Sensitivity to changes in rate of heartbeats as a measure of interoceptive ability

Dennis E O Larsson*[1,2,3], Giulia Esposito[3,4], Hugo D Critchley[3,5,6], Zoltan Dienes[1,4], Sarah N Garfinkel[3,6,7]

* Corresponding author; email: dennlarsson@gmail.com

[1] Department of Psychology, University of Sussex, Falmer, BN1 9RR, UK

[2] Leverhulme Trust

[3] Department of Neuroscience, Brighton and Sussex Medical School, University of Sussex, Falmer, BN1 9RY, UK

[4] Institute of Neuroscience, Université Catholique de Louvain, Brussels, 1200 Woluwe-Saint-Lambert, Belgium

[5] Sackler Centre for Consciousness Science, University of Sussex, Falmer, BN1 9RR, UK

[6] Sussex Partnership NHS Foundation Trust

[7] Institute of Cognitive Neuroscience, University College London, London, WC1N 3AZ, UK

Abstract

Interoception is the sensing of internal bodily signals. Individuals vary in their ability to perceive, as conscious sensations, signals like the beating of the heart. Tests of such interoceptive ability are, however, constrained in nature and reliability. Performance of the heartbeat tracking task, a widely used test of cardiac interoception, often corresponds well with individual differences in emotion and cognition, yet is susceptible to reporting bias and influenced by higher order knowledge, e.g. of expected heart rate. The present study introduces a new way of assessing cardiac interoceptive ability, focusing on sensitivity to short-term, naturalistic changes in frequency of heartbeats. Results indicate an overall tendency to report fewer heartbeats during accelerations in heart rate. This finding may be driven in part by respiration, with a reduction in heartbeat salience during inspiratory periods when heart rate typically increases. Within-participant performance was also marked by a high degree of variability relative to the heartbeat counting task. Rather than a veridical monitoring of subtle variations in physiology, cardiac interoceptive ability appears to involve interpolation, wherein interoceptive decisions are informed by a dynamic working estimate from, the integration of afferent signalling with higher order predictions.

Keywords: Interoception; Heart rate variability; Psychophysics; Metacognition

1. Introduction

Our bodies are ever-present hosts of our minds. Biological activity in both brain and body underlies and shapes psychological processes. A continuous stream of information from the body informs the brain about our current and changing state physiological functioning and of physical integrity, often independently of immediate perceptual awareness (Critchley & Harrison, 2013). Interoception describes the sense through which signals originating from within the body, especially its visceral organs, are carried, represented, integrated, and interpreted within the nervous system, across conscious and unconscious levels (Khalsa et al., 2018). Interoception is essential to the regulation of the internal milieu of the body, conveying to the brain the feedback that informs reflexive homeostatic control and allostatic adaptation (Craig, 2009; Critchley & Garfinkel, 2017). Interoception thus encompasses a multidimensional spectrum of signal processing, from physiological responses and their proximate neural representations, to the perception and awareness of these interoceptive signals and associated feelings (Critchley & Garfinkel, 2017; Khalsa et al., 2018; Quadt, Critchley, & Garfinkel, 2018). Overall, interoceptive signals are rarely consciously accessible (Khalsa et al., 2018). For exteroceptive sensations (e.g. somatosensory touch acuity), an individual's self-report, or discriminatory accuracy is tightly coupled to subjective confidence in the perception (Garfinkel et al., 2016). In contrast, for interoception, measures of perceptual accuracy vary across individuals and correspond loosely with subjective ratings of confidence or experience (e.g. in questionnaire reports of interoceptive sensitivity). On experimental tests of interoceptive ability, the correspondence between trial-by-trial accuracy and confidence can be quantified mathematically as interoceptive metacognition (insight). This metacognitive index also often diverges from measures of interoceptive task accuracy (Garfinkel, Seth, Barrett, Suzuki, & Critchley, 2015).

In addition to its important role in physiological regulation (Critchley & Harrison, 2013; Mayer, 2011), interoception is implicated in normative emotion responses (Critchley & Garfinkel, 2015; Garfinkel & Critchley, 2016; Garfinkel et al., 2014; Schulz et al., 2016), and the expression of specific clinical symptoms, including those of anxiety (Domschke, Stevens, Pfleiderer, & Gerlach, 2010; Garfinkel et al., 2016). In the present study, we sought to optimise the reliability of measuring interoceptive ability across individuals and populations. Our approach aimed for a more comprehensive understanding of how and why measures of interoceptive accuracy deviate from subjective and metacognitive measures of insight, with further implications for emotion and embodied cognition.

1.1. Cardiac interoception tasks

The cardiac signal (i.e. the heartbeat coincident with ventricular systole) stands out as a repetitive and discrete physiological event, allowing it to be recorded with some precision (Phillips, Jones, Rieger, & Snell, 1999). Moreover, heartbeats are vital and salient; their frequency and strength changes with emotion and action, when they can breach the threshold for conscious perception and contribute to affective feelings. Consequently, interoception research has tended to focus on the heartbeat. The heartbeat tracking (HBT) task, conceptualized by Schandry (1981), is a relatively straightforward procedure. Based on the notion that some people have greater sensitivity than others to their heartbeats, the task aims to quantify interoceptive ability from how well an individual can detect their own heartbeats at rest. The participant focuses attention on his/her heart and counts the number of heartbeats felt within given time periods (without directly palpating a pulse or using other external strategies). Performance is generally calculated as the error rate between reported and actual number of heartbeats (Schandry, 1981). One alternative cardiac interoception task, similarly motivated, is the heartbeat discrimination (HBD) task (Katkin, Reed, & Deroo, 1983; Whitehead, Drescher, Heiman, & Blackwell, 1977). This task also requires an interoceptive focus at rest; on individual trials, which consist of runs of heartbeats, the participant judges if a phasic exteroceptive stimulus (e.g. auditory tone or flashing light) is played either synchronously or delayed with respect to each heartbeat. Synchronicity judgements over repeated trials allows computation of performance accuracy, e.g. the percentage of correct trials, or as d’. The HBT and HBD tasks have dominated approaches to measure individual differences in cardiac interoceptive ability/accuracy, which has been presumed to be a relatively stable constitutional trait. Both tasks carry face validity, reinforced by well-documented links to other cognitive and emotional factors (Domschke et al., 2010; Garfinkel et al., 2013; Garfinkel et al., 2016; Herbert, Herbert, & Pollatos, 2011; Pollatos & Schandry, 2008; Werner, Peres, Duschek, & Schandry, 2010).

1.2. Limitations of current cardiac interoception tasks

Despite wide use, both the HBT and HBD tasks are not without controversy. Specifically, performance accuracy in the HBT task is influenced by levels of general intelligence, and relatedly by prior knowledge of one’s heart rate (Brener & Ring, 2016; Murphy et al., 2018; Phillips et al., 1999; Ring & Brener, 1996, 2018; Ring, Brener, Knapp, & Mailloux, 2015). Thus performance accuracy scores can reflect average heart rate rather than the veridical number of heartbeats within individual trials (Desmedt, Luminet, & Corneille, 2018; Zamariola, Maurage, Luminet, & Corneille, 2018). Furthermore, although the HBT task was designed as a measure of resting cardiac sensitivity, the reported number of heartbeats is observed to remain fairly constant across conditions (such as postural change) that evoke changes in heart rate: e.g. when the participant is lying down, sitting, or standing (Palmer, Ainley, & Tsakiris, 2019). Similarly, one study (requiring validation) provided some evidence to suggest that changes in heart rate induced by implanted pacemaker did not reliably change patients’ reported number of heartbeats on the HBT task (Windmann, Schonecke, Fröhlig, & Maldener, 1999).

Another point of contention is that, when performing the HBT task, people overall tend to underreport their number of heartbeats, and high accuracy scores thus appear to be, in part, related to a bias towards reporting a higher number of felt heartbeats (Desmedt et al., 2018; Zamariola et al., 2018). Since weak and diffuse interoceptive sensations (including heartbeats at rest) are generally not felt unless attention is focused and distraction removed, underreporting is perhaps expected. People are overall likely to fail to register some heartbeats even when attending to the inherently noisy cardiac signal (Ainley, Tsakiris, Pollatos, Schulz, & Herbert, 2020). However, it has also been argued that the tendency of participants to under- or (less commonly) over- report heartbeats by itself does not reliably predict HBT task performance (Zimprich, Nusser, & Pollatos, 2020). In contrast, successful completion of the HBD cannot be guided by higher order knowledge of heart rate. However, it is a more difficult task that requires multimodal integration of interoceptive and exteroceptive information, and thus utilises other processes (including general intelligence) in addition to interoception (Garfinkel et al., 2015). HBD performance typically divides populations bimodally (some that can and most that cannot do the task) limiting its application.

