This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2023.3320228

Is that my Heartbeat? Measuring and Understanding Modality-dependent Cardiac Interoception in Virtual Reality

Abdallah El Ali, Rayna Ney, Zeph M. C. van Berlo, Pablo Cesar

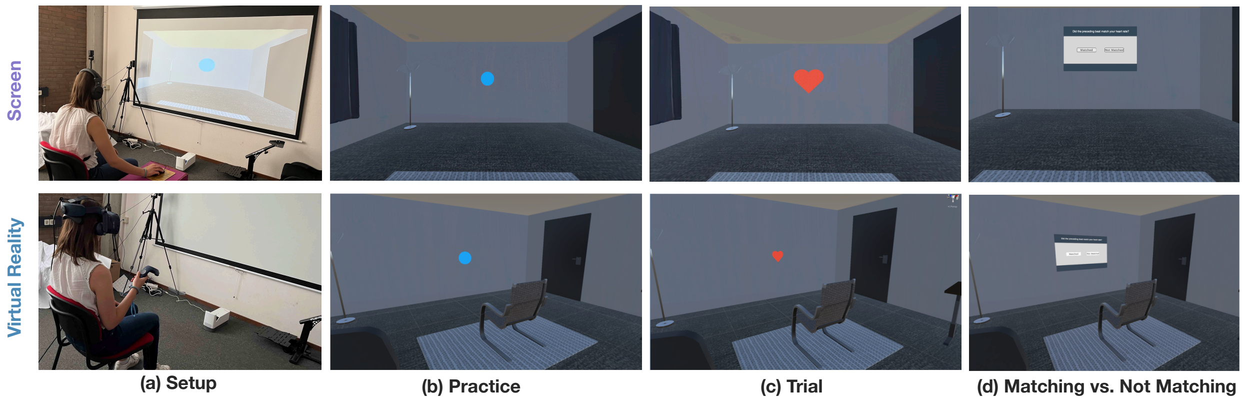


Fig. 1: Our Screen (top) and Virtual Reality (bottom) between-subjects conditions showing (a) participant interacting with the task, (b) an example practice trial, (c) an example main trial with the visual modality, and (d) prompt screen asking participant if their preceding displayed heartbeat matches theirs. *Best viewed in color.*

Abstract— Measuring interoception ('perceiving internal bodily states') has diagnostic and wellbeing implications. Since heartbeats are distinct and frequent, various methods aim at measuring cardiac interoceptive accuracy (CIAcc). However, the role of exteroceptive modalities for representing heart rate (HR) across screen-based and Virtual Reality (VR) environments remains unclear. Using a PolarH10 HR monitor, we develop a **modality-dependent cardiac recognition task** that modifies displayed HR. In a mixed-factorial design (N=50), we investigate how task environment (Screen, VR), modality (Audio, Visual, Audio-Visual), and real-time HR modifications (±15%, ±30%, None) influence CIAcc, interoceptive awareness, mind-body measures, VR presence, and post-experience responses. Findings showed that participants confused their HR with underestimates up to 30%; environment did not affect CIAcc but influenced mind-related measures; modality did not influence CIAcc, however including audio increased interoceptive awareness; and VR presence inversely correlated with CIAcc. We contribute a lightweight and extensible cardiac interoception measurement method, and implications for biofeedback displays.

Index Terms—Interoception, heart rate, virtual reality, cardiac, biofeedback, modality

1 Introduction

Interoception is the capacity to perceive one's own internal bodily states [11], and believed to influence our cognitive, affective, and interpersonal interactions [80]. It plays a crucial role in regulating homeostatic processes [57] as well as in health and disease. It influences our physical and mental health [36], with links to disorders including anxiety and depression [59], and Attention-Deficit/Hyperactivity Disorder [39]. Despite the surge of research on cardiac interoception, and with the wide availability of consumer head-mounted displays (HMDs) that enable immersive Virtual Reality (VR) applications, there is little known about the interplay between VR experiences and how we perceive our bodily states. VR experiences have been shown to influence presence in the sense of 'being there' (cf., Place Illusion [74, 75]). This has clinical implications [65], and influences emotional reactions [16] and physiology in VR [88]. Indeed, immersive therapeutic biofeedback using heart rate biofeeback has been used for lowering anxiety [38], for stress and pain reduction [42], and to foster empathy [68]. Specifically, previous work has shown that embodying an avatar, with varying fidelity and realism across (social) VR platforms can influence presence [30] and body perception measures [18].

- *Abdallah El Ali is with Centrum Wiskunde & Informatica. E-mail: aea@cwi.nl*
- *Rayna Ney is with University of Amsterdam. E-mail: raynaney@gmail.com*
- *Zeph M. C. van Berlo is with University of Amsterdam. E-mail: z.m.c.vanBerlo@uva.nl*
- *Pablo Cesar is with Centrum Wiskunde & Informatica and Delft University of Technology. E-mail: p.s.cesar@cwi.nl*

In this work, we focus on spatial presence ('being there') in VR, where we examine cardiac interoception under a disembodied VR setting (no avatar) to avoid confounds with the type of avatar embodiment. We compare such VR experiences with a screen-based environment where users have full visceral sense of their real body. Disembodied VR experiences (e.g., watching immersive 360 videos [45]) have been shown to impact how users perceive their own body, where users reported being less aware of their bodily state (skin, muscular, and cardiovascular) in comparison with desktop settings [56]. This suggests that if participants are disembodied in VR, they may have reduced interoceptive senses. Moon et al. [52] found increased physiological responses for an imagery running task, and higher reported presence when participants embodied an avatar than those who did not. Given the foregoing, this led us to investigate more closely whether (disembodied) VR experiences similarly influence how we relate to our cardiac signals, and how interoception can be measured in such environments.

While there are different kinds of interoceptive sensations (e.g., gastric, pulmonary, cardiovascular [36]), most research focuses on cardiac interoception, as heartbeats are distinct and frequent internal events, which can more easily be discriminated and measured [26]. Furthermore, non-cardiac measurements can be expensive and invasive to

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2023.3320228

administer [58]. Measurement of cardiac interoception typically spans three factors [32]: 'interoceptive accuracy', defined as the ability to perceive an internal (e.g., one's heartbeat) signal in close correspondence with a physiological measurement of it; 'interoceptive sensibility' which refers to the self-evaluation of interoceptive ability, commonly assessed through interviews or questionnaires; and 'interoceptive awareness', which is a meta-cognitive construct that reflects how well participants' beliefs about their interoceptive ability (i.e, their confidence) is matched by their actual performance on tests of interoceptive accuracy.

Given these, the most common method to date for measuring cardiac interoceptive accuracy (CIAcc) is the Heartbeat Counting Task (HCT) [67], wherein participants are asked to report the number of heartbeats they feel during pre-set time intervals. However, several works have shown that this measure is fraught with methodological problems [23, 71]. Due to this, several related tasks have recently been proposed, such as the Cardiac Recognition Task [32]. Here, heartbeats are paired with visual exteroceptive information for the biofeedback (visual icons), where such integration of interoceptive and exteroceptive information is purported to play a role in interpreting and integrating information about the state of the inner body to facilitate self-regulation [32, 37]. However, the role of visualization modality remains limited: previous work has found that audio heartbeats have the capacity to influence physiological and emotional states [8, 78, 89] and the sensation of effort [53], either on its own or when coupled with other modalities (e.g., haptic feedback). Relatedly, Chen et al. [8] explored multi-sensory heart rate (HR) representations, and found that audio-haptic feedback was most preferred while visual feedback was reported as being distracting. Yet other works found that the audio modality either did not facilitate physiological relaxation [94], or did not influence participants' physiological signals [14]. Given the importance of exteroceptive cues, and the unknown role they play in cardiac interoception and "interoceptive illusions" [69], we sought to compare exteroceptive modalities (audio, visual, or audio-visual) for measuring cardiac interoception across screen-based and VR environments.

Despite the emergence of new cardiac interoception measurement methods, these tasks either rely on expensive measurement apparatus, focus on a single modality, do not scale beyond psychophysiological experiments, or are not immediately extensible to screen-based and VR environments. Measuring cardiac interoceptive ability across devices and task environments not only has diagnostic and wellbeing implications within interoception research, but can shed light on the role of modalities for visualizing and interpreting self- and others' biosignals (cf., [22, 40, 51]) across the reality-virtuality continuum [49]. Thus, in this work, we ask: **(RQ1)** How can we design a multimodal cardiac interoception task that can be used across devices and task environments? We use a Polar H10[1] HR monitor and the Excite-O-Meter framework [63] and introduce a novel **Modality-dependent Cardiac Recognition Task**. This tasks draws on exteroceptive modality representations (Audio, Visual, Audio-Visual) and real-time modification of displayed HR (increase/decrease) to assess cardiac recognition. Our task can be executed across Screen-based and Virtual Reality (VR) environments (Figure 1), which led to our second question: **(RQ2)** What are the effects of HR presentation modality and real-time modification of displayed HR across Screen-based and VR environments on CIAcc, interoceptive awareness, mind-body subjective measures, and VR specific measures of presence? We employed a mixed-factorial design (N=50), with task environment (Screen, VR) as between-subjects factors. We investigated the effects of Modality (Audio, Visual, Audio-Visual) and displayed HR Modification ($\pm$15%, $\pm$30%, Real) on CIAcc (as binomial accuracy and d' [44]), interoceptive awareness (as ordinal confidence ratings), measures of mind and bodily perception (MAIA [47], SMS [79]), cybersickness (SSQ [35]), presence in the VR condition (IPQ [70], SUS [75, 83]), and post-experience interview responses. We furthermore tested whether our task is associated with the classic HCT [67]. Our findings showed that (a) participants across task environments underestimated their own HR up to 30%, which was indistinguishable from their actual HR (b) the task environment (Screen, VR) did not affect CIAcc but influenced mind-related measures (MAIA, SMS-Mind) (c) modality did not influence CIAcc, however including audio increased interoceptive awareness, and (d) VR IPQ General Presence inversely correlated with CIAcc.

Our exploratory work offers two primary contributions: **(1)** We introduce a lightweight and extensible method using a Polar H10 chest strap HR monitor for measuring cardiac interoception (CIAcc and awareness). This can be easily deployed across Unity-based mixed reality environments[2], thus enabling explorations between VR-specific phenomena and cardiac interoception. **(2)** We provide empirically-backed insights and considerations for assessing cardiac interoception in real and virtual environments. This has implications for designing the most suitable HR modality representation, and feeds into research directions that assess how VR-specific phenomena affect our bodily self-perception, from presence to avatar embodiment, to remote medical diagnoses for assessing interoception-related mental health outcomes.

2 Related Work

Several research strands influenced our work, which we describe below.

