Virtual reality-supported biofeedback for stress management: Beneficial effects on heart rate variability and user experience

Weibel, Raphael P.; Kerr, Jasmine I.; Naegelin, Mara; Ferrario, Andrea; Schinazi, Victor R.; La Marca, Roberto; Hoelscher, Christoph; Nater, Urs M.; von Wangenheim, Florian

Published in:
Computers in Human Behavior

DOI:
10.1016/j.chb.2022.107607

Recommended citation(APA):
Weibel, R. P., Kerr, J. I., Naegelin, M., Ferrario, A., Schinazi, V. R., La Marca, R., Hoelscher, C., Nater, U. M., & von Wangenheim, F. (2023). Virtual reality-supported biofeedback for stress management: Beneficial effects on heart rate variability and user experience. *Computers in Human Behavior*, *141*, [107607]. https://doi.org/10.1016/j.chb.2022.107607

Computers in Human Behavior 141 (2023) 107607

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Full length article

Virtual reality-supported biofeedback for stress management: Beneficial effects on heart rate variability and user experience[☆], [☆☆]

Raphael P. Weibel [a,b,*,1], Jasmine I. Kerr [a,b,1], Mara Naegelin [a,b,1], Andrea Ferrario [a,b], Victor R. Schinazi [c,d], Roberto La Marca [e,f], Christoph Hoelscher [d,g], Urs M. Nater [h,i], Florian von Wangenheim [b,d]

[a] *Mobiliar Lab for Analytics at ETH Zurich, Department of Management, Technology, and Economics, ETH Zurich, Weinbergstrasse 56/58, Zurich, 8092, Switzerland*
[b] *Chair of Technology Marketing, Department of Management, Technology, and Economics, ETH Zurich, Zurich, Switzerland*
[c] *Department of Psychology, Bond University, 14 University Drive, Robina, 4226, Queensland, Australia*
[d] *Future Health Technologies, ETH Singapore Center, Campus for Research Excellence and Technological Enterprise (CREATE), Singapore, Singapore*
[e] *Centre for Stress-Related Disorders, Clinica Holistica Engiadina, Plaz 40, Susch, 7542, Switzerland*
[f] *Chair of Clinical Psychology and Psychotherapy, Department of Psychology, University of Zurich, Binzmuehlestrasse 14, Zurich, 8050, Switzerland*
[g] *Chair of Cognitive Science, Department of Humanities, Social and Political Sciences, ETH Zurich, Clausiusstrasse 59, Zurich, 8092, Switzerland*
[h] *Chair of Clinical Psychology of Adulthood, Department for Clinical and Health Psychology, University of Vienna, Liebiggasse 5, Vienna, 1010, Austria*
[i] *Research Platform The Stress of Life (SOLE) – Processes and Mechanisms underlying Everyday Life Stress, Liebiggasse 5, Vienna, 1010, Austria*

ARTICLE INFO

Dataset link: https://osf.io/nxqs2/

Keywords:
Virtual reality
Biofeedback
Heart rate variability
Stress management
Paced breathing

ABSTRACT

Heart rate variability biofeedback (HRV-BF) is frequently used for stress management. Recently, virtual reality technology has gained attention for delivery, promising higher immersion, motivation, and attention than classical screens. However, the effects of different technologies and breathing techniques are not yet understood. In this study, 107 healthy participants completed a session in one of four conditions: HRV-BF on a desktop screen, HRV-BF via head-mounted display (HMD), standardised paced breathing without feedback (sPB) on a screen, or sPB via HMD. All setups significantly reduced perceived stress and increased heart rate variability (HRV). Practising HRV-BF, however, led to significantly greater increases in the low frequency band of HRV and cardiac coherence than sPB, and using an HMD rather than a screen also led to greater increases in cardiac coherence. As for user experience, immersion adaptation and interface quality were higher for HMDs and facilitating conditions were better for screens. While all technique and technology combinations are feasible and effective for stress management, immersing oneself in virtual reality with an HMD for HRV-BF might yield increased benefits in terms of HRV target outcomes and several user experience measures. Future research is necessary to confirm any long-term effects of such a mode of delivery.

1. Introduction

The emergence of virtual reality (VR) technology has fostered the development of VR-enhanced digital health interventions for stress management (e.g., Annerstedt et al., 2013; Gaggioli et al., 2014; Shah et al., 2015; Tong, Gromala, Choo, Amin, & Shaw, 2015). Virtual reality, which can be delivered on two-dimensional screens, multi-screen systems, or head mounted displays (HMD), allows individual adaptation of realistic stimuli and immersive situations at lower cost and effort than classical therapeutic settings (Maples-Keller, Bunnell, Kim, & Rothbaum, 2017; Rowland, Casey, Ganapathy, Cassimatis, & Clough, 2021). In recent years, VR technology has also been explored as a way to deliver heart rate variability biofeedback (HRV-BF; Lüddecke & Felnhofer, 2021). The goal of HRV-BF is to make breath-induced changes in heart rate variability (HRV) visible to clients, and to increase their bodily awareness and self-regulation skills (Lehrer et al., 2020;

☆ Funding: This study is part of a larger project and supported by an ETH Research Grant, Switzerland (ETH-09 19-2).

☆☆ Acknowledgements: We thank Dr. Stefan Wehrli from ETH Zurich's Decision Science Laboratory and Adam Fehr for their assistance in the realisation of the study.

* Correspondence to: Mobiliar Lab for Analytics at ETH Zurich, Department of Management, Technology, and Economics, ETH Zurich, Weinbergstrasse 56/58, 8092 Zurich, Switzerland.

E-mail address: raweibel@ethz.ch (R.P. Weibel).

[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.chb.2022.107607
Received 2 September 2022; Received in revised form 7 December 2022; Accepted 11 December 2022

Meichenbaum, 1976). HRV-BF is a well-established mind-body technique that is frequently and effectively used for stress management (for reviews, see Goessl, Curtiss, & Hofmann, 2017; Lehrer et al., 2020; Yu, Funk, Hu, Wang, & Feijs, 2018). In terms of user experience (UX), early studies have found evidence that employing VR with HMDs to deliver HRV-BF improves motivation, involvement and attention when compared to two-dimensional screen solutions (Blum, Rockstroh, & Göritz, 2019, 2020; Houzangbe, Christmann, Gorisse, & Richir, 2020; Rockstroh, Blum, & Göritz, 2019). However, research on VR-supported HRV-BF is still emergent and more studies are required to investigate the underlying mechanisms and replicate the stress-reducing effects (Lüddecke & Felnhofer, 2021).

2. Theoretical background

2.1. *Psychophysiological mechanisms of heart rate variability biofeedback*

Heart rate variability quantifies the variation in the time interval between two heartbeats, and is considered a marker of parasympathetic cardiac regulation (Berntson, Quigley, Norman, & Lozano, 2017). A high tonic HRV is linked to better physical and mental health and social functioning (Dishman et al., 2000; Kim, Cheon, Bai, Lee, & Koo, 2018; Porges, 2007).

Theoretical frameworks such as the cardiac coherence (McCraty & Shaffer, 2015), the resonance frequency (Lehrer, Vaschillo, & Vaschillo, 2000; Vaschillo, Vaschillo, & Lehrer, 2006) and the neurovisceral integration (Thayer & Lane, 2000) models help understand the underlying psychophysiological and neural processes and effects of HRV-BF. In essence, HRV-BF aims to improve self-regulation, increase emotional well-being, and restore and maintain autonomic balance through intentional HRV increases (Lehrer et al., 2020; Mather & Thayer, 2018; Shaffer, McCraty, & Zerr, 2014). There is some debate on whether the effects of biofeedback can be explained mainly by volitional control over biological signals or by autonomic regulation of subcognitive systems (Gaume, Vialatte, Mora-Sánchez, Ramdani, & Vialatte, 2016). Most probably, HRV-BF works through both top-down and bottom-up processes, as performing HRV-BF requires executive functions and emotion regulation skills, but it also directly increases gas exchange efficiency, strengthens baroreflexes, afferent vagal activity and immune function (Gaume et al., 2016; Lehrer et al., 2000; McCraty & Shaffer, 2015; Vaschillo et al., 2006). Mather and Thayer (2018) suggest that increases in HRV through daily HRV-BF training may increase blood flow oscillations in emotion-regulating networks of the brain (mainly prefrontal and limbic structures). They also argue that the suppression effect on the SNS and the stimulating effect on the parasympathetic nervous system of the increased heart rate oscillations could explain the stress- and anxiety-reducing properties of HRV-BF (Mather & Thayer, 2018). From a more top-down perspective and in line with the cardiac coherence model, emotional self-regulation is also thought to decrease SNS activation and vagal withdrawal (McCraty & Shaffer, 2015; Shaffer et al., 2014).

During an HRV-BF session, HRV is measured continuously and changes in HRV are fed back to clients in real-time. The feedback is intended to help clients gain voluntary control over cardiac regulation (Lehrer et al., 2020; Shaffer et al., 2014). Specifically, clients learn to increase their HRV through slow and paced breathing (PB) at their individual baroreflex resonance frequency (RF; Lehrer et al., 2020). The baroreceptor HR reflex circuit ensures that an increase in blood pressure leads to a decrease in HR and vice versa. Due to the inherent delay in this circuit, stimulating the baroreflex at a certain frequency leads to resonance, which produces very large amplitude oscillations at that frequency (i.e., high HRV; Lehrer et al., 2000; Vaschillo et al., 2006). In HRV-BF, this stimulation is achieved through slow PB at one's individual RF, which shifts respiratory sinus arrhythmia (RSA, the component of HRV attributed to breathing) from its normal, higher frequency (HF) to the baroreflex RF in the low frequency (LF) region—also described as cardiac coherence (McCraty & Shaffer, 2015). Breathing at RF is thus an integral component of HRV-BF training as it is the physiological mechanism by which HRV-BF improves HRV (Steffen, Austin, DeBarros, & Brown, 2017).

The standard protocol for RF determination (Lehrer et al., 2000) involves comparing the HRV changes of a series of breathing trials paced between 4.5 and 6.5 breaths per minute, and allows the estimation of RF at a granularity of 0.5 breaths per minute. Recently, Sakakibara, Kaneda, and Oikawa (2020) have proposed a breathing protocol that has the potential for a more granular estimation of an individual's RF. Although fairly novel, this protocol offers a promising and economic way to use exact RF pacing in HRV-BF protocols. In many studies, however, participants just follow a visual or auditory pacer at 0.1 Hz in lieu of HRV-BF at exact RF (see the recent review by Lehrer et al., 2020) since it is easier to implement and even less time-consuming. Nevertheless, some researchers (Lin et al., 2012; Steffen et al., 2017) have found evidence that HRV-BF at RF leads to better results than slow PB at a standardised, fixed rate.

2.2. *An integrative approach to biofeedback*

The goal of any biofeedback (BF) training programme is to improve self-regulation and to transfer the learnt affective, cognitive and bodily skills into daily life (Gaume et al., 2016; Meichenbaum, 1976). Therefore, BF training is generally accompanied by psychoeducational content to promote this transfer of skills. During psychoeducation for HRV-BF, clients learn about the nature of stress, stress-related symptoms, and the goal, effects and mechanisms of HRV-BF, which helps to decatastrophize symptoms and set training goals (Lehrer et al., 2020; Meichenbaum, 1976; Nanke & Rief, 2000). There are many different ways how an HRV-BF training programme can be designed and delivered. As a means to increase the standards and thereby efficacy and effectiveness of BF programmes, Gaume et al. (2016) postulate five key properties in their integrative psychoengineering model of BF that any effective BF protocol should promote: (1) perceptibility, (2) autonomy, (3) mastery, (4) motivation, and (5) learnability. This means, first, that the biosignal to be controlled has to be visualised in a perceivable and understandable way (perceptibility). For example, cognitive load should be reduced in order to avoid overwhelming clients with cues and feedback information. Second, in time, clients learn to rely on their internal feedback rather than on the BF as external feedback (autonomy). Ultimately, a client should be able to self-regulate the bodily signal without the support of technology and feedback. Third, clients should be able to control the biosignal and the level of difficulty should adapt to their progress (mastery). Fourth, BF should be experienced as extrinsically or intrinsically motivating (motivation), which is tightly connected to mastery. Biofeedback enables proficiency in a skill, which is believed to stimulate intrinsic motivational factors. Finally, clients should be able to repeatedly practise BF in order to consolidate what they have learned (learnability).

2.3. *Virtual reality-supported biofeedback*

While BF is most frequently deployed on two-dimensional displays, advances and commercialisation of VR technology have created new opportunities for feedback delivery (Yu et al., 2018). Using VR is advocated as a means to address several hindrances and improve a number of aspects of traditional BF (Rockstroh et al., 2019). Similarly, we argue that the five properties of the psychoengineering model (Gaume et al., 2016) can be further supported and promoted through the use of VR technology or, more specifically, using an HMD to experience VR instead of a classical screen. In addition, the Cognitive Affective Model of Immersive Learning (CAMIL; Makransky & Petersen, 2021) may also support the use of HMDs to deliver HRV-BF. CAMIL describes how immersion, the ability to control, and representation fidelity affect feelings of presence and agency in VR, especially with HMDs.

In return, presence and agency influence six cognitive and affective factors: interest, motivation, self-efficacy, embodiment, cognitive load, and self-regulation. Promoting these factors fosters the acquisition of factual, conceptual and procedural knowledge and the transfer of knowledge (Makransky & Petersen, 2021). CAMIL proposes that the type of technology or media used for training or learning interacts with the method used for training or learning. Methods that facilitate higher presence or agency during training will be specifically effective with VR technology (Makransky & Petersen, 2021). HRV-BF and PB as methods both require being present in the moment (e.g., following the pacer, focusing on breathing) similar to mindfulness meditation. In contrast to PB, however, HRV-BF is believed to additionally facilitate a sense of agency as clients become less and less dependent on external feedback (Gaume et al., 2016). Practising HRV-BF with an HMD might therefore be specifically effective.

