Heartbeat-evoked responses in M/EEG: A systematic review of methods with suggestions for analysis and reporting

Paul Steinfath[1,2*#], Maria Azanova[1,3#], Nikolai Kapralov[1,2#], Thomas Loesche[1], Lioba Enk[1,3], Vadim Nikulin[1##], Arno Villringer[1,3,4,5,6##]

[1]Department of Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

[2]International Max Planck Research School NeuroCom, Leipzig, Germany

[3]Max Planck School of Cognition, Leipzig, Germany

[4]LIFE – Leipzig Research Center for Civilization Diseases, University of Leipzig, Leipzig, Germany

[5]Department of Cognitive Neurology, University Hospital Leipzig, Leipzig, Germany

[6]MindBrainBody Institute, Berlin School of Mind and Brain, Humboldt University Berlin, Berlin, Germany

[#,##]These authors contributed equally to this work

*Corresponding author: steinfath@cbs.mpg.de,

Department of Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, Stephanstr. 1a, 04103, Leipzig, Germany

Abstract

Heartbeat-evoked responses (HER), as measured by electroencephalography (EEG) or magnetoencephalography (MEG), have become widely used as a marker of cardiac interoception in the study of brain-body interactions. However, HER studies report largely variable findings, at least partially due to methodological variability. To achieve consensus on HER processing and improve the reproducibility of findings, the field urgently requires a structured summary of the methods employed so far. To this end, we conducted a systematic review of 132 HER studies using non-invasive M/EEG recordings in humans. Our results reveal substantial heterogeneity across most steps of HER analysis, ranging from data acquisition and preprocessing to HER estimation and statistical approaches. The large diversity in the processing choices is accompanied by considerable proportions of unreported methodological information across reviewed studies, reaching up to 80% for key processing steps. In addition, less than 33% of studies had enough statistical power to reliably detect meta-level HER effects, while their reported spatiotemporal locations varied substantially. We provide a comprehensive reporting and quality control checklist to aid in the development of more standardized procedures, highlighting critical steps for robust HER investigations. Additionally, we share the full extracted dataset, including an interactive version, to support other researchers in answering additional specific questions they may have. We hope that these resources will improve the robustness, reproducibility, and transparency of research in the growing HER field.

Keywords

Heartbeat-evoked response, heartbeat-evoked potential, methodological variability, reporting guidelines, MEG, EEG, ECG

1. Introduction

The dynamic interplay between the central nervous system and other bodily organs is increasingly recognized as a crucial factor in understanding cognition, behavior, and overall well-being (Azzalini et al., 2019; Criscuolo et al., 2022). A growing body of research has demonstrated how the interaction between the brain and body shapes many functional aspects, ranging from the direct modulation of neuronal activity by blood pressure (Salameh et al., 2024) up to higher cognition (Umeda et al., 2016) and complex emotional states (Feldman et al., 2024; Garfinkel et al., 2014).

Among the various organs involved, the heart plays a particularly important role in brain-body interaction research. Beyond its essential function of delivering oxygenated blood and vital nutrients to the brain and the rest of the body (Hall & Hall, 2021), its rhythmic activity itself appears to serve as a pacemaker, providing an important reference frame for the brain (Tallon-Baudry et al., 2018). Accumulating evidence suggests that distinct phases of the cardiac cycle are advantageous for either perception or action; therefore, monitoring and adjustment of cardiac activity by the brain might be necessary to achieve optimal performance (Skora et al., 2022).

One key methodology for studying heart-brain interactions with electroencephalography (EEG) or magnetoencephalography (MEG) involves analyzing neural activity time-locked to the heartbeat. The resulting heartbeat-evoked responses (HERs, also known as heartbeat-evoked potentials, HEPs) reflect the processing of heartbeat-related information in the central nervous system (Dirlich et al., 1997; Schandry et al., 1986). HERs are commonly used as a marker for cardiac interoception (Pollatos & Schandry, 2004) and have been linked to a broad range of functions, such as perception (Al et al., 2020; Park et al., 2014), attention (Banellis & Cruse, 2020), motor activity (Al et al., 2023), or emotions (Di Bernardi Luft & Bhattacharya, 2015).

Despite the growing interest and number of studies, the field of HER research is facing a challenge of methodological heterogeneity and variability in reported findings (Coll et al., 2021). There appears to be no single, unified manifestation of HER. Instead, heartbeat-locked neural activity has been associated with a wide range of cognitive and emotional functions and observed across diverse topographic and temporal regions (Engelen, Solcà, et al., 2023; Park & Blanke, 2019). While some of this variability can be attributed to participant characteristics or the experimental context in which HERs are investigated, variations in data acquisition, preprocessing techniques, and statistical analyses likely also play a role.

Beyond the general methodological decisions that apply to all studies of evoked responses (such as how to filter the data and whether to use baseline correction), analysis of HERs requires special treatment, most notably due to the cardiac field artifact (CFA). CFA is caused by the instantaneous spread of magnetic and electric fields from the heart's activity due to volume conduction, and is directly captured by M/EEG sensors (Dirlich et al., 1997). To investigate the brain activity related to the heartbeat, without the CFA, researchers typically either try to avoid time ranges corresponding to high CFA, remove the CFA from the data, or include the heart's activity as a confound in the statistical analysis. Various approaches are employed for this purpose, and the outcome may depend on the selected approach and even on decisions such as the choice of settings for electrocardiogram (ECG) recording (Gray et al., 2007). To note, other confounds, such as pulse-related movement of tissue and sensors, can also introduce heartbeat-locked artifacts (Kern et al., 2013), but their influence on measured HERs remains unclear.