2.1 Measuring cardiac interoception

Measurement of interoception began flourishing around 45 years ago [67]. Earlier measurement tasks focused primarily on heartbeat perception, given their discrete nature which lends itself to discriminability. These tasks measures cardiac interoceptive accuracy, defined as the ability of an individual to explicitly and accurately identify discrete interoceptive events such as their heartbeats. Whereas some tasks [86] ask participants to discriminate between auditory tones presented either synchronously or asynchronously with their heartbeats, others, like the 'gold standard' Heartbeat Counting Task [67], ask participants to track and report their heartbeats over short periods of time. However, accumulating evidence suggests that this measure suffers from methodological problems, spanning construct validity issues [71], participant over-reporting tendencies, time interval counting instead of heartbeats, to trait-contingent characteristics of participants [23]. Researchers have since aimed to address these shortcomings, with a plethora of cardiac interoception accuracy tasks being proposed. Examples include the Heartbeat Discrimination Task [41], CARdiac Elevation Detection Task [61], and Heartbeat Matching Task [58], all of which differ in their assessment approach, yet with a recurring finding that participants routinely underestimate their heartbeats [6]. Increasing evidence from psychobiology underscores the importance of integration of interoceptive and exteroceptive information in enabling humans to interpret their inner body states to facilitate self-regulation [32, 37]. Related to the present work are Azevedo et al.'s [2] forced-choice paradigm and more recently Hodossy et al.'s [32] Cardiac Recognition Task, both of which pair visual exteroceptive information for the biofeedback, and involve a cardiac recognition task. In Azevedo's forced-choice paradigm, participants discriminated between sounds of their own heartbeat and that of another person, and found that while participants identified their own heart sound above chance, their interoceptive awareness was poorer. In their setup, they prerecorded heartbeats using a Doppler device. In Hodossy et al.'s setup, they paired interoceptive signals (i.e., heartbeats) with visual exteroceptive information for the biofeedback, mapped to a thermometer display icon. Their focus was on whether participants can count, feel, and regulate their heartbeats, when the recognition was pitted against another person's heartbeat.

Whereas these tasks focused on recognition of one's heartbeat with respect to another person's, our task draws on work within immersive virtual environments where participants' own cardiac signals are manipulated through increase/decrease of displayed heartbeats [8, 14]. Importantly however, all these previous tasks either rely on expensive measurement apparatus, focus on a single modality, or are designed specifically for psychophysiological experiments using custom software to assess cardiodynamics. Since they may not be immediately applicable for usage across screen-based and VR environments, these approaches can lack extensibility.

[1]https://www.polar.com/en/sensors/h10-heart-rate-sensor

[2]Source code: https://github.com/cwi-dis/CardioceptionVR

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2023.3320228

2.2 Understanding bodily states in real and disembodied virtual environments

Virtual Reality (VR) experiences have the capacity to influence feelings of presence in the sense of 'being there' (cf., Place Illusion [75]). Such experiences can influence our emotional reactions [16], physiology [88], and foster empathy [68]. Specifically, interoception has been shown to play a vital role in maintaining a coherent sense of self through the malleability of body representations [81], where those who exhibited low CIAcc had a stronger sense of body-ownership over a fake hand in a Rubber Hand Illusion manipulation. Importantly, changes in body ownership have been shown to influence perception of our own cardiac signals [24], where a change in body-ownership significantly improved performance of participants with lower interoceptive accuracy. This can partially be attributed to compressed time perception in VR, which was shown to influence cardiac interoception [48]. This is in line with clinical research on interoception and depersonalization-derealization disorder, where a patient exhibited impaired performance in the heartbeat detection task when compared to controls [72]. In a scoping review, Lüddecke and Felnhofer [42] found that peripheral physiological VR biofeedback can support anxiety, stress, and pain reduction, and can facilitate motivation, user experience, involvement and attentional focus among healthy adults.

In this study, we focus on disembodied VR experiences. Such disembodied experiences (e.g., watching immersive 360 videos [45]) have been shown to impact how users perceive their own body, where they reported being less aware of their bodily state (skin, muscular, and cardiovascular) in comparison with desktop settings [56]. This suggests that if participants are disembodied in VR, they may have reduced interoceptive senses. Moon et al. [52] found increased physiological responses for an imagery running task, and higher reported presence when participants embodied an avatar than those who did not. Moreover, previous work found that embodying an avatar (with varying fidelity and realism across (social) VR platforms) can influence presence [30] and body perception measures [18]. Therefore, to avoid such confounds with the type of avatar embodied, we focus on disembodied VR experiences. This lead us to examine cardiac interoception under a disembodied VR setting (no avatar), and compare it with a screen-based environment where users have full visceral sense of their real body.

2.3 Visualizing heart rate across modalities

Creating exteroceptive cues through multimodal visualization of human biosignals has a long-standing history in HCI, where biosignals are represented across different modalities and forms [43]. For heartbeats specifically, visualizations typically span visual representations, including skeuomorphic [40], text/numerical [27], screen overlays [29], or holographic displays [82]. However, they can also be represented as auditory heartbeats [15] or as haptic sensations [8, 14]. Chen et al. [8] explored multi-sensory heart rate representations, and found that audio-haptic feedback was most preferred while visual feedback was reported as being distracting. Dey et al. [14] found that providing slightly faster or slower real-time heart rate feedback using an audio-haptic representation can alter participants' emotions, but not their physiological signals. Costa et al. [10] found that providing a false feedback of a slow heart rate using two shaftless/coin vibrotactile motors for feedback can help regulate users' anxiety. When comparing audio with vibration-based haptic feedback, Zhou et al. [94] found that while heartbeat vibration influenced users' heart rate variability and aided them to physiologically relax, no effect of heartbeat sounds was observed. Tajadura-Jiménez et al. [78] explored how auditory and/or vibrotactile heartbeat stimuli influence participants' physiological state and subsequent emotional attitude to affective pictures. They found that heartbeat sounds significantly affected participants' heartbeat as well as their emotional judgments of pictures and their recall, when they were presented near the participant (headphone condition). Furthermore, Winters et al. [89] investigated how visual, auditory, or audio-visual heartbeats can elicit empathy, and found that hearing heartbeats changed participants' emotional perspective and increased their empathy self-reports. Recently, Moullec et al. [53] explored whether displaying cardiac activity across sensory modalities can significantly enhance the sensation of effort,

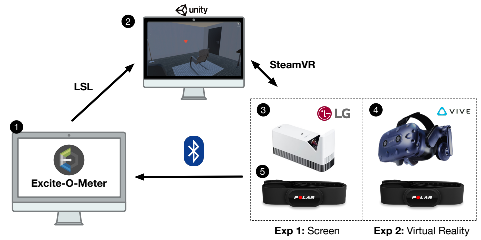


Fig. 2: System overview for both our Screen and VR conditions.

and found that a multimodal (visual, auditory, haptic) display induced more effort sensation than visual- and haptic-only displays, but not with audio-only displays. Similarly, Janssen et al. [34] explored how aurally presented heartbeats can influence how people interpret others' cardiac activity, and found that people relate increases in heart rate to increases in emotional intensity.

Despite the varied work exploring multiple modalities for representing HR and their influence on our emotional and physiological states, it remains unclear how modalities facilitate cardiac interoceptive accuracy as measured with a recognition task. To this end, our study explores how visual and auditory presentations of heartbeats can help us better understand the extent that such exteroceptive cues can aid in recognizing our own heartbeats, and the role that interoceptive technologies play in inducing such interoceptive illusions [69].

3 Methodology

3.1 Hardware and software

3.1.1 System description

Our system overview is shown in Fig. 2. The only component that differed between conditions is (3) projector and (4) VR HMD. System components are: (1) For integrating HR data into Unity in real-time, we used the Excite-o-Meter software framework [63]. (2) A virtual environment was created using the Unity 3D game engine (v.2020.3.32f1 LTS), where the SUS and IPQ questionnaires were created using the open-source VRQuestionnaireToolkit software [21]. All VR-specific hardware was integrated using SteamVR (v.1.22.12) and the XR Management plugin (v.4.2.1). (3) **Screen:** Virtual environment was projected on a 106" projection screen using a full HD (1080p) LG HF85LSR DLP short throw projector. The audio signal was sent to a closed-ear headphone. (4) **VR:** The setup consisted of an HTC Vive Pro Eye HMD and a single VIVE controller. The HMD allows for a 110° field of view with a combined resolution of 2880x1600 and a refresh rate of 90Hz. In parallel, the audio signal was sent to the headset earphones equipped in the HMD. (5) We used the Polar H10[3] heart rate monitor and chest strap to collect Electrocardiogram (ECG) and HR data.

3.1.2 Unity virtual office scene

A Unity virtual office scene was implemented to attain close correspondence between the real office environment (Screen) and its virtual counterpart (VR). Snaps - Office Pack[4] and the HDRP Furniture Pack[5] Unity assets were used for creating the office space room.

3.1.3 Real-time data streaming

Drawing on the Excite-O-Meter framework [63], the relatively low-cost Polar H10 1-lead HR monitor was deemed suitable. It enables real-time HRV analysis by providing raw ECG, HR, RRi with precision of milliseconds. Data acquisition validity tests by Quintero et al. [63] further showed a Mean Square Error of 0 for both HR and ECG signals, where signals exhibited high correlations to BrainProducts LiveAmp[6]

[3]https://www.polar.com/en/sensors/h10-heart-rate-sensor

[4]https://assetstore.unity.com/packages/3d/environments/snaps-prototype-office-137490

[5]https://assetstore.unity.com/packages/3d/props/furniture/hdrp-furniture-pack-153946

[6]https://www.brainproducts.com/solutions/liveamp/

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2023.3320228

(considered a 'gold standard' in research [60]). Recent tests by Schaffarczyk et al. [66] comparing with 12-channel ECG recordings, further confirm the validity of such measurements.

Within Excite-O-Meter, the Devices module acts as a bridge between the sensors and Unity ((1) –> (2) in Fig. 2). It streams raw physiological data via the Lab Streaming Layer (LSL)[7] protocol, a middleware that allows networking, time-synchronization, and real-time access to time-series data. In this setup, cardiac activity signals sent from the Polar H10 were sent at rates of 1Hz for displaying HR, and 130Hz for raw ECG that was used for the Heartbeat Counting Task analysis. End-to-end latency from the Polar H10 to Unity spanned a few milliseconds (given time to calculate HR from raw ECG data).

3.1.4 Modality design

We initially considered creating an individual-based mapping from the ECG signal to the heartbeat sound, by looking at cardiodynamics (cf., [2, 32]). However due to the cost of research-grade sensor data acquisition devices, and complexity of performing this calculation in real-time, we sought alternative, more scalable approaches. This is especially problematic for real-time automated ECG processing, where the ECG signal needs to be filtered / smoothed or even removed if (correctly) detected as a noisy outlier. Moreover, since we were concerned with recognizing overall heart rates, and not the specific individual cardiac cycles at a millisecond precision level, we simplified this step. This entailed a mapping from the initial trial HR to the visual and auditory modalities. Over a given time interval (e.g., three seconds), the HR dynamically updates as soon as the Polar H10 sensor registers an HR change. This is displayed within a few milliseconds in Unity. The audio and visual conditions each included representations of the first (S1) and second (S2) heart sounds. Namely of systole (contraction) "Lub-dub" and diastole (relaxation) "Pause" of the atria and ventricles opening and closing when the heart pumps blood. Systole is commonly referred to as the "lub" and the "dub" sounds (see e.g., [77]), with the pause thereafter for diastole. Illustrative videos using mock up fixed HR of 77 BPM showing each modality across HR modification conditions can be found in Supplementary Material A.