Despite such criticism (and issues to do with number of trials to attain stable measures of individual differences), these interoceptive tests remain widely used (Kleckner, Wormwood, Simmons, Barrett, & Quigley, 2015). Interoceptive accuracy, as measured through the HBT task shows high test-retest reliability across time (Ferentzi, Drew, Tihanyi, & Köteles, 2018). HBT performance is also repeatedly shown to predict to measures of cognition and emotion that fit with *a priori* theory-driven hypotheses (Critchley & Garfinkel, 2018). However, despite heuristic value and ease of implementation, HBT task limitations cannot be ignored. The present study aimed to develop a new technique to assess individual differences in interoceptive ability. We based our approach on the established HBT protocol, but our aim was to separate genuine discriminability from reporting bias. We applied a new analysis strategy that focused on sensitivity to changes in the number of heartbeats between (shorter) trials. A healthy heart does not beat at an unchanging pace, but instead manifests natural rhythmic fluctuations, such that the cardiac interbeat intervals change continuously over time. These changes, known as heart rate variability (HRV), arise from the dynamics of homeostatic regulation that work to maintain stable cardiac output in response to changing conditions (posture, action, emotion) and metabolic demand. HRV is proximately regulated through the baroreflex, where fluctuations in heart rate are matched with changes in blood pressure to maintain stable cardiac output (Smith, Thayer, Khalsa, & Lane, 2017). In short, stronger cardiac higher-pressure ejection of blood into the aorta and carotid arteries cause arterial baroreceptor to discharge. This signal, conveyed to brainstem, triggers compensatory vagal parasympathetic slowing of the subsequent heartbeat. Conversely, a weaker arterial baroreceptor signal, indicating lower ventricular ejection pressure, permits acceleration of the subsequent heartbeat (Riganello et al., 2018; Shaffer, McCraty, & Zerr, 2014; Shaffer & Venner, 2013). Importantly, the mechanical effects of breathing contribute to these changes: Inspiration results in decreased pleural pressure as the chest expands and increased abdominal pressure, which together lower right atrial pressure facilitating venous return to the right heart. Left ventricular stroke volume decreases due to increased in pulmonary blood volume, lowering aortic blood pressure (Magder, 2018). The resulting decrease in vagal tone evokes an increase in heart rate through the baroreflex to maintain overall cardiac output (Draghici & Taylor, 2016; Taylor & Eckberg, 1996). These respiration-related changes in blood pressure and heart rate cause the heart rhythm naturally fluctuate, giving rise to the naturally occurring HRV (Berntson, Quigley, Norman, & Lozano, 2016).

Natural changes in HRV occur over the course of trials of the standard HBT task. Thus, if an individual is particularly sensitive to individual heartbeats, then this sensitivity will be closely mirrored in performance i.e. the reported number of heartbeats experienced over different time-intervals. In contrast, if the individual is less sensitive to feeling her/his own heartbeats, or interpolates across fluctuations in interbeat variability to produce a number that approximates more to her/his average heart rate, then the report will be more consistent over trials when compared to veridical tracking of individual heartbeats in the context of HRV.

1.3. Hypothesis

We hypothesized that participants with high interoceptive accuracy, hence greater sensitivity to changes in the number of heartbeats, can be identified through the slope of a linear regression of reported beats against actual beats per minute (heart rate) across trials of a modified HBT task. A positive slope would indicate that participants display cardiac interoceptive sensitivity expressed as the ability to track changes in the cardiac signal; across trials the reported number of counted heartbeats increases appropriately with increased heart rate. In contrast, a negative slope would indicate a reduced ability to detect heartbeats as heart rate increases (during natural fluctuations in HRV), potentially reflecting the reduced salience of individual heartbeats (perhaps relative to heart rate) that might arise as a consequence of relatively weaker heartbeats (that are more difficult to detect) at higher heart rates. The detectability of heart rate may depend in blood pressure (Koroboki et al., 2010). The volume of blood pumped per unit time will depend on the pressure at which it is pumped and the heart rate. Thus, in the homeostasis of cardiac output, heart rate and pressure are traded against each other. This may introduce (over small time periods) a negative intra-individual correlation between heartbeat detection and heart rate under resting conditions.

The estimate of the raw slope is not biased by restrictions in range; that is, variations in HRV will not affect the expected magnitude of the slope (though the standard error of the slope will be larger the smaller the range). It is only the ability to discriminate heartbeats at different heart rates that will lead to a non-zero slope. Bias in reporting (i.e. a tendency to give high or low numbers regardless of the actual heart rate) will affect the intercept of the raw regression.

2. Methods

2.1. Participants

A total of 100 participants (76 females) between the age of 18 and 33 years (mean = 23.01, SD = 3.73) were recruited for the experiment. Participants had a mean body-mass index (BMI) of 22.52 (SD = 3.51), and 85 were right-handed. The participant number was estimated based on a power analysis of pre-existing data that determined the minimal standard error needed to achieve moderate evidence (Bayes factor) for H1 over H0 (see Appendix A). Baseline physiological data was missing from two participants due to corrupt saving of files, and another three were excluded due too much noise and movement artefacts within physiological recordings. This resulted in useable baseline data from 95 participants.

2.2. Materials

Electrocardiogram (*ECG*) recording used CED hardware and software (Cambridge Electron Design, Cambridge; 1408 signal converter and 1902 amplifier, with software Spike2 version 7.18). The ECG signals was sampled at 1000 Hz. R-waves were identified via an interactive threshold in the Spike2 software recording. Data analysis was undertaken in the R environment, version 4.0.2 (RCore, 2013), and Matlab R2020a (The MathWorks, Inc., Natick, MA); task scripts were programmed and run in Python (version 2.7.16). Baseline heart rate and non-invasive instantaneous blood pressure recordings were made using a Finometer (Finapres Medical Systems, model 1 version 1.01); analysis of beat-to-beat blood pressure used for, with accompanying software BeatScore Easy (version 02.10 build 004).

2.3. Experimental procedure

Each participant went through a baseline physiological recording period, during which heart rate and beat-to-beat blood pressure were monitored using a Finometer, with primary sensor located at the intermediate phalanx of the left middle finger. The participant rested in a sitting position for approximately two minutes before the recording started. In the first stage (*pre-stand*) of the baseline recording, the participant was asked to lie back in a supine position and relax and try not to move while heartbeats and blood pressure was measured for five minutes. In the second stage (*post-stand*) after five minutes of recording, the participant was asked to stand upright and face the wall, and stand still for another three minutes. After three minutes had passed, the recording ended and the Finometer wrist and finger cuffs were removed. This baseline physiological data was used for estimating each individual’s baseline physiology and reactivity of the autonomic and cardiovascular systems to physiological influences with postural change.

Each participant went through a modified heartbeat tracking (*HBT*) task, which we had optimised for slope analyses (describe above) measuring changes of number of counted heartbeats as a function of resting HRV. Each participant sat in front of a computer screen with a straight back and both feet on the floor with arms relaxed. The participant was instructed to pay attention to their heart and count heartbeats, without using their hands or other means to feel their pulse that might affect their performance. In the interoceptive condition, each trial started with a clear auditory signal, where the participant was required to close their eyes and start silently count felt heartbeats. After a set period of time (of which 80% of all trials were 20 seconds in duration, 10% were 18 seconds, and 10% were 22 seconds, randomly intermixed), another clear tone was played, marking the end of the trial. The participant then opened their eyes, to report: 1) how many heartbeats they counted, and 2) how confident they were that their report was correct, rated on a four-point confidence scale with accompanying descriptions: Level one was “I did not sense my heartbeats; I am completely guessing about the number of beats”; level two was “I sensed something about my heart, but I had no idea what I was counting, and I have no confidence at all in my counting”; level three was “I sporadically or faintly picked up on my heartbeat; my counting is based on something, but it may be off by a small margin"; and level four was "I clearly sensed my heartbeat, and have full confidence in my count". In an exteroceptive condition, the participant was instructed to count how many faint auditory tones were being played through the speakers over the course of individual trials of equivalent length to the interoceptive condition. The participant then rated their confidence (similar to the interoceptive condition). The volume of the tones was adjusted through an active staircase procedure aimed to match performance of the exteroceptive condition with performance of the interoceptive condition, calculated as 1 - error rate between reported and actual number of stimulus occurrences (i.e. number of heartbeats or played tones). Across participants, these two conditions were presented in an alternating order across a total of 12 blocks (with 6 for each condition), with the first condition being randomly picked for each participant. Each block contained 10 trials, for a total of 120 trials (with 60 trials for each condition). Each participant started the task with a short training block consisting of the exteroceptive condition that served both to familiarise participants with the task, and to staircase the exteroceptive task difficulty to a level where participants had an initial hit-rate of approximately 70% (between 55% and 85%) through the staircasing procedure described above. The training block continued until a hit-rate of between 55% and 85% had been maintained for two consecutive trials, for a maximum of 20 trials.

In addition to the HBT task, a subset of participants (N=47) also completed the heartbeat discrimination (*HBD*) task, where an external tone is presented synchronously or asynchronously to the participant's own heartbeats. For each trial, an externally generated tone was played 10 times, triggered by the ECG R-wave. In the synchronous condition, the tones were presented approximately 250 milliseconds (*ms*) after the R-waves (at early systole, approximately at the ECG T-wave and the ventricular ejection period, when the heart is beating). In the asynchronous condition, the tones were presented approximately 550 ms after the R-waves (late diastole, between heartbeats). At the end of each trial, the participant was required to judge whether the tone was presented synchronously or asynchronously with their heartbeats via a button press response, and then report how confident they are in their report on a four-point confidence scale (as used previously). The synchronous and asynchronous conditions were presented in a randomized order, with 30 trials per condition, for a total of 60 trials. The number of trials was determined based on a power analysis on pre-existing data (see Appendix A).