Apart from cardiac artifacts, the overlap of heartbeat- and task-evoked activity additionally complicates the analysis of HERs. When HERs are studied in task paradigms, the potential overlap with task-evoked activity must be accounted for (Azzalini et al., 2019; Steinfath et al., 2025). Furthermore, due to the cyclic nature of the heartbeat, neural responses to successive heartbeats can superimpose, especially when inter-beat intervals (IBIs) are short. The

necessity to address these problems further increases the number of analysis decisions researchers have to make.

Overall, it remains unclear to what extent methodological choices influence the HER effects reported across different studies, and thorough examination of the employed methods has been repeatedly advocated (Coll et al., 2021; Park & Blanke, 2019). This situation mirrors the reproducibility challenge in the broader field of neuroscience (Botvinik-Nezer & Wager, 2023), which is currently emphasized and addressed by newly appearing analytic approaches such as large-scale collaborative data analyses (EEGmanypipelines, Trübutschek et al., 2024), multiverse analyses (Clayson et al., 2021), or meta-methods reviews (Elson, 2019).

In light of these challenges, it is an essential first step to systematically collect and assess the methodological approaches used in HER research. To achieve improved consistency and reproducibility in HER analysis, a clear understanding of current practices is necessary. This review, therefore, focuses specifically on the methods used in HER research. To this end, we extracted and analyzed the methodological decisions, from data acquisition through preprocessing to statistical analysis, from published HER studies.

In addition, we highlight aspects that may be especially important for the HER analysis, which we hope will stimulate further HER research, discussion, increased transparency, and potential consensus. To facilitate this, we publicly share the full resulting dataset, its interactive version, the accompanying codebook, and the HER analysis checklist with this article.

2. Methods

2.1. Literature search

We conducted a systematic literature search using the PubMed database on three separate dates to identify studies investigating HERs using noninvasive neuroimaging. Table 1 specifies

the search terms used, with corresponding dates and the number of studies retrieved, as well as the total number of studies after duplicate removal.

Table 1. Search strategy

Date	**Search query**	**Studies retrieved**
15.01.2024	(eeg OR electroencephalograph*) AND (heartbeat evoked potential OR heartbeat evoked response) NOT (Review[Publication Type] OR Systematic Review[Publication Type]) AND English[Language]	647
24.07.2024 (rerun: 05.08.2024)	(eeg OR electroencephalograph* OR meg OR magnetoencephalograph*) AND (heartbeat evoked potential OR heartbeat evoked response) NOT (Review[Publication Type] OR Systematic Review[Publication Type]) AND English[Language]	682 (rerun: 683)
05.08.2024	("heartbeat evoked" OR "heart evoked") AND Journal Article[Publication Type] AND English[Language]	166
Total amount of unique studies after removing duplicate entries		**749**

After the removal of duplicates, all retrieved records (N = 749) were screened following the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Moher et al., 2009). We included studies that: (1) examined HERs, (2) applied noninvasive EEG or MEG recordings, (3) were conducted in human participants, (4) reported original data, (5) were in English language, and (6) had full text available. Exclusion criteria were: (1) intracranial recordings due to the different data analysis approaches, (2) non-original research (reviews, commentary papers, and toolboxes), and (3) studies not analyzing evoked responses (e.g., time-frequency analysis only). Because this review focuses on HER analysis methods, we also included studies analyzing data in source space and those without statistical testing. After screening titles and abstracts, we evaluated the full texts of potentially relevant articles to determine their eligibility. After applying the above-mentioned criteria, 127 articles were included. In addition, we identified seven references through the bibliographies of the reviewed publications, of which five were included, resulting in a final sample of 132 articles

used in the systematic review. The screening process is detailed in the PRISMA flow diagram (Fig. 1A).

2.2. Extraction of methodological choices

In the following, we use the term *processing steps* when referring to the various data transformations or parameters that researchers choose (for example, the cutoff frequency of a high-pass filter). A particular value (e.g., 1 Hz) that is then selected for a processing step is referred to as a *(methodological) choice*, and a combination of methodological choices for all processing steps is considered as a *pipeline* in this review. If a study uses different methodological choices for any processing step (e.g., performs the analysis with and without baseline correction), we say that several pipelines correspond to this same study.

For each article included in the review, we extracted methodological choices for a wide range of processing steps, from the acquisition of physiological data (EEG/MEG/ECG) to the calculation and statistical analysis of HERs (Table 2). Three reviewers, independently, extracted the data and then additionally checked for the consistency of the extracted information for each processing step. We developed a standardized codebook to unify and ensure consistency of labels referring to the same methodological categories (e.g., “linked mastoids” and “averaged mastoids”). The codebook is publicly available and can be found together with the final dataset at https://osf.io/znrbm.