Audio. We drew on fixed audio mappings (cf., [89]) and proceeded to render the heartbeat sound using synthesized notes similar to the sound of a real S1 and S2 using the Mixkit library[8]. The S1 and S2 sounds were pre-set at a constant interval of 0.4 seconds of atrial and ventricular systole for participants' HR, where this interval is based on an average human cardiac cycle of around 0.8 seconds[9]. The combined sounds of S1 and S2 were played to the participant at the beat of the HR value streamed from Excite-O-Meter. Similarly to Winters et al. [89], these sounds were checked for perceptual realism and verified by independent observers.

Visual. This consisted of a pulsating 2D heart icon that changed between a smaller size representing the diastole (no heart contraction), a larger size representing the onset of systole S1, and a medium form for the S2. This heart representation draws on the common iconic visualization inspired by clinical electrocardiogram monitoring devices, based on the RefVis visualization of Gradl et al. [28] and the Skeuomorphic HR design by Lee et al. [40].

Audio-Visual. The audio heartbeat and the visual heart icon were played simultaneously in synchrony, and played as such to participants.

3.2 Study design

Our study is a mixed-factorial design, where the task environment condition (IV1: Task Environment) with two levels (Screen, VR) is a between-subjects factor. In each within-subjects factor, we manipulated displayed HR such that they were either real, or 15% and 30% faster or slower than their actual HR. This resulted in five levels of our second independent variable (IV2: HR Modification). Our last variable (IV3: Modality) had three levels: Audio, Visual, and Audio-Visual. Each (modified) HR was presented across these three modalities. The 15% and 30% rate manipulations were chosen based on speed manipulations from prior work: Suzuki et al. [77] used 30% modifications, Dey et al. [15] used 20%, and Dey et al. [14] used up to 30%. In the latter, they observed that manipulations greater than 30% were too obvious and 15% was a lower threshold that still produced an effect of altering participants' emotional states. We furthermore empirically tested rates of ±10% and ±20%, however these were either not noticeable in the case of 10%, or behaved similarly to the 15% manipulation in the case of 20%. We therefore decided to adhere to ±15% and ±30%. Modality conditions were counterbalanced according to a Latin square design, as well as the initial HR trial, however only between a Real HR versus a Modified HR. Each modality was presented across 15 trials, with each HR Modification (±15%, ±30%, real HR) presented three times. The remainder of trials were subsequently randomized. At the end of the trials, participants underwent a brief exit interview.

3.2.1 Task design

Practice trials. To familiarize participants with the study, they first completed an approximately two-minute practice round. In each trial, participants underwent a 2-Alternative Forced Choice (2AFC) task, and were tasked with recognizing whether or not a pulsating blue circle (Fig. 1(b)) with a corresponding audio beat had a frequency of 40 beats per minute (BPM). After each trial, a panel popped up (see example panel in Fig. 1(d)) asking: "Did the preceding beat match 40 BPM?". Participants could select "Matched" or "Not Matched" with the Vive controllers' trigger button (VR) or mouse click (Screen). After their selection, participants were asked to rate how confident they were in their choice on an 11-point Likert scale embedded within the Unity environment (similar to Fig. 1(d)), where 0 indicates "I was not at all confident" and 10 indicates "I am extremely confident". Previous work used a 10-point scale for such ratings [26, 55], however we added an additional point to allow for greater sensitivity in case of a neutral selection preference. After these questions, the blue circle was presented again. This repeated for 5 trials, after which a panel with a "Start" button appeared to proceed to the main experiment.

Modality-dependent cardiac recognition task. Similar to a 2AFC task, our task aims at measuring how well participants can **recognize** their own HR, amidst manipulated heartbeats. Each HR in a given trial (example shown in Fig. 1(c)) was shown to the participants for seven seconds, where this duration was chosen empirically, with consideration of Azevedo et al.'s [2] and Legrand et al.'s [41] five second stimuli presentation duration, Hodossy et al.'s 10 seconds stimuli duration [32], and Winters et al.'s 20 seconds duration [89]. We did not expect physiological synchrony to occur with such durations, considering Dey et al.'s [14] finding that providing slightly faster or slower real-time HR feedback did not alter participants' physiological signals. Moreover, given the nature of our task where we compute the displayed HR at trial start, and not look at cardiodynamics (cf., [2, 32]), any effects of synchrony would be diminished. An optional two minute break was provided between modality conditions. After the HR display, participants answered the question "Did the preceding beat match your heart rate?" on a virtual panel (Fig. 1(d)). Participants used the mouse cursor (Screen) or Vive controller (VR) to select "Matched" if they believed that the heart rate was theirs, or "Not Matched" otherwise. Participants could make this choice only after the HR was presented. After selection, another panel popped (similar to Fig. 1(d)) in the same position as the previous one, asking "How confident are you in your decision?". As in the practice trials, this was on an 11-point Likert scale, where 0 indicates "I was not at all confident" and 10 indicates "I am extremely confident". Selecting "Continue" moved to the next trial.

3.2.2 Measures

We employed several measures to assess interoception and better understand participants' behavior: (a) Our primary 2AFC Cardiac Interoception Accuracy (CIAcc) measure of whether a displayed HR matched the participant's own HR (cf., [2, 32]). (b) Participants' confidence ratings in their choice on an 11-point Likert scale from 0 (not confident

[7]https://github.com/sccn/labstreaminglayer
[8]https://mixkit.co/free-sound-effects/heartbeat/
[9]https://en.wikipedia.org/wiki/Cardiac_cycle; see also Systole; Diastole

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2023.3320228

at all) to 10 (fully confident). (c) Participants' actual HR and ECG streamed from the Polar H10. (d) Responses (counted heartbeats) to the modified-instructions HCT task by Desmedt et al. [13]. Modified instructions were used to avoid contamination of time estimation strategies and knowledge of one's own HR, so participants were instructed to not count seconds or provide random guesses. (e) 37-item Likert-scale Multidimensional Assessment of Interoceptive Awareness (MAIA, v2) questionnaire [47] (Screen: Cronbach's α=0.8, VR: α=0.63) from 0 (never) to 5 (always). MAIA, commonly used across interoception research ([18]), takes a holistic approach to measuring interoception sensibility (f) 21-item self-report State Mindfulness Scale (SMS) [79] (Screen: Cronbach's α=0.91, VR: α=0.83) from 1 (strongly disagree) to 5 (strongly agree) to quantify subjective levels of attention to one's mental and bodily events. (g) Participant demographics to collect their age, sex, current occupation, any visual or auditory impairments, coffee intake prior to the study, and VR experience for the VR condition (h) Consensually audio-recorded semi-structured interviews asking participants about their overall experience with the task, whether they noticed a pattern in the displayed HR, and what strategies, if any, they employed in helping them determine their HR. Our study followed strict guidelines for ethical treatment and data protection (including COVID-19 regulations), from our institute as well as a secondary institutes' ethics and data protection committee.

VR-specific measures. To evaluate VR experiences, motion sickness and the sense of presence are two widely considered human factors [5, 7]. We chose a standardized Simulator Sickness Questionnaire (SSQ) [35] to measure the level of motion sickness on a scale from 1 (none) to 4 (severe) before (Cronbach's α=0.9) and after (Cronbach's α=0.87) the VR experience. The 14-item Likert scale Igroup Presence Questionnaire (IPQ) [70] (Cronbach's α=0.6) on a scale from 1 (fully disagree) to 7 (totally agree) was used to assess the level of presence experienced, where of specific interest was the General Presence factor of the 'sense of being there' [73, 74]. As an additional measure of presence (Place Illusion), we gave participants the Slater, Usoh, and Steed (SUS) [75, 83] 6-item Likert-scale questionnaire (Cronbach's α=0.91) from 1 (not at all) to 7 (always)[10].

3.3 Study procedure

Screen. Our Screen condition procedure is shown in Fig 3. Each session lasted approximately 60 min. Before the study, participants read and signed the data privacy, protection, and consent forms. Room lighting was kept dim to ensure maximum visibility of the projection. Participants were then asked to wear the Polar H10 HR monitor around their chest, thereafter ensuring the sensor is sufficiently wet. While the HR sensor was syncing with Excite-o-Meter, participants filled out a demographics questionnaire. Once a connection was established, participants were seated on a chair (see Fig. 1(a)) at 155 cm from the short-throw projector screen, with a mouse pad and mouse placed near their dominant hand. For this condition, the chair Unity asset was removed to allow a sensible first-person view. Participants were seated to ensure we measured resting heart rates with minimal activity-related spikes (cf., [82]). Study instructions were then explained, with a brief tutorial on the controls (namely response selection). Participants were told that their task is to recognize their own HR, however without mention of the modifications. They were asked to assess this without physically monitoring their pulse, such as touching their wrist, neck, or chest. They then entered the practice trials, and thereafter the main recognition task, which lasted approximately 18 minutes. Throughout our experiment, in case the signal between Polar H10 and Excite-o-Meter was lost (mainly to due to low sensor wetness), the experiment was paused, a wet towel applied to the HR sensor, and thereafter resumed from the last trial. This occurred mainly in our pilot tests, which were excluded from later analysis. After the main experiment, participants filled in the MAIA and SMS questionnaires, and then proceeded to the exit interview. After the interview, participants were given the HCT, and rewarded with a €10 voucher for their participation.

[10]Some items had a semantic differential response, e.g., 1 (being elsewhere) to 7 (being in the virtual environment).

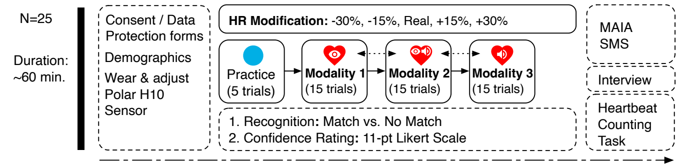


Fig. 3: Screen condition procedure.

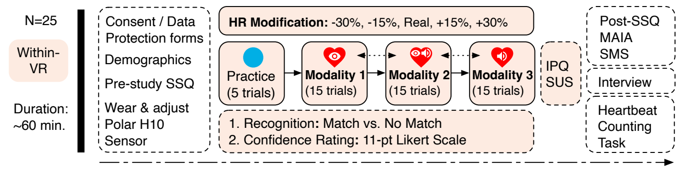


Fig. 4: VR condition procedure.