2.4. Pre-processing

All ECG recordings from the tasks were visually inspected for signal noise resulting from electrical interference, movement artefacts, or equipment failure. Inspection was performed by a trained researcher. Trials where the ECG R-waves could not be discerned from signal noise were excluded from all following analyses. For the HBD task, in trials where signal noise caused irregular stimulus presentation where tones were not presented for ten consecutive heartbeats without disruption, those trials were excluded from the analyses.

2.5. Analyses

Analyses are interpreted with respect to Bayes factors (*B*), though *p* values are provided as well (Dienes & Mclatchie, 2018). A *B* of above 3 indicates “substantial” or, better, “moderate” (Lee & Wagenmakers, 2013) evidence for the alternative hypothesis (H1) over the null hypothesis (H0); thus by symmetry a *B* below 1/3 indicates substantial (/moderate) evidence for H0 over H1 (“substantial” in the sense of just worth taking note of.) *B*s between 3 and 1/3 indicate the data collected do not sensitively distinguish H0 from H1. Thus, we report that there was no effect only when $B < 1/3$. $B_{N(0, x)}$ refers to a Bayes factor in which the predictions of H1 were modelled as a normal distribution with an SD of x where x scales the size of effect that could be expected, or half the plausible maximum (see Dienes, 2014). The distribution represents the prediction that the effect could go in either direction. $B_{H(0, x)}$ refers to a Bayes factor in which the predictions of H1 were modelled as a half-normal distribution with an SD of x; the half-normal distribution represents the prediction of an effect in only one direction. To indicate the robustness of Bayesian conclusions, for each *B*, a robustness region is reported (Dienes, 2019), giving the range of scales that qualitatively support the same conclusion (i.e. evidence as supporting H0, or as supporting H1, or there not being much evidence at all), notated as: $RR_{concluson}$ [x1, x2] where x1 is the smallest SD that gives the same conclusion and x2 is the largest. "Conclusion" means "$B < 1/3$", or "$B > 3$", or "$1/3 < B < 3$".

$IAcc_{slope}$, or sensitivity to change in number of heartbeats, was calculated from the HBT task as the raw slope from a linear regression of reported BPM against actual BPM, using the formula $Y = alpha + betaX$, where Y is the reported BPM, X is the actual BPM, *alpha* is the intercept, and *beta* is the slope. In terms of modelling H1 for a Bayes factor, as the ideal slope, and hence the maximum that could be expected, is 1 reported beat per actual beat in every minute, the SD of a normal distribution was set to half that, 0.5.

$EAcc_{slope}$, or sensitivity to change in number of exteroceptive tones, was calculated as a control measure from the exteroceptive condition in the HBT task. The raw slope from a linear regression of reported and actual number of exteroceptive tones was calculated, using the same formula as for $IAcc_{slope}$. The model of H1 was the same as for the interoceptive case, for comparability.

$Bias_{intercept}$ in the form of a tendency to over- or underreport the number of heartbeats or exteroceptive tones was estimated for the interoceptive condition. Expected bias depends on participants' beliefs about average heart rate. Were people to be informed that average heart rate is about 80 bpm and they consequently relied on this information, if the slope were zero the intercept would be 80 bpm, reflecting this pre-existing knowledge rather than discriminability. Fixing this as the mean reported heart rate, a positive slope would reduce the intercept, with an ideal slope of 1 reducing the intercept to 0. Thus, for the Bayes factor, H1 for the intercept was modelled as a half-normal with an SD of 80/2. H1 was modelled for comparability. With the intercept being expected to strongly correlate with the slope, mean reported BPM was also used as a second bias estimate ($Bias_{BPM}$) for comparison.

$IAcc_{classic}$ was calculated from the HBT task as the average error rate between actual and reported number of heart beats, using the classical interoceptive accuracy formula (Schandry, 1981):

$$IAcc_{classic} = 1 - \frac{1}{n}\sum\frac{|(HB_{actual} - HB_{reported})|}{HB_{actual}}$$

$EAcc_{classic}$ was calculated from the HBT task as the average error rate between actual and reported number of exteroceptive tones, using the same formula as for $\mathrm{IAcc}_{\mathrm{classic}}$. An ideal score is 1; maximum underreporting (saying the heart rate was 0) would give a score of 0.

Interoceptive confidence (*IC*) for the HBT task (IC_{HBT}) and for the HBD task (IC_{HBD}) was measured as the averaged trial-based confidence scores.

Exteroceptive confidence (EC_{HBT}) for the HBT task was measured as the trial-based confidence scores.

Interoceptive awareness (IAw_{HBT}, i.e. metacognitive insight into one's own objective performance) was calculated from the interoceptive condition of the HBT task as the raw slope (beta) between $\mathrm{IAcc}_{\mathrm{classic}}$ and $\mathrm{IC}_{\mathrm{HBT}}$ coded as "guess" (i.e. confidence level of 1) versus any other level of confidence (Dienes, 2015). A total of 25 participants had to be excluded from the IAw_{HBT} analysis as they did not give a single confidence report either above 1 or below 2 across the entire task, and thus a slope could not be calculated.

Exteroceptive awareness (EAw_{HBT}) was calculated from the interoceptive condition of the HBT task as the raw slope (beta) between $\mathrm{EAcc}_{\mathrm{classic}}$ and $\mathrm{EC}_{\mathrm{HBT}}$. A total of 9 participants had to be excluded from the EAw_{HBT} analysis as they did not give a single confidence report either above 1 or below 2 across the entire task, and thus a slope could not be calculated.

The results of secondary analyses of the exteroceptive HBT task condition are reported in Appendix B only for brevity.

d-prime (*d'*) was used as the sensitivity/accuracy index from the HBD task, following signal detection theory (Snodgrass & Corwin, 1988), calculated as the standardised difference between the mean of the signal-to-noise distribution, compared against the standard deviation of signal-to-noise distribution.

meta-d (Barrett, Dienes, & Seth, 2013) was used as the measure of metacognition interoceptive awareness/insight for the HDB task. Following signal detection theory (Snodgrass & Corwin, 1988), it calculates the d' (type 1) accuracy that would be expected, assuming maximum metacognitive sensitivity. Given each subject's actual type 2 performance data, one can obtain the underlying type 1 sensitivity that is expected if the subject is ideal in placing their confidence ratings. Thus, meta-d is compared to accuracy to assess metacognitive relative to the idea value.

Heart rate variability (HRV_{rest}) from the baseline Finometer recording was calculated as the root mean square of successive differences (RMSSD) of cardiac interbeat intervals using the MATLAB package HRVAS (Heart Rate Variability Analysis Software) (Ramshur, 2010).

Stroke volume (SV) was calculated from the resting-state data as the amount of blood (in millilitre) the heart pumps with each beat.

Baroreflex sensitivity (BRS) from the baseline Finometer recording was calculated using a sequence technique (Gouveia, Rocha, Van De Borne, & Lago, 2005) which entails identifying sequences of three or more contiguous heartbeats during which there is a progressive increase or decrease in systolic blood pressure (*SBP*) followed by a reduction or increase in interbeat intervals (*IBI*). Each sequence gives mean corrected values of SBP and related IBI, where the slope from a linear regression of these values gives an estimate of global BRS (BRS_{global}).

A *split-trial task reliability analysis* was performed to test the reliability of the HBT task in terms of both $\text{IAcc}_{\text{classic}}$ and $\text{IAcc}_{\text{slope}}$ scores across trials. For each participant, all trials were divided into odd and even (30 trials each), including all output of each trial. For each participant, the $\text{IAcc}_{\text{slope}}$ and $\text{IAcc}_{\text{classic}}$ scores were once again calculated for both the odd and even trials, resulting in two scores of each for every participant. Across all participants, correlational analyses were conducted between odd and even trial mean $\text{IAcc}_{\text{classic}}$ scores, as well as between odd and even trial $\text{IAcc}_{\text{slope}}$ scores. The ideal raw regression slope is 1, so H1 for the raw regression slope was modelled using a half-normal distribution with SD = 0.5.

Cardiovascular reactivity was measured from data from the baseline physiological recording period. Specifically, this was informed by an 'active stand' autonomic test procedure with concurrent non-invasive beat-to-beat monitoring of arterial blood pressure. Participants were rested in a supine position for five minutes, then actively stood up unsupported for a further three minutes. Resting HRV was calculated from the root mean square of successive differences (RMSSD) in the interbeat intervals during the five minutes of supine rest using the MATLAB package HRVAS (Heart Rate Variability Analysis Software) (Ramshur, 2010). A five minute resting period is generally considered sufficient for acquiring a stable baseline blood pressure recording, as shorter may result in an unstable baseline, and there are no significant change between 5 and 10 minutes of supine rest (Mader, Palmer, & Rubenstein, 1989). For variation in heart rate during the heartbeat counting task, the standard deviation of heart rate across all trials was also used.

2.6. Data and code availability statement

All data obtained in this study, pre-existing data used in the power analysis, HBT and HBD task code, and analysis code can be found at the project's Open Source Framework site (https://osf.io/gfc62/).