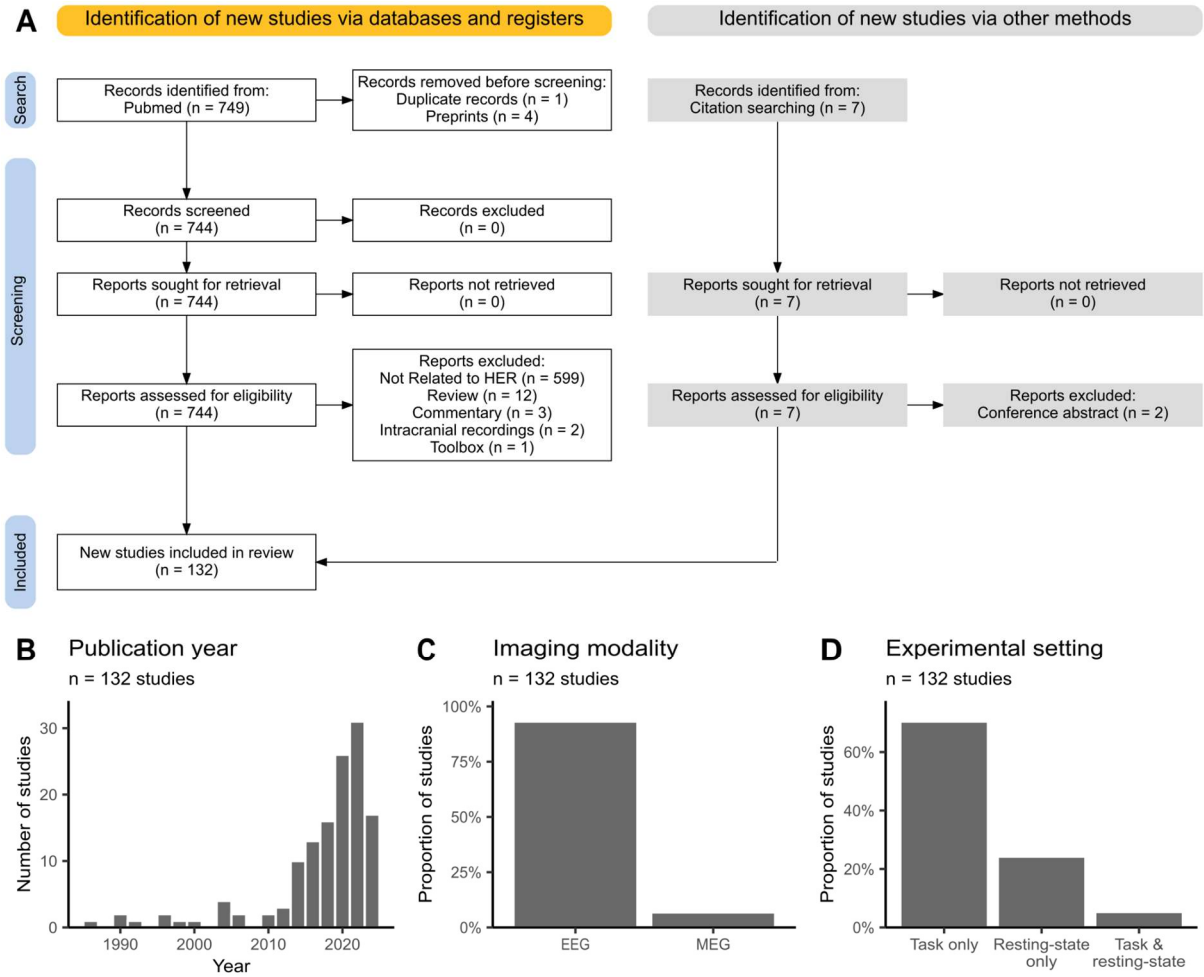

Figure 1. Overview of the studies included in the review. (A) Flowchart describing the literature selection process according to the PRISMA guidelines. (B) The majority of the included studies were published within the last 10 years. (C) Most studies use EEG data to analyze heartbeat-evoked responses (HERs). (D) HERs are typically analyzed during task performance rather than in the resting state.

145 *Table 2.* Overview of the extracted processing steps from pipelines used for studying HERs.
146 Abbreviations: HERs – heartbeat-evoked responses, M/EEG – magneto- or
147 electroencephalography, ECG – electrocardiogram, ICA – independent component analysis,
148 ICs – independent components, CFA – cardiac field artifact.

Category	**Processing step**
Experiment	Experimental setting (task or resting state), sample size, number of participant groups, number of experimental conditions, length of the recording used to analyze HERs, number of epochs used to obtain HERs (per participant and condition)
M/EEG acquisition	Imaging modality (M/EEG), number and locations (layout) of M/EEG sensors, reference electrode used during the recording (online reference)
ECG acquisition	Number and locations of ECG electrodes, ECG lead, ECG ground electrode
M/EEG preprocessing	Reference electrode used in the data analysis (offline reference), cutoff frequencies for the low- and high-pass filters, whether ICA was applied, type of ICs that were rejected, other approaches for cleaning the data
Removal of the CFA	Whether ICA was applied on epoched or continuous data, the number of rejected CFA-related ICs, the approach and criteria used for rejecting CFA-related ICs, other approaches for removal of CFA apart from ICA
HER computation	Approach to determine locations and times of interest (primary/secondary), epoch onset (R-/T-peak), start and end latencies of the analyzed time window, start and end latencies of the baseline window, selected sensors, extracted value, whether channels or time points were averaged
Statistical analysis	Statistical test, type of association tested (difference or correlation), whether cluster-based permutation test was applied, number of permutations, control variables, whether the result of the test was significant
Significant clusters	Epoch onset (R-/T-peak), channels and time-points that belong to the cluster