VR. Our VR condition procedure is shown in Fig. 4. Each session lasted approximately 60 min. The procedure was nearly identical to the Screen condition, with a few key differences. Participants had to fill in a pre-study SSQ prior to the experiment, and a post-study SSQ after the experiment session. Participants were equipped with a Vive Pro Eye HMD and remained seated for the entire duration of the VR portion (see Fig. 1(a)). The virtual office environment they were in resembled the physical room they were in, with the virtual chair position coinciding with the physical one. To avoid any confounds with the type of avatar embodied (which can influence presence [30] and body perception measures [18]), participants underwent a disembodied (no avatar first-person view) VR experience. Aside from an overview of the study and tasks, we also provided a brief tutorial on how to navigate (via head movement) within VR, and selecting questionnaire responses through ray casting with the VIVE controller. After the trials, participants filled in the SUS and IPQ questionnaires within-VR, which allows users to stay closer to the context of an ongoing exposure than outside of VR [62]. As in the Screen condition, participants could also take an optional break between modality conditions. The rest of the procedure was identical to the Screen condition, where participants were also rewarded with a €10 voucher for their participation.

3.3.1 Participants

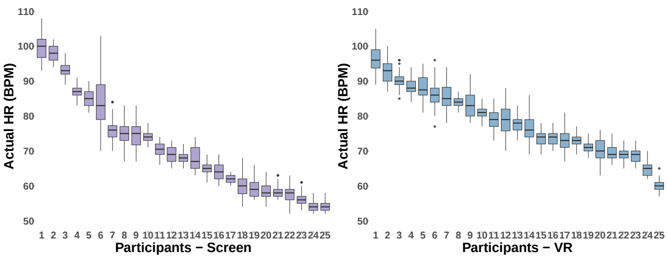
58 participants were initially recruited[11], however eight were excluded due to either sensing, recording, or logging errors. No specific requirements were set for recruiting participants for the VR condition, nor condition assignment. Participant details are described below:

Screen. 25 participants (15 m, 8 f, 2 undisclosed) aged 23-65 ($Md = 29, IQR = 15$) were recruited, primarily (though not exclusively) from our and neighboring institutes, and spanned varied occupations, most of whom were full-time employed. 19 reported having drank coffee prior to the study, although this was not deemed an exclusion measure given HR modifications were based on the current HR. None reported non-corrected visual (including testing for color blindness [33]), auditory or motor impairments.

VR. 25 other participants (16 f, 8 m, 1 undisclosed) aged 19-62 ($Md = 25, IQR = 3$) were recruited. Similarly to the Screen condition, they spanned varied occupations, with mostly students at the undergraduate and graduate level. 14 reported having drank coffee prior to participating. All but four had experience with VR, where five reported to be experienced. None reported non-corrected visual (including color blindness [33]), auditory or motor impairments.

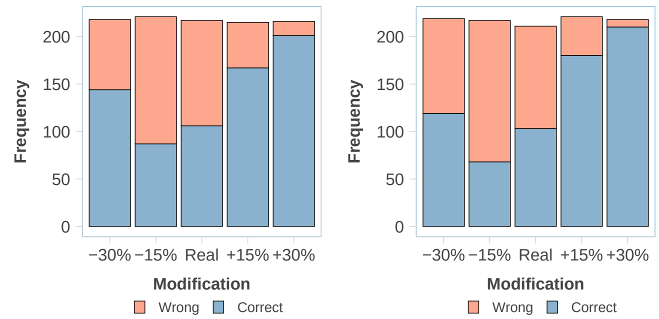
[11]Using G* Power, for effect size f=0.35 under α = 0.05 and power (1-β) = 0.92, with 90 repeated measurements within factors, and two conditions (Screen, VR) as between-subjects factors, one would need 50 participants for a mixed analysis. For analysis per within-subjects conditions (Screen or VR), one would need 10 participants given 45 repeated measurements.

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2023.3320228



(a) Screen condition. (b) VR condition.

Fig. 5: Ordered boxplots for each condition showing participants' actual HR (BPM). Participant numbers are shuffled to preserve anonymity.



(a) Screen condition. (b) VR condition.

Fig. 6: Binomial accuracy frequency distributions for each HR modification level across both conditions.

3.4 Analysis approach

3.4.1 Cardiac Interoceptive Accuracy: binomial accuracy, d'

For completeness, we treat Cardiac Interoceptive Accuracy (CIAcc) in two ways: (a) as binomial accuracy, where 1 represents match and 0 a mismatch. This is useful for initial established analyses, which allow inspecting all our IVs together per between-subjects factor, since HR Modification can itself be part of the DV under other metrics (b) as a sensitivity metric d' (d prime) used commonly in signal detection theory [44] to lower response bias given uneven trials, and previously used in interoception research [32, 37]. The formula for d' is: $z(H) - z(F)$, where $z(H)$ and $z(F)$ are the z transforms of Hit Rate ($P(Match|Real)$ and False Alarm $FA = P(Match|Fake)$, respectively. The resulting d' represents the distance between the signal (Hit Rate) and noise (False Alarm Rate), where a larger value of d' indicates greater sensitivity. For analysis that draws on d‘, HR Modification is already factored in the metric due to the varying fake HR conditions (i.e., ±15%, ±30%).

3.4.2 Interoceptive awareness: ordinal confidence ratings

We analyzed confidence as ordinal scores, given their ordinal nature and to allow for established aligned-rank transform tests [90].

3.4.3 Heartbeat Counting Task

Aside from the modified instructions to participants to lower contamination from other possibly non-interoceptive strategies [13], we used the classic formula from Schandry [67] to compute HCT scores:

$$\text{HCT Score} = \frac{1}{3}\sum\left(1 - \frac{|N_{Actual} - N_{Reported}|}{N_{Actual}}\right) \quad (1)$$

While newer methods have been developed, the differences in outcome are minuscule [93] and usage less common, so we kept the original formulation. For analysis of raw ECG data to compare with the HCT scores, we used the ‘heartPy’ library [84] for signal processing and RR interval detection, with further manual inspections that peaks were either not missed or incorrectly detected.

4 Results

4.1 Heart rate data validity

We first investigate the validity of the actual HR data calculated from the Polar H10 ECG stream. We detected and removed any HR outliers, by using the Interquartile Range (IQR) method. We removed any values that are 1.6 times the IQR greater than the third quartile or 1.6 times the IQR less than the first quartile. We chose 1.6 instead of the common 1.5 to ensure that slightly fluctuating heartbeats (which were manually inspected) are not falsely mistaken throuh faulty sensor readings. This resulted in removal of 38 records for the Screen condition (total size = 1,087 trials), and 39 for the VR condition (total size = 1,086 trials). We show the actual HR per participant[12] as ordered boxplots across both conditions (Screen, VR) in Fig. 5.

[12] Participant identifiers have been randomly shuffled for data privacy.

4.2 Practice trial performance

We first inspected how well participants performed in the five practice trials, across both conditions. Mean accuracy was at chance level for Screen (M=0.53,SD=0.50) and similarly for VR (M=0.51,SD=0.50). This was expected, as the modification rates for this task were set at ±10%, where this task only served to get participants familiarized with the task environment and controls.

4.3 Motion sickness in VR (SSQ)

A Wilcoxon signed-rank test (suitable for non-normal data) showed no significant differences between pre-study (Md=1.13,IQR=0.25) and post-study SSQ (Md=1.19,IQR=0.25) scores (Z=-1.8,p=0.08).

4.4 Effects of HR modification and modality on binomial CIAcc

Binomial accuracy frequency distributions across HR Modification for Screen and VR are shown in Fig. 6a and Fig. 6b, respectively. Binomial accuracy across participants for Modality and Task Environment are additionally shown as boxplots in Fig. 7a and Fig. 7b, respectively. Dashed line shows the mean of accuracy scores across modalities.

Screen. We first investigate separately to what extent Modality and HR Modification influence CIAcc in the Screen condition. Given CIAcc here is binomial, we run Cochran's Q tests for each of our IVs. For Modality, an asymptotic general symmetry test showed no significant differences in CIAcc among the three tested modalities ($\chi^2(2)$ = 1.99, p=0.37). For HR Modification, an asymptotic general symmetry test showed significant differences in recognition accuracy among the five tested modification rates ($\chi^2(4)$ = 38.3, p<0.001). A pairwise comparison using continuity-corrected McNemar's tests with Bonferroni correction revealed that participants scored significantly worse in -30% than 15% (p<0.001, ϕ=0.32), worse in 15% than 30% (p<0.001, ϕ=0.53), worse in Real than 30% (p<0.001, ϕ=0.24), worse in -15% than 15% (p<0.01, ϕ=0.15), worse in -15% than 30% (p<0.001, ϕ=0.33), worse in -30% than 30% (p<0.001, ϕ=0.48), worse in Real than 15% (p<0.001, ϕ=0.16), and no differences between -15% and -30%, -15% and Real, nor between -30% and Real conditions.

VR. Similarly to the Screen condition, we run Cochran's Q tests for each of our IVs. For Modality, applying an asymptotic general symmetry test showed no significant differences in binomial CIAcc among the three tested modalities ($\chi^2(2)$ = 1.12, p=0.57). For HR Modification, applying an asymptotic general symmetry test showed significant differences in CIAcc among the five tested modification rates ($\chi^2(4)$ = 64.03, p<0.001). A pairwise comparison using continuity-corrected McNemar's tests with Bonferroni correction revealed that participants scored significantly worse in -15% than -30% (p<0.05, ϕ=0.11), worse in -30% than 15% (p<0.001, ϕ=0.29), worse in 15% than 30% (p<0.001, ϕ=0.59), worse in Real than 30% (p<0.001, ϕ=0.27), worse in -15% than 15% (p<0.01, ϕ=0.12), worse in -15% than 30% (p<0.001, ϕ=0.32), worse in -15% than Real (p<0.01, ϕ=0.14), worse in -30%

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2023.3320228

Factor	F	df	p
HR Modification	59.64	4	< .001**
Modality	10.34	2	< .001**
Task Environment	42.41	1	< .001**
HR Modification x Modality	0.36	8	0.94
HR Modification x Task Environment	0.66	4	0.62
Modality x Task Environment	1.20	2	0.3
HR Modification x Modality x Task Environment	0.41	8	0.91

Table 1: Analysis of deviance on the full mixed-effects model for ordinal confidence ratings using Aligned Rank Transformed data.

than 30% (p<0.001, ϕ=0.46), worse in Real than 15% (p<0.001, ϕ=0.2), and no significant differences between -30% and Real.