3. Results

3.1. HBT task

Twelve trials across four participants from the HBT task were excluded due to noise in the ECG recordings. In terms of the main hypothesis of a relationship across trials of the HBT task between heart rate and number of reported heartbeats, the mean $\mathrm{IAcc}_{\mathrm{slope}}$ was -0.22 (SD = 0.62), this differed from zero at the group level $t(99) = -3.48$, $p < 0.001$, $B_{N(0,0.5)} = 60.65$, $RR_{B>3}[0.03, 11.20]$. Importantly, we observed across participants an average negative slope; participants tended to report fewer heartbeats as their heart rate increased.

Correlational analyses estimated only a small possible relationship between $\mathrm{IAcc}_{\mathrm{slope}}$ and $\mathrm{IAcc}_{\mathrm{classic}}$ scores ($r(98) = 0.056$, CI = [-0.14, 0.25]; and likewise for the relationship between either interoceptive accuracy and interoceptive metacognitive insight, (between $\mathrm{IAcc}_{\mathrm{slope}}$ and $\mathrm{IAw}_{\mathrm{HBT}}$: $r(73) = -0.09$, CI = [-0.31, 0.14], or between $\mathrm{IAcc}_{\mathrm{classic}}$ and $\mathrm{IAw}_{\mathrm{HBT}}$, ($r(73) = -0.19$, CI = [-0.40, 0.04]) (see Figure 1). $\mathrm{Bias}_{\mathrm{intercept}}$ was positively correlated with $\mathrm{IAcc}_{\mathrm{classic}}$ ($r(98) = 0.20$, $p = 0.045$, $b = 46.64$, $SE = 22.94$, $B_{N(0,80)} = 1.86$, $RR_{1/3<B<3}[0, 546.6]$, although the BF indicated only anecdotal evidence for this effect. $\mathrm{Bias}_{\mathrm{intercept}}$ was strongly negatively correlated with $\mathrm{IAcc}_{\mathrm{slope}}$ ($r(98) = -.95$, $p < 0.001$, $b = -69.12$, $SE = 2.37$, $B_{N(0,80)} > 100$, $RR_{B>3}[0, \infty]$). Thus, the bias to over or under -report heartbeats was related to interoceptive accuracy quantified both by the classic measure and, particularly, our new slope-based index. Specifically, under-reporting related to a higher $\mathrm{IAcc}_{\mathrm{classic}}$ score, and over-reporting bias inversely related to positive slope scores. For fixed means, the negative relation between intercept and slope for a regression line follows mathematically. So for additional comparison, $\mathrm{IAcc}_{\mathrm{slope}}$ was also compared with bias_{BPM} as an additional estimate of bias not confounded by the regression, revealing evidence for no correlation (r(73) = -0.09, p = 0.432, b = -0.02, SE = 0.03, $B_{N(0,80)} < 0.01$, $RR_{B<1/3}$[0.12, ∞]), though bias_{BPM} did correlate with $\text{IAcc}_{\text{classic}}$ (r(98) = 0.92, p < 0.001, b = 67.69, SE = 2.96, $B_{N(0,80)} > 100$, $RR_{B>3}$[0.20, ∞]). The estimate of the correlation between $\text{bias}_{\text{intercept}}$ and IAw_{HBT} was r(73) = 0.06, CI = [-0.17, 0.28], and between bias_{BPM} and IAw_{HBT}, r(73) = -0.08, CI = [-0.30, 0.15] (see Figure 1).

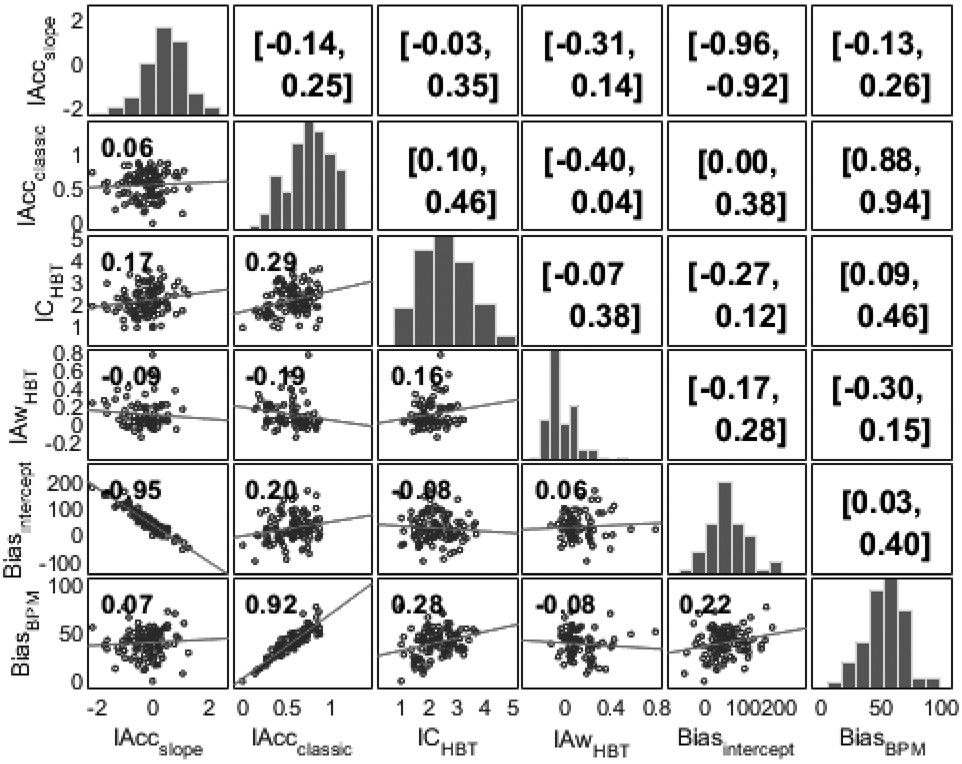

Figure 1. HBT task correlogram. Scatter plots of HBT task scores between row and column variables, and variable distribution along diagonal axis. Each scatter plot of the lower half of the correlogram contains the corresponding correlation coefficient (r), and the boxes in the upper half contains the corresponding confidence intervals (i.e. [lower, upper]).

3.2. Accuracy split-trial reliability analysis

From the split-trial task reliability analysis, results showed positive correlations between odd and even trial IAcc_{slope} ($r(98) = 0.22$, $p = 0.031$, CI = [0.02, 0.40], $b = 0.21$, $SE = 0.09$, $B_{H(0,0.5)} = 4.90$, $RR_{B>3}[0.058, 0.874]$) as well as between odd and even trial $\mathrm{IAcc}_{classic}$ ($r(98) = 0.99$, $p < 0.001$, CI = [0.982, 0.992], $b = 1.00$, $SE = 0.02$, $B_{H(0,0.5)} > 100$, $RR_{B>3}[0, \infty]$) (see Figure 2), signifying task reliability of both accuracy measures, although the BF related to the IAcc_{slope} measure shows the evidence to be inconclusive. Comparing the correlations using the Silver, Hittner, and May (2004) modification of the Dunn and Clark (1969's) z estimated the difference between the correlations to be $z = -16.34$, 95% CI[-0.97, -0.59].

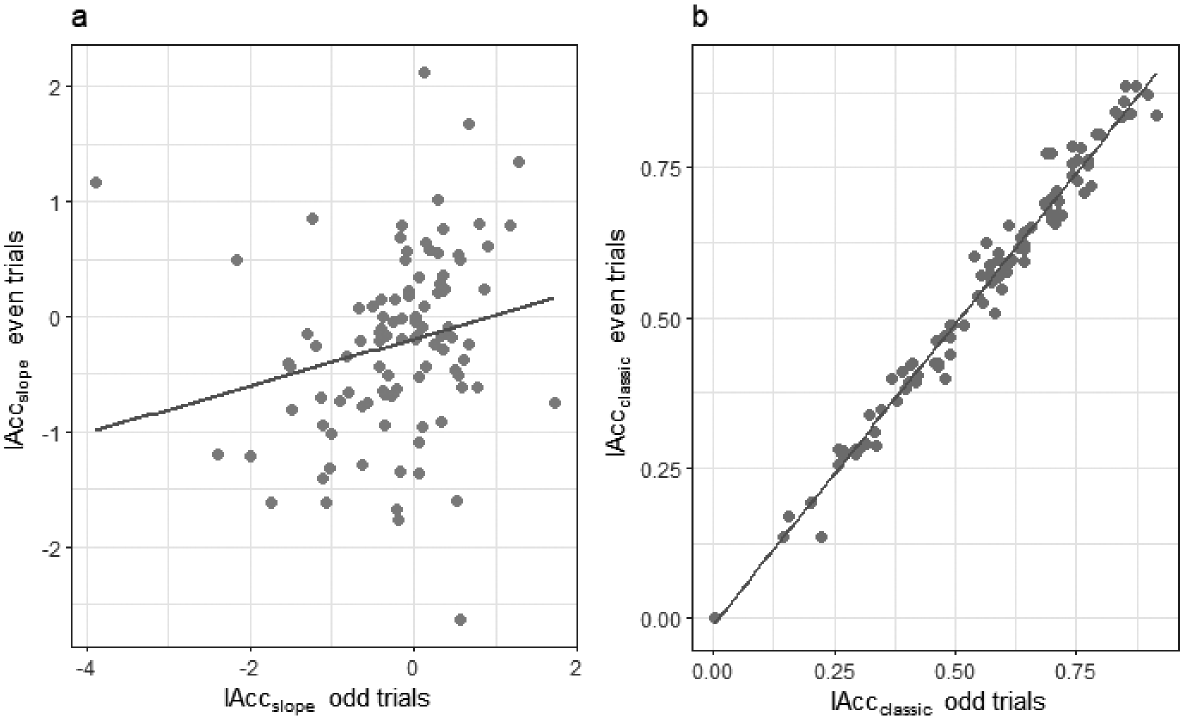

Figure 2. Accuracy split-trial task reliability plot. Comparing scores from odd and even trials for each participant showed significant correlation between IAcc_{slope} scores (2a), and a significant correlation between $\mathrm{IAcc}_{classic}$ scores (2b).