Moreover, a virtual feedback may be more beneficial for a client's ability to perceive and control their biosignal (i.e., perceptibility and mastery; Gaume et al., 2016) than a classical screen. The feedback can be incorporated seamlessly into the virtual environment, which allows clients to directly alter their surroundings through successful physiological regulation. Indeed, Blum et al. (2019) found that VR-supported HRV-BF with HMD can increase relaxation self-efficacy more strongly than a standard HRV-BF. Rockstroh et al. (2019) argue that the high level of control may make the feedback more intuitive and more powerful than abstract and sometimes complex graphical visualisations on a two-dimensional screen. Simplified and intuitive feedback might also reduce cognitive load (Sun, Cao, & Ma, 2017; Wollmann et al., 2016). Several studies found that experiencing BF in VR via HMD leads to greater attention and less mind-wandering than practising BF on a two-dimensional screen (Blum et al., 2019, 2020; Rockstroh et al., 2019). This may be attributed to the strong sense of presence induced by VR (Cummings & Bailenson, 2016; Rockstroh, Blum, & Göritz, 2020) and the elimination of external distractions through the use of an HMD. In fact, CAMIL (Makransky & Petersen, 2021) postulates that increased control of an environment positively influences presence and agency, which in turn increases interest, reduces cognitive load and supports self-regulation.

Fostering engagement and motivation is viewed as one of the key challenges of traditional BF (Lüddecke & Felnhofer, 2021). Here, CAMIL (Makransky & Petersen, 2021) proposes that motivation in a VR setting, especially with HMDs, can be reinforced by a sense of immersion, higher representational fidelity and the ability to control the environment. Indeed, previous research has found that immersive, interactive VR environments can help make the training experience more vivid, interesting and attractive, and has been found to elicit high levels of engagement and motivation (Houzangbe et al., 2020; Rockstroh et al., 2019). In the long-run, these properties might help to increase adherence to a stress management intervention in VR, and thereby support learnability (i.e., long-term memory formation) as described by Gaume et al. (2016). Lüddecke and Felnhofer (2021) argue that realism of virtual environments may also improve the transfer of learnt skills into everyday life because clients can practise self-regulation in more realistic environments (in contrast to a medical setting), and thus support clients' autonomy and learnability (see Gaume et al., 2016). Studies have found that user agency has also been linked directly to the training success of VR-supported BF training (Houzangbe et al., 2020). Indeed, CAMIL (Makransky & Petersen, 2021) suggests that representational fidelity and the ability to control the virtual environment directly influence user agency.

While the results from early studies are promising, the high variability in study protocols—for example in terms of number, duration and content of HRV-BF sessions, exact type of explicit or implicit feedback and visualisation (Kennedy & Parker, 2019)—has made it difficult to inform the choice of specific aspects when designing an HRV-BF training. Moreover, most previous studies were exploratory in nature and focused on feasibility and user experience in within-subject designs (Gradl, Wirth, Zillig, & Eskofier, 2018; Houzangbe et al., 2020; Maarsingh, Bos, Tuijn, & Renard, 2019; van Rooij, Lobel, Harris, Smit, & Granic, 2016). In addition, previous works on VR-supported HRV-BF (Blum et al., 2019; Rockstroh et al., 2019) used a fixed breathing rate of 0.1 Hz in their HRV-BF conditions rather than exact and individual RF estimates. They also compared rich natural environments with embedded feedback in VR via HMD against simplistic abstract visualisations on two-dimensional screens, which does not allow the isolation of specific effects of display technologies.

2.4. Hypotheses

To the best of our knowledge, no previous research on VR-supported HRV-BF has compared the use of an HMD to a two-dimensional screen displaying the same virtual environment, or compared the effects of HRV-BF at exact RF to breathing at a fixed rate in VR. Our aim is thus to determine the most suited technique (i.e., HRV-BF at RF vs. standardised PB) and technology (i.e., HMD vs. two-dimensional screen) combination for a VR-supported stress management single-session protocol. Based on the theoretical frameworks and previous studies discussed above, we have derived the following hypotheses:

H1a: Both HRV-BF and standardised PB (sPB; i.e., 0.1 Hz) lead to significant changes in psychological measures (i.e., decrease perceived stress, increase perceived calmness, relaxation, good mood and wakefulness) during a training session.
H1b: Both HRV-BF and sPB lead to significant changes in cardiac measures (i.e., decrease in HR, increase in HRV features) during a training session.
H2a: HRV-BF leads to even greater changes in psychological measures than sPB.
H2b: HRV-BF leads to even greater changes in cardiac measures than sPB.
H3a: Using an HMD for HRV-BF and sPB leads to greater changes in psychological measures than a classical screen.
H3b: Using an HMD for HRV-BF and sPB leads to greater changes in cardiac measures than a classical screen.
H4a: Using an HMD leads to even greater changes in psychological measures during HRV-BF than sPB.
H4b: Using an HMD leads to even greater changes in cardiac measures during HRV-BF than sPB.
H5a: User experience ratings (e.g., hedonic motivation, involvement, behavioural intention) are more favourable for HRV-BF than for sPB.
H5b: User experience ratings (e.g., hedonic motivation, involvement, behavioural intention) are more favourable for using an HMD than a classical screen.

3. Methods and materials

3.1. Participants

Participants were recruited via the University's online recruitment website and were required to be between the ages of 18 and 40, have normal or corrected-to-normal vision, exhibit no disability of arms or hands, and have obtained at least a secondary school diploma. Exclusion criteria consisted of self-reported acute and chronic somatic diseases or psychiatric disorders, regular medication, medication in the last two months to treat acute illnesses, any cardioactive medication, the consumption of psychoactive substances in the last three months, heavy drinking (≥ 15 and ≥ 8 standard drinks per week for men and women, respectively) or consumption of tobacco (> 5 cigarettes per week, excluding weekends). Taking into account the effects of the menstrual cycle, hormonal contraceptives, pregnancy and breastfeeding on the autonomic nervous system and in particular on HRV (see, e.g., Hirshoren et al., 2002; Laborde, Mosley, & Thayer, 2017; Mezzacappa, Kelsey, & Katkin, 2005), only women with regular menstrual cycles, who did not take hormonal contraception, were not pregnant and did not lactate were included. Participants were also asked to follow a normal sleep routine and to refrain from intense physical training and drinking alcohol the day before the experiment. Participants were told not to drink caffeinated beverages or eat in the last two hours leading up the experiment. Participants gave written informed consent and were compensated for participation with an equivalent of around $50. The experiment was approved by the University's ethics commission and conducted in accordance with the Declaration of Helsinki.

An a priori power analysis using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) for a two-way independent ANOVA revealed that a sample size of 112 participants would be sufficient to achieve an α of 0.05, a power of 0.95 and uncover a large effect size (f = 0.4) of UX measures. In order to detect medium effect sizes (f = 0.25) for psychological and cardiac measures, power analyses for a three-way mixed ANOVA with three repeated psychological and for one with four repeated cardiac measures revealed samples sizes of 60 and 52, respectively. To be able to detect all expected effect sizes and to balance out the participants per conditions, we recruited 120 participants (60 female). In total, 115 participants took part in the experiment. As the result of technical difficulties, 107 (48 female, mean age 22.52 ± 3.33 years) participants were included in the final analyses.

3.2. Procedure

Participants were tested in small groups and each session lasted for 120 min. Upon arrival, participants were randomly assigned to one of four conditions determined by the two two-level factors, technique and technology, manipulated in the experiment: HRV biofeedback via head-mounted display (HRV-BF$_{\text{HMD}}$), HRV biofeedback on a two-dimensional desktop screen (HRV-BF$_{\text{screen}}$), standardised paced breathing via head-mounted display (sPB$_{\text{HMD}}$), standardised paced breathing on a two-dimensional desktop screen (sPB$_{\text{screen}}$).

In the introductory part of the experiment, participants received a pre-recorded psychoeducational presentation on the upcoming experimental procedure, including their assigned type of technique and technology. They were also introduced to and practised slow and paced breathing for a couple of minutes. Next, participants put on a chest belt for continuous monitoring of their cardiac activity. They were seated and remained seated for the entirety of the experiment in front of a desktop computer, where they acclimatised to the experimental setting. Here, they answered a first round of questionnaires to assess their psychological state and some of the baseline characteristics, namely, trait questionnaires we expected not to affect their emotional state negatively (i.e., BPNSFS, ACTA, WHO-5 and ITQ; see Table 1). All other baseline characteristics, which could be potentially emotionally confronting, were assessed post-training. The following training phase consisted of three blocks and lasted approximately 30 min. In each block, participants engaged with a virtual environment either via HMD or via their already assigned desktop screen, depending on their experimental condition. The first block of 6 min served to determine the breathing frequency (corresponding to their exact RF) for participants in the HRV-BF$_{\text{HMD}}$ and HRV-BF$_{\text{screen}}$ conditions. This block was followed by two training blocks of 10 min each. All three blocks were separated by short breaks of approximately 2 min. During the second break, participants answered a second round of questionnaires on the desktop screen. After the training phase, participants completed a final round of psychological state questionnaires followed by UX and the mentioned psychological trait questionnaires (see Fig. 1).

3.3. Virtual environment

The virtual environment used for the training blocks depicted a vast mountainous landscape. Participants experienced the virtual environment from the perspective of sitting on a large tree trunk overlooking an even, grassy meadow (see Fig. 2). The meadow was surrounded by a small number of trees and bushes with a few boulders at the far end, and mountain peaks rising in the distance. The sun was either in a fixed position (first block) or rising gradually (the two training blocks). The soundscape consisted of a mix of bird songs and sounds of water flowing in a small creek. A breathing pacer was placed in the middle of the meadow. It consisted of a semi-transparent white cylinder with a small disc moving up and down inside the cylinder. To emphasise the direction of movement, a small trail was visualised following the moving disc.

3.4. Technologies and technical specifications

The virtual environment was developed with the game engine Unity (Unity Technologies, San Francisco, USA) and ran on HMDs or desktop computers depending on the conditions. The Oculus Quest (Reality Labs, Menlo Park, USA) was used to deliver the trainings in the HMD conditions (HRV-BF$_{\text{HMD}}$ and sPB$_{\text{HMD}}$). The virtual soundscape was audible directly through the in-built speakers of the Oculus Quest. The HRV-BF$_{\text{screen}}$ and sPB$_{\text{screen}}$ conditions received their training on a Lenovo IdeaCentre AOI 700 desktop computer with a 23.8″ screen and wore headphones attached to the desktop computer. The chest belt used to record cardiac activity and enable the HRV-BF was a Polar H10 device (Polar Electro Oy, Kempele, Finland). All data from the chest belts was streamed to a custom-made Windows application (based on the Polar SDK) and sent to a centralised database, where the RF and HRV score were computed. The resulting values were then accessed by the virtual environment application, which adapted the environment accordingly in the HRV-BF conditions.

3.5. Techniques

Biofeedback. We used a custom Python script that relied on the `hrv` and `scipy` libraries (Version 0.2.8 and 1.5.3, respectively) to automatically process the collected R–R interval (RRI) signal during the experiment. For all calculations, segments of the raw RRI signal streamed to the database were filtered for ectopic beats and motion artefacts using the filtering functionalities of `hrv`. Fast Fourier Transform of the signal interpolated with cubic splines at 4 Hz was used to estimate the segment's power spectral density. During the first block of the training phase, we followed Sakakibara et al.'s 2020 protocol in order to determine the exact RF breathing frequency for HRV-BF$_{\text{HMD}}$ and HRV-BF$_{\text{screen}}$ conditions. This protocol uses the peak frequency of the low frequency component of the resting HRV under respiratory control at 0.25 Hz as an estimate for exact RF. Specifically, we calculated RF as the argmax in the region of 0.075 Hz to 0.10833 Hz (i.e., between 4.5 and 6.5 breaths per minute) of the power spectral density from the last 5 min of the collected RRI signal during the first block, see Fig. A.1 in Appendix. For the HRV-BF$_{\text{HMD}}$ and HRV-BF$_{\text{screen}}$ conditions only, the calculated RF was set as the frequency of the breathing pacer for the following two HRV-BF training blocks, in which participants aimed to increase their HRV through slow PB. During the two training blocks, increases in HRV were computed regularly at short time intervals (i.e., every 10 s, from the last 90 s of collected data) and quantified with a cumulative HRV score based on the coherence ratio (CR, see Section 3.9; McCraty & Shaffer, 2015). A CR ≥ 1 increased the HRV score by one unit (i.e., a positive feedback), whereas a CR < 1 had no influence on the HRV score. Changes in the HRV score led to discrete changes in the virtual environment. Specifically, increases in HRV led to flower growth in the grassy meadow, the sun rising on the horizon and the natural soundscape intensifying.

Paced breathing. Participants in the sPB$_{\text{HMD}}$ and sPB$_{\text{screen}}$ conditions were also shown a breathing pacer with the same frequency of 0.25 Hz in the first block, but their exact RF was only calculated in the background and not used to adjust the breathing pacer in the training blocks. In the second two blocks they performed slow breathing guided

Table 1

Baseline characteristics (mean and standard deviation) and p-values of F-tests or χ^2-tests of group differences.

	Total sample (107)	HRV-BF$_{screen}$ (25)	HRV-BF$_{HMD}$ (25)	sPB$_{screen}$ (29)	sPB$_{HMD}$ (28)	p
Age [years]	22.52 (3.33)	22.42 (2.08)	22.25 (3.50)	23.52 (4.51)	21.82 (2.50)	.35
BMI [kg/m^2]	21.63 (2.18)	21.72 (2.42)	22.40 (2.51)	21.40 (2.05)	21.12 (1.64)	.30
Chronic stress [PSS-10]	16.75 (5.74)	18.19 (6.03)	16.21 (5.32)	15.59 (6.51)	17.07 (4.86)	.33
Chronic stress [DASS-21$_S$]	9.64 (7.19)	11.38 (8.52)	10.42 (7.27)	8.00 (6.61)	9.07 (6.22)	.36
Depression [DASS-21$_D$]	7.81 (7.55)	10.08 (9.09)	8.33 (6.82)	5.72 (7.13)	7.43 (6.71)	.05
Anxiety [DASS-21$_A$]	4.95 (5.10)	5.69 (5.98)	5.33 (4.16)	3.86 (4.69)	5.07 (5.43)	.38
Wellbeing [WHO-5]	43.78 (15.23)	48.62 (16.23)	41.67 (14.82)	40.00 (16.28)	45.00 (12.73)	.17
General health perception [SF-36$_{ghp}$]	79.63 (14.41)	80.81 (14.31)	76.75 (12.72)	83.86 (14.58)	76.61 (15.16)	.18
Autonomy satisfaction [BPNSFS$_{as}$]	15.38 (2.40)	14.81 (2.55)	15.88 (2.35)	15.72 (2.55)	15.14 (2.12)	.26
Autonomy frustration [BPNSFS$_{af}$]	8.50 (2.78)	9.31 (3.08)	8.25 (3.01)	7.83 (2.51)	8.64 (2.47)	.48
Competence satisfaction [BPNSFS$_{cs}$]	16.02 (2.54)	15.58 (2.94)	16.50 (2.11)	16.17 (2.62)	15.86 (2.46)	.71
Competence frustration [BPNSFS$_{cf}$]	7.79 (2.95)	8.19 (3.38)	7.00 (1.89)	7.97 (3.10)	7.93 (3.14)	.78
Immersive tendency [ITQ]	76.15 (10.63)	73.77 (10.68)	77.78 (10.77)	76 (11.75)	76.64 (9.28)	.68

Header row includes sample sizes in parentheses. Abbreviations: BMI = Body mass index; kg/m^2 = kilogramme per metre squared; PSS-10 = Perceived Stress Scale with 10 items; DASS-21 = Depression Stress and Anxiety Scales with 21 items; WHO-5 = World Health Organization Five Well-being Index; SF-36 = Short Form Health Survey; BPNSFS = Basic Psychological Needs Satisfaction and Frustration Scale; ITQ = Immersive Tendency Questionnaire; HRV-BF$_{HMD}$ = heart rate variability biofeedback via head-mounted display; HRV-BF$_{screen}$ = heart rate variability biofeedback on a desktop screen; sPB$_{HMD}$ = standardised paced breathing via head-mounted display; sPB$_{screen}$ = standardised paced breathing on a desktop screen.