4.5 Effects of task environment and modality on CIAcc (d')

To assess if there were any significant effects of Task Environment (Screen, VR) and Modality (Audio, Visual, Audio-Visual) on CIAcc (d'), we ran a two-way mixed-design ANOVA with Task Environment as between-subjects factor, and Modality as within-subjects. Recall that since we use d' as a measure, HR Modification is already factored in our dependent variable. All assumptions (normality, homogeneity of variances and covariances, ensuring no extreme outliers, sphericity) were met. We found no significant effect of Screen (M=0.42,SD=0.9) and VR (M=0.52,SD=0.51) Task Environment (F(1,48) = 0.24, p=0.63) nor of Modality (F(2,96) = 0.34, p=0.71) on CIAcc (d'), nor an interaction effect (F(2,96) = 1.1, p=0.34). The findings for Modality are in line with the Cochran Q tests run on binomial CIAcc earlier (see 4.4), where an additional Chi-squared test with Yates' continuity correction further confirmed no significant differences between Task Environment conditions on binomial CIAcc ($\chi^2(1)$ = 1.1, p = 0.3).

4.6 Effects of modality on ordinal confidence ratings

Confidence ratings across participants for each Modality for both Screen (N=25) and VR (N=25) conditions are shown as boxplots in Fig. 7c and Fig. 7d, respectively. Dashed line shows the mean of confidence ratings across modalities, and lines with asterisks indicate pairwise (Bonferroni corrected) significance. Fig. 6 shows descriptive statistics for HR Modification conditions. We analyzed the combined effects of HR Modification, Modality, and Task Environment on confidence ratings, by fitting a full linear mixed-effects model on our Confidence Ratings data. Since our data distribution is not normal, we applied the aligned rank transform prior to fitting [90]. Post-hoc contrast tests were performed using the Aligned Rank Transform tool ART-C [20].

Analysis of deviance table is shown in Table 1. A full mixed-effects model showed significance for all main effects: HR Modification ($p < 0.001$), Modality ($p < 0.001$), and Task Environment ($p < 0.001$). No significant interaction effects were found. Contrast tests for the main effect of Task Environment revealed significant differences (β=157.5, $p < 0.01$) between Screen (M=6.3,SD=2.4) and VR (M=5.7,SD=2.3). Contrast tests for the main effect of HR Modification revealed significant differences between all levels ($p < 0.001$), except between -15% and Real ($p = .14$), and between -30% and Real ($p = .24$). This indicates that participants' confidence ratings did not vary accordingly between their real HR, and manipulations of up to -30% slower. Contrast tests for the main effect of Modality revealed significant differences in Confidence Ratings between Audio and Visual (β=70.83, $p < 0.001$), and Audio-Visual and Visual (β=29.75, $p < 0.05$), as well as between Audio and Audio-Visual (β=-64.42, $p < 0.05$). This highlights that including audio gives participants' higher confidence that their performance is better than relying on a visual exteroceptive modality alone. Full contrasts tests tables are shown in Supplementary Material B.

4.7 Relationship between HCT and CIAcc (d')

We assessed whether there was a relationship between HCT scores and CIAcc (d') across the Screen and VR conditions. We found no significant correlation between HCT and Screen (ρ=-0.18, p=0.38) nor between HCT and VR (ρ=0.03, p=0.88). As an additional measure, we removed participants whose performance on the HCT fell below 35%. This resulted in the removal of five participants in the Screen condition, and three in the VR condition. However the correlations both remained low and statistically not significant.

4.8 Presence, mindfulness, and interoceptive awareness

4.8.1 Presence

To assess if the VR condition (N=25) had any unique effect (namely presence) on participants, we inspected the IPQ and SUS questionnaire responses. Sum of scores (normalized where appropriate) across participants for the SUS and each factor of IPQ are shown as boxplots in Fig. 8a. To explore the relationship between these constructs and interoceptive accuracy (d'), we ran Spearman correlation analyses to allow assessment of monotonic relationships. Since we conduct multiple correlation comparisons however, our results may be prone to a higher number of false positives (Type I errors) [76]. Using Bonferroni adjustment however is too conservative: while it lowers Type I errors, it can also increase Type II errors [4]. Therefore, we adjusted our tests using the False Discovery Rate (FDR) [3] method, which has been shown to be more balanced and follows prior work within HCI [91]. We follow recommendations from psychological research [1] for determining correlation coefficient strength: (low: $0.1 < |corr| < 0.3$; moderate; $0.3 < |corr| < 0.6$; high: $0.6 < |corr| < 1.0$). Our correlation results are shown in Fig. 8b, where significant correlations are highlighted by asterisks. Many of the IPQ constructs exhibit significant correlations to each other but also to SUS (p<0.01), which highlights that they are indeed measuring similar factors. Surprisingly however, we find that the IPQ-G ("sense of being there") construct exhibits a significant moderate inverse correlation (p<0.05) with accuracy (d') scores.

4.8.2 Mindfulness and interoceptive awareness measures across task environments

MAIA. To assess differences between Screen and VR MAIA responses, we computed a Wilcoxon rank sum test on the independent samples of participant-aggregated total MAIA scores (cf., [18]. This was deemed appropriate given ordinal ratings, which showed a non-normal distribution. We find that scores for Screen (Md=32.4,IQR=4.1) and VR (Md=29,IQR=4.15) significantly differed ($Z = 2.25$, $p < 0.05$, $r = 0.32$).

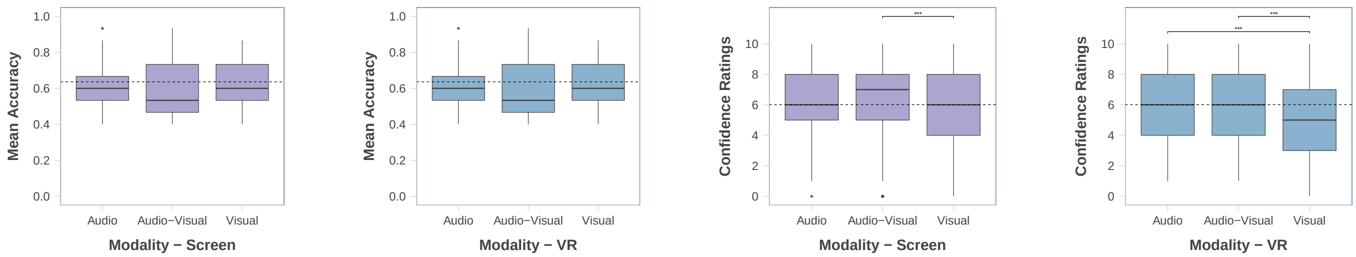
SMS. Similarly to above, we assessed if there were differences between Screen and VR SMS responses, split between SMS-Mind and SMS-Body given they are separate sub-scales [79]. A Wilcoxon rank sum test between Screen (Md=5,IQR=10) and VR (Md=48,IQR=10) on SMS-Mind showed significant differences ($Z = 2.4$, $p < 0.05$, $r = 0.34$), whereas no significant differences between Screen (Md=19,IQR=5) and VR (Md=17,IQR=5) were found for SMS-Body ($Z = 1.5$, $p = 0.13$).

4.9 Participant interviews

We analyzed the interviews with inductive thematic analysis [9]. First we coded it according to evoked topics. Then within each topic we analyzed emerging themes: task difficulty, modality preference, perceived patterns and strategies for determining HR, and impact of VR.

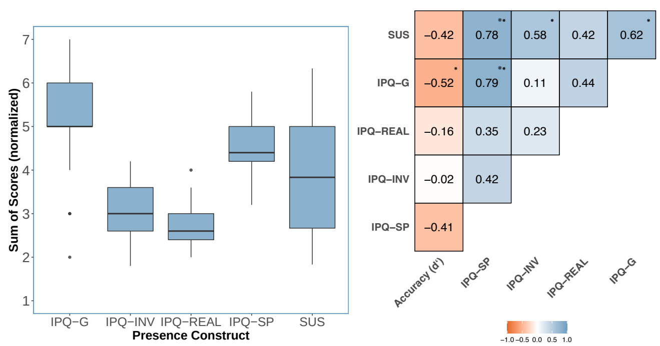
Screen: Most participants (19) found determining their own HR to be very difficult (P10: *"Oh, no, I'm horrible at this. Very, very hard with the exceptions of when it was very, very fast and very, very slow, I would say I have absolutely no clue."*). Most participants (14) stated they did not notice any pattern in the displayed HR (P5: *"I did not notice a pattern in the heart rate itself."*), a few (3) stated they noticed a pattern within a block of trials (P8: *"thought that when we're nearing the end of the trial, it's increasing"*), whereas the rest (4) stated they noticed a pattern between blocks (P15: *"I had the feeling that the starting heart rate is almost always the same in each of the individual cycles"*). Regarding strategies on determining their HR, most participants (15) stated they tried to sense their HR within their body (P8: *"I was just like feeling it pulsing in my heart"*), whereas others (5) used their breathing as an aid (P9: *"I tried to breathe in and out and to focus again while doing that"*), with the rest (4) relying on prior knowledge and counting (P4: *"I just noticed that some are a lot faster than others. Those ones I knew weren't mine"*). Participants mostly found the audio only (15) condition easiest (P16: *"Well perhaps because you sometimes can hear your own heartbeat so it's more*

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2023.3320228



(a) Accuracy - Screen. (b) Accuracy - VR. (c) Confidence Ratings - Screen. (d) Confidence Ratings - VR.

Fig. 7: Boxplots showing effects of Modality (Audio, Visual, Audio-Visual) and Task Environment (Screen, VR) on (a-b) mean HR recognition binomial accuracy, and (c-d) ordinal confidence ratings. Dashed line indicates mean score.



(a) Presence constructs (IPQ and SUS). (b) Correlations with CIAcc (d').

Fig. 8: Presence construct analyses showing (a) Normalized sum of scores boxplot for each presence construct (IPQ and SUS) (b) Spearman correlations with presence constructs (** p<0.01, * p<0.05).

in your mind how you experience it."), followed by audio-visual (5) modality (P15: *"...both [Audio-Visual] was the easiest. I felt more immersed with the sound and the visual because also it resonates more with you and your body and your perception..."*, and visual (5) modality (P9: *"It was easier when there was no sound because I could kind of hear my heart rate in a weird way."*).