3.3. HBD task

A subset of the participants ($n = 47$) performed the HBD (heartbeat discrimination) task in addition to the HBT (heartbeat tracking) task (see Figure 3). 120 trials across 25 participants were excluded due to noise in the ECG recording causing incorrect or irregular stimulus presentation timings relative to the R-peaks.

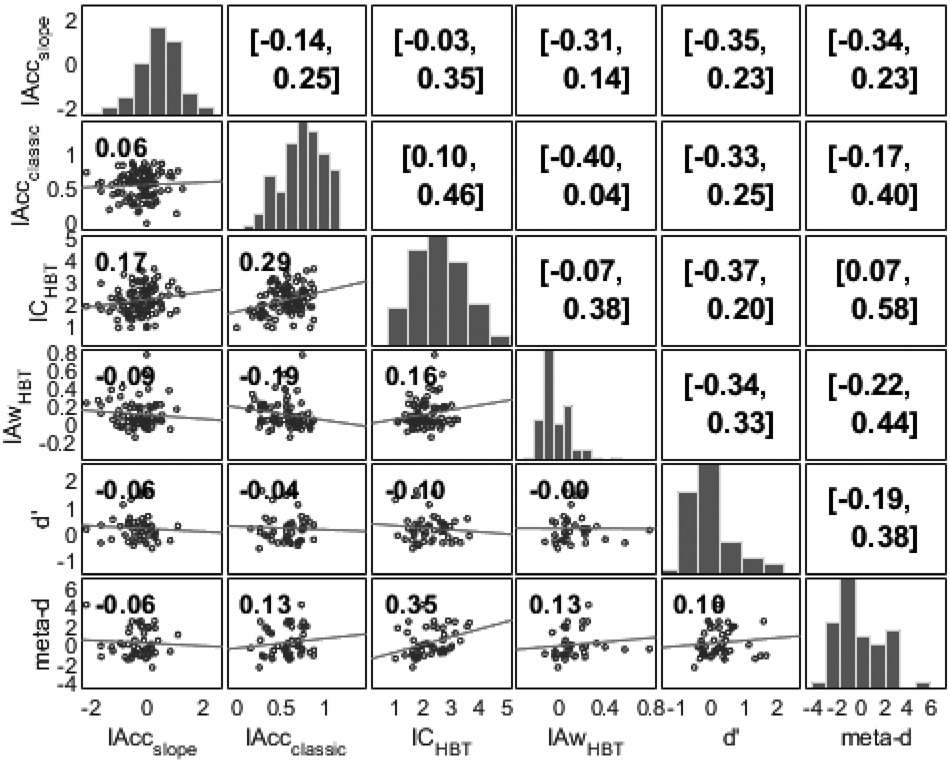

Figure 3. HBD task correlogram. Scatter plots of HBD task and interoceptive HBT task scores between row and column variables, and variable distribution along diagonal axis. Each scatter plot of the lower half of the correlogram contains the corresponding correlation coefficient (r), and the boxes in the upper half contains the corresponding confidence intervals (i.e. [lower, upper]).

3.4. Baseline physiology

Correlations between physiological measures and HBT task performance revealed a notable negative correlation between within-task heart rate (HR) and IAcc$_{classic}$ score (r(98) = -0.26, CI = [-0.43, -0.06]), suggesting the IAcc$_{classic}$ measure to be affected by, or related to, participant's heart rate, with higher heart rate being associated with lower IAcc$_{classic}$ scores (see Figure 4).

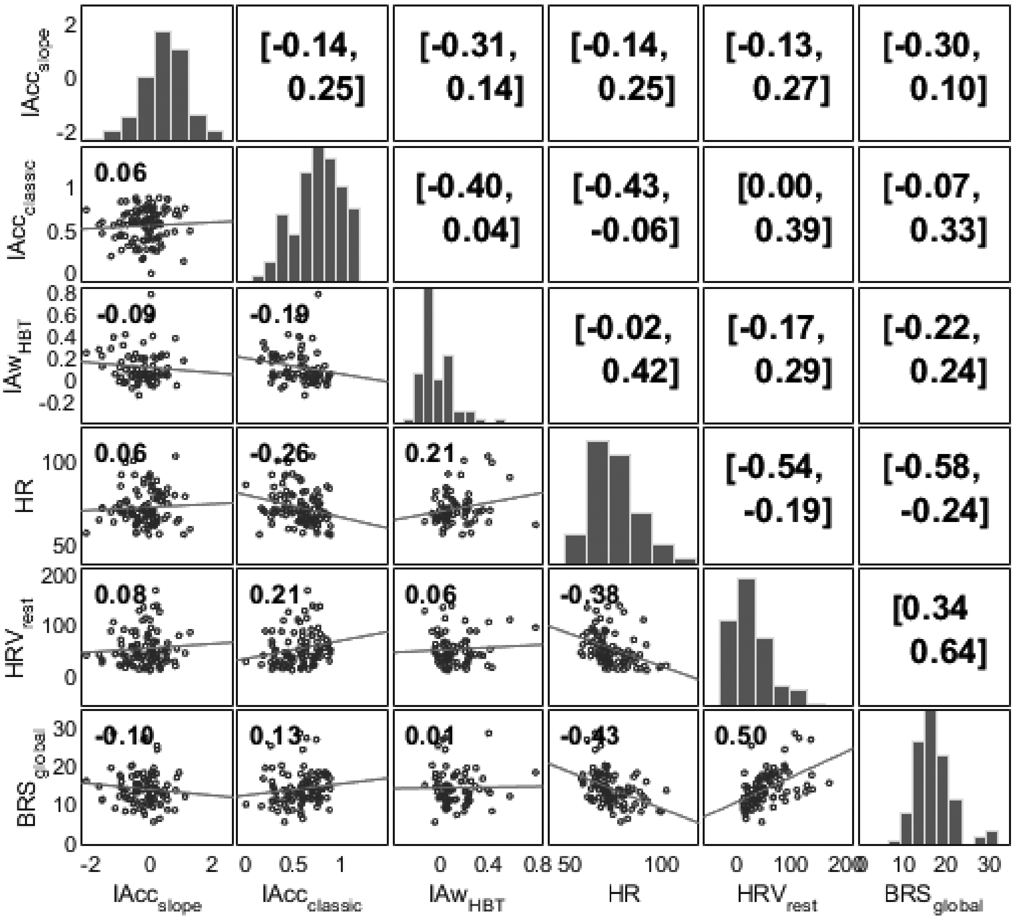

Figure 4. Baseline physiology correlogram. Scatter plots of baseline data and interoceptive HBT task scores between row and column variables, and variable distribution along diagonal axis. Each scatter plot of the lower half of the correlogram contains the corresponding correlation coefficient (r), and the boxes in the upper half contains the corresponding confidence intervals (i.e. [lower, upper]).

4. Discussion

This study introduced a new interoceptive accuracy measure, based on how changes in number of counted heartbeats corresponds to natural variations in heart rate. Our aim was to characterise this measure systematically, quantify its internal reliability and evaluate it in comparison to classical measures of interoceptive accuracy (Schandry, 1981). The contrast between this new measure and the classical Schandry (1981) accuracy calculation does shed valuable light on cardiac interoception, and particularly the use of the HBT task.

In testing the main hypothesis, our results revealed a negative mean $\text{IAcc}_{\text{slope}}$ across participants. This indicates that overall, heartbeat counting in the HBT task is indeed affected by changes in number of heartbeats. However, this was not in the intuitively expected direction: Participants tended to report fewer heartbeats on trials with an increase in heart rate (i.e. on trials with a higher number of heartbeats). Nonetheless, this outcome of our rigorous methodology indicates that cardiac interoceptive perception through counting is affected by changes in the number of heartbeats, arguably contradicting previous findings (Palmer et al., 2019; Windmann et al., 1999). The negative slope may be driven by other physiological changes that accompany increases in heart rate at rest. One plausible mechanism relates to the respiratory cycle and its coupling to the baroreflex: Inspiration decreases pressure in the chest, decreasing left ventricular filling and stroke volume. Lower baroreceptor activation triggers a baroreflexive increase in heart rate; when heartbeats are weaker, interbeat intervals shorten. This mechanism, interpolated over brief time periods, helps stabilise cardiac output (Draghici & Taylor, 2016; Riganello et al., 2018; Shaffer et al., 2014; Shaffer & Venner, 2013; Taylor & Eckberg, 1996). In the context of cardiac interoception and the HBT task, instances of increased heart rate occur during periods of weaker heartbeats. This itself could make heartbeats harder to detect. Interpolation over these periods of reduced signal would result in perhaps lower estimated counts of heartbeat occurrence. One interesting related observation in our study is that, rather than the participants simply reporting generally constant value across all trials (as they might if they were for example reporting their average heart rate), performance was instead affected by fluctuations in the frequency of heartbeats coupled reflexively to heartbeat strength. Thus, the observed negative slope between reported and actual heartbeats when performing the task, suggests that participants are sensitive to the strength of heartbeats, rather than the overall number of heartbeats.

