Systematic review and meta-analysis of the relationship between the heartbeat-evoked potential and interoception.

Michel-Pierre Coll*
McGill University
michel-pierre.coll@mcgill.ca

Hannah Hobson†
University of York
hannah.hobson@york.ac.uk

Geoffrey Bird‡
University of Oxford
geoff.bird@psy.ox.ac.uk

Jennifer Murphy§
Royal Holloway, University of London
jennifer.murphy@rhul.ac.uk

December 15, 2020

Abstract

The Heartbeat Evoked Potential (HEP) has been proposed as a neurophysiological marker of interoceptive processing. Despite its use to validate interoceptive measures and to assess interoceptive functioning in clinical groups, the empirical evidence for a relationship between HEP amplitude and interoceptive processing, including measures of such processing, is scattered across several studies with varied designs. The aim of this systematic review and meta-analysis was to examine the body of HEP-interoception research, and consider the associations the HEP shows with various direct and indirect measures of interoception, and how it is affected by manipulations of interoceptive processing. Specifically, we assessed the effect on HEP amplitude of manipulating attention to the heartbeat; manipulating participants' arousal; the association between the HEP and behavioural measures of cardiac interoception; and comparisons between healthy and clinical groups. Following database searches and screening, 45 studies were included in the systematic review and 42 in the meta-analyses. We noted variations in the ways individual studies have attempted to address key confounds, particularly the cardiac field artefact. Meta-analytic summaries indicated there were moderate to large effects of attention, arousal, and clinical status on the HEP, and a moderate association between HEP amplitude and behavioural measures of interoception. Problematically, the reliability of the meta-analytic effects documented here remain unknown, given the lack of standardised protocols for measuring the HEP. Thus, it is possible effects are driven by confounds such as cardiac factors or somatosensory effects.

Keywords: Heartbeat evoked potential — Interoceptive accuracy — Meta-analysis — HEP

Acknowledgements: MP Coll is supported by a fellowship from the Canadian Institutes of Health Research. Geoffrey Bird is supported by the UK's Economic & Social Research Council ES/R007527/1. The authors declare no conflict of interest.

* **Corresponding author,** Department of Psychology, 2001 McGill College, Montreal, Canada
† Department of Psychology, University of York, York, UK, YO10 5DD
‡ Department of Experimental Psychology, University of Oxford, Anna Watts Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford, UK, OX2
§ Department of Psychology, Royal Holloway, University of London, Egham, Surrey, TW20 0EX

Highlights

- The heartbeat-evoked potential (HEP) has been used as a measure of interoception in experimental and clinical contexts
- The evidence for a relationship between the HEP and interoception is scattered across multiple small studies with varied designs.
- Here, we performed a systematic review and meta-analysis of the evidence for a link between HEP amplitude and interoception.
- We found evidence for moderate to large effects of various interoceptive manipulations on HEP amplitude.
- However, we highlight various issues in the measurement and interpretation of the HEP that need to be addressed in future research.

1 Introduction

Interoception is defined as the perception of the internal state of the body, including hunger, thirst, cardiac and respiratory signals (Craig, 2003). Atypical interoception (both atypically high and low interoceptive abilities) has been theoretically and empirically linked with physical health-related problems including diabetes and obesity (Barrett & Simmons, 2015; Lemche et al., 2014; Simmons & DeVille, 2017) as well as a number of psychiatric conditions including depression and anxiety (Pollatos et al., 2009). Interoception has also been shown to play a role in our emotional experience, with atypical interoception associated with problems with emotion regulation (Füstös et al., 2013), emotion recognition (Terasawa et al., 2014) and increased levels of alexithymia (Murphy, Catmur, et al., 2018).

The increasing interest in the role of interoception in physical and mental health and emotional processes has been paralleled by discussion regarding interoception's measurement (Murphy, Brewer, et al., 2018). Both behavioural and self-report measures of interoception have been developed. Most behavioural measures of interoception are based on cardiac signals, including heartbeat counting (Dale & Anderson, 1978; Rainer Schandry, 1981) and heartbeat discrimination tasks (Whitehead et al., 1977; for a discussion of different variants see Brener & Ring, 2016), though other non-cardiac based tasks have also been developed (Murphy et al., 2018; van Dyck et al., 2016; for a discussion see Khalsa et al., 2018). However, behavioural tasks that require explicit judgements may be difficult for developing populations (e.g., younger children), or certain clinical populations (e.g. those with Autism Spectrum Disorder, a population in which interoception is being actively investigated; see Hatfield et al., 2019). Furthermore, even in typical adult populations, some research questions may require a measure of interoceptive processing that is implicit: for example, when researching the effect of another process (such as attention, arousal or increased cognitive load) or task on interoception. To circumvent the potential limitations of behavioural and self-reported interoception measures, researchers have sought a neurophysiological measure of interoceptive processing. One candidate is the Heartbeat Evoked Potential (HEP). The HEP is a scalp-recorded event-related potential (ERP), time-locked to participants' heartbeats (typically to the R-wave seen in the ECG, though the time interval between the R-wave peak and the onset of the HEP varies across studies). The HEP purportedly reflects the cortical processing of cardiac activity, and has been argued to represent a neurophysiological marker of interoception (Pollatos & Schandry, 2004). The HEP has already been used to validate other measures, including new behavioural tasks for interoceptive processing in infants (Maister et al., 2017) and brain stimulation techniques aimed at disrupting interoceptive processes (Pollatos et al., 2016). However, while several studies have used the HEP as a measure of interoception (Park & Blanke, 2019), there has yet to be a systematic meta-analysis of HEP research. This is crucial, both if we are to consider the HEP to be a reliable and valid indicator of cortical interoceptive processing of cardiac signals, and to refine methodologies that allow us to use the HEP in research and clinical settings. Indeed, Park and Blanke (2019) argued from their consideration of the HEP literature that there needed to be a standardised approach to HEP studies, and that currently the studies in the field were highly heterogeneous in their analytic approaches.

Here, we performed a systematic review and meta-analysis of studies relating HEPs to any direct or indirect measure or manipulation of interoception. This includes: manipulation of attention to interoceptive signals, manipulations of arousal, associations with behavioural measures of interoceptive ability, and comparisons between healthy controls and clinical groups hypothesised to have abnormal interoceptive processing. We review the relevant methodological aspects that vary across studies and report meta-analytical evidence for a link between interoception and the HEP. We finally discuss this evidence in addition to suggestions to improve future research using the HEP.

2 Methods

2.1 Literature search strategy

The literature search was conducted following the Preferred Reporting Items for Systematic review and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009) but since we expected studies to be highly heterogeneous in their designs and measurements, we did not register the systematic review protocol. We searched for articles on PubMed for studies published in English that were available online before the date the literature search was performed (November 5th, 2019) using different combinations of keywords (e.g. "heartbeat", "potential", "evoked", "ERP, "cortical", see Supplementary Table S1). The reference lists and citation reports of eligible studies were also consulted. We additionally performed a search of the Open Access Theses and Dissertations and the Open Grey database and on Google Scholar using the same keywords to search for documents not indexed in PubMed, but this search did not identify any additional eligible documents.

2.2 Criteria for considering studies for this review

To be included in this review, studies had to report scalp ERP data time-locked to heartbeats measured via ECG and at least one of the following 1) an objective measure or manipulation of interoception 2) assessment in clinical groups or 3) an arousal manipulation. Studies that met these criteria were included independently of the age, gender and clinical status of participants. One study that solely assessed interoception using a self-reported questionnaire was excluded from the quantitative meta-analysis due to the debate about what precisely is measured by these questionnaires (Mehling, 2016; Murphy et al., 2020). We additionally excluded 2 studies from the quantitative meta-analysis that did not report sufficient information about the results or analysed a variable other than HEP amplitude. The number of studies included in each step of the search process is shown in Figure 1.

2.3 Data collection and analysis

The screening of full text articles for their eligibility and data extraction from included articles was performed independently by two reviewers. A third independent reviewer compared the extracted data and flagged any inconsistencies. Inconsistencies were resolved through a group discussion between the three reviewers. We did not calculate the Kappa coefficient for the agreement between reviewers as agreement was extremely high and most inconsistencies between reviewers reflected mistakes or misunderstandings in the data extraction that were resolved through a second look at the articles. This procedure led to the selection of 45 studies for the methodological review and 42 for the quantitative meta-analysis. Included articles were published between 1991 and 2019 in 28 different journals (see Table S1 and asterisks in the reference list). The data reported in this review were manually extracted from the text of the published articles or accompanying materials and are available in Supplementary Table S1.

2.4 Types of interventions

To facilitate the compilation and comparison of results, we divided the statistical tests reported in each study into four main categories depending on the question addressed by the study; Attention, Performance, Clinical and Arousal. The Attention category includes studies manipulating interoception by directing attention towards or away from the body. Studies in the Performance category are those which related the amplitude of the HEP to performance on an interoceptive task either using correlation with behavioural measures, comparison of groups formed on the basis of their interoceptive performance (e.g. good or bad) or by comparing the HEP before and after an intervention that significantly impacted behavioural performance on an interoceptive task (e.g. training, neurostimulation). The Clinical category includes all studies using cross-sectional designs to compare typical comparison participants to clinical participants assumed to have atypical interoception. Finally, the Arousal category included studies in which the HEP amplitude was measured and compared across conditions of low and high arousal.

2.5 Methodological review

We reviewed the different methodological approaches used for the recording and processing of EEG data with the goal of measuring the HEP. To this end, we collected information on various methodological aspects (see Table S1). Here we report information regarding the participants included in these studies and the preprocessing steps used. We did not assess the risk of bias or the quality of the evidence since no standard guidelines exist for the types of studies included in this review.