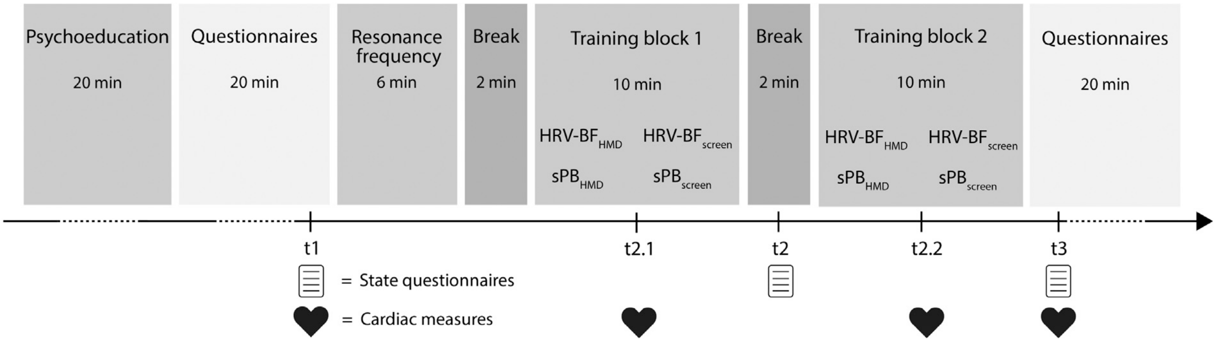


Fig. 1. Study protocol. Cardiac activity was measured and processed in real-time during resonance frequency determination and both training blocks in order to provide heart rate variability feedback to the HRV-BF$_{HMD}$ and HRV-BF$_{screen}$ conditions. Psychological state was measured at t1, t2 and t3. For statistical analysis, cardiac measures were calculated post-hoc on 10 min segments at t1, t2.1, t2.2 and t3. Abbreviations: HRV-BF$_{HMD}$ = Heart rate variability biofeedback via head-mounted display; HRV-BF$_{screen}$ = heart rate variability biofeedback on a desktop screen; sPB$_{HMD}$ = standardised paced breathing via head-mounted display; sPB$_{screen}$ = standardised paced breathing on a desktop screen.

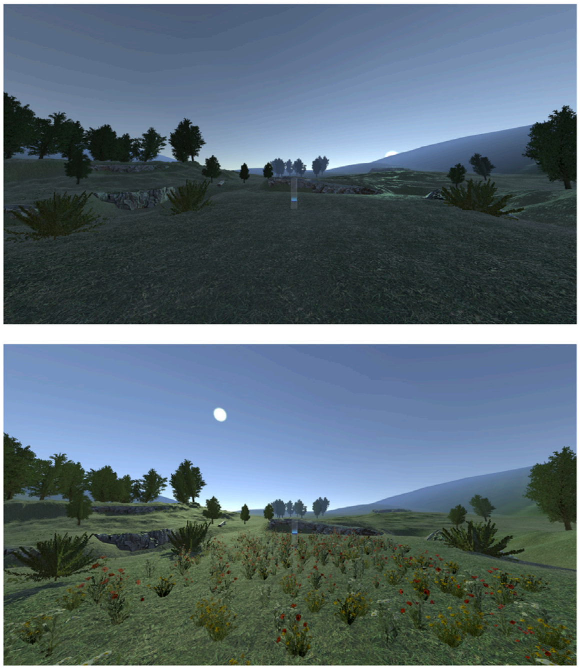


Fig. 2. These images depict the virtual environment at the beginning (top) of the HRV-BF training and towards the end (bottom). A brighter and more colourful environment reflects higher increases in heart rate variability, assessed via coherence ratio. Abbreviations: HRV-BF = Heart rate variability biofeedback.

by a non-personalised pacer set at 6 breaths per minute (0.1 Hz) in the VR environment. Additionally, the meadow was filled with flowers from the beginning and did not change over time, the sun rose at a constant speed and the soundscape intensified unrelated to their progress, simulating a day going by.

3.6. Baseline characteristics

Baseline characteristics were assessed in order to characterise our sample and to compare it to other samples.

Health status was assessed with the Perceived Stress Scale on a 5-point Likert scale (PSS-10; Cohen, Kamarck, & Mermelstein, 1983), the Depression Anxiety Stress Scales on a 4-point Likert scale (DASS-21; Lovibond & Lovibond, 1995), the World Health Organization Five Well-being Index on a 6-point Likert scale (WHO-5; Topp, Østergaard, Søndergaard, & Bech, 2015) and the subscale General Health Perception of the Short Form Health Survey consisting of Likert scales with varying lengths (SF-36; Morfeld, Kirchberger, & Bullinger, 2011).

Basic psychological needs (i.e., autonomy and competence) were measured with the Basic Psychological Needs Satisfaction and Frustration Scale (BPNSFS Heissel et al., 2018). The BPNSFS consists of 12 items on a 5-point Likert scale and is based on self-determination theory (SDT; Ryan & Deci, 2000). SDT is a theory of human motivation which includes autonomy and competence as basic psychological needs.

Immersive tendency was assessed with the Immersive Tendency Questionnaire (ITQ; Witmer & Singer, 1998), which captures participants' abilities to immerse themselves into an environment. Participants answered 18 items on a 7-point Likert scale.

3.7. Psychological state measures

Subjective responses to the training were measured with the Multidimensional Mood Questionnaire (MDMQ; Steyer, Schwenkmezger, Notz, & Eid, 1997) and two self-developed Visual Analogue Scales (VAS). The MDMQ comprises the subscales good mood–bad mood ($MDMQ_{mood}$), calmness–nervousness ($MDMQ_{calmness}$) and wakefulness–sleepiness ($MDMQ_{wakefulness}$). Each subscale is assessed with four items on a 5-point Likert scale, resulting in a range of 4 to 20. Note that higher scores on the MDMQ subscales indicate good mood, calmness or wakefulness. The two VAS were presented to participants on a slider ranging from 0 to 100 (low to high) in order to assess feelings of relaxation ($VAS_{relaxed}$) and stress ($VAS_{stressed}$). One person did not receive the questionnaire at time point t2 due to a technical error, and was excluded from subsequent analyses of psychological state measures (i.e., N = 106).

3.8. Measures of user experience

The *Autonomy and Competence in Technology Adoption Questionnaire* (ACTA; Peters, Calvo, & Ryan, 2018) is also based on SDT (Ryan & Deci, 2000). The ACTA is a new UX tool used to measure the degree to which people feel autonomous or controlled, and competent in adopting a technology. The ACTA has four indices: the Autonomy Regulation Score ($ACTA_{ARS}$), the Controlled Regulation Score ($ACTA_{CRS}$), the Relative Autonomy Index ($ACTA_{RAI}$) and the Perceived Competence Score ($ACTA_{PCS}$). The $ACTA_{ARS}$ is made up of six items on a 5-point Likert scale that capture intrinsic (“It is going to be fun to use”) and identified regulation (“I believe it could improve my life”), while the $ACTA_{CRS}$ comprises introjected (“I want others to know I use it”) and external regulation (“I feel pressured to use it”), also measured with six items. The $ACTA_{RAI}$ is the subtraction of the $ACTA_{CRS}$ from $ACTA_{ARS}$. Finally, the $ACTA_{PCS}$ was assessed with the original two items and an additional self-developed one (“It will be challenging for me to use”), also on a 5-point Likert scale. Higher scores on all subscales reflect higher values of the measured construct.

The *Presence Questionnaire* (PQ) by Witmer and Singer (1998) was developed to assess “presence” in virtual environments. It consists of four subscales with 7-point Likert scales: involvement (PQ_{Invo}), sensor fidelity (PQ_{SensFi}), immersion adaptation ($PQ_{ImrsAdpt}$) and interface quality ($PQ_{IntQual}$). Involvement is defined as the degree to which participants felt involved in the virtual experience. Immersion adaptation assesses participants’ ease of adapting to the system in order to get immersed. Interface quality is the degree to which the interface quality has hindered participants from getting immersed (i.e., a lower value is better).

The *Flow Short Scale* (FSS) by Rheinberg, Vollmeyer, and Engeser (2003) assesses the experience of flow in any activity. The FSS consists of two subscales with 7-point Likert scales, the experience of flow (FSS_{flow}) and the perceived outcome importance ($FSS_{anxiety}$). The *System usability scale* (SUS) by Brooke (1996) is a ten-item scale that assesses the complexity, ease of use, need for training and other usability related aspects of a system. The aggregated score ranges from 0–100 and can be reduced to an adjusted rating (excellent ≥ 80.3, good > 68, ok $= 68$, poor < 68, awful < 51).

The *Unified Theory of Acceptance and Use of Technology Questionnaire* (UTAUT) by Venkatesh, Morris, Davis, and Davis (2003) assesses how different constructs affect the use of a technology. We only included some of the subscales of the UTAUT, based on their applicability to our experimental setting. Specifically, we assessed effort expectancy ($UTAUT_{EffExp}$), facilitating conditions ($UTAUT_{FacilCond}$), behavioural intention ($UTAUT_{BhvInt}$) and hedonic motivation ($UTAUT_{HedMotv}$). Each subscale was evaluated with 7-point Likert scales. $UTAUT_{EffExp}$ is defined as the degree of ease that participants associate with the use of a system. $UTAUT_{FacilCond}$ assesses whether users believe that there is a support infrastructure (organisational and technical) that helps them to use the system. $UTAUT_{BhvInt}$ assesses if participants would use a system in the future (if they could) and $UTAUT_{HedMotv}$ assesses whether the use of the system is fun or creates pleasure.

The *Simulator Sickness Questionnaire* (SSQ; Kennedy, Lane, Berbaum, & Lilienthal, 1993) assesses the degree of simulator sickness that participants suffer from while interacting with a virtual environment. It is split into three factors based on clusters of symptoms with 4-point Likert scales: nausea (SSQ_{Nas}), effects on the oculomotor systems (SSQ_{Ocu}) and amount of disorientation (SSQ_{Dis}).

3.9. Cardiac measures

Heart rate and HRV measures were derived from the recorded RRI in accordance with the standards of measurement established by the Task Force on HRV (Malik et al., 1996). Similar to the real-time processing, raw recordings were filtered for artefacts and ectopic beats using the Python library `hrv` and its threshold-based filtering algorithm. In addition, filtering results were visually inspected and manually corrected, if necessary. One participant with more than 15% of missing beats was excluded from further analyses (i.e., N = 106). After artefact removal, mean heart rate (HR) and HRV measures were determined on 10-minute intervals. The first such interval is set during the first questionnaire round and begins one minute after the start of the pre-training questionnaire (t1), the second two intervals correspond to the two 10-minute training blocks (t2.1 and t2.2), and the final interval is set during the second questionnaire round and begins one minute after the start of the post training questionnaire (t3). To quantify HRV in the time-domain, we used the measure SDNN (i.e., the standard deviation of normal-to-normal heartbeat intervals), which has been shown to increase during (HRV-BF induced) PB at around 0.1 Hz (Lin et al., 2012). In addition, we considered the power in the low frequency (LF) band (i.e., 0.04 Hz–0.15 Hz) of the power spectral density, as well as the segment average of the coherence ratio (CR), as calculated for the determination of the feedback score in the HRV-BF conditions. CR is defined as peak power/(total spectral power – peak power), where peak power = power within a 0.03 Hz band around the highest peak between 0.04 Hz and 0.26 Hz, and total power = power between 0.0033 Hz and 0.4 Hz. As in the real-time processing, the power spectral density was estimated via Fast Fourier Transform from the filtered RRI series interpolated at 4 Hz using cubic splines.

3.10. Statistical analysis

Data analysis was performed using R (version 4.0.3) and RStudio (version 1.4.1103). A significance level of .05 was used for the omnibus tests. The Shapiro–Wilk test and QQ-plots indicated violations of the assumption of normality and Levene’s test revealed heteroscedasticity across several measures. Moreover, boxplots with an interquartile range ± 1.5 identified a number of outliers. Therefore, we used the non-parametric aligned rank transform analysis of variance (ART) procedure provided by the `ARTool` R package for the analyses (Wobbrock, Findlater, Gergle, & Higgins, 2011). Eighteen dependent UX variables were analysed with fixed-effects only models based on the ART procedure with Technique (2 levels: HRV-BF vs. sPB) and Technology (2 levels: HMD vs. screen) as between-subjects factors. One UX measure (i.e., ACTA) was only collected pre training (t1) and all the others were assessed post training (t3). The effects on the eight dependent psychological state and four dependent cardiac measures were analysed with mixed-effects models based on the ART procedure with the fixed effect of the within-subject factor Time (three levels: t1, t2, t3 for psychological measures; four levels: t1, t2.1, t2.2, t3 for cardiac measures) and additional random effects of participants. Effect size η_p^2’s were computed for the main and interaction effects of all models. Omnibus tests were followed by planned contrast tests for comparisons of interest. Significance levels of planned contrasts were also set at .05 and adjusted using the Benjamini–Hochberg method. The ART-C procedure, an extension of ART, includes an align-and-rank procedure for single-factor and multifactor contrasts (Elkin, Kay, Higgins, & Wobbrock, 2021). The `ARTool` package further provides effect size Cohen's d of single-factor pairwise comparisons.

