress relief was observed when both audio and visual breathing cues were presented in a congruent manner as likely because participants could already follow the breathing pattern sufficiently well just by using unimodal cues. This is not surprising as both visual and auditory cues have already been reported to be successful in guiding breathing and reduce perceived stress level in the literature.

On the other hand, the null effect of congruency on reported stress level and stress relief in the incongruent audiovisual trials was surprising. One explanation is that participants may have simply ignored the incongruent visual modality and kept breathing based on the auditory stimulus. Remember that the trial started with the stimuli in sync before the visual stimulus gradually drifted out of sync. Participants may therefore have noticed the growing inconsistency in the visual cues. In response to this, they may have chosen to sync their breath with the auditory cues that kept providing a steady, reliable, and familiar breathing pattern. An alternative explanation may be that participants attributed increased weight to the auditory cues or even that the auditory cues may have an increased salience in the case of this guided breathing task. Based on the results of this experiment, it is not possible to argue in favour of one or the other hypothesis.

Concluding, we did not observe a constructive or a destructive interference in perceived stress levels due to exposure to multimodal audiovisual guided breathing stimuli. Future work will look into verifying subjective responses with objective data such as the accuracy of following the cues in the guided breathing stimulus but also verifying subjective stress reports with physiological data. It will also look into testing out different incongruency patterns to investigate the reason behind the null effect of stimulus congruency in the results.

6. CONCLUSIONS

Motivated by several applications that use multimodal cues for guided breathing, we performed a controlled evaluation study in order to understand better the effect of multimodal guided breathing cues and cue congruency. The experiment manipulated guided breathing stimulus modality to be unimodal visual, auditory, audiovisual congruent and audiovisual incongruent. We measured reported stress level after exposure to the test stimuli. Exposure to the test stimuli led to a similar amount of reduction in reported stress level (stress relief) in all trials but the effect of stimuli modality was not significant. Both the unimodal and the congruent audiovisual cues used in the study guided breathing sufficiently well. Furthermore, it appears that the visual incongruent cue in the respective experiment condition was either ignored or was not salient enough to significantly reduce stress relief.

7. REFERENCES

[1] M. Prpa, E. R. Stepanova, T. Schiphorst, B. E. Riecke, and P. Pasquier, "Inhaling and exhaling: How technologies can perceptually extend our breath awareness," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–15, 2020.

[2] L. Kennedy and S. H. Parker, "Biofeedback as a stress management tool: a systematic review," *Cognition, Technology & Work*, vol. 21, no. 2, pp. 161–190, 2019.

[3] A. Steptoe and M. Kivimäki, "Stress and cardiovascular disease," *Nature Reviews Cardiology*, vol. 9, no. 6, p. 360, 2012.

[4] T. Sonne and M. M. Jensen, "Chillfish: A respiration game for children with adhd," in *Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*, pp. 271–278, 2016.

[5] J. Costa, F. Guimbretière, M. F. Jung, and T. Choudhury, "Boostmeup: Improving cognitive performance in the moment by unobtrusively regulating emotions with a smartwatch," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 3, no. 2, pp. 1–23, 2019.

[6] D. Plans, D. Morelli, S. Sütterlin, L. Ollis, G. Derbyshire, and M. Cropley, "Use of a biofeedback breathing app to augment poststress physiological recovery: Randomized pilot study," *JMIR formative research*, vol. 3, no. 1, p. e12227, 2019.

[7] S. Zepf, N. El Haouij, J. Lee, A. Ghandeharioun, J. Hernandez, and R. W. Picard, "Studying personalized just-in-time auditory breathing guides and potential safety implications during simulated driving," in *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization*, pp. 275–283, 2020.

[8] T. Sonne, T. Merritt, P. Marshall, J. J. Lomholt, J. Müller, and K. Grønbæk, "Calming children when drawing blood using breath-based biofeedback," in

Proceedings of the 2017 Conference on Designing Interactive Systems, pp. 725–737, 2017.

[9] V. Deschodt-Arsac, R. Lalanne, B. Spiluttini, C. Bertin, and L. M. Arsac, "Effects of heart rate variability biofeedback training in athletes exposed to stress of university examinations," *PloS one*, vol. 13, no. 7, p. e0201388, 2018.

[10] S. J. Morgan and J. A. M. Mora, "Effect of heart rate variability biofeedback on sport performance, a systematic review," *Applied psychophysiology and biofeedback*, vol. 42, no. 3, pp. 235–245, 2017.