2.3. Data analysis

Data analysis and visualization were performed using R (version 4.4.3; R Core Team, 2025). For most processing steps, we assessed the overall scope of methodological choices and the proportion of studies or pipelines that utilized each choice. Since the information relevant for this review is not reported for some studies and pipelines, we specify the exact number of data points above each plot. Some studies perform several analyses with different selections of methods — if this is the case for a particular plot, we refer to the data points as pipelines; otherwise, we refer to them as studies.

Entropy. To quantify and compare the variability of methodological choices for different processing steps, we used the concept of Shannon entropy (Shannon, 1948). Larger entropy values indicate a greater heterogeneity in the choices. Since entropy depends on the number of available options, we normalized it to reflect whether a predominant choice exists or multiple choices appear with similar frequencies for a processing step. Additionally, we binned some continuous parameters (e.g., the number of trials) that contained more than 10 unique values to make them comparable to categorical variables before computing entropy.

Aggregation of EEG locations. To analyze the spatial distribution of observed HER effects, we collected information about the EEG channels included in the HER analysis and, if applicable, about those belonging to at least one of the reported significant clusters. To combine studies that used different EEG layouts during data recording, we calculated the number of occurrences for each channel and layout separately. The occurrence counts were interpolated to form topographies independent of actual sensor locations. The topographies were then summed across the considered layouts, that is, variations of the 10-20 system with 19, 32, and 64 channels, as well as the 128-channel BioSemi layout.

Minimal detectable effect sizes. To compare minimal effect sizes that could be robustly detected across pipelines, we (1) computed minimal detectable effect size for one- and two-sample t-tests, ANOVAs, and correlations using pwr R package (version 1.3-0; Champely et

al., 2020) based on the information about sample sizes, numbers of groups, conditions, and required power of 80% with significance level set at 5%, (2) converted these effect sizes to Cohen's d, when required, and finally (3) converted Cohen's d into Hedges' g using the exact correction approach (Hedges & Olkin, 1985). The resulting Hedges' g estimates of minimal detectable effect size were compared to previously reported effect sizes in HER literature (Coll et al., 2021). Details on the estimation of Hedges' g are provided in Section A of the Supplementary Material. Finally, we examined how detectable effect sizes could change if within-participant variability in HER values across epochs was taken into account. To do this, we simulated variances of hypothetical HER signals within and between participants (see Section B of the Supplementary Material for more details on these estimations).

Estimation of the number of epochs. As M/EEG signals are prone to high amounts of environmental and physiological noise, obtaining a clear evoked response often requires averaging over a substantial number of epochs. For each pipeline, we extracted the minimum (preferred) or mean number of heartbeat-locked epochs per participant and condition, where possible. Those were not readily available in some studies. For such cases, we estimated approximate numbers using strategies such as (1) dividing the overall reported number of epochs by the number of analyzed participants, (2) multiplying the reported length of the recording in minutes within which HERs were analyzed by the average heart rate, when reported, or otherwise by 60 beats per minute. If reported, those were adjusted by the number of epochs removed during cleaning.

HER peak timing comparison. To statistically compare the timing of effects observed with different HER analysis strategies (e.g., clustering vs. averaging), we implemented a permutation test. For each strategy, the distribution of analysis time window start and end times was aggregated as a cumulative sum, and the peak was identified. Next, the peak time difference between the two strategies was calculated. To test if the differences between strategies were statistically significant, all analysis time windows from both strategies were pooled and randomly assigned to two groups of the original sizes. The peak timing differences

were recalculated, and the permutation procedure was repeated 1000 times. Lastly, p-values were obtained by comparing the original peak timing difference to the permutation distribution.

3. Results

The results of the review are structured as follows. We start with an overview of included studies (Section 3.1), followed by an overview of processing steps that were extracted and analyzed (Section 3.2). Sections 3.3-3.9 provide a detailed description of methodological choices that occur in the literature for various processing steps from data acquisition to statistical analysis.

3.1. Overview of the included studies

In total, 132 studies published between 1986 and 2024 were included in the review (Fig. 1B). The vast majority of studies used EEG recordings (123/132, 93%), while only nine studies (6.8%) employed MEG to measure HERs (Fig. 1C). As shown in Fig. 1D, most studies (93/132; 70%) focused on the analysis of HERs during performance of various tasks involving perception of own heartbeats (see Coll et al., 2021 for a review) or external stimuli (e.g., Al et al., 2020; Marshall et al., 2020; Park et al., 2014). A smaller proportion of studies (32/132; 24%) performed resting-state analyses, testing the potential relationship between the amplitude of HER and diverse psycho-physiological states (e.g., Flasbeck et al., 2021; Giusti et al., 2024; Yoris et al., 2024) as well as diverse clinical conditions (e.g., Birba et al., 2022; Kumral et al., 2022; Schmitz et al., 2021; Solcà et al., 2020). Seven studies (5.3%) analyzed HERs both during task and resting state.