4. Results

4.1. Baseline characteristics

Randomisation checks showed that baseline characteristics were well-balanced across conditions (all $p's \geq .05$). An overview of all means and standard deviations of these characteristics can be found in Table 1.

4.2. Repeated psychological state measures

Omnibus tests revealed a significant main effect of the within-subjects factor Time on $\text{MDMQ}_{\text{mood}}$, $\text{MDMQ}_{\text{wakefulness}}$, and $\text{MDMQ}_{\text{calmness}}$. Similarly, we found a significant main effect of Time on $\text{VAS}_{\text{relaxed}}$ and $\text{VAS}_{\text{stressed}}$ (all $p's < .001$; see Tables A.1 and A.3 in the Appendix for detailed summary and test statistics). Finally, we found significant two-way interaction effects of Technology × Technique ($p = .04$) and Technology × Time ($p = .03$) on $\text{VAS}_{\text{stressed}}$, and a significant three-way interaction of Technology × Technique × Time on $\text{VAS}_{\text{stressed}}$ ($p < .001$). No other significant main or interaction effects on repeated psychological measures were found (all $p's \geq .05$).

Single-factor contrasts of significant main effects of Time revealed that mean values of $\text{MDMQ}_{\text{mood}}$, $\text{MDMQ}_{\text{calmness}}$, $\text{MDMQ}_{\text{wakefulness}}$ and $\text{VAS}_{\text{relaxed}}$ all increased significantly from t1 to t3 across conditions (all $p's < .003$).

Multifactor contrasts of the significant three-way interaction Technology × Technique × Time revealed that perceived stress levels (i.e., $\text{VAS}_{\text{stressed}}$) significantly decreased individually for all technology and technique combinations from t1 to t3 (all $p's < .04$).

Comparisons at specific measurement time points revealed that the $\text{VAS}_{\text{stressed}}$ was significantly higher for sPB_{HMD} than for $\text{sPB}_{\text{screen}}$ at t2, $t(188) = 2.65, p = .03$. However, there were no other significant differences between factor level combinations at specific measurement time points ($p \geq .05$), which would explain the significant interactions effect.

Taken together, these results confirm hypothesis H1a that both HRV-BF and sPB lead to significant changes in psychological state over time (see Figs. 3 and 4). The results do, however, not confirm hypothesis H2a that practising HRV-BF compared to sPB leads to greater changes in an individual's psychological state. Similarly, the results do also not confirm hypotheses H3a that using an HMD leads to greater psychological changes compared to a screen while practising both HRV-BF or sPB, and H4a that using an HMD leads to greater psychological changes during HRV-BF than during sPB.

4.3. Repeated cardiac measures

Omnibus tests revealed a significant main effect of the within-subjects factor Time on HR, SDNN, LF, and CR (all $p's < .001$; see Tables A.4 and A.5 in the Appendix for detailed summary and test statistics). There were additional significant main effects of the between-subjects factor Technique on LF ($p < .001$) and CR ($p < .001$), and a main effect of the between-subjects factor Technology on CR ($p < .001$). Moreover, we found significant two-way interaction effects of Technique × Time on LF ($p < .001$) and CR ($p < .001$), and a significant two-way interaction effect of Technology × Time on CR ($p < .001$). Finally, the analysis revealed a significant three-way interaction of Technology × Technique × Time on CR ($p < .001$). No other significant effects on cardiac measures were found (all $p's \geq .05$).

Single-factor contrasts of the significant main effect of Time on HR revealed that HR steadily decreased significantly from t1 to t3, while SDNN significantly increased from t1 to t2.1, and significantly decreased again from t2.2 to t3 across conditions (all $p's < .001$).

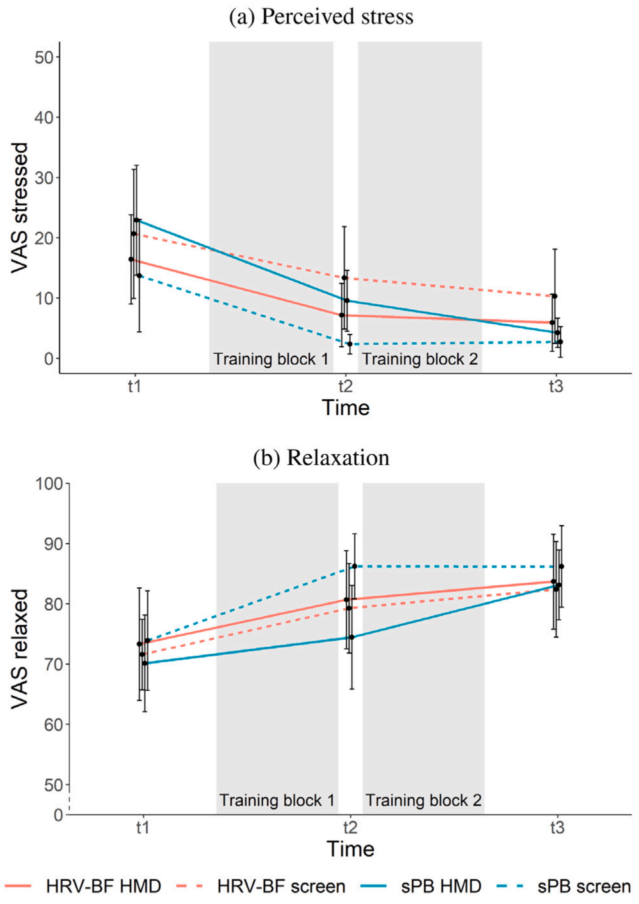


Fig. 3. Condition means of repeated psychological state measures for the VAS stressed (a) and relaxed subscales (b). Error bars indicate 95% confidence intervals. Abbreviations: VAS = Visual Analogue Scale; HRV-BF = heart rate variability biofeedback; sPB = standardised paced breathing; HMD = head-mounted display.

Single-factor contrasts of the significant main effect of Technique revealed that practising HRV-BF rather than sPB led to higher increases in both LF, $t(102) = 47.4, p = .02$, and CR, $t(102) = 69.7, p < .001$. For the significant main effect of Technology, contrast testing revealed that wearing an HMD while practising HRV-BF or sPB led to higher increases in CR, $t(102) = 51.20, p = .001$.

Multifactor contrasts of the significant interaction effect of Technique × Time revealed significant increases of LF for both Technique factor levels HRV-BF and sPB from t1 to t2.1 and significant decreases for both from t2.2 to t3 and still remains significantly higher post-training at t3 compared to pre-training at t1 (all $p's < .05$). However, contrast tests revealed no differences between the two techniques at specific measurement time points (all $p's \geq .05$).

For the significant three-way interaction of Technology × Technique × Time on CR, multifactor contrasts revealed significant increases individually for all technology and technique combinations (i.e., $\text{HRV-BF}_{\text{HMD}}$, $\text{HRV-BF}_{\text{screen}}$, sPB_{HMD} and $\text{sPB}_{\text{screen}}$) from t1 to t2.1, followed by significant decreases from t2.2 to t3 (all $p's < .05$). Planned comparisons at specific measurement time points revealed that CR was significantly higher for $\text{HRV-BF}_{\text{screen}}$ than for $\text{HRV-BF}_{\text{HMD}}$ at t1, $t(327) = -2.44, p = .04$, and significantly lower for $\text{sPB}_{\text{screen}}$ than for $\text{HRV-BF}_{\text{HMD}}$ at t2.1, $t(327) = 2.69, p = .02$. There were no other significant differences between other factor level combinations at specific measurement time points (all $p's \geq .05$).

Overall, these results confirm hypothesis H1b that both HRV-BF and sPB lead to significant changes in all target cardiac measures over time (i.e., HR, SDNN, LF, and CR). Moreover, the results partially confirm hypothesis H2b that changes in cardiac activity are greater when practising HRV-BF rather than practising sPB (i.e., only for LF and CR). Results also show that increases in CR are highest while wearing an HMD instead of using a classical screen during HRV-BF and sPB (see Fig. 5), which partially confirms hypothesis H3b. However, we do not find evidence that confirms hypothesis H4b that using an HMD for HRV-BF leads to greater cardiac changes than using an HMD for sPB.

4.4. Measures of user experience

Omnibus tests revealed significant main effects of the factor Technique on ACTA_{PCS}, $F(1,103) = 9.54, p = .003, \eta_p^2 = 0.08$, PQ_{Invo}, $F(1,103) = 10.99, p = .001, \eta_p^2 = 0.1$, $\text{UTAUT}_{\text{FacilCond}}$, $F(1,103) = 6.45, p = .01, \eta_p^2 = 0.06$, and SSQ_{Nas}, $F(1,103) = 5.55, p = .02, \eta_p^2 = 0.05$. Furthermore, we found significant main effects of the factor Technology on $\text{PQ}_{\text{IntQual}}$, $F(1,103) = 5.44, p = .022, \eta_p^2 = 0.05$, $\text{PQ}_{\text{ImrsAdpt}}$, $F(1,103) = 3.96, p < .05, \eta_p^2 = 0.04$, $\text{UTAUT}_{\text{FacilCond}}$, $F(1,103) = 16.3, p < .01, \eta_p^2 = 0.14$, $\text{UTAUT}_{\text{HedMotv}}$, $F(1,103) = 6.19, p = .01, \eta_p^2 = 0.06$, and SSQ_{Dis}, $F(1,103) = 4.87, p = .03, \eta_p^2 = 0.05$. Finally, we found significant Technology × Technique interaction effects for both $\text{UTAUT}_{\text{FacilCond}}$, $F(1,103) = 6.84, p = .01, \eta_p^2 = 0.06$, and $\text{UTAUT}_{\text{HedMotv}}$ $F(1,103) = 6.23, p = .01, \eta_p^2 = 0.06$. No other significant main or interaction effects where found for the measures of user experience (all $p's \geq .05$; see Table A.2 in the Appendix for detailed summary).

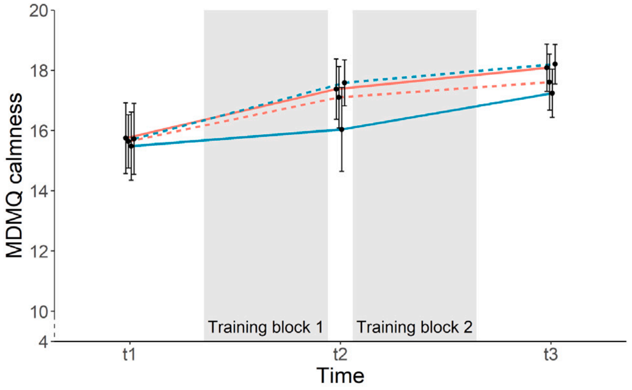
Single-factor contrasts of the significant main effect of Technique showed that anticipating HRV-BF was associated with lower perceived competence (ACTA_{PCS}) than anticipating sPB, $t(103) = -3.09, p = .003, d = -0.37$ and that PQ_{Invo} was higher for HRV-BF compared to sPB, $t(103) = 3.32, p = .001, d = 0.64$. Similarly, SSQ_{Nas} was significantly higher for the $\text{HRV-BF}_{\text{HMD}}$ and $\text{HRV-BF}_{\text{screen}}$ conditions, $t(103) = 2.36, p = .02, d = 0.46$.

For Technology, the single-factor contrasts showed that the interface quality ($\text{PQ}_{\text{IntQual}}$) interfered less in the HMD than in the screen conditions $t(103) = -2.33, p = .02, d = 0.45$, and that it was easier for participants to adapt and immerse ($\text{PQ}_{\text{ImrsAdpt}}$) themselves in the HMD system than it was in the screen setup, $t(103) = 1.99, p = .05, d = 0.39$. Additionally, single-factor pairwise comparisons show that SSQ_{Dis} was significantly higher for the $\text{HRV-BF}_{\text{HMD}}$ and sPB_{HMD} conditions, $t(103) = 2.21, p = .03, d = 0.43$.

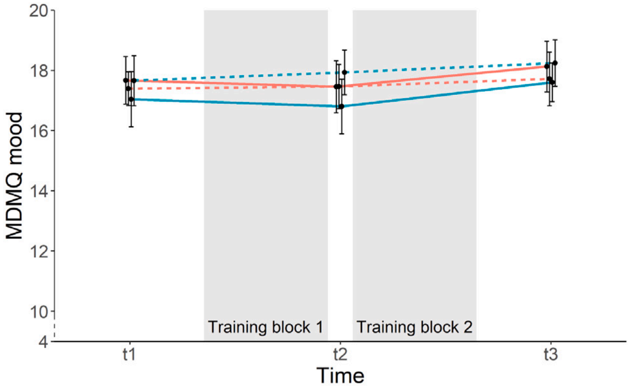
Multifactor contrasts of $\text{UTAUT}_{\text{FacilCond}}$ revealed that for sPB the screen setup outperformed the HMD, $t(103) = -4.18, p < .01$, whereas there was no significant difference in the case of HRV-BF. Overall the sPB on the screen was the combination with the highest value of $\text{UTAUT}_{\text{FacilCond}}$. Multifactor contrasts of $\text{UTAUT}_{\text{HedMotv}}$ shows that for sPB the HMD outperformed the screen setup, $t(103) = 3.62, p = .003$, whereas for HRV-BF there was no significant difference between the technologies.

Although there were no significant effects for $\text{UTAUT}_{\text{EffExp}}$, the mean of 6.25 out of 7 shows that participants assessed all systems as easy to use. For $\text{UTAUT}_{\text{BhvInt}}$, a mean of 3.88 out of 7 reveals a medium degree of intent of using the system in the future. Furthermore, with a mean of 81.36 (*SD* 10.19) of the SUS score, all the systems have achieved an adjusted rating of "excellent". Even though there were significant effects for both SSQ_{Dis} and SSQ_{Nas} the mean values were low on both scales, with an overall mean of 123.72 (*SD* 29.39) out of 682 for SSQ_{Dis}, and of 85.95 (*SD* 15.37) out of 467 for SSQ_{Nas}. Similarly, the mean of SSQ_{Ocu} (90.25 out of 371) shows that the system only led to minimal amounts of simulator sickness.