3 Quantitative meta-analysis

We performed separate quantitative meta-analyses to assess the relationship between interoception and the amplitude of the HEP within each of the Attention, Performance, Arousal and Clinical categories described above. For each category, we manually extracted data from the manuscripts describing the statistical results for the test of interest with as much spatial and temporal precision as possible. However, when analyses reported a single statistic for several scalp locations and time points, this statistic was attributed to all

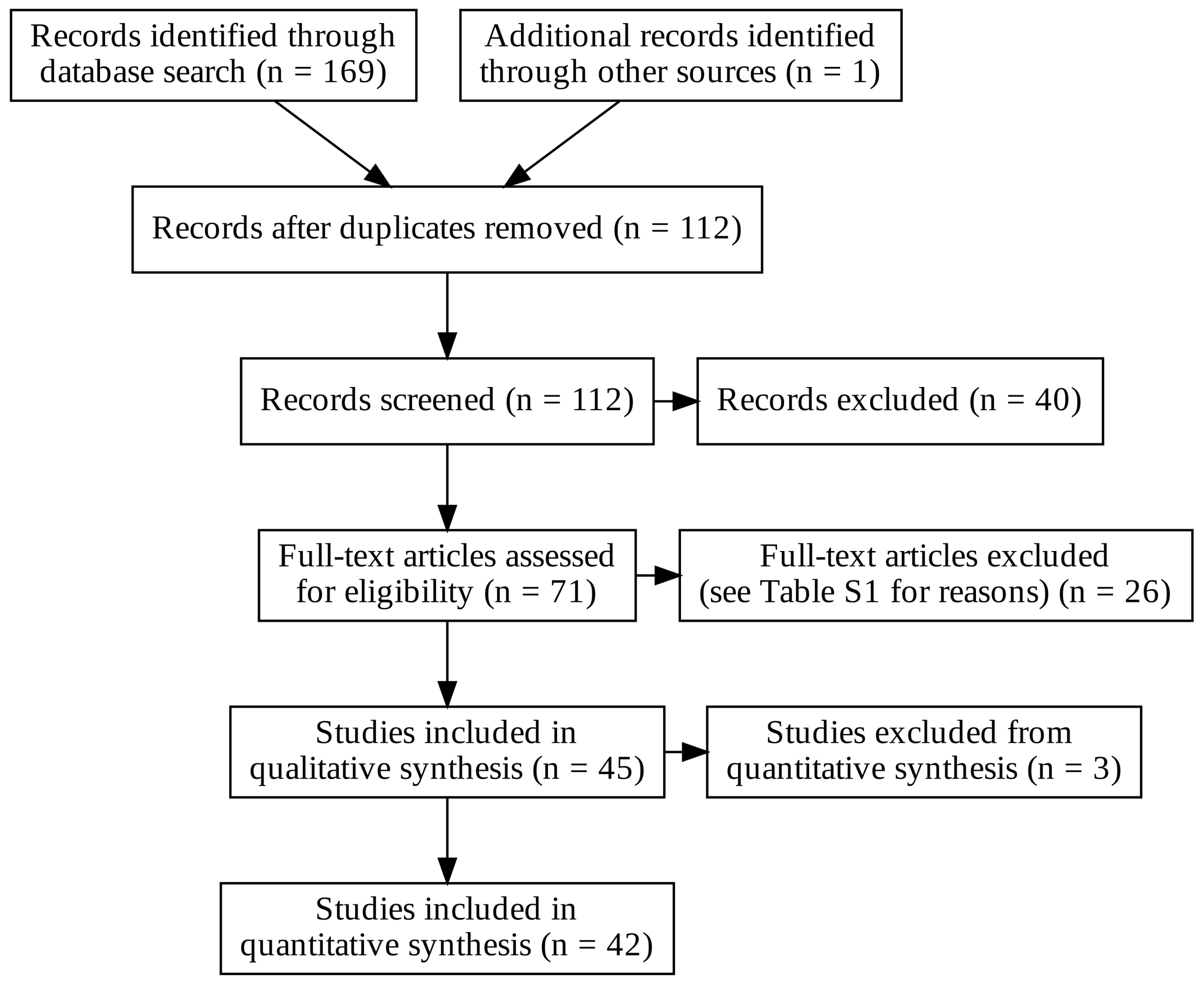

Figure 1: PRISMA chart illustrating the literature review and study selection process.

these scalp locations and time-points. When only the exact p-value was reported, the corresponding statistic was found using the appropriate distribution. When the effect was reported as significant or non-significant without test statistics, the effect size was calculated assuming $p = 0.05$ or $p = 0.5$, respectively (Cooper & Hedges, 1994; Moran et al., 2017). When only mean and standard deviations were reported for paired conditions, the t-statistic was calculated assuming a correlation of 0.7 between measurements. Results from non-parametric tests were interpreted using the corresponding parametric distributions. All test statistics were converted to Hedges' g, a standardized measure of difference that is less biased than Cohen's d, especially for small samples (Hedges, 1981). The Hedges' g was always calculated from the classical Cohen's d, meaning that the calculations were the same for within and between subjects design, which allowed the comparison of effect sizes across studies independently of the type of design used (Westfall, 2016). We also conducted heterogeneity analyses, reporting Q and I^2 for each meta-analytic effect reported.

Importantly, since the direction of HEP effects (increase or decrease in amplitude) can vary according to the EEG reference used and the scalp locations analysed, we chose to report the absolute effect size for all tests. Note that this provides a more liberal assessment of the meta-analytic effects and tests the null hypothesis there is no relationship between interoception and HEP without specifying a direction for this relationship.

We used different strategies to summarise effect sizes across studies for each category. First, we plotted the number of studies analysing each time point and scalp location. We then selected the most analysed time-window and scalp locations for each category and performed a random-effect meta-analysis of all studies reporting effects in this region of interest using the *metafor* package in R (Viechtbauer, 2010) with a restricted maximum-likelihood estimator. If a single study reported more than one statistic in the region of interest, we averaged the statistics prior to calculating the effect size. Second, to assess the spatio-temporal distribution of the effects, we performed a mass-univariate analysis for each category. In this analysis, we fitted a random effect model on the effect sizes for each channel of a 64 channel standard 10-20 layout and each time point between 0 and 700 ms post heartbeat if at least 3 studies reported effects at this time/location. The summary effect size at each time/location is reported and we highlight the times/locations for which the random-effects model reached the traditional significance threshold ($p < 0.05$). However, these values are indicative and should be interpreted with caution as they are not independent (same data spanned multiple locations/time-windows) and not corrected for multiple comparisons.

3.1 Data availability

All data and scripts used to process data and generate the figures, the PRISMA guidelines checklist and supplementary information and figures are available online at osf.io/mrac3/. All data processing and analyses were performed using custom Python and R scripts. EEG plots were produced using the *MNE-Python* package (Gramfort et al., 2013, 2014). Random effect meta-analyses were performed using the *metafor* R package (Viechtbauer, 2010).

4 Methodological review

4.1 Participants and design.

On average, the studies reviewed included 34.77 participants with an average of 21.80 participants per experimental group and 1.77 participants excluded from the analyses. Approximately half of the studies employed within-participants designs (N = 22) while the others employed a between group design (N = 22) or examined individual differences (N = 1). Of the 22 studies using a between-participants design, 15 of these included at least one clinical group. These groups included patients with major depression, cardiac problems, bipolar disorder, epilepsy, insomnia, obsessive compulsive disorder, panic disorder, substance abuse, diabetes, generalized anxiety disorder, multiple sclerosis, anorexia nervosa, depersonalization disorder, hypertension and nightmare disorder.

4.2 EEG and ECG recording and preprocessing

4.2.1 Recordings

On average, the EEG was recorded from 58.02 scalp electrodes (SD = 49.92, range: 2-256) and the ECG from 2.28 electrodes (SD=1.56, range: 1-12). The majority of studies measured the HEP referenced to the mastoids (N = 20), while other studies used the average reference (N = 14), earlobes (N = 5), vertex (N =3), nose (N =2) or did not report the reference used (N =1).

4.2.2 Epochs

All studies segmented epochs time-locked to the peak of the ECG r-wave for each heartbeat. All but one study (N = 41) corrected the HEP using the pre-stimulus baseline which started on average -168.18 ms before the r-peak (SD=57.15, range: -200- -50). Interestingly, of the 30 studies reporting the baseline period used to correct the HEP, 8 used a baseline period ending -125 to -25 ms relative to the r-peak to avoid including the onset of the r-wave in the baseline. The average epoch length was 755 ms (SD=200, range: 500-1300 ms) post r-peak. Only 15 studies reported the number of epochs included in each condition of interest which was on average 526.34 (SD=477.25, range: 70-1600).

4.2.3 Filtering and artifacts

Forty-two studies reported using a high-pass filter with an average cutoff frequency of 0.39 Hz (SD=0.36, range: 0.01-1) and 43 studies reported using a low-pass filter with an average cutoff frequency of 37.80 Hz (SD=17.65, range: 15-100). To remove noise from the EEG signal, 26 studies used independent component analysis, 23 reported visually inspecting the EEG to remove noisy epochs, 15 used an EOG correction algorithm and 14 automatically rejected trials exceeding a specific threshold. Note that some studies combined several of these approaches (see Table S1).

The HEP is time-locked to the electrical activity of the heart and it is therefore necessarily contaminated by this activity. Since the goal of recording the HEP is to measure the "cortical processing" of the heartbeat and not the activity of the heart muscle itself, several studies (N = 39) employed various strategies to attempt to remove or mitigate the influence of the cardiac field artifact (CFA) on the HEP. Among these strategies, the most popular (N = 17) was using independent component analysis to attempt to remove components associated with the CFA from the EEG signal. It should be noted that the studies using ICA varied with regards to the detail they supplied, with some simply noting that ICA was used to remove the CFA (Yoris et al., 2017), while others included the criteria used to identify components representing the CFA (e.g. Gentsch et al., 2019; Mai et al., 2018) or used packages to assist with semi-automatic detection of the CFA (e.g. Terhaar et al., 2012). Uniquely, Villena-González et al., (2017) did not include participants' data if a component matching the properties of CFA could not be identified. Other studies used the Hjorth source derivation (Hjorth, 1975; N = 3), current source density estimates (Perrin et al., 1989; N = 7) subtracted the ECG itself from the EEG (N=3) or subtracted the activity recorded from the nose (N = 2). To mitigate the effects of the CFA on the HEP, several studies used control analyses (N = 6) which consisted of either performing the analyses of interest on a second time-window assumed to be less contaminated by the CFA, performing analysis on the ECG itself to show that the effects of interest were not driven by changes in the ECG or analysing both CFA-corrected and CFA-uncorrected data to examine the impact of the CFA on the HEP. Finally, some studies claimed to avoid the CFA by focusing on a specific time-window that is thought to be less contaminated by the CFA (N = 13). This time window however was not consistent across the studies claiming to be selecting a time window to avoid the CFA: some examined windows starting from as early as 200 ms post- r-peak (Huang et al., 2018; Petzschner et al., 2019; Adrián Yoris et al., 2018) or 300 ms (de la Fuente et al., 2019), while others looked after 400 ms (Pollatos et al., 2016) or 455 ms (Schulz et al., 2013, 2018; Schulz, Ferreira de Sá, et al., 2015; Schulz, Köster, et al., 2015). This would seem to indicate that there is not an agreed time window in which the CFA can be assumed not to affect the data.