3.2. Overview of methodological choices

For each processing step of Table 2, we extracted methodological choices made by researchers (e.g., which value of a parameter to use or whether to apply a certain method). In total, we extracted 410 pipelines (note that some studies included multiple pipelines due to

226 the use of different parameters or strategies). In the subsequent sections, we only consider
227 distinct pipelines with respect to each processing step we discuss. For example, when
228 discussing preprocessing steps, multiple pipelines that differ only in terms of statistical
229 analysis are treated as one. This should ensure that our analyses reflect methodological
230 variability within each processing step without being inflated by differences in unrelated steps.
231 The total number of pipelines (or studies, when only one pipeline per study was reported) is
232 specified above each figure panel or in parentheses after each reported value.

233 Based on the information we extracted for data processing steps, from acquisition to statistics,
234 we observe substantial lack of consensus in most methodological choices, as indicated by
235 large entropy (Fig. 2A). For example, frequency cutoffs of high- and low-pass filters used in
236 preprocessing show great variability (Fig. 3E). Similarly, researchers select widely varying time
237 windows of interest for their HER analysis (see Section 3.7 for more details). Few steps show
238 consistency across the field: most studies use EEG data, select the R-peak as the reference
239 event for time-locking HERs, and analyze HER amplitude.

240 During acquisition and preprocessing, most studies apply exactly one choice per step, e.g.,
241 they either perform ICA artifact correction or do not, but rarely both (Fig. 2B). In contrast, using
242 several approaches for HER computation and statistical analysis within the same study is quite
243 common: 56% of studies (74/132) use multiple statistical tests, and 45% (59/132) vary the
244 start or the end latency of the time window of interest. Many such cases involve primary and
245 secondary data analysis, or several related hypotheses within the same study, and therefore,
246 typically do not require additional statistical correction.

247 However, almost no study systematically compares several methodological choices per
248 processing step to assess the effects of researcher-dependent decisions on HER results.
249 Among the reviewed studies, Candia-Rivera, Catrambone, et al. (2021) is a single such study
250 that varies a preprocessing step (offline reference). Future studies should address the
251 outstanding question of how processing decisions affect HER findings. Additionally, HER

research articles could show if findings are robust to decisions such as baseline correction or different CFA cleaning methods (Box 1).

Adding to that, a substantial proportion of methodological decisions is unreported (Fig. 2C). For 71% of studies reporting CFA-related independent component rejection (49/69), information on the number of rejected components is not provided. For 80% of studies (105/132), the ECG ground location is not explicitly stated. As natural contamination with CFA makes HERs distinct from other event-related responses, it is critical to provide as much information as possible about heartbeat-related processing and acquisition decisions. In addition, we observe a high percentage of unreported values for the following critical methodological choices: the number of epochs averaged to obtain HERs (73/132 studies, 55%), recording length (78/132, 59%), and the start or end of the baseline window (59/132, 45%). In 28% of studies (37/132), ECG locations are not reported. Overall, these results highlight that even for parameters critical to HER computation and analysis, there is a lack of transparency in the field, which contributes to reduced reproducibility of results. To remedy this issue, we provide a comprehensive reporting and quality control checklist with this review (Box 1).