There are of course other plausible explanations for why phasic increases in heart rate, typically driven at rest by cardiac acceleration during inspiration, may compromise the perception of heartbeats: The active movement of diaphragm and chest wall on inspiration may inhibit cardiorespiratory afferent signalling (likely within brainstem but potentially also at other levels of the neuraxis) or, alternatively, may generate competing respiratory interoceptive signals and sensations that overshadow weaker cardiac signals. Cardiac interoception in states of low arousal may be driven more by the *relative strength* or *salience* of heartbeats rather than the *amount* of heartbeats.

While the negative mean slope shows a highly significant effect across participants, the split-trial reliability measure showed that the $IAcc_{slope}$ measure did show a markedly high amount of noise in its scoring within participants compared to the $IAcc_{classic}$ measure (based on the classical Schandry (1981) calculation). Thus, the measure falls short in reliably identifying people strongly sensitive to changes in their number of heartbeats, which may be due to task difficulty coupled with between-subject variance in heartbeat perception and task performance. Future studies should consider screening for a minimal level of task performance to see if more nuanced trends in $IAcc_{slope}$ scores could be found.

Directly comparing the two cardiac accuracy measures (traditional Schandry (1981) calculation $IAcc_{classic}$ and the slope $IAcc_{slope}$ measure) revealed strong evidence against a relation between them. The $IAcc_{classic}$ measure is usually relatively consistent over time (Ferentzi et al., 2018; Herbert et al., 2011; Parkin et al., 2014; Pollatos, Traut-Mattausch, Schroeder, & Schandry, 2007), in line with the within-subject reliability that we also observed here. This lack of relation could imply $IAcc_{classic}$ scores to be unrelated to the measures ability to detect changes in the number of heartbeats. It remains unclear to what extent this would be due to a limitation of the $\mathrm{IAcc}_{\mathrm{slope}}$ measure, or the $\mathrm{IAcc}_{\mathrm{classic}}$ measure not being sensitive to changes in cardiac events.

There was also limited correspondence between accuracy and interoceptive metacognitive awareness when comparing accuracy score with metacognitive measures of the HBT task ($\mathrm{IAw}_{\mathrm{HBT}}$) and HBD task (meta-d'). However, this was also true for comparisons between awareness and $\mathrm{IAcc}_{\mathrm{classic}}$ scores. In terms of accuracy, performance in the HBT and HBD tasks may not necessarily correlate (Ring & Brener, 2018), hence the observed small relationships are not too unexpected.

Interoception is the signalling and processing of physiological signals, and as such estimates of interoceptive ability should be expected to map relatively well onto underlying physiological events. In relation to physiology, within-task heart rate (HR) showed a negative correlation with $\mathrm{IAcc}_{\mathrm{classic}}$ score, showing that participants with a lower HR tended to display higher accuracy scores, which is in line with $\mathrm{IAcc}_{\mathrm{classic}}$ being positively correlated with $\mathrm{bias}_{\mathrm{intercept}}$ (a measure showing the baseline tendency to over- or underreport the number of heartbeats). Given that people tend to under-report the number of heartbeats (Desmedt et al., 2018; Zamariola et al., 2018), and that lower HR was here associated with greater heartbeat detection accuracy in terms of $\mathrm{IAcc}_{\mathrm{classic}}$ scores, this would suggest that those with a lower HR either have an easier time perceiving their heartbeats, or that their lower average HR means that their true number of heartbeats is generally closer to their reported number of heartbeats, which would support the view that differences in accuracy are driven by differences in report bias (Zamariola et al., 2018).

At the core of this experiment is a reliance on the occurrence of natural heart rate variability (HRV) which gives rise to fluctuations in the number of heartbeats on a trial-by-trial basis. HRV has a cardiac cycle-length dependence: Higher heart rate yields a shorter IBI, which leaves less opportunity for variation, resulting in an overall lower HRV. In contrast, a slower heart rate yields longer IBI with greater room for variation, resulting in an overall greater HRV (Berntson et al., 2016). For this reason, it can be expected that task performance in the form of sensitivity to changes in number of heartbeats will be confounded by average HR in virtue of the task being more difficult the lower the HRV, given that there is less variation to pick up on. However, within these data, it could not be determined if there is a relation between baseline HRV (HRV_{rest}) and either $IAcc_{classic}$ or $IAcc_{slope}$, or between baseline baroreflex sensitivity (BRS_{global}) and either of the two accuracy measures. However, it is important to note that these HRV and BRS recordings were both taking during rest, as opposed to during task performance. As these measures involved applying pressure to the participant's wrist to record changes in blood pressure, the added pressure gives the participant a very clear sensation of their pulse for the duration of the recording, which would greatly affect the participant's task performance if done during the behavioural tasks (Murphy et al., 2019).

In conclusion, participants vary in the number of heartbeats counted in a specified timeframe as a function of changes in their heart rate (due to natural fluctuations in heart rate). Arguably, this was in the opposite to the expected direction, since participants reported fewer heartbeats as their heart rate increases, a finding that may be driven by changes in heartbeat strength in terms of moment-to-moment changes in blood pressure. Compared to other interoceptive tasks, within-participant performance was marked by a high degree of noise. This inverse slope suggests that participants may be more sensitive to relative strength and salience of heartbeats when performing cardiac interoception tasks, rather than the amount of heartbeats. Future work could usefully track beat-to-beat measures of ventricular ejection strength to test how detection parameters vary. Since we observed that the reported number of heartbeats did not increase with increases in heart rate, our findings also suggest that interoceptive monitoring at rest may not be a finely tuned process of tracking and detecting subtle internal events. Rather, available interoceptive information guides estimations that in turn may be partly shaped by higher order beliefs.

Declaration of Conflicting Interests

The authors declared they have no conflicts of interest with respect to the authorship or publication of this article.

Author contributions

D.E.O. Larsson, S.N. Garfinkel, and Z. Dienes conceived the study. D.E.O. Larsson and G. Esposito collected the data and conducted the statistical analyses. D.E.O. Larsson drafted the paper, and wrote together with S.N. Garfinkel, Z. Dienes and H.D. Critchley. All authors approved the final version of this manuscript for submission.

Acknowledgements

This work was supported by the Leverhulme Trust.

Appendix A: Analysis on pre-existing data

The extent that HB counting speed is modulated by subtle changes in heart rate on a trial-by-trial basis was assessed using pre-existing data from 219 healthy participants who performed the HB counting task. Each participant went through six trials of six different trial lengths (25, 30, 35, 40, 45, and 50 seconds) in a randomized order. For each participant, a general linear regression was plotted between actual and reported beats per minute (*BPM*), where sensitivity to changes in heart rate is reflected in the slope of each participant (see Figure 5).

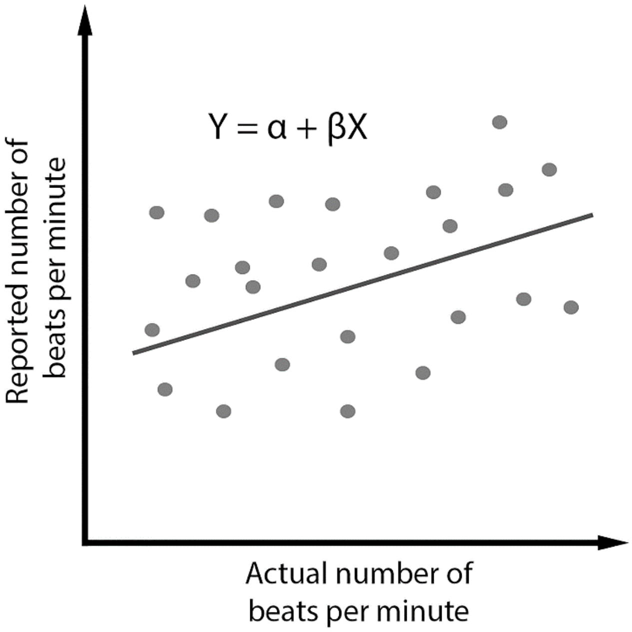

Figure A1. Plotting heart rate sensitivity. Illustration of a linear model plotted between actual and reported beats per minute, where α (the intercept) signifies the bias (i.e. base number of beats per minute reported when the actual number is zero), and β (the slope) used as an interoceptive accuracy estimate based on the slope ($IAcc_{slope}$).