VR: Here again (19), most participants found cardiac interoception to be very difficult (P8: *"I think I was basing my answers on my intuition, because it was pretty difficult. I was trying to listen to my heartbeat, like for my body in a way, but it was almost impossible."*). Most participants (16) stated they did not notice any pattern in the displayed HR (P25: *"I thought the faster and slower heartbeats were random or were mixed"*), a few (6) stated they noticed a pattern within a block of trials (P6: *"It's a little bit random. But it's always included like, something is very very fast"*), whereas the rest (3) stated they noticed a pattern between blocks (P10: *"I had the idea that in the last round I had less heartbeats that I felt matched like that."*). Regarding strategies on determining their HR, here as well most participants (14) stated they tried to sense their HR within their body (P15: *"I like just focused my attention inside and then on my chest"*), whereas others (7) used their breathing as an aid (P20: *"I was trying to breathe and trying to feel the sensation of my body, if my beat rate was kind of accelerated or calm."*), with the rest (4) relying on prior knowledge and counting (P17: *"Yeah, it was obvious for some, like too fast or too slow heart rate. I felt like those are not my heart rate. And then some of the middle ones."*). In this condition, participants also found the audio only condition (14) easiest (P22: *"...audio was the easiest because there was no distraction from the visual component."*), followed by audio-visual (9) modality (P26: *"I would say the combination, but that was mostly down to the audio."*), and visual (2) only (P25: *"Yeah because the visual and the sound [Audio-Visual] both like drowned out my own internal feelings...when there's no sounds I can better guess my own heart rate than when there is."*). Lastly, while some participants indicated that being in VR helped them focus (P12: *"Well maybe because I didn't have any other stimulations from the environment that made it a bit easier..."*), a few reported a diminished sense of bodily awareness (P24: *"...something I think a little bit different from what I can feel in the reality world."*; P29: *"I felt like it was harder to feel my heartbeat in VR than like sitting here right now."*; P31: *"So yeah, I felt like it was just a general feeling of being less aware of my heartbeat and maybe like it was I don't know, being influenced by this."*).

5 Discussion

5.1 Study limitations

First, while assessing the validity of the HCT was not in scope of this work, we caution hasty interpretations of our HCT findings. This is because we relied on the raw ECG data from the Polar H10 for a higher temporal resolution analysis. Prior work has shown that HCT measures can differ widely depending on the measurement device used [54], where we expect that data acquisition quality using research and medical grade pulse oximeters [41] or ECG devices [32] are higher than a single-lead Polar H10 chest strap device. Second, given our lightweight, extensible method, we used fixed modality mappings from the (trial start) HR signal to the exteroceptive modality designs (Audio, Visual), even though the streamed HR dynamically updated the visualization. For our purposes, this was sufficient for displaying HR at 1Hz, which allowed us to understand how users perceive their own cardiac signals across different task environments. However, this is less precise than higher sampling rates using continuous measurement Doppler devices [2] or continuous treatment of the raw ECG signal using real-time systolic peak detection algorithms and post-processing techniques (interpolation, filtering) [41]. The latter however are not immediately applicable for cross-environment interaction settings, due to more expensive and less scalable setups than ours. If the aim is to better understand human interoceptive processing by examining fine-grained cardiodynamics (e.g., in biological psychology research), then one would need to factor in such within-trial measurement precision.

Lastly, our VR setup may not have been immersive enough, whereby ensuring users experience both the Plausibility and the Place Illusion [73, 74] can lead to more realistic behaviors (cf., VR learning scenarios [50]). Since there were no other virtual characters, with no action towards and response back from VR elements [25], our VR environment was low on immersiveness (cf., INV in Fig. 8a). In this regard, our work leaves open to what extent our developed task, in measuring CIAcc and interoceptive awareness, can influence our sense of embodiment should users be represented as avatars in a virtual environment (cf., [18, 85]). While a more immersive setup is a logical next step, in this work we wanted to firstly develop a task that can be used across environments (from screen-based to VR), and to test this in environments that are as close to each other as possible. This allowed us to focus on the role of modality representations, with a specific focus on self-reported 'being there' presence in the disembodied VR condition.

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2023.3320228

5.2 Towards cardiac interoception measurement across mixed reality environments

Given the crucial role that interoception plays in regulating our physical and mental health [36], it is unsurprising that several methods have been proposed to circumvent the HCT limitations [71]) for measuring CIAcc. This primarily stems from the private nature of interoceptive sensations, which continue to elude measurement by researchers and clinicians. One goal of this work was to provide a lightweight and non-invasive means of measuring cardiac interoception, by drawing on accumulating evidence that exteroceptive information [32, 37, 77] can aid interoceptive processes. Despite the emergence of new cardiac interoception assessment methods, these tasks either rely on expensive measurement apparatus, focus on a single modality, are designed specifically for psychophysiological experiments, and thus may not be immediately adapted for usage across screen-based and VR environments. To this end, we developed and tested a novel modality-dependent cardiac recognition task **(RQ1)**. This task allowed us to examine more closely, through quantitative (Sec 4.6) as well as qualitative (Sec 4.9) analysis, to what extent representing users' cardiac signals using audio and visual modalities can support them in recognizing their own heartbeats. Moreover, our task is extensible and can be easily deployed across Unity-based mixed reality environments. This invites future HCI explorations between VR-specific phenomena and cardiac interoception, where it warrants further investigation how starting with a biofeedback training session or participants' utilizing deep breathing techniques can influence performance. Furthermore, it invites further exploration on how multimodal biofeedback displays can support human interoceptive processes, and for creating interoceptive illusions [69].

5.3 Exteroceptive modalities and task environment did not influence heart rate underestimates

We investigated the effects of HR presentation modality and real-time modification of displayed HR across Screen-based and VR environments **(RQ2)** on CIAcc, interoceptive awareness, mind-body subjective measures, and VR specific measures of presence. For CIAcc, we found no statistical difference between the screen-based and VR conditions, however ordinal confidence ratings were higher for the Screen-based condition (Fig. 7c,7d). On the one hand, we found our results surprising given that participants confused their actual displayed HR with underestimates of up to 30%, irrespective of the role of modality and task environment condition. Moreover, we cannot attribute a potential lower performance stemming from exteroceptive information if we consider that interoceptive processes do not compete with visual and auditory attention (cf., Wicken's Multiple Resource Theory [87]). On the other hand, earlier work by Schandry [67] found that participants underestimated their HR by 26%, Ring et al. [64] found participants underestimated their HR by 37% while sitting, and more recently Legrand et al. [41] using a HR discrimination task found underestimates by around 7 BPM on average (based on psychometric parameter estimates). However in all these tasks, they used either the HCT or a psychophysical variant, which involved counting or discrimination, and not cardiac recognition with the support of modalities. Despite that Azevedo et al. [2] report above-chance recognition accuracy of one's own heart sounds, their comparisons were recorded offline and with another person's heartbeat, and included a wider cardiovascular fingerprint that is better captured using a Doppler device. Moreover, such high recognition was found only for the high interoceptor group. To this end, cardiac underestimation should be taken as a starting point for designing such displays, whether for understanding body perceptions [31], or relaxation [92] and meditative [12] training. Furthermore, while the form of mediation (modality) did not influence participants' CIAcc ability, our confidence rating analyses and participants' audio-only preferences in their interview responses (Sec 4.9) highlight the important role of audio-based HR visualization on cardiac recognition performance. This parallels previous findings on the superior role of audio heartbeats, in their capacity to influence physiological and emotional states [8, 78, 89] and the sensation of effort [53], even should they be coupled with other modalities such as vibrotactile [78] or thermal feedback [19]. Given this, our findings suggest that an audio-based HR representation, whether presented solely or in combination with other modalities, can support cardiac interoceptive awareness.

5.4 Why does the Place Illusion in VR inversely correlate with cardiac interoceptive accuracy?

We revisit our finding that reported presence ('being there') [73, 74] inversely correlated with CIAcc (see Fig. 8b). If we tread less cautiously, while IPQ-GP exhibited a significant inverse correlation, both SUS and IPQ-SP still showed relatively high inverse correlations (even if not statistically significant). Why might this be so? One explanation is that since participants were not embodied in VR, they would be subjected to conditions where their bodily signals can become less pronounced since they suddenly lack a body (their sense of embodiment diminishes). In this case, those who reported truly being in another place could have been caught up in the place illusion at the expense of sensitivity towards their bodily signals. Indeed, prior work by Murray et al. [56] found that such VR experiences can affect body perceptions, where participants reported being less attuned to their bodily signals when compared with non-VR settings. This is further supported by Moon et al.'s [52] study that found increased physiological responses for an imagery running task, with higher reported presence only by participants who embodied an avatar. This may not necessarily rule out other explanations where this can be due to a modification in how time is perceived in VR [48]. Furthermore, we can consider that earlier work has shown that interoception plays a vital role in maintaining a sense of self through the malleability of body representations [81], where this suggests the importance of inhabiting an avatar when we are transported to immersive virtual spaces. However, the inverse could also be true, where better interoceptors are better at suspending their belief in the place illusion (see Fig. 8b). The latter gains credence if we consider that Döllinger et al. [17] did not find significant relationships between the sense of embodiment in VR and interoceptive accuracy. However, their measure of accuracy was based on the HCT, which itself requires further investigation. Lastly, if we factor in participants' scores on MAIA and SMS, while we find that being in VR significantly affects subjective measures of interoceptive awareness and mindfulness, this seems to be confined to measures of mind and attention (cf., MAIA's assessment [46]) rather than bodily measures (given that SMS-Body did not show significant differences). This suggests that if and whatever distortions in cardiac interoception there were (as measured with our task), they pertained largely to mindfulness dispositions and cognitive perceptions of one's self. This would have implications for facilitating mindfulness practices and training in XR environments (cf., [17]), despite users not being embodied as avatars.

6 Conclusion

This work aimed at addressing how we might measure cardiac interoception across screen-based and VR environments using a Polar H10 HR monitor. This was done by devising a modality-dependent cardiac recognition task that modifies displayed heartbeats (±15%, ±30%, None) and using different exteroceptive modalities (Audio, Visual, Audio-Visual) for visualizing HR. In a mixed-factorial design (N=50) with Screen and VR as between-subjects factors, we looked at how HR modality and the displayed real-time HR modifications influence cardiac interoceptive accuracy, interoceptive awareness, mind-body measures, and VR presence. We found that participants routinely confused their actual HR with underestimates up to 30%; the task environment did not affect CIAcc but influenced mind-related subjective measures; the choice of modality did not influence CIAcc, however there was evidence that audio increased interoceptive awareness; and IPQ General Presence in VR inversely correlated with CIAcc. Our work contributes a lightweight and extensible cross-environment method for measuring cardiac interoception (`https://github.com/cwi-dis/CardioceptionVR`), and empirically-backed insights and considerations for assessing cardiac interoception in real and virtual environments. This furthers opportunities for HCI research to support cardiac interoception research through designing multimodal biofeedback across mixed-reality environments.