(a) Calmness



(b) Mood



(c) Wakefulness

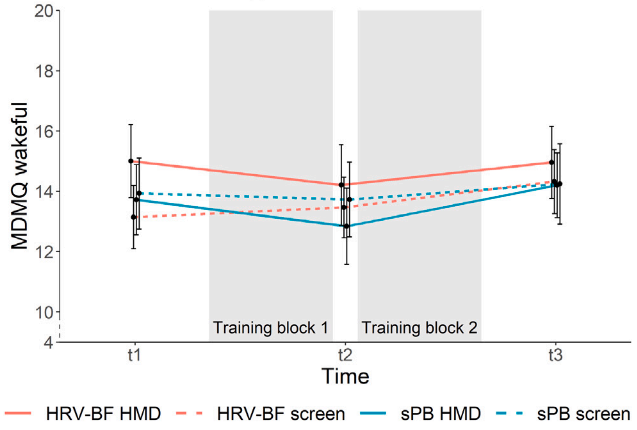


Fig. 4. Condition means of repeated psychological state measures for the MDMQ calmness (a), mood (b), and wakefulness subscales (b). Error bars indicate 95% confidence intervals. Abbreviations: MDMQ = Multidimensional Mood Questionnaire; HRV-BF = heart rate variability biofeedback; sPB = standardised paced breathing; HMD = head-mounted display.

The results of the user experience measures were mixed (hypotheses H5a and H5b). Namely, although PQ_{Invo} was higher and therefore more favourable for HRV-BF (see Fig. 6), the other measures either showed no significant differences or that HRV-BF was rated less favourable (ACTA_{PCS} and SSQ_{Nas}). The use of HMD was only rated favourably with respect to $\text{PQ}_{\text{IntQual}}$ and $\text{PQ}_{\text{ImrsAdpt}}$, but there were no other significant differences. In addition, while $\text{UTAUT}_{\text{HedMotv}}$ was moderately high when using an HMD independent of the technique that is practiced, it was negatively affected when using a screen for PB compared to HRV-BF (see Fig. 6). On the contrary, $\text{sPB}_{\text{screen}}$ was the best configuration with respect to $\text{UTAUT}_{\text{FacilCond}}$.

5. Discussion

In this study, we set out to determine the most suited technique and technology combination for VR-supported stress management. To that

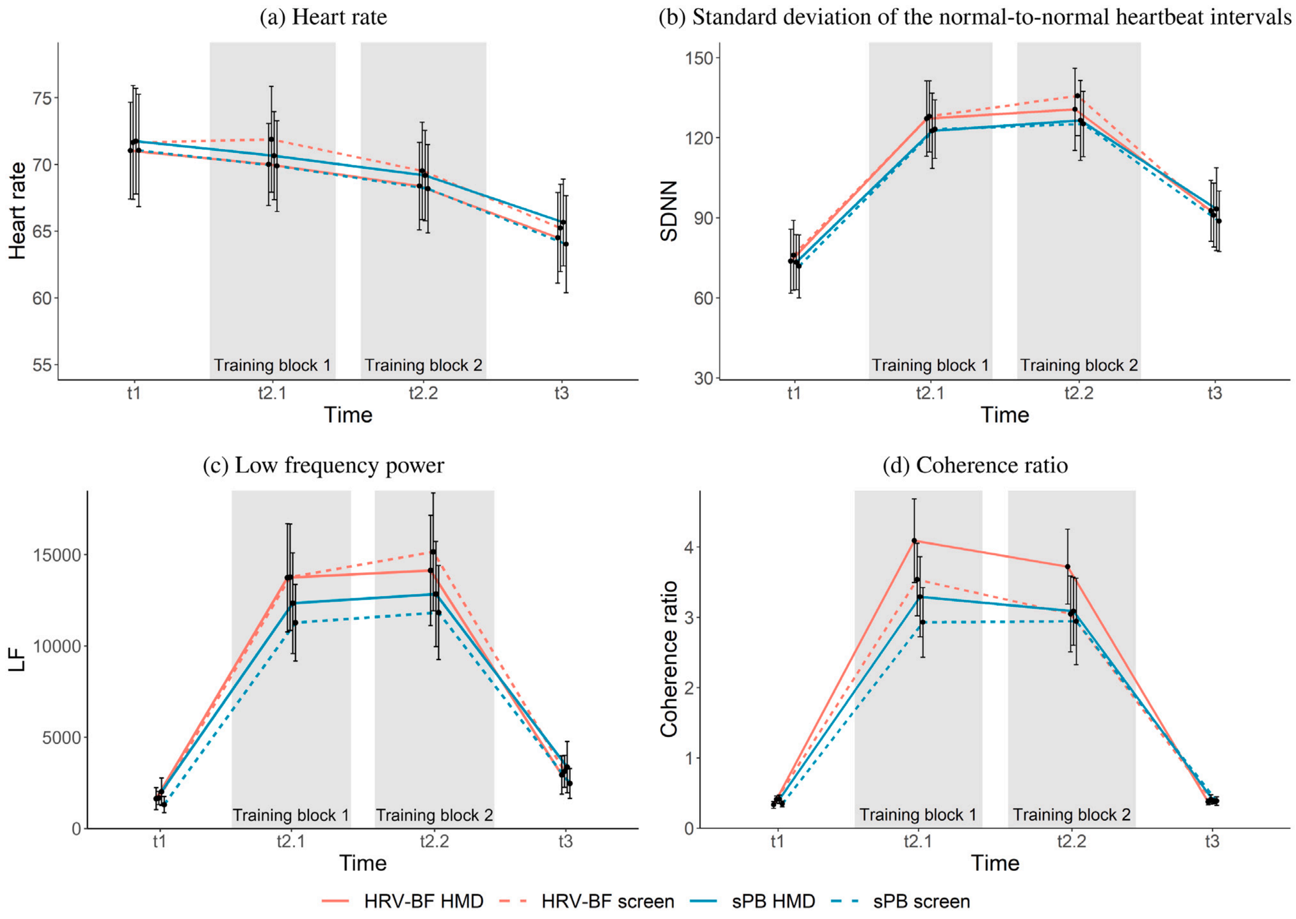


Fig. 5. Condition means of repeated cardiac measures. Error bars indicate 95% confidence intervals. Abbreviations: SDNN = Standard deviation of normal-to-normal heartbeat intervals; LF = power in the low frequency band of heart rate variability; HRV-BF = heart rate variability biofeedback; sPB = standardised paced breathing; HMD = head-mounted display.

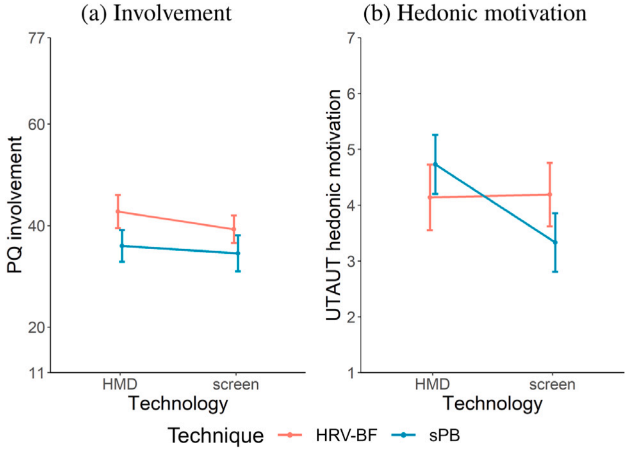


Fig. 6. Condition means of the user experiences measures for the involvement scale of the PQ (a) and the hedonic motivation scale of the UTAUT (b). Error bars represent 95% confidence intervals. Abbreviations: PQ = Presence Questionnaire; HMD = head-mounted display; UTAUT = Unified Theory of Acceptance and Use of Technology Questionnaire; HRV-BF = heart rate variability biofeedback; sPB = standardised paced breathing.

end, we compared practising heart rate variability biofeedback at exact resonance frequency (HRV-BF) to standardised paced breathing only (sPB) in the same virtual environment either delivered with an head-mounted display (HMD) or on a classical screen. Our results confirm hypotheses H1a and H1b that a training session with every technology and technique combination was able to significantly decrease perceived stress, improve mood, increase calmness, wakefulness and relaxation, while also slowing down heart rate (HR) and increasing all heart rate variability measures (HRV). Furthermore, results partially confirm hypothesis H2b, showing that the two HRV target parameters low frequency power (LF) and the coherence ratio (CR) were significantly higher during HRV-BF than during sPB, independent of the technology used. HRV-BF did not, however, lead to greater changes in heart rate and SDNN compared to sPB during the training session. Similarly, our results indicate that one of the HRV features, CR, was significantly higher overall for both technique types when using an HMD rather than a screen, while no other differences were found, partially confirming hypothesis H3b. Our results do not confirm hypotheses H4a and H4b, which stated that using an HMD during HRV-BF compared to during sPB would lead to even greater changes in psychological and cardiac measures. With respect to user experience (UX) and hypotheses H5a and H5b, results were mixed. More specifically, all technique and technology combinations were rated positively, while HRV-BF increased feelings of involvement more than sPB, and HMDs received more favourable ratings than screens with regard to interface quality and immersion adaption. In contrast, facilitating conditions for either technique were rated higher for screens and significantly higher for sPB_{screen} in comparison to sPB_{HMD}. Hedonic motivation was moderate to high for all combinations but significantly lower when sPB was practiced on a classical screen rather than with an HMD. Similarly, perceived competence was high for all combinations but significantly lower for HRV-BF than for sPB.

Our sample can be considered to be fairly healthy. The characteristics of our sample are mostly within the range of samples and norms found in the literature (Heissel et al., 2018; Hinz, Daig, Petrowski, & Brähler, 2012; Kerr et al., 2020; Lovibond & Lovibond, 1995). Chronic stress, as measured by the PSS, is slightly elevated, yet still within standard deviations of comparable samples (González-Ramírez, Rodríguez-Ayán, & Hernández, 2013; Klein et al., 2016). Well-being scores, as measured by the WHO-5 questionnaire, can be situated in the category "fair" right at the borderline to "good" (Bech, Olsen, Kjoller, & Rasmussen, 2003).

On a psychological level, we observed no increased benefits of HRV-BF over sPB, or HMDs over screens, as all setups led to significant yet indistinguishable changes over time.

These results suggest that every technique and technology combination significantly affects psychological state in the desired direction. Moreover, although participants did not start out as stressed before the session, the experimental manipulations were still able to induce significant changes. The effect sizes ranged from large for relaxation and calmness, to medium for mood and wakefulness, and small to medium for perceived stress confirming that HRV-BF as well as sPB affect an individual's perceived arousal and emotional state regardless of the technology used (Goessl et al., 2017; Mather & Thayer, 2018; Thayer & Lane, 2000). Contrary to our other hypotheses (i.e., H1a, H2a, H3a, H4a), we did not observe that HRV-BF outperforms sPB, or that HMDs outperform screens in terms of affecting an individual's psychological state beneficially. Perhaps, psychologically noticeable differences only become apparent with repeated practice of techniques and use of technologies. The lack of significant differences between conditions might also be attributed to the fact that self-reports are affected by different problems. They rely on participants' recall abilities, their level of self-awareness and willingness to report their genuinely perceived emotional state (Epel et al., 2018). In addition, descriptive statistics also suggest potential floor and ceiling effects for some of the psychological state measures (e.g., perceived stress, calmness, mood).

On a physiological level, the results similarly show that all technique and technology combinations are able to achieve the desired effect on autonomic activity, reflected by decreased sympathetic and increased parasympathetic cardiac activity in line with theory and findings from existing studies on HRV-BF (Goessl et al., 2017; Lehrer et al., 2020, 2000; McCraty & Shaffer, 2015; Shaffer et al., 2014). Moreover, our results also support other findings that HRV-BF at exact RF breathing produces greater power in the LF range than not breathing at exact RF (Lin et al., 2012; Steffen et al., 2017) as proposed by Vaschillo et al. (2006)'s resonance frequency model. In fact, an exploratory analysis of the calculated RF values for the HRV-BF conditions revealed that the mean RF was significantly different from the 6 breaths per minute (0.1 Hz) used in the sPB conditions. Specifically, a two-sided one sample t-test with $\mu_0 = 6$ resulted in $t(59) = -6.76, p < .001$ and an estimated mean of 5.44 breaths per minutes for the HRV-BF conditions (see also Fig. A.1 in the Appendix).

The expected increases in LF, CR and SDNN during the training session in all conditions are reflective of resonance effects (see Karavidas et al., 2007). The inverted U-shaped pattern of LF and CR clearly shows that soon as the paced breathing is terminated and individual switch back to spontaneous breathing, RSA shifts back from the LF to the high frequency (HF) band, decreasing LF power and thereby CR. Post-training, CR returns to baseline levels, which is expected as it requires a certain amount of LF power to achieve a ratio that greater or equal to 1. Compared to CR, LF and SDNN decrease more slowly post-training and values are still significantly higher post-training compared to levels pre-training. This might indicate increased sustained vagal activity immediately after the training. Slower spontaneous breathing rates post-training compared to pre-training might also have additionally increased potential sustained effects of vagal activity resulting in these observations. Indeed, Karavidas and colleagues (2007) did observe that the number of breaths per minute remained lower post-training compared to pre-training when participants returned to their normal breathing pattern. The observed continuous decreases in HR are in line with other studies that report a slight decrease of HR of around 5 beats per minute to mean values between 65 and 70 beats per minute (e.g., Karavidas et al., 2007; Rockstroh et al., 2019; Van Diest et al., 2014). From a theoretical perspective, PB at slow rates such as 6 breaths per minute results in bursts of vagal efferent traffic, which, in turn, lead to large HR fluctuations and decreases of mean HR (Grossman & Taylor, 2007; Shaffer & Ginsberg, 2017). The potentially heightened vagal activity post-training, as suggested by levels of LF and SDNN, might also explain the sustained decrease of HR post-training that we observed.

As to the extent of observed changes, it is challenging to compare absolute values of cardiac features across studies due to differences in terms of number and duration of blocks and different experimental protocols for HRV-BF and sPB in reported studies. Nonetheless, the previously described study by Blum et al. (2019) on HMD-supported HRV-BF observed near identical large effect sizes of Time on CR ($\eta_p^2 = 0.80$; our study: $\eta_p^2 = 0.85$). As for LF and SDNN, we found equal effect sizes of Time (SDNN: $\eta_p^2 = 0.74$; LF: $\eta_p^2 = 0.89$) as reported in the validation study of Sakakibara et al.'s 2020 novel protocol (SDNN: $\eta_p^2 = 0.80$; LF: $\eta_p^2 = 0.86$), which situates the effects of our results also within the more established standard protocol used to determine exact RF (see Lehrer et al., 2000; Sakakibara et al., 2020).