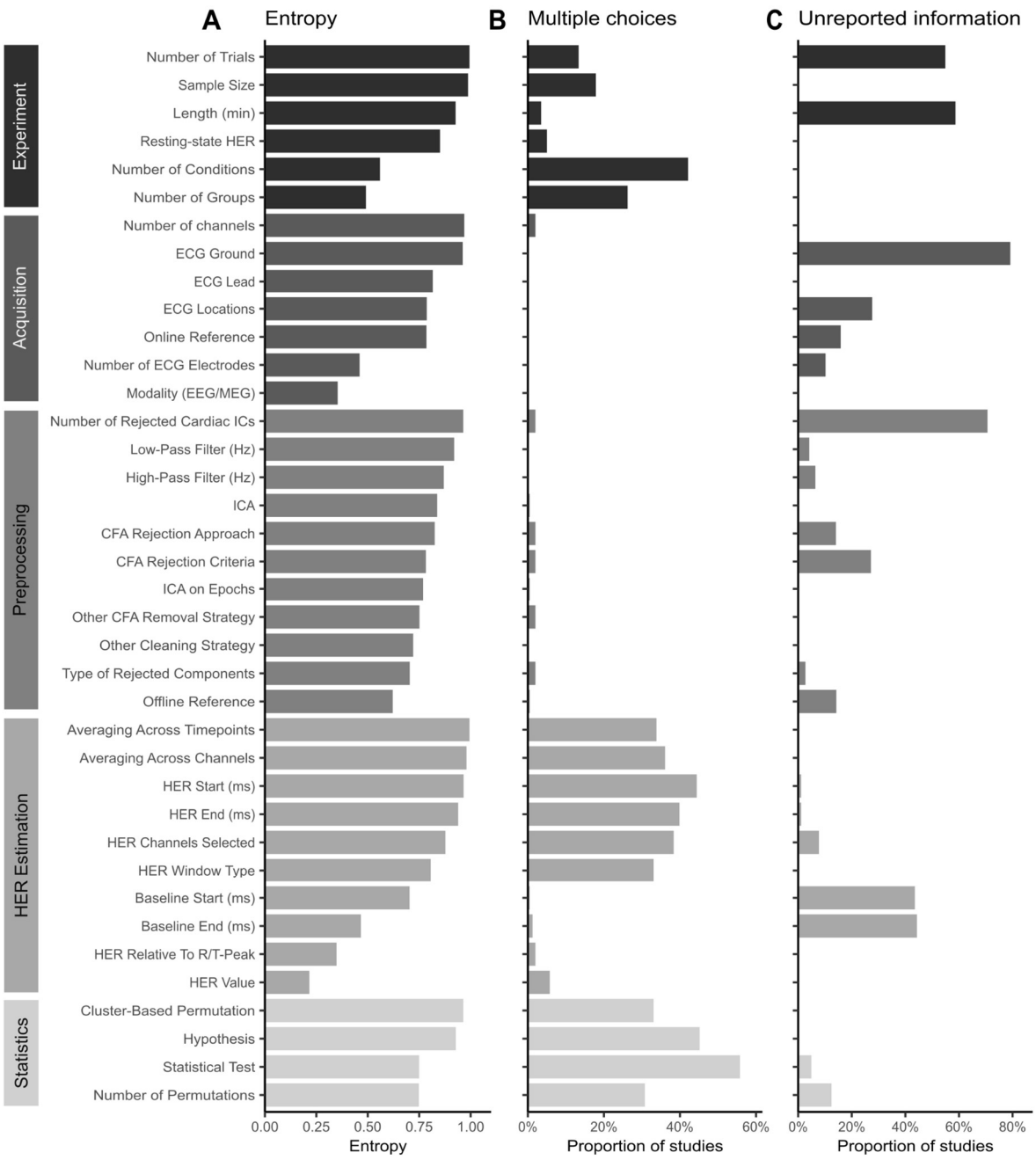

Figure 2. Overview of the extracted processing steps, heterogeneity in methodological choices, and unreported information. (A) Entropy was calculated to illustrate the diversity in processing choices across studies. Within categories, the processing steps are sorted by entropy in descending order. (B) Some studies use several analysis approaches, specifically to estimate heartbeat-evoked response (HER) values and perform the statistics. (C) The highest percentage of unreported information was observed for ECG ground location and the number of rejected cardiac independent components (ICs), a relevant parameter for artifact

correction in the HER field. Abbreviations: ICA – independent component analysis, CFA – cardiac field artifact, ECG – electrocardiography, EEG – electroencephalography, MEG – magnetoencephalography.

3.3. M/EEG acquisition and preprocessing of evoked responses

Since HERs typically occur in the 0.5–2 μV range (Pérez et al., 2005; Pollatos & Schandry, 2004; Schandry & Weitkunat, 1990), M/EEG recordings of these signals often have a relatively low signal-to-noise ratio (SNR), requiring careful data preprocessing and artifact removal strategies. At the same time, the preprocessing itself can affect the shape and amplitude of ERP components (Clayson et al., 2021). In this section, we focus on the processing steps that are common to all ERPs, while the specific methods for suppressing CFA are discussed in more detail in Section 3.5.

In EEG, the choice of the reference electrode is important as it affects the shape and distribution of the ERP (Luck, 2014), making direct comparisons across studies with different references difficult. As shown in Fig. 3A, we observe some variability with respect to the online reference used during data acquisition, with linked mastoids (23%, 28/123 EEG studies) and Cz (17%, 21/123) as the most common options. This variability is reduced during data processing, where the common average reference (CAR, Fig. 3B) emerges as the most dominant choice (45%, 57/127 pipelines). This observation is complemented by Candia-Rivera, Catrambone, et al. (2021), who compared the influence of different re-referencing schemes on HERs and concluded that CAR leads to the most consistent results across a range of brain-heart interaction measures. However, they also note that the optimal re-referencing choice depends on the experimental settings and can vary across studies.

When analyzing relatively slow evoked responses, such as the HER, optimal filter settings are an important consideration to avoid removing or inducing effects (Luck, 2014; Tanner et al., 2015). We observe that, across studies, filtering is generally done between 0.1 and 40 Hz, with the most dominant high-pass filter cutoffs at 0.1, 0.5, and 1 Hz (Fig. 3E). In standard ERP

302 analyses, frequencies above 40 Hz are typically excluded because they often are
303 contaminated with different artifacts like muscle or line noise (Luck, 2014). Meanwhile, the
304 high-pass filter cutoff must balance the removal of low-frequency noise while preserving the
305 slowly evolving components of the heartbeat-locked signal. Given that some neurons track
306 changes in inter-beat intervals (De Falco et al., 2024) and assuming the resting heart rate in
307 humans is approximately 1 Hz, a high-pass filter of 1 Hz might inadvertently remove genuine
308 HER activity. Ultimately, the filter selection represents a trade-off between the specific aims of
309 the HER analysis and the noise characteristics of the dataset. Future studies are needed to
310 shed light on the frequency components of the HER across contexts, their relationship to heart
311 rate, and the impact of different filter cutoffs on the observed associations between HERs and
312 behavior or cardiac artifacts. Approaches such as multiverse analysis could be used for the
313 systematic exploration of such effects (Clayson et al., 2021). In the meantime, we recommend
314 reporting the filter settings used to obtain HER (Box 1).