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2023.3320228

References

[1] H. Akoglu. User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine*, 18(3), 2018. doi: 10.1016/j.tjem.2018.08.001 7

[2] R. T. Azevedo, S. M. Aglioti, and B. Lenggenhager. Participants' above-chance recognition of own-heart sound combined with poor metacognitive awareness suggests implicit knowledge of own heart cardiodynamics. *Sci. Rep.*, 6:26545, May 2016. 2, 4, 8, 9

[3] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B: Methodological*, 57(1):289–300, 1995. 7

[4] Y. Benjamini and D. Yekutieli. The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.*, 29, 08 2001. 7

[5] M. Bessa, M. Melo, D. Narciso, L. Barbosa, and J. Vasconcelos-Raposo. Does 3d 360 video enhance user's vr experience? an evaluation study. In *Proc. Int. Conf. on HCI*. ACM, NY, USA, 2016. 5

[6] J. Brener and C. Ring. Towards a psychophysics of interoceptive processes: the measurement of heartbeat detection. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 371(1708):20160015, Nov. 2016. 2

[7] E. Chang, H. T. Kim, and B. Yoo. Virtual reality sickness: A review of causes and measurements. *IJHCI*, 36(17):1658–1682, 2020. doi: 10.1080/10447318.2020.1778351 5

[8] H. Chen, A. Dey, M. Billinghurst, and R. W. Lindeman. Exploring the design space for multi-sensory heart rate feedback in immersive virtual reality. In *Proc. OZCHI '17*, pp. 108–116. ACM, NY, USA, 2017. 2, 3, 9

[9] V. Clarke, V. Braun, and N. Hayfield. Thematic analysis. *Qualitative psychology: A practical guide to research methods*, 222(2015). 7

[10] J. Costa, A. T. Adams, M. F. Jung, F. Guimbretière, and T. Choudhury. Emotioncheck: Leveraging bodily signals and false feedback to regulate our emotions. In *Proc. UbiComp '16*, pp. 758–769. ACM, USA, 2016. 3

[11] A. D. Craig. How do you feel? interoception: the sense of the physiological condition of the body. *Nat. Rev. Neurosci.*, 3(8):655–666. 1

[12] C. Daudén Roquet and C. Sas. Interoceptive interaction: An embodied metaphor inspired approach to designing for meditation. In *Proc. CHI '21*. ACM, NY, USA, 2021. doi: 10.1145/3411764.3445137 9

[13] O. Desmedt, O. Corneille, O. Luminet, J. Murphy, G. Bird, and P. Maurage. Contribution of time estimation and knowledge to heartbeat counting task performance under original and adapted instructions. *Biological Psychology*, 154, 2020. doi: 10.1016/j.biopsycho.2020.107904 5, 6

[14] A. Dey, H. Chen, M. Billinghurst, and R. W. Lindeman. Effects of manipulating physiological feedback in immersive virtual environments. In *Proc. CHI PLAY '18*, pp. 101–111. ACM, NY, USA, 2018. 2, 3, 4

[15] A. Dey, H. Chen, A. Hayati, M. Billinghurst, and R. W. Lindeman. Sharing manipulated heart rate feedback in collaborative virtual environments. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 248–257, 2019. doi: 10.1109/ISMAR.2019.00022 3, 4

[16] J. Diemer, G. W. Alpers, H. M. Peperkorn, Y. Shiban, and A. Mühlberger. The impact of perception and presence on emotional reactions: a review of research in virtual reality. *Front. Psychol.*, 6:26, Jan. 2015. 1, 3

[17] N. Döllinger, C. Wienrich, and M. E. Latoschik. Challenges and opportunities of immersive technologies for mindfulness meditation: A systematic review. *Frontiers in Virtual Reality*, 2:29, 2021. 9

[18] N. Döllinger, E. Wolf, D. Mal, N. Erdmannsdörfer, M. Botsch, M. E. Latoschik, and C. Wienrich. Virtual reality for mind and body: Does the sense of embodiment towards a virtual body affect physical body awareness? In *Proc. CHI EA '22*. ACM, NY, USA, 2022. 1, 3, 5, 7, 8

[19] A. El Ali, X. Yang, S. Ananthanarayan, T. Röggla, J. Jansen, J. Hartcher-O'Brien, K. Jansen, and P. Cesar. Thermalwear: Exploring wearable on-chest thermal displays to augment voice messages with affect. In *Proc. CHI '20*, pp. 1–14. ACM, USA, 2020. doi: 10.1145/3313831.3376682 9

[20] L. A. Elkin, M. Kay, J. J. Higgins, and J. O. Wobbrock. An aligned rank transform procedure for multifactor contrast tests, 2021. 7

[21] M. Feick, N. Kleer, A. Tang, and A. Krüger. The virtual reality questionnaire toolkit. In *Adjunct Publication of the 33rd Annual ACM Symposium on User Interface Software and Technology*, UIST '20 Adjunct, p. 68?69. ACM, New York, NY, USA, 2020. doi: 10.1145/3379350.3416188 3

[22] M. A. Feijt, J. H. Westerink, Y. A. D. Kort, and W. A. IJsselsteijn. Sharing biosignals: An analysis of the experiential and communication properties of interpersonal psychophysiology. *Human–Computer Interaction*, 0(0):1–30, 2021. doi: 10.1080/07370024.2021.1913164 2

[23] E. Ferentzi, O. Wilhelm, and F. Köteles. What counts when heartbeats are counted. *Trends in Cognitive Sciences*, 2022. 2

[24] M. L. Filippetti and M. Tsakiris. Heartfelt embodiment: Changes in body-ownership and self-identification produce distinct changes in interoceptive accuracy. *Cognition*, 159:1–10, Feb. 2017. 3

[25] M. Garau, D. Friedman, H. R. Widenfeld, A. Antley, A. Brogni, and M. Slater. Temporal and spatial variations in presence: Qualitative analysis of interviews from an experiment on breaks in presence. 2008. 8

[26] S. N. Garfinkel, A. K. Seth, A. B. Barrett, K. Suzuki, and H. D. Critchley. Knowing your own heart: Distinguishing interoceptive accuracy from interoceptive awareness. *Biological Psychology*, 104:65–74, 2015. doi: 10.1016/j.biopsycho.2014.11.004 1, 4

[27] C. George and M. Hassib. Towards augmenting ivr communication with physiological sensing data. In *Proc. CHI EA '19*, pp. 1–6. ACM, NY, USA, 2019. doi: 10.1145/3290607.3313082 3

[28] S. Gradl, M. Wirth, T. Zillig, and B. M. Eskofier. Visualization of heart activity in virtual reality: A biofeedback application using wearable sensors. In *2018 IEEE International Conference on Wearable and Implantable BSN*, pp. 152–155, 2018. doi: 10.1109/BSN.2018.8329681 4

[29] M. Hassib, D. Buschek, P. W. Wozniak, and F. Alt. Heartchat: Heart rate augmented mobile chat to support empathy and awareness. In *Proc. CHI '17*, pp. 2239–2251. ACM, NY, USA, 2017. 3

[30] P. Heidicker, E. Langbehn, and F. Steinicke. Influence of avatar appearance on presence in social vr. In *2017 IEEE Symposium on 3DUI*, pp. 233–234, 2017. doi: 10.1109/3DUI.2017.7893357 1, 3, 5

[31] L. Hirsch. Bodystorming exteroceptive stimulation to train body awareness. In *Mensch und Computer 2019 - Workshopband*. Gesellschaft für Informatik e.V., Bonn, 2019. doi: 10.18420/muc2019-ws-407 9

[32] L. Hodossy, V. Ainley, and M. Tsakiris. How do we relate to our heart? neurobehavioral differences across three types of engagement with cardiac interoception. *Biological Psychology*, 165:108198, 2021. doi: 10.1016/j.biopsycho.2021.108198 2, 4, 6, 8, 9

[33] S. Ishihara et al. Tests for color blindness. *American Journal of Ophthalmology*, 1(5):376, 1918. 5

[34] J. H. Janssen, W. A. Ijsselsteijn, J. H. D. M. Westerink, P. Tacken, and G.-J. de Vries. The tell-tale heart. *Int. J. Synth. Emot.*, 4(1):65–91. 3

[35] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, 3(3):203–220, 1993. 2, 5

[36] S. S. Khalsa and et al. Interoception and mental health: A roadmap. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 3(6):501–513, 2018. Interoception and Mental Health. 1, 9

[37] S. S. Khalsa, D. Rudrauf, J. S. Feinstein, and D. Tranel. The pathways of interoceptive awareness. *Nature Neuroscience*, 12(12):1494–1496, Dec 2009. doi: 10.1038/nn.2411 2, 6, 9

[38] O. D. Kothgassner, A. Goreis, I. Bauda, A. Ziegenaus, L. M. Glenk, and A. Felnhofer. Virtual reality biofeedback interventions for treating anxiety. *Wiener klinische Wochenschrift*, 134(1):49–59, Jan. 2022. doi: 10.1007/s00508-021-01991-z 1

[39] K. Kutscheidt, T. Dresler, J. Hudak, B. Barth, F. Blume, T. Ethofer, A. J. Fallgatter, and A.-C. Ehlis. Interoceptive awareness in patients with attention-deficit/hyperactivity disorder (ADHD). *Atten. Defic. Hyperact. Disord.*, 11(4):395–401, Dec. 2019. 1

[40] S. Lee, A. El Ali, M. Wijntjes, and P. Cesar. Understanding and designing avatar biosignal visualizations for social virtual reality entertainment. In *Proc. CHI '22*. ACM, USA, 2022. doi: 10.1145/3491102.3517451 2, 3, 4

[41] N. Legrand, N. Nikolova, C. Correa, M. Brændholt, A. Stuckert, N. Kildahl, M. Vejlø, F. Fardo, and M. Allen. The heart rate discrimination task: A psychophysical method to estimate the accuracy and precision of interoceptive beliefs. *Biological Psychology*, 168:108239, 2022. 2, 4, 8, 9

[42] R. Lüddecke and A. Felnhofer. Virtual Reality Biofeedback in Health: A Scoping Review. *Applied Psychophysiology and Biofeedback*, 47(1):1–15, Mar. 2022. doi: 10.1007/s10484-021-09529-9 1, 3

[43] E. Lux, M. Adam, V. Dorner, S. Helming, M. T. Knierim, and C. Weinhardt. Live biofeedback as a user interface design element: A review of the literature. *Commun. Assoc. Inf. Syst.*, 43:18, 2018. 3

[44] N. Macmillan and C. Creelman. *Detection Theory: A User's Guide*. Taylor & Francis, 2004. 2, 6

[45] A. MacQuarrie and A. Steed. Cinematic virtual reality: Evaluating the effect of display type on the viewing experience for panoramic video. In *2017 IEEE VR*, pp. 45–54, 2017. doi: 10.1109/VR.2017.7892230 1, 3

[46] W. Mehling. Differentiating attention styles and regulatory aspects of self-reported interoceptive sensibility. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 371(1708):20160013, Nov. 2016. 9

[47] W. E. Mehling, M. Acree, A. Stewart, J. Silas, and A. Jones. The multi-

dimensional assessment of interoceptive awareness, version 2 (maia-2). *PLOS ONE*, 2018. doi: 10.1371/journal.pone.0208034 2, 5

[48] K. Meissner and M. Wittmann. Body signals, cardiac awareness, and the perception of time. *Biol. Psychol.*, 86(3):289–297, Mar. 2011. 3, 9

[49] P. Milgram, H. Takemura, A. Utsumi, and F. Kishino. Augmented reality: a class of displays on the reality-virtuality continuum. In H. Das, ed., *Telemanipulator and Telepresence Technologies*, vol. 2351, pp. 282 – 292. International Society for Optics and Photonics, SPIE, 1995. 2