[11] J. Frey, M. Grabli, R. Slyper, and J. R. Cauchard, "Breeze: Sharing biofeedback through wearable technologies," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–12, 2018.

[12] R. Gervais, J. Frey, A. Gay, F. Lotte, and M. Hachet, "Tobe: Tangible out-of-body experience," in *Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*, pp. 227–235, ACM, 2016.

[13] J. Kim, Y.-W. Park, and T.-J. Nam, "Breathingframe: An inflatable frame for remote breath signal sharing," in *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction*, pp. 109–112, 2015.

[14] J. Harris, S. Vance, O. Fernandes, A. Parnandi, and R. Gutierrez-Osuna, "Sonic respiration: controlling respiration rate through auditory biofeedback," in *CHI'14 Extended Abstracts on Human Factors in Computing Systems*, pp. 2383–2388, ACM, 2014.

[15] G. McCaig and S. Fels, "Playing on heart-strings: experiences with the 2hearts system," *System*, 2001.

[16] I. Bergstrom, S. Seinfeld, J. Arroyo-Palacios, M. Slater, and M. V. Sanchez-Vives, "Using music as a signal for biofeedback," *International Journal of Psychophysiology*, vol. 93, no. 1, pp. 140–149, 2014.

[17] B. Yu, M. Funk, J. Hu, and L. Feijs, "Unwind: a musical biofeedback for relaxation assistance," *Behaviour & Information Technology*, vol. 37, no. 8, pp. 800–814, 2018.

[18] M. Schein, B. Gavish, M. Herz, D. Rosner-Kahana, P. Naveh, B. Knishkowy, E. Zlotnikov, N. Ben-Zvi, and R. Melmed, "Treating hypertension with a device that slows and regularises breathing: a randomised, double-blind controlled study," *Journal of human hypertension*, vol. 15, no. 4, pp. 271–278, 2001.

[19] G. Leslie, A. Ghandeharioun, D. Zhou, and R. W. Picard, "Engineering music to slow breathing and invite relaxed physiology," in *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, pp. 1–7, IEEE, 2019.

[20] G. Marentakis, D. Borthakur, P. Batchelor, J. P. Andersen, and V. Grace, "Using breath-like cues for guided breathing," in *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–7, 2021.

[21] B. Yu, N. Bongers, A. Van Asseldonk, J. Hu, M. Funk, and L. Feijs, "Livingsurface: biofeedback through shape-changing display," in *Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*, pp. 168–175, ACM, 2016.

[22] S. Balters, E. L. Murnane, J. A. Landay, and P. E. Paredes, "Breath booster! exploring in-car, fast-paced breathing interventions to enhance driver arousal state," in *Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare*, pp. 128–137, 2018.

[23] C. D. Wickens, "Multiple resources and performance prediction," *Theoretical issues in ergonomics science*, vol. 3, no. 2, pp. 159–177, 2002.

[24] D. R. Perrott, K. Saberi, K. Brown, and T. Z. Strybel, "Auditory psychomotor coordination and visual search performance," *Perception & Psychophysics*, vol. 48, no. 3, pp. 214–226, 1990.

[25] B. A. Lewis, B. Penaranda, D. M. Roberts, and C. L. Baldwin, "Effectiveness of bimodal versus unimodal alerts for distracted drivers," 2013.

[26] J.-H. Lee and C. Spence, "Assessing the benefits of multimodal feedback on dual-task performance under demanding conditions," in *Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction-Volume 1*, pp. 185–192, British Computer Society, 2008.

[27] P. Janata and E. Childs, "Marketbuzz: Sonification of real-time financial data," in *The International Conference on Auditory Display*, Georgia Institute of Technology, 2004.

[28] A. J. Hornof, Y. Zhang, and T. Halverson, "Knowing where and when to look in a time-critical multimodal dual task," in *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 2103–2112, ACM, 2010.

[29] F.-X. Lesage, S. Berjot, and F. Deschamps, "Clinical stress assessment using a visual analogue scale," *Occupational medicine*, vol. 62, no. 8, pp. 600–605, 2012.

[30] A. Sawai, T. Motomura, T. Oshima, S. Sawai, T. Fujikawa, H. Fujii, Y. Bannai, Y. Takeda, M. Ohno, and O. Tochikubo, "Influence of acute mental arithmetic stress on taste and pungency," *Journal of nutritional science and vitaminology*, vol. 65, no. 3, pp. 224–232, 2019.