Returning to the differences between conditions, an additional explanation for the higher CR and LF during HRV-BF compared to sPB could be the significantly higher levels of involvement reported by individuals practising HRV-BF compared to individuals who practised sPB (i.e., medium effect size on the PQ_{Invo} subscale of the Presence Questionnaire). Similarly, the higher levels of CR achieved while using an HMD compared to a classical screen, independent of the technique practised, might be partly explained by higher levels of immersion adaptation and interface quality (i.e., other subscales of the Presence Questionnaire) reported for the use of HMDs (both with small effect size). Indeed, involvement, immersion adaptation and interface quality are important factors of presence (Witmer & Singer, 1998), which in the Cognitive Affective Model of Immersive Learning (CAMIL; Makransky & Petersen, 2021) positively affect the training outcome (target parameters CR and LF in our case). Taken together, it seems that HRV-BF, but also sPB, is able to facilitate at least one of the two psychological affordances (i.e., presence) of learning with an HMD, as proposed by CAMIL (Makransky & Petersen, 2021), which is assumed to make the use of an HMD for these techniques specifically effective.

With respect to autonomy and competence in technology adoption, all technique and technology combinations were rated positively. Although mean values on perceived competence for both training types are situated in the upper third of the score range, adopting sPB was associated with significantly higher perceived competence compared to adopting HRV-BF (medium effect size). Such differences could be attributed to the novelty and complexity of HRV-BF and it being a skill that has to be learned over time in order to feel more competent about it. Nonetheless, the high mean values suggest that participants' feelings of mastery and autonomy were promoted to a satisfactory extent through the means of psychoeducation on HRV-BF. These starting positions may hint at the potential of all setups to satisfy basic psychological needs (i.e., competence and autonomy) over a longer period of time and thereby increase intrinsic motivation (Ryan & Deci, 2000). With regards to hedonic motivation, we only found significant differences in the sPB conditions, where using an HMD was experienced as significantly more fun than using a classical screen (small to medium effect size). When comparing the two technologies, hedonic motivation was equally high for HMDs and screens while practising HRV-BF. Experiencing agency through the practice of HRV-BF in an immersive virtual environment, be it on screen or via HMD, might have facilitated participants' motivation, as conceptually proposed by CAMIL (Makransky & Petersen, 2021).

5.1. Limitations and future research

There are a series of limitations associated with this study. First, the BF was based on a comparatively global HRV score. Although such a cumulative score can act as positive reinforcement and might be a motivating external feedback, there is no potential for a transfer from external to internal feedback, since it holds no information on instantaneous changes in HRV, which also makes the signal less perceptible and controllable. Therefore, we believe that in this study breathing at exact RF during HRV-BF is the primary reason for the significantly different increases of CR and LF during HRV-BF compared to sPB only, rather than the BF learning process, i.e., the visualised changes of HRV as a means of BF. However, we decided against a more fine-grained continuous HRV feedback in this single-session experiment in order to reduce participants' cognitive load and limit frustration (see Makransky & Petersen, 2021; Sun et al., 2017; Wollmann et al., 2016).

Second, we did not include an additional control condition without any type of breathing. Adding such a condition would have provided unequivocal evidence that the observed changes over time were induced by HRV-BF and sPB and not by the setting or by chance. Similarly, keeping the virtual environment and objects as constant as possible (e.g., flower growth over time) across conditions would further increase the comparability of UX results. Relating to the first limitation, we also did not include a condition that would have enabled us to study the specific effects of both exact RF breathing and of feedback since we compared HRV-BF at exact RF to sPB without feedback. Future studies could, for example, compare the effects of different breathing patterns and the specific effects of biofeedback in a factorial design by delivering both breathing at exact RF and breathing at 0.1 Hz with and without feedback. Additionally, the study might have profited from subjecting participants to a stress test either before or after the training in order to study direct stress-reducing effects of HRV-BF. Inducing stress would have potentially also avoided floor effects of measures (i.e., perceived stress).

Third, this study is cross-sectional in nature and only allows conclusions on efficacy and system feasibility of a single-session protocol but not on the effectiveness regarding long-term stress management. Similarly, changes in UX measures over time with repeated use of a VR-supported HRV-BF should be investigated in order to rule out a novelty effect commonly observed for new technologies. It may be particularly interesting to see how involvement and motivation change and whether clients can increase their perceived competence and autonomy over time through the repeated use. Perhaps, as a result, participants may feel more competent to manage their stress and learn to choose how and when to apply these techniques in their everyday lives.

5.2. Conclusion

This study shows that practising either HRV-BF or sPB on a classical screen or an HMD in a single-session leads to beneficial psychological and cardiac effects in terms of stress management. The comparison of the two different techniques further revealed that practising HRV-BF at exact resonance frequency leads to greater changes in cardiac coherence and in the low frequency band of heart rate variability than sPB as proposed by the resonance frequency model. Moreover, this study is the first to show that using an HMD instead of a classical screen might hold more potential in terms of increases in cardiac coherence. In terms of UX, HRV-BF seems to enable greater feelings of involvement than sPB, but leaves clients feelings less competent in adapting the technique, regardless of the technology they are delivered with. As far as the choice of technology goes, HMDs might provide better interface quality and foster hedonic motivation and immersion adaption more than classical screens, regardless of the technique used. Overall, it might be specifically efficacious to immerse oneself in virtual reality via an HMD to practice breath-related techniques such as paced breathing or HRV-BF. Nevertheless, a virtual reality-supported HRV-BF training should be well accompanied by psychoeducational content and future research should investigate such a training's effectiveness, and look into underlying mechanisms, clients' user experience and adherence to such a training over a longer period of time.

CRediT authorship contribution statement

Raphael P. Weibel: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization, Project administration, Funding acquisition. **Jasmine I. Kerr:** Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization, Project administration, Funding acquisition. **Mara Naegelin:** Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization, Project administration, Funding acquisition. **Andrea Ferrario:** Conceptualization, Methodology, Investigation, Writing – review & editing, Funding acquisition. **Victor R. Schinazi:** Conceptualization, Methodology, Writing – review & editing, Funding acquisition. **Roberto La Marca:** Conceptualization, Resources, Writing – review & editing, Funding acquisition. **Christoph Hoelscher:** Writing – review & editing, Funding acquisition. **Urs M. Nater:** Methodology, Resources, Writing – review & editing, Supervision. **Florian von Wangenheim:** Conceptualization, Methodology, Resources, Writing – review & editing, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Our data and scripts for the analysis are available at: https://osf.io/nxqs2/.

Appendix

See Tables A.1–A.5 and Fig. A.1.

Table A.1
Sample size and means (standard deviations) of repeated psychological state measures of each condition.

		t1		t2		t3	
MDMQ$_{mood}$							
HRV-BF$_{HMD}$	24	17.67	(1.88)	17.46	(2.04)	18.13	(1.99)
HRV-BF$_{screen}$	28	17.39	(1.45)	17.46	(1.90)	18.28	(2.03)
sPB$_{HMD}$	25	17.04	(2.23)	16.80	(2.20)	17.60	(1.56)
sPB$_{screen}$	29	17.66	(2.19)	17.93	(1.94)	18.24	(2.03)
MDMQ$_{calmness}$							
HRV-BF$_{HMD}$	24	15.75	(2.79)	17.38	(2.37)	18.08	(1.86)
HRV-BF$_{screen}$	28	15.64	(2.28)	17.11	(2.63)	17.61	(2.39)
sPB$_{HMD}$	25	15.48	(2.74)	16.04	(3.37)	17.24	(1.94)
sPB$_{screen}$	29	15.72	(3.09)	17.59	(2.01)	18.21	(1.72)
MDMQ$_{wakefulness}$							
HRV-BF$_{HMD}$	24	15.00	(2.86)	14.21	(3.18)	14.96	(2.84)
HRV-BF$_{screen}$	28	13.14	(2.70)	13.46	(2.59)	14.32	(2.74)
sPB$_{HMD}$	25	13.72	(2.82)	12.84	(3.06)	14.20	(2.60)
sPB$_{screen}$	29	13.93	(3.09)	13.72	(3.26)	14.24	(3.50)
VAS$_{relaxed}$							
HRV-BF$_{HMD}$	24	73.29	(22.04)	80.67	(19.27)	83.67	(18.64)
HRV-BF$_{screen}$	28	71.57	(15.08)	79.25	(19.12)	82.39	(20.48)
sPB$_{HMD}$	25	70.12	(19.43)	74.44	(20.84)	83.12	(14.01)
sPB$_{screen}$	29	73.90	(21.73)	86.21	(14.24)	86.17	(17.75)

(*continued on next page*)

Table A.1 (*continued*).

		t1	t2	t3
$\text{VAS}_{\text{stressed}}$				
$\text{HRV-BF}_{\text{HMD}}$	24	16.42 (17.52)	7.17 (12.44)	5.92 (11.21)
$\text{HRV-BF}_{\text{screen}}$	28	20.64 (27.61)	13.36 (21.92)	10.29 (20.20)
sPB_{HMD}	25	22.92 (22.10)	9.56 (12.26)	4.24 (5.83)
$\text{sPB}_{\text{screen}}$	29	13.69 (24.52)	2.35 (4.19)	2.72 (6.72)

Abbreviations: MDMQ = Multidimensional Mood Questionnaire; VAS = Visual Analogue Scale. $\text{HRV-BF}_{\text{HMD}}$ = heart rate variability biofeedback via head-mounted display; $\text{HRV-BF}_{\text{screen}}$ = heart rate variability biofeedback on a desktop screen; sPB_{HMD} = standardised paced breathing via head-mounted display; $\text{sPB}_{\text{screen}}$ = standardised paced breathing on a desktop screen.

Table A.2
Sample size and means (standard deviations) of user experience measures of each condition.

	$\text{HRV-BF}_{\text{HMD}}$ (n = 24)	$\text{HRV-BF}_{\text{screen}}$ (n = 28)	sPB_{HMD} (n = 26)	$\text{sPB}_{\text{screen}}$ (n = 29)
ACTA_{ARS}	19.50 (3.53)	20.50 (3.59)	19.31 (4.16)	19.93 (3.40)
ACTA_{CRS}	9.63 (3.31)	9.71 (3.58)	9.69 (3.52)	9.00 (2.96)
ACTA_{RAI}	9.88 (4.96)	10.79 (5.14)	9.62 (4.25)	10.93 (4.79)
ACTA_{PCS}	9.38 (2.14)	9.97 (1.90)	10.19 (1.67)	11.17 (2.74)
PQ_{Invo}	42.75 (7.71)	39.25 (7.03)	35.96 (7.79)	34.52 (9.30)
$\text{PQ}_{\text{SensFi}}$	20.04 (4.21)	17.89 (4.20)	20.12 (3.29)	19.97 (4.93)
$\text{PQ}_{\text{ImrsAdpt}}$	36.67 (5.37)	34.82 (6.09)	37.85 (6.01)	34.59 (8.39)
$\text{PQ}_{\text{IntQual}}$	9.00 (3.06)	9.21 (2.81)	7.00 (2.50)	9.17 (3.86)
FSS_{flow}	50.38 (8.92)	44.57 (10.17)	48.00 (9.82)	48.83 (10.82)
$\text{FSS}_{\text{anxiety}}$	5.54 (3.39)	6.46 (3.80)	5.88 (2.89)	4.72 (2.71)
SUS	80.52 (9.70)	80.62 (9.64)	81.35 (9.41)	82.76 (12.01)
$\text{UTAUT}_{\text{EffExp}}$	6.21 (0.74)	6.20 (0.78)	6.39 (0.52)	6.25 (0.94)
$\text{UTAUT}_{\text{FacilCond}}$	4.90 (1.42)	5.30 (0.88)	4.94 (1.23)	6.24 (0.67)
$\text{UTAUT}_{\text{BhvInt}}$	3.35 (1.40)	4.15 (1.46)	3.92 (1.83)	4.05 (1.62)
$\text{UTAUT}_{\text{HedMotv}}$	4.14 (1.39)	4.19 (1.47)	4.73 (1.31)	3.33 (1.38)
SSQ_{Nas}	85.46 (13.34)	91.31 (18.44)	86.23 (16.07)	80.93 (11.58)
SSQ_{Ocu}	90.33 (22.79)	94.75 (23.93)	91.83 (21.26)	84.43 (23.43)
SSQ_{Dis}	132.24 (29.60)	122.79 (30.79)	128.49 (31.33)	113.28 (23.74)

Abbreviations: ACTA = Autonomy and Competence in Technology Adoption Questionnaire; ACTA_{ARS} = Autonomy Regulation Score; ACTA_{CRS} = Controlled Regulation Score; ACTA_{RAI} = Relative Autonomy Index; ACTA_{PCS} = Perceived Competence Score; PQ = Presence Questionnaire; PQ_{Invo} = Involvement; $\text{PQ}_{\text{SensFi}}$ = Sensor fidelity; $\text{PQ}_{\text{ImrsAdpt}}$ = Immersion adaptation; $\text{PQ}_{\text{IntQual}}$ = Interface quality; FSS = Flow Short Scale; FSS_{flow} = Flow; $\text{FSS}_{\text{anxiety}}$ = Perceived outcome importance; SUS = System Usability Scale; UTAUT = Unified Theory of Acceptance and Use of Technology Questionnaire; $\text{UTAUT}_{\text{EffExp}}$ = Effort expectancy; $\text{UTAUT}_{\text{FacilCond}}$ = Facilitating conditions; $\text{UTAUT}_{\text{BhvInt}}$ = Behavioural intention; $\text{UTAUT}_{\text{HedMotv}}$ = Hedonic motivation; SSQ = Simulator Sickness Questionnaire; SSQ_{Nas} = Nausea; SSQ_{Ocu} = Oculomotor; SSQ_{Dis} = Disorientation; $\text{HRV-BF}_{\text{HMD}}$ = heart rate variability biofeedback via head-mounted display; $\text{HRV-BF}_{\text{screen}}$ = heart rate variability biofeedback on a desktop screen; sPB_{HMD} = standardised paced breathing via head-mounted display; $\text{sPB}_{\text{screen}}$ = standardised paced breathing on a desktop screen.

Table A.3
Repeated psychological state measures: Results of mixed-effects models based on the ART procedure.