5 Quantitative meta-analysis

5.1 Effect of attention on the HEP

We identified 11 studies that assessed the effect of attention to the body on the HEP. As shown in Table 1, most of these studies compared the HEP during the heartbeat counting task (Schandry, 1981) with a rest or control condition. This comparison was performed in most cases in a 350-500 ms post r-peak time-window and at a fronto-central location (Figure 1A).

The mass-univariate analysis performed at each time point and location indicated that the strongest effects emerged at approximately 350 ms and peaked at 400 ms in central and fronto-central electrodes (Figure 1B). The random-effect meta-analysis carried-out in a region of interest covering the 350-500 ms time-window and locations Cz, C3, C4, Fz, F3, F4, FC3, FCz, FC4 included 10 studies and indicated that attention to the heart had a moderate and significant influence on the HEP amplitude ($g = 0.37$ [90% CI: 0.24-0.49], $p < 0.001$). No significant heterogeneity was observed across studies ($Q = 9.94$, $df = 9$, $p = 0.36$, $I^2 = 0.01\%$).

Table 1: Sample size, manipulation and statistical tests for studies included in the Attention category.

Study	N	Attention condition	Comparison condition	Statistical test
Montoya et al. 1993	26	HB counting	Tone counting	ANOVA
Couto et al. 2013	5	HB counting	Tone counting	Permutation t-test
García-Cordero et al. 2017	50	HB tapping	Tone tapping	Permutation t-test
Judah et al. 2018	37	False feedback	Rest	ANOVA
Leopold et al. 2001	50	HB counting	Tone counting	MANOVA
Mai et al. 2018	46	HB counting	Rest	T-test
Petzschner et al. 2019	19	Attend to heart	Attend to white noise	Statistical parametric mapping
Salamone et al. 2018	46	HB counting	Tone counting	Permutation t-test
Schulz et al. 2015b	46	HB counting	Rest	ANOVA
Terhaar et al. 2012	31	HB counting	Tone counting	ANOVA
Villena-González et al. 2018	8	HB counting	Visual counting	ANOVA

HB: Heartbeat

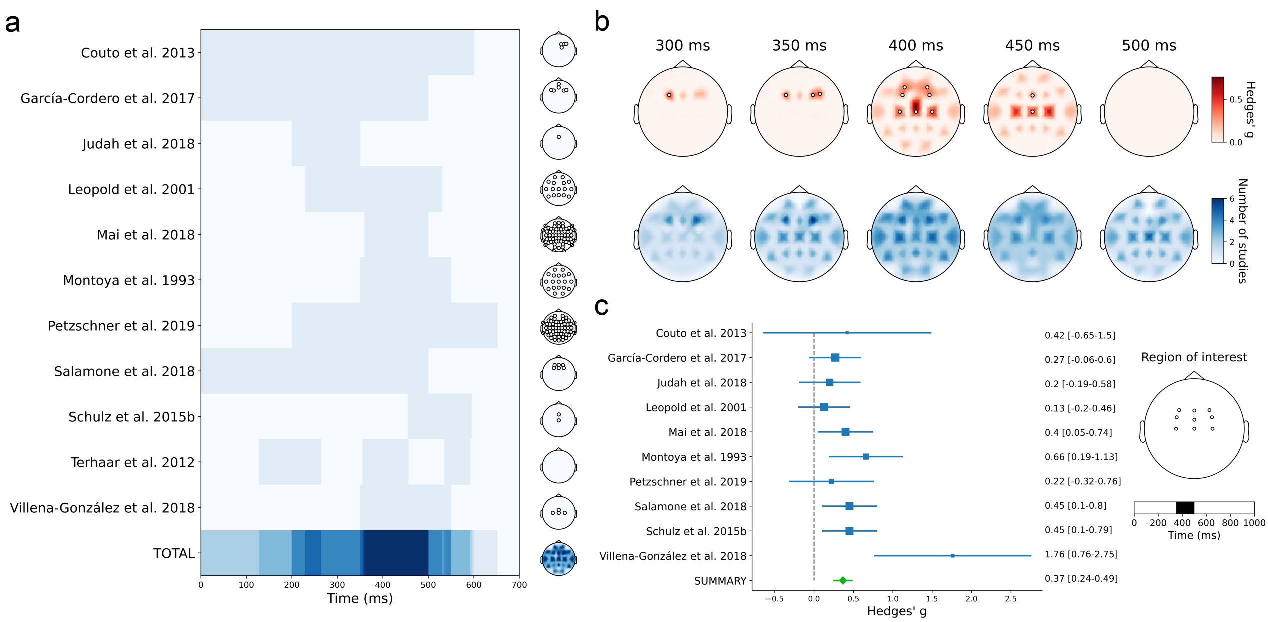

Figure 2: Meta-analysis of studies included in the Attention category. (a) Time-windows and scalp locations analysed in each study and total frequencies for each location and time-point. Empty scalp maps show studies that did not report locations or used the global field potential. **(b)** Results from the mass-univariate analyses performed at each time-point and location. The red scalp maps show the summary effect size at each location and the blue scalp maps show the number of studies considered at each location. Highlighted locations show significant effects at $p < 0.05$ uncorrected. **(c)** Forest-plot of the average effect size (+/- 90% confidence intervals) reported in the region of interest depicted on the right and the summary effect size from the random-effect meta-analysis (green). The size of the blue squares reflects the sample size in each study.

5.2 Relationship between interoceptive performance and HEP

We identified 20 relevant tests across 14 studies that related HEP amplitude to behavioural performance on an interoceptive task (Table 2). The majority of studies correlated the performance on the heartbeat counting task with the HEP. Other studies classified participants as good or bad heartbeat perceivers on the basis of their interoceptive accuracy and compared the HEP across these two groups. Two studies compared the HEP before and after a successful interoceptive training intervention, and one study compared the HEP after participants received transcranial magnetic stimulation to both a target structure thought to be involved in interoception (insula and somatosensory cortex) and a control stimulation site. The time-window of interest was more widespread than in the Attention category but the majority of studies in the Performance category investigated effects in a 200-300 ms post r-peak time-window and at fronto-central locations (Figure 2A).

The mass-univariate analysis performed at each time point and location indicated that the strongest effects peaked at 250 ms in central and fronto-central electrodes (Figure 2B). The random-effect meta-analysis conducted in a region of interest covering the 200-300 ms time-window and locations Cz, C1, C2, C3, C4, FCz, FC1, FC2, FC3, FC4, FC5, FC6 included 9 studies and indicated that performance on interoceptive tasks was moderately related to the HEP amplitude ($g = 0.39$ [90% CI: 0.23-0.54], $p < 0.001$). No significant heterogeneity was observed across studies ($Q = 6.99$, $df = 8$, $p = 0.54$, $I^2 = 0.01\%$). Since studies in this category were almost evenly split between a 200-300 ms time-window and a 400-500 ms time window, we also performed a region of interest analysis in this later time-window at the same locations. This analysis included 9 studies and also indicated a moderate effect size ($g = 0.35$ [90% CI: 0.19-0.52], $p < 0.001$) and no evidence of heterogeneity across studies ($Q = 12.18$, $df = 8$, $p = 0.2$, $I^2 = 26.54\%$).