315 Due to the low amplitude of HERs, additional artifact-reduction strategies are often used to
316 clean the data and obtain robust HERs that can be compared across groups or conditions.
317 Our review shows that 55% of the studies (72/132) removed noisy epochs, while 11% (15/132)
318 explicitly mention that bad channels were excluded. The majority of pipelines (73%, 97/133;
319 Fig. 3C) applied independent component analysis (ICA), a commonly used decomposition
320 method designed to separate the recorded M/EEG signal into its underlying brain and non-
321 brain sources.

322 The primary artifact sources that are addressed by ICA include eye movement (89%, 85 out
323 of 96 pipelines that reported the types of rejected components), cardiac (72%, 69/96), muscle
324 (23%, 22/96), or other nuisance sources (Fig. 3D). While the removal of cardiac artifacts is
325 discussed in Section 3.5, it is important to recognize that other physiological activity such as
326 eye movements (Ohl et al., 2016), muscle activity (Niizeki et al., 1993), or respiration (Yasuma
327 & Hayano, 2004) also co-vary with the heartbeat. The extent to which these factors influence

the measured HERs, and whether ICA can mitigate their effects, remains an open question that should be addressed in future studies.

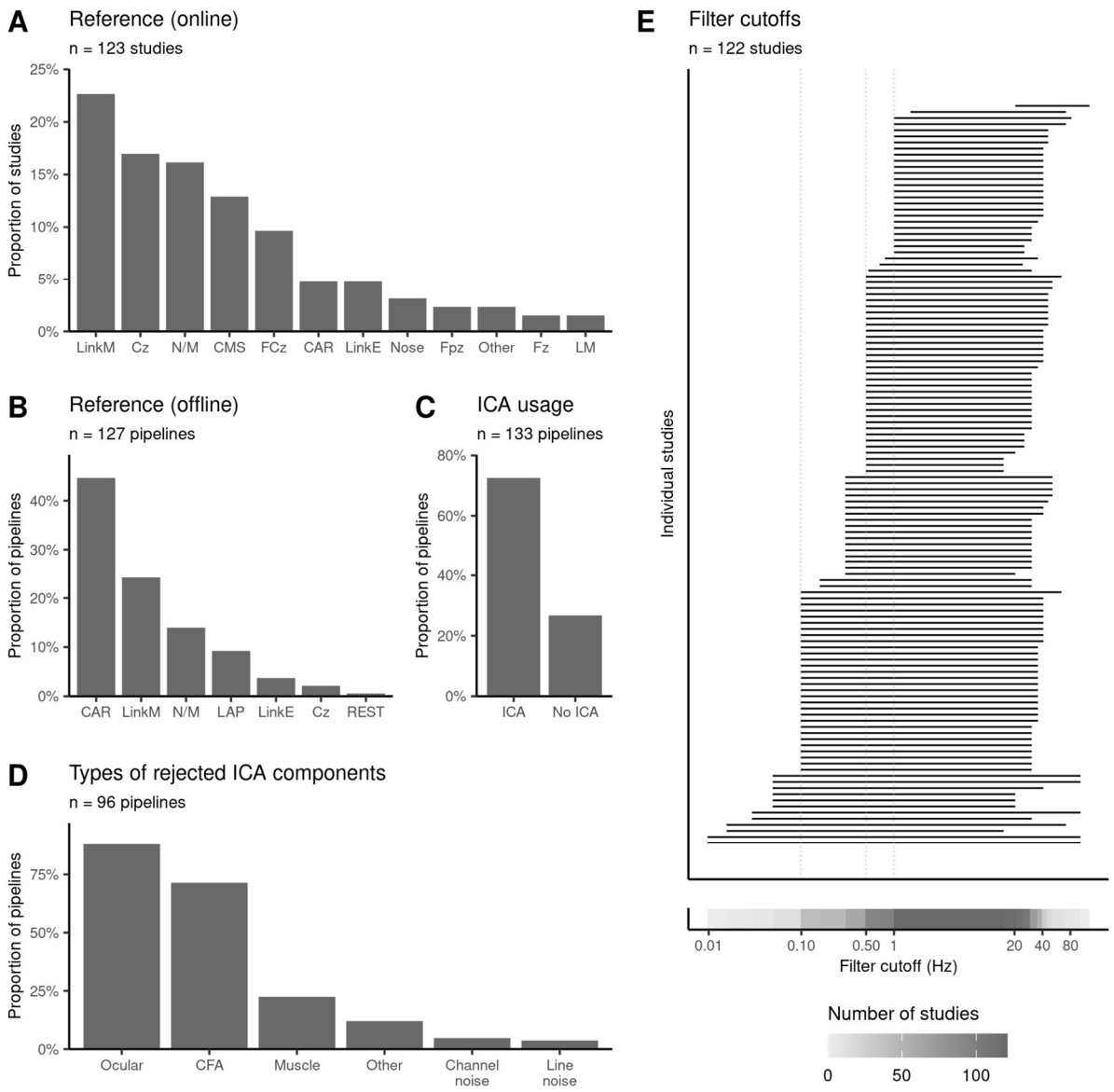

Figure 3. Heterogeneity in M/EEG acquisition settings and preprocessing methods. (A) Linked mastoids and Cz are most commonly used during the recording as reference. Only studies that used EEG are shown. The category Other denotes references that appeared only in a single study. (B) Common average is the most common reference used in the data analysis. Only studies that used EEG are shown. (C) Most studies use ICA for artifact removal. (D) ICA components related to eye movements (ocular) and CFA are rejected most frequently. (E) Frequency ranges analyzed by the reviewed studies. Popular high-pass filter choices include