Measure	*df1, df2*	*F*	*p*	η^2	95% CI
$\text{MDMQ}_{\text{mood}}$					
Technology	1, 102	1.30	.26	0.01	[0.00, 0.09]
Technique	1, 102	0.01	.91	0.00	[0.00, 0.02]
Time	**2, 204**	**7.48**	**.00**	**0.07**	**[0.01, 0.14]**
Technology × technique	1, 102	3.83	.05	0.04	[0.00, 0.13]
Technology × time	2, 204	0.71	.49	0.00	[0.00, 0.04]
Technique × time	2, 204	0.23	.79	0.00	[0.00, 0.02]
Technol. × techniq. × time	2, 204	0.05	.95	0.00	[0.00, 0.01]
$\text{MDMQ}_{\text{calmness}}$					
Technology	1, 102	0.93	.34	0.00	[0.00, 0.08]
Technique	1, 102	0.45	.50	0.00	[0.00, 0.06]
Time	**2, 204**	**48.74**	**.00**	**0.32**	**[0.22, 0.41]**
Technology × technique	1, 102	1.47	.23	0.01	[0.00, 0.09]
Technology × time	2, 204	0.40	.67	0.00	[0.00, 0.03]
Technique × time	2, 204	0.91	.41	0.00	[0.00, 0.04]
Technol. × techniq. × time	2, 204	0.56	.57	0.00	[0.00, 0.03]
$\text{MDMQ}_{\text{wakefulness}}$					
Technology	1, 102	0.19	.67	0.00	[0.00, 0.05]
Technique	1, 102	0.41	.53	0.00	[0.00, 0.06]
Time	**2, 204**	**8.45**	**.00**	**0.08**	**[0.02, 0.15]**
Technology × technique	1, 102	2.32	.13	0.02	[0.00, 0.11]
Technology × time	2, 204	1.19	.31	0.01	[0.00, 0.05]
Technique × time	2, 204	0.17	.84	0.00	[0.00, 0.02]
Technol. × techniq. × time	2, 204	1.40	.25	0.01	[0.00, 0.05]
$\text{VAS}_{\text{relaxed}}$					
Technology	1, 102	0.35	.55	0.00	[0.00, 0.06]
Technique	1, 102	0.00	1.0	0.00	[0.00, 0.00]
Time	**2, 204**	**37.45**	**.00**	**0.27**	**[0.17, 0.36]**
Technology × technique	1, 102	2.67	.11	0.03	[0.00, 0.11]
Technology × time	2, 204	1.51	.22	0.01	[0.00, 0.06]
Technique × time	2, 204	0.11	.90	0.00	[0.00, 0.01]
Technol. × techniq. × time	2, 204	0.82	.44	0.00	[0.00, 0.04]
$\text{VAS}_{\text{stressed}}$					
Technology	1, 102	1.55	.22	0.02	[0.00, 0.09]
Technique	1, 102	0.19	.66	0.00	[0.00, 0.05]
Time	**2, 204**	**43.20**	**.00**	**0.30**	**[0.20, 0.39]**
Technology × technique	**1, 102**	**4.47**	**.04**	**0.04**	**[0.00, 0.14]**
Technology × time	**2, 204**	**3.64**	**.03**	**0.03**	**[0.00, 0.09]**
Technique × time	2, 204	1.18	.31	0.01	[0.00, 0.05]
Technol. × techniq. × time	**2, 204**	**7.18**	**.00**	**0.07**	**[0.01, 0.14]**

Significant effects ($p < .05$) are marked in bold. Abbreviations: ART = Aligned rank transform analysis of variance procedure; *df1, df2* = degrees of freedom 1 and 2; F = F-statistic; p = probability value; η^2 = partial eta squared; 95% CI = 95% confidence interval; MDMQ = Multidimensional Mood Questionnaire; VAS = Visual Analogue Scale.

Table A.4
Repeated cardiac measures: Results of mixed-effects models based on the ART procedure.

Measure	*df1, df2*	*F*	*p*	η^2	95% CI
Heart rate					
Technology	1, 102	0.15	.70	0.00	[0.00, 0.05]
Technique	1, 102	0.05	.83	0.00	[0.00, 0.04]
Time	**3, 306**	**87.77**	**.00**	**0.46**	**[0.38, 0.53]**
Technology × technique	1, 102	0.51	.48	0.00	[0.00, 0.07]
Technology × time	3, 306	0.21	.89	0.00	[0.00, 0.01]
Technique × time	3, 306	0.32	.81	0.00	[0.00, 0.02]
Technol. × techniq. × time	3, 306	0.37	.77	0.00	[0.00, 0.02]
SDNN					
Technology	1, 102	0.02	.88	0.00	[0.00, 0.03]
Technique	1, 102	0.90	.34	0.00	[0.00, 0.08]
Time	**3, 306**	**287.39**	**.00**	**0.74**	**[0.69, 0.77]**
Technology × technique	1, 102	0.19	.66	0.00	[0.00, 0.05]
Technology × time	3, 306	0.07	.98	0.00	[0.00, 0.00]
Technique × time	3, 306	1.29	.28	0.01	[0.00, 0.04]
Technol. × techniq. × time	3, 306	0.11	.96	0.00	[0.00, 0.00]
Low frequency power					
Technology	1, 102	0.75	.39	0.00	[0.00, 0.07]
Technique	**1, 102**	**5.79**	**.02**	**0.05**	**[0.00, 0.16]**
Time	**3, 306**	**802.42**	**.00**	**0.89**	**[0.87, 0.90]**
Technology × technique	1, 102	1.90	.17	0.02	[0.00, 0.10]
Technology × time	3, 306	1.28	.28	0.01	[0.00, 0.04]
Technique × time	**3, 306**	**9.49**	**.00**	**0.09**	**[0.03, 0.14]**
Technol. × techniq. × time	3, 306	1.13	.34	0.01	[0.00, 0.04]
Coherence ratio					
Technology	**1, 102**	**10.95**	**.00**	**0.10**	**[0.02, 0.22]**
Technique	**1, 102**	**21.18**	**.00**	**0.17**	**[0.06, 0.30]**
Time	**3, 306**	**574.69**	**.00**	**0.85**	**[0.82, 0.87]**
Technology × technique	1, 102	2.27	.13	0.02	[0.00, 0.11]
Technology × time	**3, 306**	**10.44**	**.00**	**0.09**	**[0.04, 0.15]**
Technique × time	**3, 306**	**16.57**	**.00**	**0.14**	**[0.07, 0.21]**
Technol. × techniq. × time	**3, 306**	**4.65**	**.00**	**0.04**	**[0.01, 0.09]**

Significant effects ($p < .05$) are marked in bold. Abbreviations: ART = Aligned rank transform analysis of variance procedure; *df1, df2* = degrees of freedom 1 and 2; F = F-statistics; p = probability value; η^2 = partial eta squared; 95% CI = 95% confidence interval; bpm = beats per minute; SDNN = standard deviation of normal-to-normal heartbeat intervals.

Table A.5
Sample size and means (standard deviations) of repeated cardiac measures of each condition.

		t1		t2.1		t2.2		t3	
Heart rate									
HRV-BF$_{\text{HMD}}$	24	71.03	(8.60)	70.00	(7.29)	68.38	(7.78)	64.50	(8.06)
HRV-BF$_{\text{screen}}$	27	71.65	(10.77)	71.88	(10.04)	69.52	(9.24)	65.24	(8.29)
sPB$_{\text{HMD}}$	26	71.74	(9.81)	70.66	(8.19)	69.17	(8.39)	65.65	(8.06)
sPB$_{\text{screen}}$	29	71.05	(11.07)	69.88	(8.96)	68.19	(8.71)	64.02	(9.57)
SDNN									
HRV-BF$_{\text{HMD}}$	24	73.73	(28.43)	127.19	(33.41)	130.64	(36.50)	92.61	(27.10)
HRV-BF$_{\text{screen}}$	27	75.95	(33.12)	127.99	(33.66)	135.71	(37.74)	90.98	(30.29)
sPB$_{\text{HMD}}$	26	73.35	(25.60)	122.65	(35.04)	126.50	(37.12)	93.26	(38.42)
sPB$_{\text{screen}}$	29	71.82	(31.05)	123.22	(28.87)	125.16	(32.25)	88.73	(29.77)
Low frequency power									
HRV-BF$_{\text{HMD}}$	24	1633.44	(1426.51)	13736.67	(7019.15)	14134.11	(7158.09)	2936.87	(2484.86)
HRV-BF$_{\text{screen}}$	27	1674.87	(902.59)	13766.80	(7374.12)	15151.55	(8160.79)	3133.37	(2216.19)
sPB$_{\text{HMD}}$	26	2021.54	(1862.09)	12339.68	(6827.13)	12841.53	(7122.09)	3363.62	(3483.61)
sPB$_{\text{screen}}$	29	1312.00	(1164.59)	11273.88	(5512.30)	11830.12	(6765.28)	2470.73	(2156.81)
Coherence ratio									
HRV-BF$_{\text{HMD}}$	24	0.33	(0.12)	4.09	(1.41)	3.72	(1.26)	0.37	(0.10)
HRV-BF$_{\text{screen}}$	27	0.41	(0.13)	3.53	(1.30)	3.05	(1.36)	0.41	(0.17)
sPB$_{\text{HMD}}$	26	0.41	(0.15)	3.29	(1.41)	3.09	(1.20)	0.38	(0.09)
sPB$_{\text{screen}}$	29	0.35	(0.11)	2.93	(1.30)	2.94	(1.62)	0.39	(0.16)

Abbreviations: SDNN = Standard deviation of normal-to-normal heartbeat intervals; HRV-BF$_{\text{HMD}}$ = heart rate variability biofeedback via head-mounted display; HRV-BF$_{\text{screen}}$ = heart rate variability biofeedback on a desktop screen; sPB$_{\text{HMD}}$ = standardised paced breathing via head-mounted display; sPB$_{\text{screen}}$ = standardised paced breathing on a desktop screen.

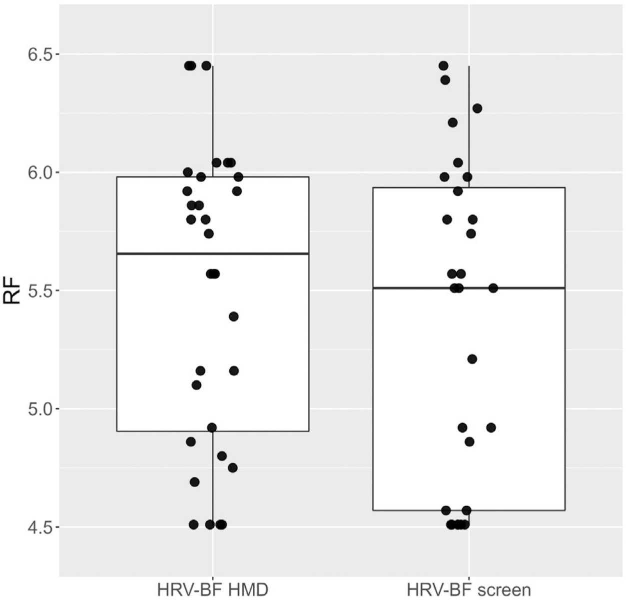


Fig. A.1. Calculated resonance frequency values for the HRV-BF conditions. Abbreviations: RF = resonance frequency in breaths per minute; HMD = head-mounted display; HRV-BF = heart rate variability biofeedback.

References

Annerstedt, M., Jönsson, P., Wallergård, M., Johansson, G., Karlson, B., Grahn, P., et al. (2013). Inducing physiological stress recovery with sounds of nature in a virtual reality forest — Results from a pilot study. *Physiology & Behavior*, *118*, 240–250. http://dx.doi.org/10.1016/j.physbeh.2013.05.023.

Bech, P., Olsen, L. R., Kjoller, M., & Rasmussen, N. K. (2003). Measuring well-being rather than the absence of distress symptoms: A comparison of the SF-36 Mental Health subscale and the WHO-Five well-being scale. *International Journal of Methods in Psychiatric Research*, *12*(2), 85–91. http://dx.doi.org/10.1002/mpr.145.

Berntson, G. G., Quigley, K. S., Norman, G. J., & Lozano, D. L. (2017). Cardiovascular psychophysiology. In *Handbook of Psychophysiology* (4th ed.). (pp. 183–216). New York: Cambridge University Press, http://dx.doi.org/10.1017/9781107415782.009.

Blum, J., Rockstroh, C., & Göritz, A. S. (2019). Heart rate variability biofeedback based on slow-paced breathing with immersive virtual reality nature scenery. *Frontiers in Psychology*, *10*, 2172. http://dx.doi.org/10.3389/fpsyg.2019.02172.

Blum, J., Rockstroh, C., & Göritz, A. S. (2020). Development and pilot test of a virtual reality respiratory biofeedback approach. *Applied Psychophysiology and Biofeedback*, *45*(3), 153–163. http://dx.doi.org/10.1007/s10484-020-09468-x.

Brooke, J. (1996). SUS: A 'quick and dirty' usability scale. In *Usability Evaluation in Industry* (1st ed.). London, UK: CRC Press, http://dx.doi.org/10.1201/9781498710411.

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, *24*(4), 385–396. http://dx.doi.org/10.2307/2136404.

Cummings, J. J., & Bailenson, J. N. (2016). How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology*, *19*(2), 272–309. http://dx.doi.org/10.1080/15213269.2015.1015740.

Dishman, R., Nakamura, Y., Garcia, M., Thompson, R., Dunn, A., & Blair, S. (2000). Heart rate variability, trait anxiety, and perceived stress among physically fit men and women. *International Journal of Psychophysiology*, *37*(2), 121–133. http://dx.doi.org/10.1016/s0167-8760(00)00085-4.

Elkin, L. A., Kay, M., Higgins, J. J., & Wobbrock, J. O. (2021). An aligned rank transform procedure for multifactor contrast tests. In *The 34th annual ACM symposium on user interface software and technology* (pp. 754–768). http://dx.doi.org/10.1145/3472749.3474784.

Epel, E. S., Crosswell, A. D., Mayer, S. E., Prather, A. A., Slavich, G. M., Puterman, E., et al. (2018). More than a feeling: A unified view of stress measurement for population science. *Frontiers in Neuroendocrinology*, *49*, 146–169. http://dx.doi.org/10.1016/j.yfrne.2018.03.001.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191. http://dx.doi.org/10.3758/BF03193146.

Gaggioli, A., Pallavicini, F., Morganti, L., Serino, S., Scaratti, C., Briguglio, M., et al. (2014). Experiential virtual scenarios with real-time monitoring (interreality) for the management of psychological stress: A block randomized controlled trial. *Journal of Medical Internet Research*, *16*(7), http://dx.doi.org/10.2196/jmir.3235.

Gaume, A., Vialatte, A., Mora-Sánchez, A., Ramdani, C., & Vialatte, F. B. (2016). A psychoengineering paradigm for the neurocognitive mechanisms of biofeedback and neurofeedback. *Neuroscience & Biobehavioral Reviews*, *68*, 891–910. http://dx.doi.org/10.1016/j.neubiorev.2016.06.012.