[50] H. L. Miller and N. L. Bugnariu. Level of immersion in virtual environments impacts the ability to assess and teach social skills in autism spectrum disorder. *Cyberpsychol. Behav. Soc. Netw.*, 19(4):246–256. 8

[51] C. Moge, K. Wang, and Y. Cho. Shared user interfaces of physiological data: Systematic review of social biofeedback systems and contexts in hci. In *Proc. CHI '22*. ACM, USA, 2022. doi: 10.1145/3491102.3517495 2

[52] K. Moon, M. Lee, U. Hong, S. Noh, C. Park, and K. Han. Physiological response of imagery running with or without an avatar in 3d virtual reality: A preliminary study. *The Asian Journal of Kinesiology*, 23, 2021. 1, 3, 9

[53] Y. Moullec, J. Saint-Aubert, J. Manson, M. Cogné, and A. Lécuyer. Multisensory display of self-avatar's physiological state: virtual breathing and heart beating can increase sensation of effort in vr. *IEEE TVCG*, 28(11):3596–3606, 2022. doi: 10.1109/TVCG.2022.3203120 2, 3, 9

[54] J. Murphy, R. Brewer, M.-P. Coll, D. Plans, M. Hall, S. S. Shiu, C. Catmur, and G. Bird. I feel it in my finger: Measurement device affects cardiac interoceptive accuracy. *Biol Psychol*, 148:107765, Sept. 2019. 8

[55] J. Murphy, R. Brewer, D. Plans, S. S. Khalsa, C. Catmur, and G. Bird. Testing the independence of self-reported interoceptive accuracy and attention. *Quarterly Journal of Experimental Psychology*, 73(1):115–133, 2020. PMID: 31519137. doi: 10.1177/1747021819879826 4

[56] C. D. Murray and M. S. Gordon. Changes in bodily awareness induced by immersive virtual reality. *Cyberpsychol. Behav.*, 4(3), 2001. 1, 3, 9

[57] S. Oppenheimer and D. Cechetto. The insular cortex and the regulation of cardiac function. *Compr. Physiol.*, 6(2):1081–1133, Mar. 2016. 1

[58] C. Palmer, V. Ainley, and M. Tsakiris. Fine tuning your heart: a novel method for measuring interoceptive accuracy. *PsyArXiv*, 2019. 2

[59] M. P. Paulus and M. B. Stein. Interoception in anxiety and depression. *Brain Struct. Funct.*, 214(5-6):451–463, June 2010. 1

[60] D. J. Plews, B. Scott, M. Altini, M. Wood, A. E. Kilding, and P. B. Laursen. Comparison of heart-rate-variability recording with smartphone photoplethysmography, polar H7 chest strap, and electrocardiography. *Int. J. Sports Physiol. Perform.*, 12(10):1324–1328, Nov. 2017. 4

[61] S. Ponzo, D. Morelli, C. Suksasilp, M. Cairo, and D. Plans. Measuring interoception: The cardiac elevation detection task. *Frontiers in Psychology*, 12, 2021. doi: 10.3389/fpsyg.2021.712896 2

[62] S. Putze, D. Alexandrovsky, F. Putze, S. Höffner, J. D. Smeddinck, and R. Malaka. Breaking the experience: Effects of questionnaires in vr user studies. In *Proc. CHI '20*, pp. 1–15. ACM, NY, USA, 2020. 5

[63] L. Quintero, J. E. Munoz, J. Mooij, and M. Gaebler. Excite-o-meter: Software framework to integrate heart activity in virtual reality. In *2021 IEEE ISMAR*, pp. 357–366, 2021. 2, 3

[64] C. Ring, J. Brener, K. Knapp, and J. R. Mailloux. Effects of heartbeat feedback on beliefs about heart rate and heartbeat counting: A cautionary tale about interoceptive awareness. *Biological Psychology*, 104, 2015. 9

[65] G. Riva and F. Mantovani. Being there: Understanding the feeling of presence in a synthetic environment and its potential for clinical change. In C. Eichenberg, ed., *Virtual Reality in Psychological, Medical and Pedagogical Applications*, chap. 1. IntechOpen, Rijeka, 2012. 1

[66] M. Schaffarczyk, B. Rogers, R. Reer, and T. Gronwald. Validity of the polar h10 sensor for heart rate variability analysis during resting state and incremental exercise in recreational men and women. *Sensors*, 22(17), 2022. doi: 10.3390/s22176536 4

[67] R. Schandry. Heart beat perception and emotional experience. *Psychophysiology*, 18(4):483–488. 2, 6, 9

[68] F. Schoeller, P. Bertrand, L. J. Gerry, A. Jain, A. H. Horowitz, and F. Zenasni. Combining virtual reality and biofeedback to foster empathic abilities in humans. *Frontiers in Psychology*, 9:2741, 2019. 1, 3

[69] F. Schoeller, A. H. Horowitz, P. Maes, A. Jain, N. Reggente, L. C. Moore, M. Trousselard, A. Klein, L. Barca, G. Pezzulo, and et al. Interoceptive technologies for clinical neuroscience. *PsyArXiv*, 2022. 2, 3, 9

[70] T. Schubert, F. Friedmann, and H. Regenbrecht. The experience of presence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments*, 10(3):266–281, 2001. 2, 5

[71] A. Schulz, S. N. Back, V. K. Schaan, K. Bertsch, and C. Vögele. On the construct validity of interoceptive accuracy based on heartbeat counting: Cardiovascular determinants of absolute and tilt-induced change scores. *Biological Psychology*, 164:108168, 2021. 2, 9

[72] L. Sedeño, B. Couto, M. Melloni, A. Canales-Johnson, A. Yoris, S. Baez, S. Esteves, M. Velásquez, P. Barttfeld, M. Sigman, R. Kichic, D. Chialvo, F. Manes, T. A. Bekinschtein, and A. Ibanez. How do you feel when you can't feel your body? interoception, functional connectivity and emotional processing in depersonalization-derealization disorder. *PLoS One*, 9(6):e98769, June 2014. 3

[73] M. Slater. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 364(1535):3549–3557, Dec. 2009. 5, 8, 9

[74] M. Slater, D. Banakou, A. Beacco, J. Gallego, F. Macia-Varela, and R. Oliva. A separate reality: An update on place illusion and plausibility in virtual reality. *Frontiers in Virtual Reality*, 3, 2022. 1, 5, 8, 9

[75] M. Slater, M. Usoh, and A. Steed. Depth of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, 3(2):130–144, Jan. 1994. doi: 10.1162/pres.1994.3.2.130 1, 2, 3, 5

[76] G. D. Smith and S. Ebrahim. Data dredging, bias, or confounding: They can all get you into the bmj and the friday papers, 2002. 7

[77] K. Suzuki, S. N. Garfinkel, H. D. Critchley, and A. K. Seth. Multisensory integration across exteroceptive and interoceptive domains modulates self-experience in the rubber-hand illusion. *Neuropsychologia*, 51(13):2909–2917, 2013. doi: 10.1016/j.neuropsychologia.2013.08.014 4, 9

[78] A. Tajadura-Jiménez, A. Väljamäe, and D. Västfjäll. Self-representation in mediated environments: The experience of emotions modulated by auditory-vibrotactile heartbeat. *CyberPsychology & Behavior*, 11(1):33–38, 2008. PMID: 18275310. doi: 10.1089/cpb.2007.0002 2, 3, 9

[79] G. Tanay and A. Bernstein. State mindfulness scale (sms): Development and initial validation. *Psychological assessment*, 25, 09 2013. 2, 5, 7

[80] M. Tsakiris and H. Critchley. Interoception beyond homeostasis: affect, cognition and mental health. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 371(1708), Nov. 2016. 1

[81] M. Tsakiris, A. Tajadura-Jiménez, and M. Costantini. Just a heartbeat away from one's body: interoceptive sensitivity predicts malleability of body-representations. *Proc. Biol. Sci.*, 278(1717):2470–2476. 3, 9

[82] R. Ueoka, A. AlMutawa, and H. Katsuki. Emotion hacking vr (eh-vr): Amplifying scary vr experience by accelerating real heart rate using false vibrotactile biofeedback. In *SIGGRAPH ASIA '16 Emerging Technologies*. ACM, NY, USA, 2016. doi: 10.1145/2988240.2988247 3, 5

[83] M. Usoh, E. Catena, S. Arman, and M. Slater. Using presence questionnaires in reality. *Presence: Teleoper. Virtual Environ.*, 2000. 2, 5

[84] P. van Gent, H. Farah, N. van Nes, and B. van Arem. Heartpy: A novel heart rate algorithm for the analysis of noisy signals. *Transportation Research Part F: Traffic Psychology and Behaviour*, 66:368–378, 2019. 6

[85] T. Waltemate, D. Gall, D. Roth, M. Botsch, and M. E. Latoschik. The impact of avatar personalization and immersion on virtual body ownership, presence, and emotional response. *IEEE TVCG*, 24(4), 2018. 8

[86] W. E. Whitehead, V. M. Drescher, P. Heiman, and B. Blackwell. Relation of heart rate control to heartbeat perception. *Biofeedback and Self-regulation*, 2:371–392, 1977. 2

[87] C. D. Wickens. Multiple resources and mental workload. *Human factors*, 50(3):449–455, 2008. 9

[88] B. Wiederhold, R. Davis, and M. Wiederhold. The effect of immersiveness on physiology. *Studies in health technology and informatics*, 58:52–60, 02 1998. doi: 10.3233/978-1-60750-902-8-52 1, 3

[89] R. M. Winters, B. N. Walker, and G. Leslie. Can you hear my heartbeat?: Hearing an expressive biosignal elicits empathy. In *Proc. CHI '21*. ACM, New York, NY, USA, 2021. 2, 3, 4, 9

[90] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins. *The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures*, pp. 143–146. ACM, New York, NY, USA, 2011. 6, 7

[91] T. Xue, A. E. Ali, G. Ding, and P. Cesar. Investigating the relationship between momentary emotion self-reports and head and eye movements in hmd-based 360° vr video watching. In *Proc. CHI EA '21*. ACM, 2021. 7

[92] B. Yu, J. Hu, M. Funk, and L. Feijs. Delight: biofeedback through ambient light for stress intervention and relaxation assistance. *Personal and Ubiquitous Computing*, 22(4):787–805, Aug 2018. 9

[93] G. Zamariola, P. Maurage, O. Luminet, and O. Corneille. Interoceptive accuracy scores from the heartbeat counting task are problematic: Evidence from simple bivariate correlations. *Biological Psychology*, 137:12–17, 2018. doi: 10.1016/j.biopsycho.2018.06.006 6

[94] Y. Zhou, A. Murata, and J. Watanabe. The calming effect of heartbeat vibration. In *2020 IEEE HAPTICS*, pp. 677–683, 2020. 2, 3