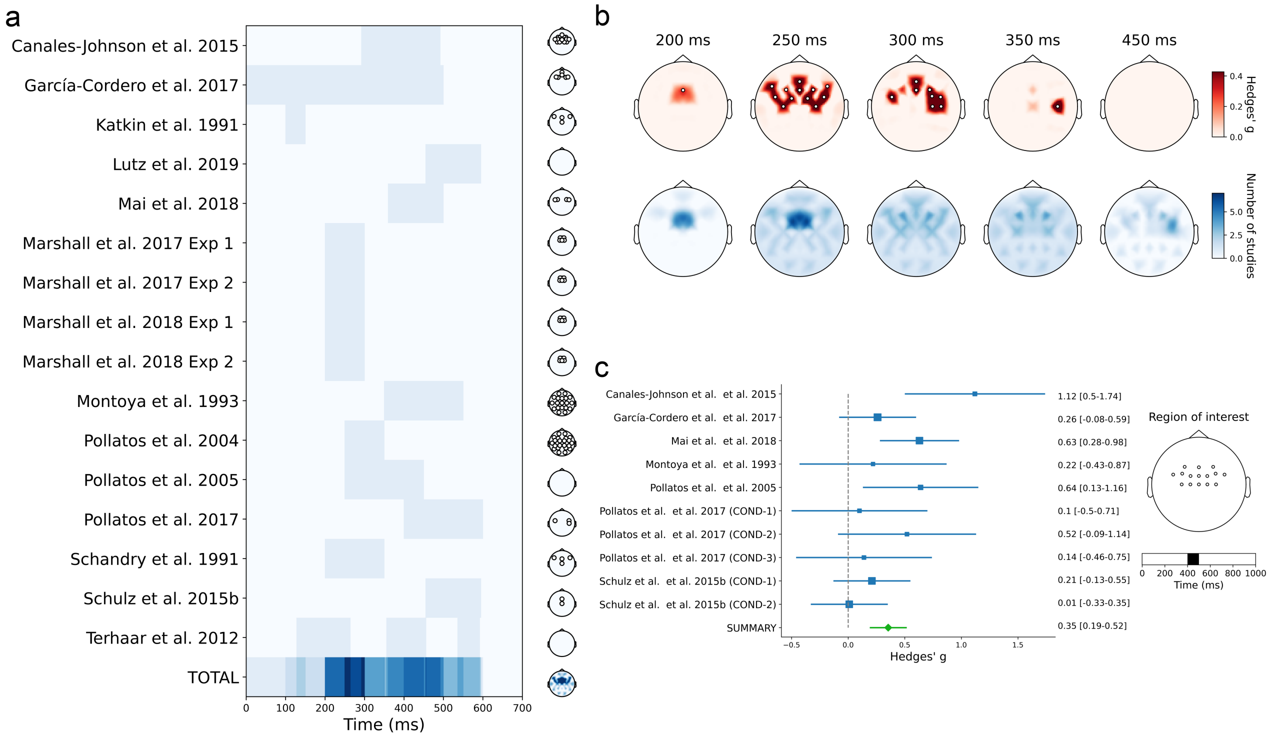

Figure 3: Meta-analysis of studies included in the Performance category. (a) Time-windows and scalp locations analysed in each study and total frequencies for each location and time-point. Empty scalp maps show studies that did not report locations or used the global field potential. **(b)** Results from the mass-univariate analyses performed at each time-point and location. The red scalp maps show the summary effect size at each location and the blue scalp maps show the number of studies considered at each location. Highlighted locations show significant effects at $p < 0.05$ uncorrected. **(c)** Forest-plot of the average effect size (+/- 90% confidence intervals) reported in the region of interest depicted on the right and the summary effect size from the random-effect meta-analysis (green). The size of the blue squares reflects the sample size in each study.

Table 2: Sample size(s), variable/manipulation and statistical tests for studies included in the Performance category.

Study	N 1	N 2	HEP variable	Performance variable /manipulation	Statistical test
Canales-Johnson et al. 2015	17	16	Mean amplitude	Good vs Bad at HB counting	Between groups (ANOVA)
García-Cordero et al. 2017	50		Mean amplitude	Pre vs Post training	Within group (t-test)
Katkin et al. 1991	12		Peak amplitude	HB detection standard deviation	Correlation (Spearman)
Lutz et al. 2019 (COND-1)	38		Mean amplitude HB counting	HB counting accuracy	Correlation (Pearson)
Lutz et al. 2019 (COND-2)	38		Mean amplitude rest	HB counting accuracy	Correlation (Pearson)
Mai et al. 2018	46		Mean amplitude	HB counting accuracy	Correlation (Spearman)
Marshall et al. 2017 Exp 1	25		Mean amplitude	HB counting accuracy	Correlation (Spearman)
Marshall et al. 2017 Exp 2	24		Mean amplitude	HB counting accuracy	Correlation (Spearman)
Marshall et al. 2018 Exp 1	25		Mean amplitude	HB counting accuracy	Correlation (Spearman)
Marshall et al. 2018 Exp 2	25		Mean amplitude	HB counting accuracy	Correlation (Spearman)
Montoya et al. 1993	11	16	Mean amplitude	Good vs Bad at HB counting	Between groups (ANOVA)
Pollatos et al. 2004	18	26	Mean amplitude	Good vs Bad at HB counting	Between groups (ANOVA)
Pollatos et al. 2005	22	22	Mean amplitude	Good vs Bad at HB counting	Between groups (ANOVA)
Pollatos et al. 2017 (COND-1)	15		Mean amplitude	HB counting accuracy	Correlation (Pearson)
Pollatos et al. 2017 (COND-2)	15		Mean amplitude	TMS insula vs TMS occipital	Within group (t-test)
Pollatos et al. 2017 (COND-3)	15		Mean amplitude	TMS somatosensory vs TMS occipital	Within group (t-test)
Schandry et al. 1991	20		Mean amplitude	Pre vs Post training	Between groups (ANOVA)
Schulz et al. 2015b (COND-1)	47		Mean amplitude	HB counting accuracy	Correlation (Pearson)
Schulz et al. 2015b (COND-2)	47		Mean amplitude	HB detection accuracy	Correlation (Pearson)
Terhaar et al. 2012	30		Mean amplitude	HB counting accuracy	Correlation (Pearson)

COND: Condition, HB: Heartbeat, TMS: Transcranial magnetic stimulation

5.3 Effect of Arousal on the HEP

Fourteen studies compared the HEP amplitude between high and low arousal conditions for a total of 22 tests (with some studies reporting more than one comparison). High arousal was induced using a variety of methods such as presenting affective cues, delivering pain stimulation, depriving participants of food or injecting cortisol (Table 3). Note that when a single study tested several similar conditions, we selected the comparison that was thought to maximize the difference in arousal. The analyses performed in the studies included in the arousal category were mostly focused on a 200-300 ms post r-peak time-window and at a fronto-central location (Figure 4A).

The mass-univariate analysis performed at each time point and location indicated that the strongest effects peaked at 250 ms in central and fronto-central electrodes (Figure 4B). The random-effect meta-analysis carried-out in a region of interest covering the 200-300 ms time-window and locations Cz, C1, C2, C3, C4, FCz, FC1, FC2, FC3, FC4, FC5, FC6 and AFz included 19 tests (Figure 4C) and indicated that changes in arousal had a large effect on the HEP amplitude (g = 0.72 [90% CI: 0.6-0.83], $p < 0.001$). No significant heterogeneity was observed across studies (Q = 17.49, df = 18, p = 0.5, I^2 ¡ 0.01%).

Table 3: Sample size(s), variable/manipulation and statistical tests for studies included in the Arousal category.

Study	N	Low arousal condition	High arousal condition	Statistical test
Fukushima et al. 2011	21	Physical judgement	Affective Judgement	Permutation t-test
Gentsch et al. 2018	17	Neutral faces repetition	Emotional faces repetition	Permutation t-test
Gray et al. 2007	10	Low cognitive effort	High cognitive effort	Statistical parametric mapping
Ito et al. 2019	27	Positive thoughts	Negative thoughts	ANOVA
Luft et al. 2015	16	Neutral cues	Affective cues	Permutation t-test
MacKinnon et al. 2013 (COND-1)	26	Rest eyes closed	Positive memory	T-test
MacKinnon et al. 2013 (COND-2)	26	Rest eyes closed	Negative memory	T-test
MacKinnon et al. 2013 (COND-3)	26	Rest eyes closed	Breathing	T-test
Marshall et al. 2017 Exp 1	25	Neutral faces repetition	Angry faces repetition	T-test
Marshall et al. 2017 Exp 2 (COND-1)	24	Neutral faces cued repetition	Angry faces cued repetition	T-test
Marshall et al. 2017 Exp 2 (COND-2)	24	Neutral faces uncued repetition	Angry faces uncued repetition	T-test
Marshall et al. 2018 Exp 1 (COND-1)	25	Neutral faces repetition	Angry faces repetition	T-test
Marshall et al. 2018 Exp 1 (COND-2)	25	Neutral faces repetition	Pain faces repetition	T-test
Marshall et al. 2018 Exp 2 (COND-1)	25	Neutral faces repetition	Sad faces repetition	T-test
Marshall et al. 2018 Exp 2 (COND-2)	25	Neutral faces repetition	Happy faces repetition	T-test
Marshall et al. 2019	30	Neutral faces repetition	Angry faces repetition	Permutation t-test
Park et al. 2016 (COND-1)	16	No stroking	Asynchronous stroking	Permutation t-test
Park et al. 2016 (COND-2)	16	Synchronous stroking	Asynchronous stroking	Permutation t-test
Schulz et al. 2013	16	Placebo infusion	Cortisol infusion	T-test
Schulz et al. 2015a	16	Food deprivation	Satiated	ANOVA
Sel et al. 2018	25	Control condition	Faces presented in synchrony with heartbeat	Permutation t-test
Shao et al. 2011	21	No-pain control	Painful stimulation	ANOVA