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

González-Ramírez, M. T., Rodríguez-Ayán, M. N., & Hernández, R. L. (2013). The Perceived Stress Scale (PSS): Normative data and factor structure for a large-scale sample in Mexico. *The Spanish Journal of Psychology*, *16*, E47. http://dx.doi.org/10.1017/sjp.2013.35.

Gradl, S., Wirth, M., Zillig, T., & Eskofier, B. M. (2018). Visualization of heart activity in virtual reality: A biofeedback application using wearable sensors. In *2018 IEEE 15th international conference on wearable and implantable body sensor networks* (pp. 152–155). http://dx.doi.org/10.1109/BSN.2018.8329681.

Grossman, P., & Taylor, E. W. (2007). Toward understanding respiratory sinus arrhythmia: Relations to cardiac vagal tone, evolution and biobehavioral functions. *Biological Psychology*, *74*(2), 263–285. http://dx.doi.org/10.1016/j.biopsycho.2005.11.014.

Heissel, A., Pietrek, A., Flunger, B., Fydrich, T., Rapp, M. A., Heinzel, S., et al. (2018). The validation of the German basic psychological need satisfaction and frustration scale in the context of mental health. *European Journal of Health Psychology*, *25*(4), 119–132. http://dx.doi.org/10.1027/2512-8442/a000017.

Hinz, A., Daig, I., Petrowski, K., & Brähler, E. (2012). Mood in the German population: Norms of the multidimensional mood questionnaire MDBF. *Psychotherapie, Psychosomatik, Medizinische Psychologie*, *62*(2), 52–57. http://dx.doi.org/10.1055/s-0031-1297960.

Hirshoren, N., Tzoran, I., Makrienko, I., Edoute, Y., Plawner, M. M., Itskovitz-Eldor, J., et al. (2002). Menstrual cycle effects on the neurohumoral and autonomic nervous systems regulating the cardiovascular system. *The Journal of Clinical Endocrinology and Metabolism*, *87*(4), 1569–1575. http://dx.doi.org/10.1210/jcem.87.4.8406.

Houzangbe, S., Christmann, O., Gorisse, G., & Richir, S. (2020). Effects of voluntary heart rate control on user engagement and agency in a virtual reality game. *Virtual Reality*, *24*(4), 665–681. http://dx.doi.org/10.1007/s10055-020-00429-7.

Karavidas, M. K., Lehrer, P. M., Vaschillo, E., Vaschillo, B., Marin, H., Buyske, S., et al. (2007). Preliminary results of an open label study of heart rate variability biofeedback for the treatment of major depression. *Applied Psychophysiology and Biofeedback*, *32*, 19–30. http://dx.doi.org/10.1007/s10484-006-9029-z.

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology*, *3*(3), 203–220. http://dx.doi.org/10.1207/s15327108ijap0303_3.

Kennedy, L., & Parker, S. H. (2019). Biofeedback as a stress management tool: A systematic review. *Cognition, Technology & Work*, *21*(2), 161–190. http://dx.doi.org/10.1007/s10111-018-0487-x.

Kerr, J. I., Naegelin, M., Weibel, R. P., Ferrario, A., La Marca, R., von Wangenheim, F., et al. (2020). The effects of acute work stress and appraisal on psychobiological stress responses in a group office environment. *Psychoneuroendocrinology*, *121*, 104837. http://dx.doi.org/10.1016/j.psyneuen.2020.104837.

Kim, H. -G., Cheon, E. -J., Bai, D. -S., Lee, Y. H., & Koo, B. -H. (2018). Stress and heart rate variability: A meta-analysis and review of the literature. *Psychiatry Investigation*, *15*(3), 235–245. http://dx.doi.org/10.30773/pi.2017.08.17.

Klein, E., Brähler, E., Dreier, M., Reinecke, L., Müller, K., Schmutzer, G., et al. (2016). The German version of the perceived stress scale – Psychometric characteristics in a representative German community sample. *BMC Psychiatry*, *16*, 159. http://dx.doi.org/10.1186/s12888-016-0875-9.

Laborde, S., Mosley, E., & Thayer, J. F. (2017). Heart rate variability and cardiac vagal tone in psychophysiological research – Recommendations for experiment planning, data analysis, and data reporting. *Frontiers in Psychology*, *8*, 213. http://dx.doi.org/10.3389/fpsyg.2017.00213.

Lehrer, P., Kaur, K., Sharma, A., Shah, K., Huseby, R., Bhavsar, J., et al. (2020). Heart rate variability biofeedback improves emotional and physical health and performance: A systematic review and meta-analysis. *Applied Psychophysiology and Biofeedback*, *45*(3), 109–129. http://dx.doi.org/10.1007/s10484-020-09466-z.

Lehrer, P. M., Vaschillo, E., & Vaschillo, B. (2000). Resonant frequency biofeedback training to increase cardiac variability: Rationale and manual for training. *Applied Psychophysiology and Biofeedback*, *25*, 177–191. http://dx.doi.org/10.1023/A:1009554825745.

Lin, G., Xiang, Q., Fu, X., Wang, S., Wang, S., Chen, S., et al. (2012). Heart rate variability biofeedback decreases blood pressure in prehypertensive subjects by improving autonomic function and baroreflex. *The Journal of Alternative and Complementary Medicine*, *18*(2), 143–152. http://dx.doi.org/10.1089/acm.2010.0607.

Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the depression anxiety stress scales* (2nd ed.). Sidney: Psychology Foundation, http://dx.doi.org/10.1037/t01004-000.

Lüddecke, R., & Felnhofer, A. (2021). Virtual reality biofeedback in health: A scoping review. *Applied Psychophysiology and Biofeedback*, 1–15. http://dx.doi.org/10.1007/s10484-021-09529-9.

Maarsingh, B. M., Bos, J., Tuijn, C. F. J. V., & Renard, S. B. (2019). Changing stress mindset through stressjam: A virtual reality game using biofeedback. *Games for Health Journal*, *8*(5), 326–331. http://dx.doi.org/10.1089/g4h.2018.0145.

Makransky, G., & Petersen, G. B. (2021). The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality. *Educational Psychology Review*, *33*(3), 937–958. http://dx.doi.org/10.1007/s10648-020-09586-2.

Malik, M., Bigger, J. T., Camm, A. J., Kleiger, R. E., Malliani, A., Moss, A. J., et al. (1996). Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *European Heart Journal*, *17*(3), 354–381. http://dx.doi.org/10.1093/oxfordjournals.eurheartj.a014868.

Maples-Keller, J. L., Bunnell, B. E., Kim, S. -J., & Rothbaum, B. O. (2017). The use of virtual reality technology in the treatment of anxiety and other psychiatric disorders. *Harvard Review of Psychiatry*, *25*(3), 103–113. http://dx.doi.org/10.1097/hrp.0000000000000138.

Mather, M., & Thayer, J. F. (2018). How heart rate variability affects emotion regulation brain networks. *Current Opinion in Behavioral Sciences*, *19*, 98–104. http://dx.doi.org/10.1016/j.cobeha.2017.12.017.

McCraty, R., & Shaffer, F. (2015). Heart rate variability: New perspectives on physiological mechanisms, assessment of self-regulatory capacity, and health risk. *Global Advances in Health and Medicine*, *4*(1), 46–61. http://dx.doi.org/10.7453/gahmj.2014.073.

Meichenbaum, D. (1976). Cognitive factors in biofeedback therapy. *Biofeedback and Self-Regulation*, *1*, 201–216. http://dx.doi.org/10.1007/BF00998587.

Mezzacappa, E. S., Kelsey, R. M., & Katkin, E. S. (2005). Breast feeding, bottle feeding, and maternal autonomic responses to stress. *Journal of Psychosomatic Research*, *58*, 351–365. http://dx.doi.org/10.1016/j.jpsychores.2004.11.004.

Morfeld, M., Kirchberger, I., & Bullinger, M. (2011). SF-36 Fragebogen zum Gesundheitszustand: Deutsche Version des Short Form-36 Health Survey. In *SF-36 Fragebogen zum Gesundheitszustand. Deutsche Version des Short Form-36 Health Survey* (2nd ed.). (pp. 1–121). Hogrefe.

Nanke, A., & Rief, W. (2000). Biofeedback-Therapie bei somatoformen Störungen. *Verhaltenstherapie*, *10*(4), 238–248. http://dx.doi.org/10.1159/000056633.

Peters, D., Calvo, R. A., & Ryan, R. M. (2018). Designing for motivation, engagement and wellbeing in digital experience. *Frontiers in Psychology*, *9*, 797. http://dx.doi.org/10.3389/fpsyg.2018.00797.

Porges, S. W. (2007). The polyvagal perspective. *Biological Psychology*, *74*(2), 116–143. http://dx.doi.org/10.1016/j.biopsycho.2006.06.009.

Rheinberg, F., Vollmeyer, R., & Engeser, S. (2003). Die Erfassung des Flow-Erlebens. *Diagnostik von Motivation und Selbstkonzept* (pp. 261–279). Hogrefe.

Rockstroh, C., Blum, J., & Göritz, A. S. (2019). Virtual reality in the application of heart rate variability biofeedback. *International Journal of Human-Computer Studies*, *130*, 209–220. http://dx.doi.org/10.1016/j.ijhcs.2019.06.011.

Rockstroh, C., Blum, J., & Göritz, A. S. (2020). Combining VR and biofeedback: The effects on perceived restorativeness and presence. *Journal of Media Psychology*, *32*(4), 176–186. http://dx.doi.org/10.1027/1864-1105/a000270.

van Rooij, M., Lobel, A., Harris, O., Smit, N., & Granic, I. (2016). DEEP: A biofeedback virtual reality game for children at-risk for anxiety. In *Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems* (pp. 1989–1997). http://dx.doi.org/10.1145/2851581.2892452.

Rowland, D. P., Casey, L. M., Ganapathy, A., Cassimatis, M., & Clough, B. A. (2021). A decade in review: A systematic review of virtual reality interventions for emotional disorders. *Psychosocial Intervention*, *31*(1), 1–20. http://dx.doi.org/10.5093/pi2021a8.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78. http://dx.doi.org/10.1037//0003-066x.55.1.68.

Sakakibara, M., Kaneda, M., & Oikawa, L. O. (2020). Efficacy of paced breathing at the low-frequency peak on heart rate variability and baroreflex sensitivity. *Applied Psychophysiology and Biofeedback*, *45*(1), 31–37. http://dx.doi.org/10.1007/s10484-019-09453-z.

Shaffer, F., & Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms. *Frontiers in Public Health*, *5*, 258. http://dx.doi.org/10.3389/fpubh.2017.00258.

Shaffer, F., McCraty, R., & Zerr, C. L. (2014). A healthy heart is not a metronome: An integrative review of the heart's anatomy and heart rate variability. *Frontiers in Psychology*, *5*, 1040. http://dx.doi.org/10.3389/fpsyg.2014.01040.

Shah, L. B. I., Torres, S., Kannusamy, P., Chng, C. M. L., Hong-Gu, H., & Klainin-Yobas, P. (2015). Efficacy of the virtual reality-based stress management program on stress-related variables in people with mood disorders: The feasibility study. *Archives of Psychiatric Nursing*, *29*(1), 6–13. http://dx.doi.org/10.1016/j.apnu.2014.09.003.

Steffen, P. R., Austin, T., DeBarros, A., & Brown, T. (2017). The impact of resonance frequency breathing on measures of heart rate variability, blood pressure, and mood. *Frontiers in Public Health*, *5*, 222. http://dx.doi.org/10.3389/fpubh.2017.00222.

Steyer, R., Schwenkmezger, P., Notz, P., & Eid, M. (1997). *Der Mehrdimensionale Befindlichkeitsfragebogen MDBF [Multidimensional Mood Questionnaire]*. Goettingen, Germany: Hogrefe.

Sun, Z., Cao, N., & Ma, X. (2017). Attention, comprehension, execution: Effects of different designs of biofeedback display. In *Proceedings of the 2017 CHI conference extended abstracts on human factors in computing systems* (pp. 2132–2139). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/3027063.3053082.

Thayer, J. F., & Lane, R. D. (2000). A model of neurovisceral integration in emotion regulation and dysregulation. *Journal of Affective Disorders*, *61*(3), 201–216. http://dx.doi.org/10.1016/s0165-0327(00)00338-4.

Tong, X., Gromala, D., Choo, A., Amin, A., & Shaw, C. (2015). The virtual meditative walk: An immersive virtual environment for pain self-modulation through mindfulness-based stress reduction meditation. In R. Shumaker, & S. Lackey (Eds.), *Virtual, Augmented and Mixed Reality* (pp. 388–397). Cham: Springer International Publishing, http://dx.doi.org/10.1007/978-3-319-21067-4_40.

Topp, C. W., Østergaard, S. D., Søndergaard, S., & Bech, P. (2015). The WHO-5 well-being index: A systematic review of the literature. *Psychotherapy and Psychosomatics*, *84*(3), 167–176. http://dx.doi.org/10.1159/000376585.

Van Diest, I., Verstappen, K., Aubert, A. E., Widjaja, D., Vansteenwegen, D., & Vlemincx, E. (2014). Inhalation/exhalation ratio modulates the effect of slow breathing on heart rate variability and relaxation. *Applied Psychophysioloogy Biofeedback*, *39*, 171–180. http://dx.doi.org/10.1007/s10484-014-9253-x.

Vaschillo, E. G., Vaschillo, B., & Lehrer, P. M. (2006). Characteristics of resonance in heart rate variability stimulated by biofeedback. *Applied Psychophysiology and Biofeedback*, *31*(2), 129–142. http://dx.doi.org/10.1007/s10484-006-9009-3.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 425–478. http://dx.doi.org/10.2307/30036540.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*, *7*(3), 225–240. http://dx.doi.org/10.1162/105474698565686.

Wobbrock, J. O., Findlater, L., Gergle, D., & Higgins, J. J. (2011). The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 143–146). http://dx.doi.org/10.1145/1978942.1978963.

Wollmann, T., Abtahi, F., Eghdam, A., Seoane, F., Lindecrantz, K., Haag, M., et al. (2016). User-centred design and usability evaluation of a heart rate variability biofeedback game. *IEEE Access*, *4*, 5531–5539. http://dx.doi.org/10.1109/ACCESS.2016.2601882.

Yu, B., Funk, M., Hu, J., Wang, Q., & Feijs, L. (2018). Biofeedback for everyday stress management: A systematic review. *Frontiers in ICT*, *5*, 23. http://dx.doi.org/10.3389/fict.2018.00023.