This was performed as proof of principle, on data taken from a design that was not optimized for the present analysis (i.e. only 6 trials of differing length were performed per participant). The analysis yielded an average slope of 0.136 and standard deviation of 2.396 (see Figure 6). A t-test against the null hypothesis (a value of zero representing no effect) showed the effect to be non-significant (t(218) = 0.841, SE = 0.16, p = 0.402). The maximum slope that could be achieved if people detected their heart rate without error would be 1 BPM / BPM. Thus setting the model of the hypothesis (H1) for a Bayes factor as a half-normal with SD = 0.5 (i.e. half the maximum), gives $B_{H(0,\ 0.5)}$ = 0.67, which is insensitive (B >1/3), indicating that more data from each participant is necessary to find strong evidence for an effect. As the standard deviation is expected to shrink according to the square root of number of trials, the trial number was increased 3^2 times, rounded up to 60 trials. Thus, the new expected standard deviation was estimated as 2.396*sqrt(6/60) = 0.76. H1 was modelled as a half-normal with an SD of 0.136. Based on a sample mean of 0.136 and standard deviation of 0.76, a sample size of 100 participants would give a Bayes factor of 3.11 if the sample mean were 0.136, which would be enough to give moderate evidence for an effect.

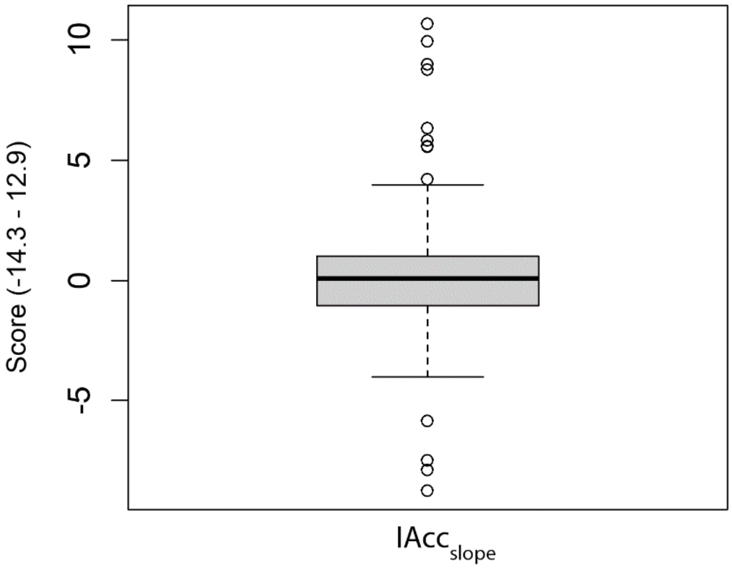

Figure A2. Mean $IAcc_{slope}$. Y-axis label signifies the range (min – max) of $IAcc_{slope}$ scores. Whiskers correspond to the 1.5 inter-quartile range of the lower and upper quartiles.

Appendix B: HBT task interoceptive and exteroceptive conditions

Comparing the different accuracy measures of the HBT task, Shapiro-Wilk test showed both the interoceptive ($IAcc_{slope}$, $IAcc_{classic}$) and exteroceptive ($EAcc_{slope}$, $EAcc_{classic}$) accuracy scores to be normally distributed. Comparing interoceptive and exteroceptive accuracy scores, a paired-sample t-test revealed a notable difference (t(99) = 2.60, CI = [0.01, 0.06]) between IAcc$_{classic}$ (mean = 0.56, SD = 0.20) and EAcc$_{classic}$ (mean = 0.53, SD = 0.16). There was also as a strong difference (t(99) = -10.36, CI = [-0.80, -0.54]) between IAcc$_{slope}$ (mean = -0.21, SD = 0.62) and EAcc$_{slope}$ (mean = 0.45, SD = 0.27), indicating that participants had significantly better performance in terms of the slope measure in the exteroceptive compared to interoceptive task, suggesting they had an easier time detecting changes in number of auditory tones compared to changes in number of heartbeats (see Figure 7). Participants did not show a difference in task confidence (t(99) = 0.39, CI = [-0.10, 0.15]) between the interoceptive (IC$_{HBT}$, mean = 2.25, SD = 0.63) and exteroceptive (EC$_{HBT}$, mean = 2.23, SD = 0.49) conditions. However, they did show greater metacognitive awareness (t(69) = -9.49, CI = [-0.38, -0.25]) in the exteroceptive (EAw$_{HBT}$, mean = 0.47, SD = 0.25) compared to interoceptive (IAw$_{HBT}$, mean = 0.13, SD = 0.15) condition.

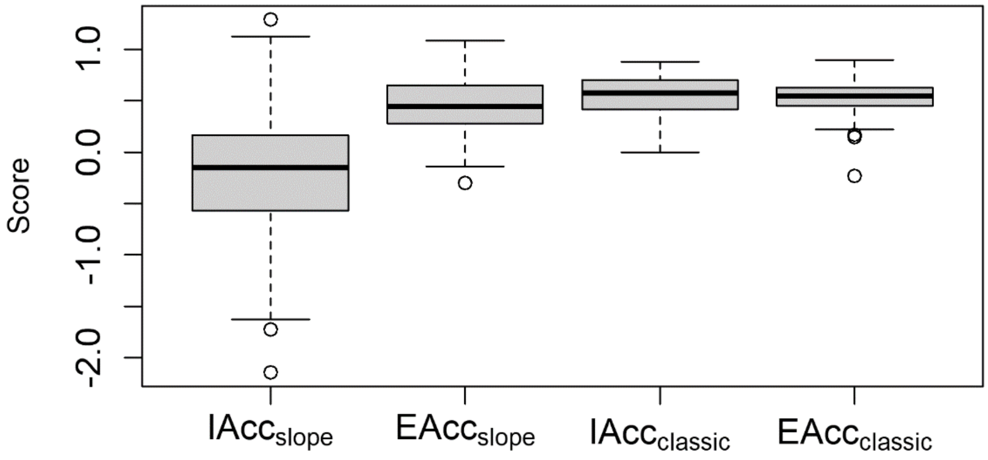

Figure B1. Mean accuracy scores from the HBT task in the form of interoceptive and exteroceptive slope (IAcc$_{slope}$ and EAcc$_{slope}$ respectively) and interoceptive and exteroceptive error ratio (IAcc$_{classic}$ and EAcc$_{classic}$ respectively). Whiskers correspond to the 1.5 inter-quartile range of the lower and upper quartiles.

When investigating the relationship between interoceptive and exteroceptive conditions, strong correlations were observed between IAcc$_{classic}$ and EAcc$_{classic}$ (r(98) = 0.78, CI = [0.69, 0.85]), IC$_{HBT}$ and EC$_{HBT}$ (r(98) = 0.41, CI = [0.23, 0.56]), indicating performance accuracy and confidence were related across interoceptive and exteroceptive domains. Meanwhile, the relationship between interoceptive and exteroceptive measures for the slope analysis ($IAcc_{slope}$ and $EAcc_{slope}$) ($r(98) = 0.15$, CI = [-0.06, 0.32]) nor for metacognitive awareness (IAw_{HBT} and EAw_{HBT}) ($r(68) = 0.20$, CI = [-0.03, 0.42]) were not as strong.

Bibliography

Ainley, V., Tsakiris, M., Pollatos, O., Schulz, A., & Herbert, B. M. (2020). Comment on “Zamariola et al. (2018), Interoceptive Accuracy Scores are Problematic: Evidence from Simple Bivariate Correlations”—The empirical data base, the conceptual reasoning and the analysis behind this statement are misconceived and do not support the authors’ conclusions. *Biological psychology, 152*, 107870. doi:https://doi.org/10.1016/j.biopsycho.2020.107870

Barrett, A., Dienes, Z., & Seth, A. (2013). Measures of Metacognition on Signal-Detection Theoretic Models. *Psychological methods, 18*. doi:10.1037/a0033268

Berntson, G. G., Quigley, K. S., Norman, G. J., & Lozano, D. (2016). *Cardiovascular psychophysiology* (Vol. 4). Cambridge: Cambridge University Press.

Brener, J., & Ring, C. (2016). Towards a psychophysics of interoceptive processes: the measurement of heartbeat detection. *Philosophical Transactions of the Royal Society B: Biological Sciences, 371*(1708), 20160015.

Craig, A. D. (2009). How do you feel--now? The anterior insula and human awareness. *Nature reviews neuroscience, 10*(1).

Critchley, H. D., & Garfinkel, S. N. (2015). Interactions between visceral afferent signaling and stimulus processing. *Frontiers in neuroscience, 9*, 286.

Critchley, H. D., & Garfinkel, S. N. (2017). Interoception and emotion. *Current opinion in psychology, 17*, 7-14.

Critchley, H. D., & Garfinkel, S. N. (2018). The influence of physiological signals on cognition. *Current opinion in behavioral sciences, 19*, 13-18.

Critchley, H. D., & Harrison, N. A. (2013). Visceral influences on brain and behavior. *Neuron, 77*(4), 624-638.

Desmedt, O., Luminet, O., & Corneille, O. (2018). The heartbeat counting task largely involves non-interoceptive processes: Evidence from both the original and an adapted counting task. *Biological psychology, 138*, 185-188.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in psychology, 5*, 781.

Dienes, Z. (2015). How Bayesian statistics are needed to determine whether mental states are unconscious.

Dienes, Z. (2019). How do I know what my theory predicts? *Advances in Methods and Practices in Psychological Science, 2*(4), 364-377.

Domschke, K., Stevens, S., Pfleiderer, B., & Gerlach, A. L. (2010). Interoceptive sensitivity in anxiety and anxiety disorders: an overview and integration of neurobiological findings. *Clinical psychology review, 30*(1), 1-11.