COND: Condition

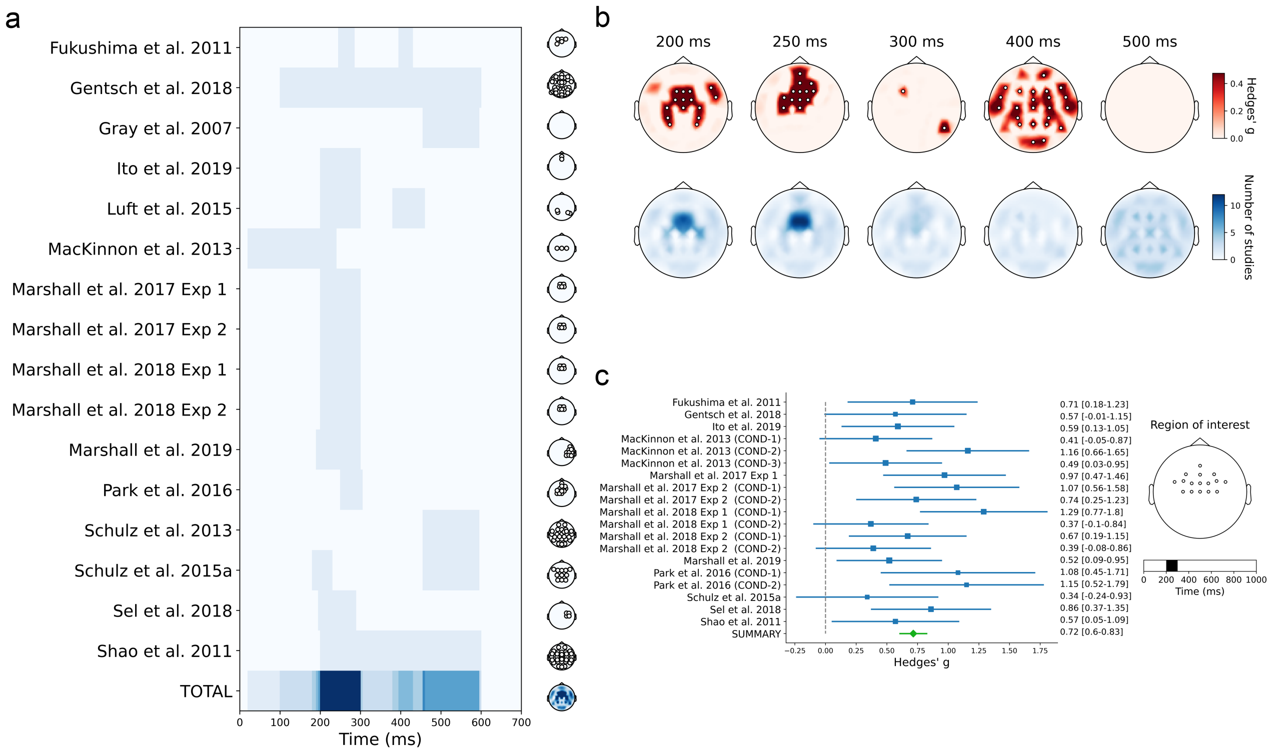

Figure 4: Meta-analysis of studies included in the Arousal category. **(a)** Time-windows and scalp locations analysed in each study and total frequencies for each location and time-point. Empty scalp maps show studies that did not report locations or used the global field potential. **(b)** Results from the mass-univariate analyses performed at each time-point and location. The red scalp maps show the summary effect size at each location and the blue scalp maps show the number of studies considered at each location. Highlighted locations show significant effects at $p < 0.05$ uncorrected. **(c)** Forest-plot of the average effect size (+/- 90% confidence intervals) reported in the region of interest depicted on the right and the summary effect size from the random-effect meta-analysis (green). The size of the blue squares reflects the sample size in each study.

5.4 Differences in HEP between clinical and control groups

We identified 14 studies comparing the HEP amplitude between control and clinical participants. The HEP was measured at rest in some studies while others measured the HEP during an interoceptive task or sleep (see Table 4). There was substantial variability in the time-windows and locations used for analyses but most studies in this category focused on the 400-500 ms time-window and fronto-central locations (Figure 5A).

The mass-univariate analysis performed at each time point and location indicated that the strongest effects peaked at 400 ms in right fronto-central electrodes (Figure 5B). The random-effect meta-analysis carried-out in a region of interest covering the 400-500 ms time-window and locations C4, F4, Cz, Fz, FC4, FCz, FPz, FC6, C1, C2, FC1, FC2, AFz, FC8, F8, AF4 and AF8 included 13 tests (Figure 5) and indicated that there was a moderate effect of clinical group on the HEP amplitude (g = 0.49 [90% CI: 0.35-0.63], $p < 0.001$) and no significant heterogeneity was observed across studies (Q = 9.94, df = 12, p = 0.62, $I^2 < 0.01$

Table 4: Sample size(s), variable/manipulation and statistical tests for studies included in the Clinical category.

Study	N low	N high	Low inter group	High inter group	Condition	Statistical test
de la Fuente et al. 2019	25	25	Substance abuse	Control	HTT pre + post feedback	Permutation T-test
Judah et al. 2018	30	19	Control	Social anxiety	Other	ANOVA
Leopold et al. 2001	25	25	Diabetic	Control	HCT + tone counting	T-test
Lutz et al. 2019	19	19	Control	Anorexia	HCT	ANOVA
Müller et al. 2016	34	31	Borderline personality disorder	Control	Rest	ANOVA
Pang et al. 2019 (COND-1)	25	15	Control	Generalized anxiety disorder	Rest eyes open	ANOVA
Pang et al. 2019 (COND-2)	25	15	Control	Generalized anxiety disorder	Rest eyes closed	ANOVA
Perogamvros et al. 2019 (COND-1)	11	11	Control	Nightmare disorder	REM sleep	T-test
Perogamvros et al. 2019 (COND-2)	11	11	Control	Nightmare disorder	Awake	T-test
Perogamvros et al. 2019 (COND-3)	11	11	Control	Nightmare disorder	NREM sleep	T-test
Salamone et al. 2018	34	46	Multiple sclerosis	Control	HTT + beat tapping	Permutation T-test
Schulz et al. 2015b (COND-1)	23	24	Depersonalization disorder	Control	Rest	T-test
Schulz et al. 2015b (COND-2)	23	24	Depersonalization disorder	Control	HCT	T-test
Schulz et al. 2018	30	25	Deceased from cardiac arrest	Survived cardiac arrest	Rest	ANOVA
Terhaar et al. 2012	16	16	Depressed	Control	HCT + tone counting	ANOVA
Wei et al. 2017	32	32	Control	Insomnia	Rest	Permutation T-test
Yoris et al. 2017	15	25	Control	Obsessive compulsive disorder	HTT + beat tapping	Permutation T-test
Yoris et al. 2018	24	26	Hypertension	Control	HTT + beat tapping	Permutation T-test

COND: Condition, HCT: Heartbeat counting task, HTT: Heartbeat tapping task, N/REM: Non/Rapid eye movement

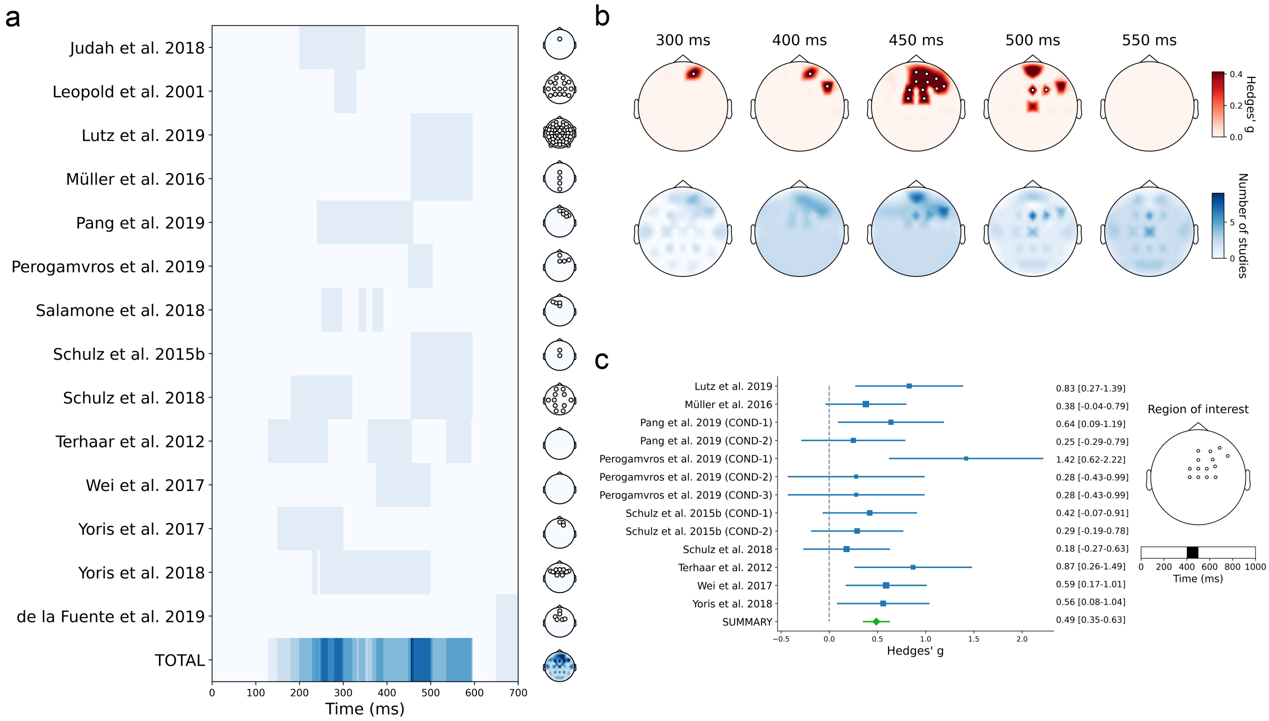

Figure 5: Meta-analysis of studies included in the Clinical category. (a) Time-windows and scalp locations analysed in each study and total frequencies for each location and time-point. Empty scalp maps show studies that did not report locations or used the global field potential. **(b)** Results from the mass-univariate analyses performed at each time-point and location. The red scalp maps show the summary effect size at each location and the blue scalp maps show the number of studies considered at each location. Highlighted locations show significant effects at $p < 0.05$ uncorrected. **(c)** Forest-plot of the average effect size (+/- 90% confidence intervals) reported in the region of interest depicted on the right and the summary effect size from the random-effect meta-analysis (green). The size of the blue squares reflects the sample size in each study.

6 Discussion

The HEP is regularly used as a measure of interoceptive processing yet the evidence for a link between the HEP and interoceptive processes remains scattered across multiple small studies. Furthermore, the strength of this evidence is potentially affected by the variability of the methods used to measure the HEP and interoceptive processes (Park & Blanke, 2019). Here we performed a systematic review and meta-analysis of studies linking the HEP to different types of interoceptive processing. We found significant meta-analytic evidence for a moderate to large relationship between HEP amplitude and various measures/manipulations of interoception. However, we found substantial variability in the methods used to process and measure the HEP. Furthermore, the tasks and manipulations used to link the HEP to interoception might be subject to confounds not adequately addressed by most studies measuring the HEP. We detail these points below.

We found evidence for a moderate effect of orienting attention towards the heart on the HEP, which was strongest at fronto-central locations in a 400 to 500 ms post r-peak time-window. Although the increased HEP amplitude with heart-focussed attention may reflect an increase in interoceptive processing, it must also be acknowledged that it might be the case that attention towards the heart increases attention towards the somatosensory sensations associated with heart beats (Khalsa et al., 2009; Park & Blanke, 2019). Therefore, it remains unclear if changes in the HEP in these studies is really due to increased interoceptive processing, or an increase in somatosensory processing. Future work should attempt to rule out the potential somatosensory contribution to the HEP.

Furthermore, across 16 studies we found a moderate relationship between performance on interoceptive tasks and the HEP amplitude in a 400-500 ms time window and at fronto-central locations. However, most studies reviewed used the HCT which has received considerable criticism, notably due to the fact that HCT scores are tied to heart rates (Ring et al., 1994; Zamariola et al., 2018) beliefs about heart rate (Brener & Ring, 2016; Murphy et al., 2018; Windmann et al., 1999), and may be subject to response bias (e.g., Desmedt et al., 2018). While there is a clear need to develop better tasks of interoceptive accuracy, variants of the heartbeat detection task (Brener & Ring, 2016) have been designed which address issues with the HCT and should be considered for use in future HEP studies. Interestingly, the only study reviewed using both the HCT and the HDT found a strong effect with the HCT, but not with the HDT (Schulz, Köster, et al., 2015).

The suggestion that the HEP at least partly reflects differences in cardiac dynamics rather than differences in interoceptive processes (Dirlich et al., 1997) is supported by the fact that we found that studies inducing various states of arousal (which changes cardiac dynamics) had the largest effect on the HEP. As noted, cardiac dynamics have also been found to influence performance on behavioural tasks of cardiac interoceptive accuracy (Knapp-Kline & Kline, 2005; O'Brien et al., 1998; Ring et al., 1994; Zamariola et al., 2018). Whether the influence of cardiac dynamics on interoception (as measured by both the HEP and measures of cardiac interoceptive accuracy) should be cause for concern depends somewhat on one's definition of interoception and what we seek to infer when comparing individuals; for example, if we care only whether individuals can perceive their heartbeat, then it is perhaps unimportant if differences between individuals are driven by physiology (e.g., blood pressure or cardiac output). However, if we wish to infer that individual variations reflect individual differences at a higher-order (i.e., more cognitive) level (e.g., 'the cortical processing of heartbeats'), then the influence of cardiac dynamics becomes problematic. In any case, additional work is necessary to validate the link between the HEP and interoception using well-controlled interoceptive tasks and by taking into account how changes in cardiac dynamics influences the HEP and tasks of interoception.

To move forward however, the field of interoception will need to consider various methodological issues with HEP measurements. Our methodological review indicates that, as is common in the ERP and EEG literature (Coll, 2018; Hobson & Bishop, 2017; Luck & Gaspelin, 2017), the HEP literature is characterized by studies with small samples, considerable analytical variability, and no direct and pre-registered replications. Of specific interest to the HEP and in contrast to other well-established ERPs, is the fact that there is an apparent lack of consensus regarding the preprocessing and measurement of the HEP (Park & Blanke, 2019). For example, the reference scheme used for analysis varied considerably which makes the direct comparison of the HEP shape and amplitude across studies difficult. Additionally, several studies attempted to remove the CFA artifact from the HEP using various methods with the rationale that the HEP should reflect the cortical processing of heartbeats and not the heartbeats themselves. The variability in the use of correction procedures introduces significant discrepancy in the shape of the HEP as can be easily noted by comparing the ERP time-course plots from studies using different correction methods. Some studies reported using a later time-window to avoid influence from the CFA but, across all studies, analyses were distributed across most time-points in a 200-600 ms post r-peak window raising questions regarding the optimal time-window to use for HEP measurements. Future work needs to address these issues by comparing the effect of different CFA correction methods in different time-windows to establish which method and time-window best serve the goal of separating the HEP measured on the scalp from the ECG. Adopting the same measurement and processing method in all HEP studies would drastically reduce the analytical heterogeneity in the HEP literature, and lead to more reliable and robust results provide standards to assess the quality of the evidence and risk of bias in future meta-analyses.

Given the issues highlighted above and the usual caveats of cross-sectional studies, it is unclear what can be concluded from the studies comparing the HEP between healthy participants and those with clinical diagnoses. While the meta-analytic effect of the clinical group on the HEP was of moderate size, multiple (non-interoceptive) factors could explain this difference and it seems clear that further validation of the measure is needed before it can be used reliably to infer anything about interoceptive processing in clinical populations.

Several limitations to this study need to be acknowledged. First, we did not consider publication bias (Rosenthal, 1979) and within-study biases created by the fact that most studies reviewed only reported significant effects from several tests. This means that all the effect sizes obtained are necessarily inflated. Large pre-registered studies are necessary to accurately estimate the effect size of the relationship between interoception and the HEP. Furthermore, we did not consider the direction of the HEP effects since this direction (more negative or more positive amplitude) depends on various preprocessing steps and measurement choices, making its meaning unclear. Again, this probably leads to inflated effect sizes since it is possible that some studies found effects in opposite directions. This is of special importance for the comparison between clinical groups and healthy participants, as some clinical conditions have been claimed to result in increased interoceptive ability and/or increased attention to interoceptive signals (Domschke et al., 2010). Adopting a consensus on how to measure the HEP would allow future research to consider and interpret the direction of changes in HEP amplitude across conditions.

In conclusion, we found meta-analytic evidence for a relationship between the HEP and interoception but methodological concerns raise questions regarding the validity of this relationship. Additional work is needed to assess, 1) how to best measure the HEP, 2) the basic characteristics of this ERP, 3) its relationship to cardiac dynamics, and 4) the link between the HEP and interoception which should be investigated using multiple, carefully controlled, interoceptive tasks. In the absence of a consensus on how to measure the HEP and clear evidence for its validity as an interoceptive measure, the use of the HEP to gain insights on the interoceptive functioning in clinical populations or across groups, or as a way to validate new measures of interoception, is premature.

Studies included in systematic review only

(Baranauskas et al., 2017; Schandry et al., 1986; Yuan et al., 2007)

Studies included in both systematic review and meta-analysis

(Canales-Johnson et al., 2015; de la Fuente et al., 2019; Fukushima et al., 2011; García-Cordero et al., 2017; Gentsch et al., 2019; Gray et al., 2007; Huang et al., 2018; Ito et al., 2019; Judah et al., 2018; Katkin et al., 1991; Leopold Schandry, 2001; Luft Bhattacharya, 2015; Lutz et al., 2019; MacKinnon et al., 2013; Mai et al., 2018; Marshall et al., 2017, 2018, 2019; Montoya et al., 1993; Müller et al., 2015; Pang et al., 2019; Park et al., 2016; Perogamvros et al., 2019; Petzschner et al., 2019; Pollatos et al., 2005, 2016; Pollatos Schandry, 2004; Salamone et al., 2018; R. Schandry Weitkunat, 1990; Schulz et al., 2013, 2018; Schulz, Ferreira de Sá, et al., 2015; Schulz, Köster, et al., 2015; Sel et al., 2017; Shao et al., 2011; Terhaar et al., 2012; Villena-González et al., 2017; Wei et al., 2016; A. Yoris et al., 2017; Adrián Yoris et al., 2018)

References

* Studies included in the systematic review only
* Studies included in both systematic review and meta-analysis

*Baranauskas, M., Grabauskaitė, A., & Griškova-Bulanova, I. (2017). Brain responses and self-reported indices of interoception: Heartbeat evoked potentials are inversely associated with worrying about body sensations. *Physiology & Behavior*, 180, 1–7. https://doi.org/10.1016/j.physbeh.2017.07.032

Barrett, L. F., & Simmons, W. K. (2015). Interoceptive predictions in the brain. *Nature Reviews Neuroscience*, 16(7), 419–429.https://doi.org/10.1038/nrn3950

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. https://doi.org/10.1098/rstb.2016.0015

Craig, A. (2003). Interoception: The sense of the physiological condition of the body. *Current Opinion in Neurobiology*, 13(4), 500–505. https://doi.org/10.1016/S0959-4388(03)00090-4

**Canales-Johnson, A., Silva, C., Huepe, D., Rivera-Rei, Á., Noreika, V., Garcia, M. del C., Silva, W., Ciraolo, C., Vaucheret, E., Sedeño, L., Couto, B., Kargieman, L., Baglivo, F., Sigman, M., Chennu, S., Ibáñez, A., Rodríguez, E., & Bekinschtein, T. A. (2015). Auditory Feedback Differentially Modulates Behavioral and Neural Markers of Objective and Subjective Performance When Tapping to Your Heartbeat. *Cerebral Cortex (New York, N.Y.: 1991)*, 25(11), 4490–4503. https://doi.org/10.1093/cercor/bhv076

Coll, M.-P. (2018). Meta-analysis of ERP investigations of pain empathy underlines methodological issues in ERP research. *Social Cognitive Affective Neuroscience*, 15.

Cooper, H., & Hedges, L. V. (1994). *The handbook of research synthesis.* xvi, 573.

Dale, A., & Anderson, D. (1978). Information Variables in Voluntary Control and Classical Conditioning of Heart Rate: Field Dependence and Heart-Rate Perception. Perceptual and Motor Skills, 47(1), 79–85. https://doi.org/10.2466/pms.1978.47.1.79

**de la Fuente, A., Sedeño, L., Vignaga, S. S., Ellmann, C., Sonzogni, S., Belluscio, L., García-Cordero, I., Castagnaro, E., Boano, M., Cetkovich, M., Torralva, T., Cánepa, E. T., Tagliazucchi, E., Garcia, A. M., & Ibañez, A. (2019). Multimodal neurocognitive markers of interoceptive tuning in smoked cocaine. *Neuropsychopharmacology*, 44(8), 1425–1434. https://doi.org/10.1038/s41386-019-0370-3

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. https://doi.org/10.1016/j.biopsycho.2018.09.004

Dirlich, G., Vogl, L., Plaschke, M., & Strian, F. (1997). Cardiac field effects on the EEG. *Electroencephalography and Clinical Neurophysiology*, 102(4), 307–315. https://doi.org/10.1016/s0013-4694(96)96506-2

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. https://doi.org/10.1016/j.cpr.2009.08.008

**Fukushima, H., Terasawa, Y., & Umeda, S. (2011). Association between interoception and empathy: Evidence from heartbeat-evoked brain potential. International Journal of *Psychophysiology: Official Journal of the International Organization of Psychophysiology*, 79(2), 259–265. https://doi.org/10.1016/j.ijpsycho.2010.10.015