Draghici, A. E., & Taylor, J. A. (2016). The physiological basis and measurement of heart rate variability in humans. *Journal of physiological anthropology, 35*(1), 22.

Dunn, O. J., & Clark, V. (1969). Correlation coefficients measured on the same individuals. *Journal of the American Statistical Association, 64*(325), 366-377.

Ferentzi, E., Drew, R., Tihanyi, B. T., & Köteles, F. (2018). Interoceptive accuracy and body awareness–Temporal and longitudinal associations in a non-clinical sample. *Physiology & behavior, 184*, 100-107.

Garfinkel, S. N., Barrett, A. B., Minati, L., Dolan, R. J., Seth, A. K., & Critchley, H. D. (2013). What the heart forgets: Cardiac timing influences memory for words and is modulated by metacognition and interoceptive sensitivity. *Psychophysiology, 50*(6), 505-512.

Garfinkel, S. N., & Critchley, H. D. (2016). Threat and the body: how the heart supports fear processing. *Trends in cognitive sciences, 20*(1), 34-46.

Garfinkel, S. N., Minati, L., Gray, M. A., Seth, A. K., Dolan, R. J., & Critchley, H. D. (2014). Fear from the heart: sensitivity to fear stimuli depends on individual heartbeats. *Journal of Neuroscience, 34*(19), 6573-6582.

Garfinkel, S. N., Seth, A. K., Barrett, A. B., Suzuki, K., & Critchley, H. D. (2015). Knowing your own heart: distinguishing interoceptive accuracy from interoceptive awareness. *Biological psychology, 104*, 65-74.

Garfinkel, S. N., Tiley, C., O'Keeffe, S., Harrison, N. A., Seth, A. K., & Critchley, H. D. (2016). Discrepancies between dimensions of interoception in autism: Implications for emotion and anxiety. *Biological psychology, 114*, 117-126.

Gouveia, S., Rocha, A. P., Van De Borne, P., & Lago, P. (2005). *Assessing baroreflex sensitivity in the sequences technique: local versus global approach.* Paper presented at the Computers in Cardiology, 2005.

Herbert, B. M., Herbert, C., & Pollatos, O. (2011). On the relationship between interoceptive awareness and alexithymia: is interoceptive awareness related to emotional awareness? *Journal of personality, 79*(5), 1149-1175.

Katkin, E. S., Reed, S., & Deroo, C. (1983). *A methodological analysis of 3 techniques for the assessment of individual-differences in heartbeat detection.* Paper presented at the Psychophysiology.

Khalsa, S. S., Adolphs, R., Cameron, O. G., Critchley, H. D., Davenport, P. W., Feinstein, J. S., . . . Mehling, W. E. (2018). Interoception and mental health: a roadmap. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 3*(6), 501-513.

Kleckner, I. R., Wormwood, J. B., Simmons, W. K., Barrett, L. F., & Quigley, K. S. (2015). Methodological recommendations for a heartbeat detection-based measure of interoceptive sensitivity. *Psychophysiology, 52*(11), 1432-1440.

Koroboki, E., Zakopoulos, N., Manios, E., Rotas, V., Papadimitriou, G., & Papageorgiou, C. (2010). Interoceptive awareness in essential hypertension. *International Journal of Psychophysiology, 78*(2), 158-162.

Lee, M., & Wagenmakers, E.-J. (2013). Bayesian data analysis for cognitive science: A practical course. In: New York, NY: Cambridge University Press.

Mader, S. L., Palmer, R. M., & Rubenstein, L. Z. (1989). Effect of timing and number of baseline blood pressure determinations on postural blood pressure response. *Journal of the American Geriatrics Society, 37*(5), 444-446.

Magder, S. (2018). Heart-Lung interaction in spontaneous breathing subjects: the basics. *Annals of translational medicine, 6*(18).

Mayer, E. A. (2011). Gut feelings: the emerging biology of gut–brain communication. *Nature reviews neuroscience, 12*(8), 453.

Murphy, J., Brewer, R., Coll, M.-P., Plans, D., Hall, M., Shiu, S. S., . . . Bird, G. (2019). I feel it in my finger: Measurement device affects cardiac interoceptive accuracy. *Biological psychology, 148*, 107765.

Murphy, J., Millgate, E., Geary, H., Ichijo, E., Coll, M.-P., Brewer, R., . . . Bird, G. (2018). Knowledge of resting heart rate mediates the relationship between intelligence and the heartbeat counting task. *Biological psychology, 133*, 1-3.

Palmer, C., Ainley, V., & Tsakiris, M. (2019). Fine Tuning Your Heart: a novel method for measuring interoceptive accuracy.

Parkin, L., Morgan, R., Rosselli, A., Howard, M., Sheppard, A., Evans, D., . . . Dalgleish, T. (2014). Exploring the relationship between mindfulness and cardiac perception. *Mindfulness, 5*(3), 298-313.

Phillips, G. C., Jones, G. E., Rieger, E. J., & Snell, J. B. (1999). Effects of the presentation of false heart-rate feedback on the performance of two common heartbeat-detection tasks. *Psychophysiology, 36*(4), 504-510.

Pollatos, O., & Schandry, R. (2008). Emotional processing and emotional memory are modulated by interoceptive awareness. *Cognition & Emotion, 22*(2), 272-287.

Pollatos, O., Traut-Mattausch, E., Schroeder, H., & Schandry, R. (2007). Interoceptive awareness mediates the relationship between anxiety and the intensity of unpleasant feelings. *Journal of Anxiety Disorders, 21*(7), 931-943.

Quadt, L., Critchley, H. D., & Garfinkel, S. N. (2018). The neurobiology of interoception in health and disease. *Annals of the New York Academy of Sciences, 1428*(1), 112-128.

Ramshur, J. T. (2010). *Design, evaluation, and application of heart rate variability analysis software (HRVAS).* University of Memphis Memphis, TN,

RCore, T. (2013). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing [Internet]. Vienna, Austria. In.

Riganello, F., Larroque, S. K., Bahri, M. A., Heine, L., Martial, C., Carrière, M., . . . Chatelle, C. (2018). A Heartbeat Away From Consciousness: Heart Rate Variability Entropy can discriminate disorders of consciousness and is correlated with resting-state fMRI brain connectivity of the Central Autonomic Network. *Frontiers in neurology, 9*.

Ring, C., & Brener, J. (1996). Influence of beliefs about heart rate and actual heart rate on heartbeat counting. *Psychophysiology, 33*(5), 541-546.

Ring, C., & Brener, J. (2018). Heartbeat counting is unrelated to heartbeat detection: A comparison of methods to quantify interoception. *Psychophysiology, 55*(9), e13084.

Ring, C., Brener, J., Knapp, K., & Mailloux, J. (2015). Effects of heartbeat feedback on beliefs about heart rate and heartbeat counting: a cautionary tale about interoceptive awareness. *Biological psychology, 104*, 193-198.

Schandry, R. (1981). Heart beat perception and emotional experience. *Psychophysiology, 18*(4), 483-488.

Schulz, A., Matthey, J. H., Vögele, C., Schaan, V., Schächinger, H., Adler, J., . . . Michal, M. (2016). Cardiac modulation of startle is altered in depersonalization-/derealization disorder: evidence for impaired brainstem representation of baro-afferent neural traffic. *Psychiatry research, 240*, 4-10.

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.

Shaffer, F., & Venner, J. (2013). Heart rate variability anatomy and physiology. *Biofeedback, 41*(1), 13-25.

Silver, N. C., Hittner, J. B., & May, K. (2004). Testing dependent correlations with nonoverlapping variables: a monte carlo simulation. *The Journal of Experimental Education, 73*(1), 53-69.

Smith, R., Thayer, J. F., Khalsa, S. S., & Lane, R. D. (2017). The hierarchical basis of neurovisceral integration. *Neuroscience & Biobehavioral Reviews, 75*, 274-296.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117*(1), 34.

Taylor, J. A., & Eckberg, D. L. (1996). Fundamental relations between short-term RR interval and arterial pressure oscillations in humans. *Circulation, 93*(8), 1527-1532.

Werner, N. S., Peres, I., Duschek, S., & Schandry, R. (2010). Implicit memory for emotional words is modulated by cardiac perception. *Biological psychology, 85*(3), 370-376.

Whitehead, W. E., Drescher, V. M., Heiman, P., & Blackwell, B. (1977). Relation of heart rate control to heartbeat perception. *Biofeedback and Self-regulation, 2*(4), 371-392.

Windmann, S., Schonecke, O. W., Fröhlig, G., & Maldener, G. (1999). Dissociating beliefs about heart rates and actual heart rates in patients with cardiac pacemakers. *Psychophysiology, 36*(3), 339-342.

Zamariola, G., Maurage, P., Luminet, O., & Corneille, O. (2018). Interoceptive accuracy scores from the heartbeat counting task are problematic: Evidence from simple bivariate correlations. *Biological psychology, 137*, 12-17.

Zimprich, D., Nusser, L., & Pollatos, O. (2020). Are interoceptive accuracy scores from the heartbeat counting task problematic? A comment on Zamariola et al.(2018). *Biological psychology, 152*, 107868.