Füstös, J., Gramann, K., Herbert, B. M., & Pollatos, O. (2013). On the embodiment of emotion regulation: Interoceptive awareness facilitates reappraisal. *Social Cognitive and Affective Neuroscience*, 8(8), 911–917. https://doi.org/10.1093/scan/nss089

**García-Cordero, I., Esteves, S., Mikulan, E. P., Hesse, E., Baglivo, F. H., Silva, W., García, M. D. C., Vaucheret, E., Ciraolo, C., García, H. S., Adolfi, F., Pietto, M., Herrera, E., Legaz, A., Manes, F., García, A. M., Sigman, M., Bekinschtein, T. A., Ibáñez, A., & Sedeño, L. (2017). Attention, in and Out: Scalp-Level and Intracranial EEG Correlates of Interoception and Exteroception. *Frontiers in Neuroscience*, 11, 411. https://doi.org/10.3389/fnins.2017.00411

**Gentsch, A., Sel, A., Marshall, A. C., & Schütz-Bosbach, S. (2019). Affective interoceptive inference: Evidence from heart-beat evoked brain potentials. *Human Brain Mapping*, 40(1), 20–33. https://doi.org/10.1002/hbm.24352

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., & Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, 7. https://doi.org/10.3389/fnins.2013.00267

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Parkkonen, L., & Hämäläinen, M. S. (2014). MNE software for processing MEG and EEG data. *NeuroImage*, 86, 446–460. https://doi.org/10.1016/j.neuroimage.2013.10.027

**Gray, M. A., Taggart, P., Sutton, P. M., Groves, D., Holdright, D. R., Bradbury, D., Brull, D., & Critchley, H. D. (2007). A cortical potential reflecting cardiac function. *Proceedings of the National Academy of Sciences of the United States of America*, 104(16), 6818–6823. https://doi.org/10.1073/pnas.0609509104

Hatfield, T. R., Brown, R. F., Giummarra, M. J., & Lenggenhager, B. (2019). Autism spectrum disorder and interoception: Abnormalities in global integration? *Autism*, 23(1), 212–222. https://doi.org/10.1177/1362361317738392

Hedges, L. V. (1981). Distribution Theory for Glass's Estimator of Effect size and Related Estimators. *Journal of Educational Statistics*, 6(2), 107–128. https://doi.org/10.3102/10769986006002107

Hjorth, B. (1975). An on-line transformation of EEG scalp potentials into orthogonal source derivations. *Electroencephalography and Clinical Neurophysiology*, 39(5), 526–530. https://doi.org/10.1016/0013-4694(75)90056-5

Hobson, H. M., & Bishop, D. V. M. (2017). The interpretation of mu suppression as an index of mirror neuron activity: Past, present and future. *Royal Society Open Science*, 4(3), 160662. https://doi.org/10.1098/rsos.160662

**Huang, C., Gevirtz, R. N., Onton, J., & Criado, J. R. (2018). Investigation of vagal afferent functioning using the Heartbeat Event Related Potential. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, 131, 113–123. https://doi.org/10.1016/j.ijpsycho.2017.06.007

**Ito, Y., Shibata, M., Tanaka, Y., Terasawa, Y., & Umeda, S. (2019). Affective and temporal orientation of thoughts: Electrophysiological evidence. Brain Research, 1719, 148–156. https://doi.org/10.1016/j.brainres.2019.05.041

**Judah, M. R., Shurkova, E. Y., Hager, N. M., White, E. J., Taylor, D. L., & Grant, D. M. (2018). The relationship between social anxiety and heartbeat evoked potential amplitude. *Biological Psychology*, 139, 1–7. https://doi.org/10.1016/j.biopsycho.2018.09.013

**Katkin, E. S., Cestaro, V. L., & Weitkunat, R. (1991). Individual differences in cortical evoked potentials as a function of heartbeat detection ability. *The International Journal of Neuroscience*, 61(3–4), 269–276.

Khalsa, S. S., Adolphs, R., Cameron, O. G., Critchley, H. D., Davenport, P. W., Feinstein, J. S., Feusner, J. D., Garfinkel, S. N., Lane, R. D., Mehling, W. E., Meuret, A. E., Nemeroff, C. B., Oppenheimer, S., Petzschner, F. H., Pollatos, O., Rhudy, J. L., Schramm, L. P., Simmons, W. K., Stein, M. B., … Paulus, M. P. (2018). Interoception and Mental Health: A Roadmap. *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging,* 3(6), 501–513. https://doi.org/10.1016/j.bpsc.2017.12.004

Khalsa, S. S., Rudrauf, D., Feinstein, J. S., & Tranel, D. (2009). The pathways of interoceptive awareness. *Nature Neuroscience*, 12(12), 1494–1496. https://doi.org/10.1038/nn.2411

Knapp-Kline, K., & Kline, J. P. (2005). Heart rate, heart rate variability, and heartbeat detection with the method of constant stimuli: Slow and steady wins the race. *Biological Psychology*, 69(3), 387–396. https://doi.org/10.1016/j.biopsycho.2004.09.002

Lemche, A. V., Chaban, O. S., & Lemche, E. (2014). Alexithymia as a risk factor for type 2 diabetes mellitus in the metabolic syndrome: A cross-sectional study. *Psychiatry Research*, 215(2), 438–443. https://doi.org/10.1016/j.psychres.2013.12.004

**Leopold, C., & Schandry, R. (2001). The heartbeat-evoked brain potential in patients suffering from diabetic neuropathy and in healthy control persons. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 112(4), 674–682.

Luck, S. J., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology*, 54(1), 146–157. https://doi.org/10.1111/psyp.12639

**Luft, C. D. B., & Bhattacharya, J. (2015). Aroused with heart: Modulation of heartbeat evoked potential by arousal induction and its oscillatory correlates. *Scientific Reports*, 5, 15717. https://doi.org/10.1038/srep15717

**Lutz, A. P. C., Schulz, A., Voderholzer, U., Koch, S., van Dyck, Z., & Vögele, C. (2019). Enhanced cortical processing of cardio-afferent signals in anorexia nervosa. *Clinical Neurophysiology*, 130(9), 1620–1627. https://doi.org/10.1016/j.clinph.2019.06.009

**MacKinnon, S., Gevirtz, R., McCraty, R., & Brown, M. (2013). Utilizing heartbeat evoked potentials to identify cardiac regulation of vagal afferents during emotion and resonant breathing. *Applied Psychophysiology and Biofeedback*, 38(4), 241–255. https://doi.org/10.1007/s10484-013-9226-5

**Mai, S., Wong, C. K., Georgiou, E., & Pollatos, O. (2018). Interoception is associated with heartbeat-evoked brain potentials (HEPs) in adolescents. *Biological Psychology*, 137, 24–33. https://doi.org/10.1016/j.biopsycho.2018.06.007

Maister, L., Tang, T., & Tsakiris, M. (2017). Neurobehavioral evidence of interoceptive sensitivity in early infancy. *ELife*, 6, e25318. https://doi.org/10.7554/eLife.25318

**Marshall, A. C., Gentsch, A., Jelinčić, V., & Schütz-Bosbach, S. (2017). Exteroceptive expectations modulate interoceptive processing: Repetition-suppression effects for visual and heartbeat evoked potentials. *Scientific Reports*, 7(1), 16525. https://doi.org/10.1038/s41598-017-16595-9

**Marshall, A. C., Gentsch, A., Schröder, L., & Schütz-Bosbach, S. (2018). Cardiac interoceptive learning is modulated by emotional valence perceived from facial expressions. *Social Cognitive and Affective Neuroscience*, 13(7), 677–686. https://doi.org/10.1093/scan/nsy042

**Marshall, A. C., Gentsch, A., & Schütz-Bosbach, S. (2019). Interoceptive cardiac expectations to emotional stimuli predict visual perception. *Emotion.* https://doi.org/10.1037/emo0000631

Mehling, W. (2016). Differentiating attention styles and regulatory aspects of self-reported interoceptive sensibility. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 371(1708). https://doi.org/10.1098/rstb.2016.0013

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, T. P. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLOS Medicine*, 6(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

**Montoya, P., Schandry, R., & Müller, A. (1993). Heartbeat evoked potentials (HEP): Topography and influence of cardiac awareness and focus of attention. *Electroencephalography and Clinical Neurophysiology*, 88(3), 163–172.

Moran, T. P., Schroder, H. S., Kneip, C., & Moser, J. S. (2017). Meta-analysis and psychophysiology: A tutorial using depression and action-monitoring event-related potentials. International Journal of Psychophysiology, 111, 17–32. https://doi.org/10.1016/j.ijpsycho.2016.07.001

**Müller, L. E., Schulz, A., Andermann, M., Gäbel, A., Gescher, D. M., Spohn, A., Herpertz, S. C., & Bertsch, K. (2015). Cortical Representation of Afferent Bodily Signals in Borderline Personality Disorder: Neural Correlates and Relationship to Emotional Dysregulation. *JAMA Psychiatry*, 72(11), 1077–1086. https://doi.org/10.1001/jamapsychiatry.2015.1252

Murphy, J., Brewer, R., Hobson, H., Catmur, C., & Bird, G. (2018). Is alexithymia characterised by impaired interoception? Further evidence, the importance of control variables, and the problems with the Heartbeat Counting Task. *Biological Psychology*, 136, 189–197. https://doi.org/10.1016/j.biopsycho.2018.05.010

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

Murphy, J., Catmur, C., & Bird, G. (2018). Alexithymia is associated with a multidomain, multidimensional failure of interoception: Evidence from novel tests. *Journal of Experimental Psychology. General*, 147(3), 398–408. https://doi.org/10.1037/xge0000366

Murphy, J., Millgate, E., Geary, H., Ichijo, E., Coll, M.-P., Brewer, R., Catmur, C., & Bird, G. (2018). Knowledge of resting heart rate mediates the relationship between intelligence and the heartbeat counting task. *Biological Psychology*, 133, 1–3. https://doi.org/10.1016/j.biopsycho.2018.01.012

O'Brien, W. H., Reid, G. J., & Jones, K. R. (1998). Differences in heartbeat awareness among males with higher and lower levels of systolic blood pressure. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, 29(1), 53–63. https://doi.org/10.1016/s0167-8760(98)00004-x

**Pang, J., Tang, X., Li, H., Hu, Q., Cui, H., Zhang, L., Li, W., Zhu, Z., Wang, J., & Li, C. (2019). Altered Interoceptive Processing in Generalized Anxiety Disorder—A Heartbeat-Evoked Potential Research. *Frontiers in Psychiatry*, 10. https://doi.org/10.3389/fpsyt.2019.00616

**Park, H.-D., Bernasconi, F., Bello-Ruiz, J., Pfeiffer, C., Salomon, R., & Blanke, O. (2016). Transient Modulations of Neural Responses to Heartbeats Covary with Bodily Self-Consciousness. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 36(32), 8453–8460. https://doi.org/10.1523/JNEUROSCI.0311-16.2016

Park, H.-D., & Blanke, O. (2019). Heartbeat-evoked cortical responses: Underlying mechanisms, functional roles, and methodological considerations. *NeuroImage*, 197, 502–511. https://doi.org/10.1016/j.neuroimage.2019.04.081

**Perogamvros, L., Park, H.-D., Bayer, L., Perrault, A. A., Blanke, O., & Schwartz, S. (2019). Increased heartbeat-evoked potential during REM sleep in nightmare disorder. *NeuroImage: Clinical*, 22, 101701. https://doi.org/10.1016/j.nicl.2019.101701

Perrin, F. E., Echallier, J. F., Bertrand, O., & Pernier, J. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology*. https://doi.org/10.1016/0013-4694(89)90180-6

**Petzschner, F. H., Weber, L. A., Wellstein, K. V., Paolini, G., Do, C. T., & Stephan, K. E. (2019). Focus of attention modulates the heartbeat evoked potential. *NeuroImage*, 186, 595–606. https://doi.org/10.1016/j.neuroimage.2018.11.037

**Pollatos, O., Herbert, B. M., Mai, S., & Kammer, T. (2016). Changes in interoceptive processes following brain stimulation. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 371(1708). https://doi.org/10.1098/rstb.2016.0016

**Pollatos, O., Kirsch, W., & Schandry, R. (2005). Brain structures involved in interoceptive awareness and cardioafferent signal processing: A dipole source localization study. *Human Brain Mapping*, 26(1), 54–64. https://doi.org/10.1002/hbm.20121

**Pollatos, O., & Schandry, R. (2004). Accuracy of heartbeat perception is reflected in the amplitude of the heartbeat-evoked brain potential. *Psychophysiology*, 41(3), 476–482. https://doi.org/10.111/1469-8986.2004.00170.x

Pollatos, O., Traut-Mattausch, E., & Schandry, R. (2009). Differential effects of anxiety and depression on interoceptive accuracy. *Depression and Anxiety*, 26(2), 167–173. https://doi.org/10.1002/da.20504

Ring, C., Liu, X., & Brener, J. (1994). Cardiac stimulus intensity and heartbeat detection: Effects of tilt-induced changes in stroke volume. *Psychophysiology*, 31(6), 553–564. https://doi.org/10.1111/j.1469-8986.1994.tb02348.x

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86(3), 638–641. https://doi.org/10.1037/0033-2909.86.3.638

**Salamone, P. C., Esteves, S., Sinay, V. J., García-Cordero, I., Abrevaya, S., Couto, B., Adolfi, F., Martorell, M., Petroni, A., Yoris, A., Torquati, K., Alifano, F., Legaz, A., Cassará, F. P., Bruno, D., Kemp, A. H., Herrera, E., García, A. M., Ibáñez, A., & Sedeño, L. (2018). Altered neural signatures of interoception in multiple sclerosis. *Human Brain Mapping*, 39(12), 4743–4754. https://doi.org/10.1002/hbm.24319

*Schandry, R., Sparrer, B., & Weitkunat, R. (1986). From the heart to the brain: A study of heartbeat contingent scalp potentials. *The International Journal of Neuroscience*, 30(4), 261–275.

**Schandry, R., & Weitkunat, R. (1990). Enhancement of heartbeat-related brain potentials through cardiac awareness training. *The International Journal of Neuroscience*, 53(2–4), 243–253.

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

**Schulz, A., Ferreira de Sá, D. S., Dierolf, A. M., Lutz, A., van Dyck, Z., Vögele, C., & Schächinger, H. (2015). Short-term food deprivation increases amplitudes of heartbeat-evoked potentials. *Psychophysiology*, 52(5), 695–703. https://doi.org/10.1111/psyp.12388

**Schulz, A., Köster, S., Beutel, M. E., Schächinger, H., Vögele, C., Rost, S., Rauh, M., & Michal, M. (2015). Altered patterns of heartbeat-evoked potentials in depersonalization/derealization disorder: Neurophysiological evidence for impaired cortical representation of bodily signals. *Psychosomatic Medicine*, 77(5), 506–516. https://doi.org/10.1097/PSY.0000000000000195

**Schulz, A., Stammet, P., Dierolf, A. M., Vögele, C., Beyenburg, S., Werer, C., & Devaux, Y. (2018). Late heartbeat-evoked potentials are associated with survival after cardiac arrest. *Resuscitation*, 126, 7–13. https://doi.org/10.1016/j.resuscitation.2018.02.009

**Schulz, A., Strelzyk, F., Ferreira de Sá, D. S., Naumann, E., Vögele, C., & Schächinger, H. (2013). Cortisol rapidly affects amplitudes of heartbeat-evoked brain potentials: Implications for the contribution of stress to an altered perception of physical sensations? *Psychoneuroendocrinology*, 38(11), 2686–2693. https://doi.org/10.1016/j.psyneuen.2013.06.027

**Sel, A., Azevedo, R. T., & Tsakiris, M. (2017). Heartfelt Self: Cardio-Visual Integration Affects Self-Face Recognition and Interoceptive Cortical Processing. *Cerebral Cortex (New York, N.Y.: 1991)*, 27(11), 5144–5155. https://doi.org/10.1093/cercor/bhw296

**Shao, S., Shen, K., Wilder-Smith, E. P. V., & Li, X. (2011). Effect of pain perception on the heartbeat evoked potential. C*linical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 122(9), 1838–1845. https://doi.org/10.1016/j.clinph.2011.02.014

Simmons, W. K., & DeVille, D. C. (2017). Interoceptive contributions to healthy eating and obesity. *Current Opinion in Psychology*, 17, 106–112. https://doi.org/10.1016/j.copsyc.2017.07.001

Terasawa, Y., Moriguchi, Y., Tochizawa, S., & Umeda, S. (2014). Interoceptive sensitivity predicts sensitivity to the emotions of others. *Cognition & Emotion*, 28(8), 1435–1448. https://doi.org/10.1080/02699931.2014.888988

**Terhaar, J., Viola, F. C., Bär, K.-J., & Debener, S. (2012). Heartbeat evoked potentials mirror altered body perception in depressed patients. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 123(10), 1950–1957. https://doi.org/10.1016/j.clinph.2012.02.086

van Dyck, Z., Vögele, C., Blechert, J., Lutz, A. P. C., Schulz, A., & Herbert, B. M. (2016). The Water Load Test As a Measure of Gastric Interoception: Development of a Two-Stage Protocol and Application to a Healthy Female Population.*PLOS ONE*, 11(9), e0163574. https://doi.org/10.1371/journal.pone.0163574

Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software*, 36(1), 1–48. https://doi.org/10.18637/jss.v036.i03

**Villena-González, M., Moënne-Loccoz, C., Lagos, R. A., Alliende, L. M., Billeke, P., Aboitiz, F., López, V., & Cosmelli, D. (2017). Attending to the heart is associated with posterior alpha band increase and a reduction in sensitivity to concurrent visual stimuli. *Psychophysiology*, 54(10), 1483–1497. https://doi.org/10.1111/psyp.12894

**Wei, Y., Ramautar, J. R., Colombo, M. A., Stoffers, D., Gómez-Herrero, G., van der Meijden, W. P., Te Lindert, B. H. W., van der Werf, Y. D., & Van Someren, E. J. W. (2016). I Keep a Close Watch on This Heart of Mine: Increased Interoception in Insomnia. *Sleep*, 39(12), 2113–2124. https://doi.org/10.5665/sleep.6308

Westfall, J. (2016, March 25). Five different "Cohen's d" statistics for within-subject designs. Cookie Scientist. http://jakewestfall.org/blog/index.php/2016/03/25/five-different-cohens-d-statistics-for-within-subject-designs/

Whitehead, W. E., Drescher, V. M., Heiman, P., & Blackwell, B. (1977). Realtion of heart rate control to heartbeat perception. *Biofeedback and Self-Regulation*, 2(4), 317–392.

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. https://doi.org/10.1017/s0048577299980381

**Yoris, A., García, A. M., Traiber, L., Santamaría-García, H., Martorell, M., Alifano, F., Kichic, R., Moser, J. S., Cetkovich, M., Manes, F., Ibáñez, A., & Sedeño, L. (2017). The inner world of overactive monitoring: Neural markers of interoception in obsessive-compulsive disorder. *Psychological Medicine*, 47(11), 1957–1970. https://doi.org/10.1017/S0033291717000368

**Yoris, Adrián, Abrevaya, S., Esteves, S., Salamone, P., Lori, N., Martorell, M., Legaz, A., Alifano, F., Petroni, A., Sánchez, R., Sedeño, L., García, A. M., & Ibáñez, A. (2018). Multilevel convergence of interoceptive impairments in hypertension: New evidence of disrupted body-brain interactions. *Human Brain Mapping*, 39(4), 1563–1581. https://doi.org/10.1002/hbm.23933

*Yuan, H., Yan, H.-M., Xu, X.-G., Han, F., & Yan, Q. (2007). Effect of heartbeat perception on heartbeat evoked potential waves. *Neuroscience Bulletin*, 23(6), 357–362. https://doi.org/10.1007/s12264-007-0053-7

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. https://doi.org/10.1016/j.biopsycho.2018.06.006