0.1, 0.5, or 1 Hz. Abbreviations: LinkM – linked mastoids, CMS – common mode sense, CAR – common average reference, LinkE – linked earlobes, LM – left mastoid, LAP – laplacian or current source density analysis, N/M – not mentioned, REST – reference electrode standardization technique, ICA – independent component analysis, CFA – cardiac field artifact.

3.4. ECG acquisition

ECG recordings are primarily used to extract cardiac markers for time-locking the HER (typically, either R-peak or T-peak). At the same time, they are also often (by 59/132 studies, 45%) employed in control analyses to ensure that observed HER effects are not accompanied and thus cannot be explained by differences in cardiac activity. Placement of the ECG electrodes can play an important role in such analyses, as the presence of differences in ECG between experimental conditions has been shown to depend on the ECG derivation of choice (Gray et al., 2007). Moreover, using just one ECG lead to control for the spread of CFA might not be sufficient to capture the three-dimensional CFA in full (Pérez et al., 2005). With these motivations in mind, we evaluated the consistency of ECG electrode setups used in the literature.

Most studies (67%, 89/132) dedicate two electrodes to record ECG (Fig. 4A), but the placement of the electrodes on the body varies substantially. Although lead II is used most frequently (by 36% of studies, 48/132), some studies also use lead I (11%, 15/132) and lead III (2.3%, 3/132), or locations that do not correspond to any of the standard leads (27%, 35/132; marked as ‘N/C’ in Fig. 4B). Even with standard lead configurations, electrode locations vary (e.g., torso vs. limb placement; Fig. 4C) though the significance of this variability remains unclear and may differ between clinical and non-clinical populations (Sheppard et al., 2011).

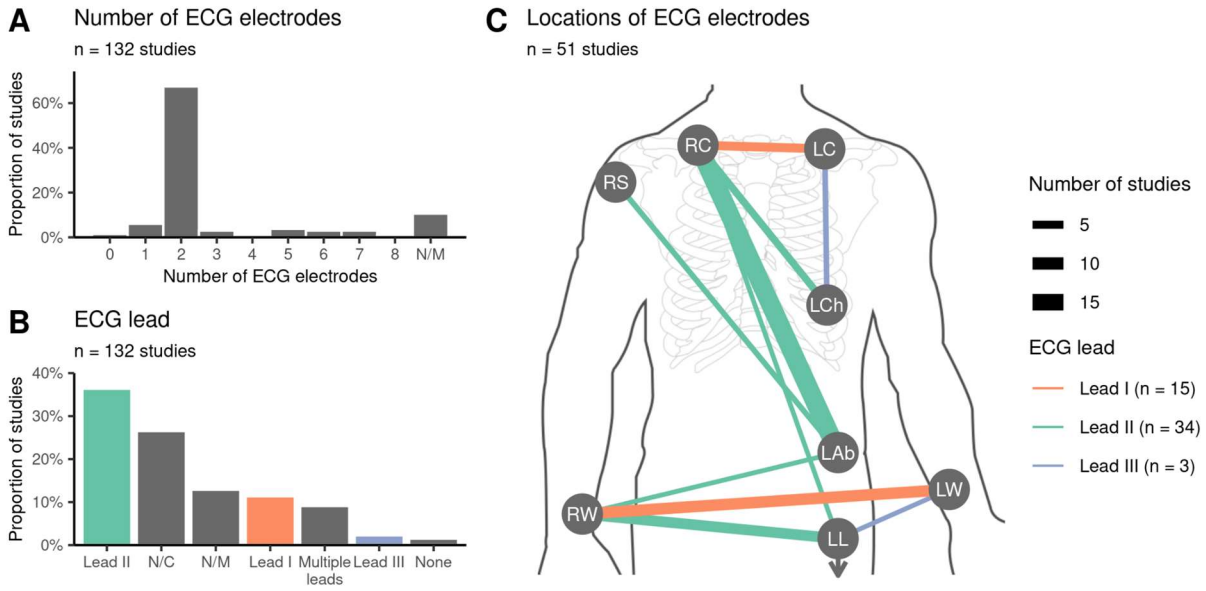

Figure 4. Various ECG leads and electrode locations are used in the literature. (A) Most studies use two electrodes to record ECG. (B) While lead II is the most commonly used derivation, many studies also use electrode placements that do not correspond to any of the standard leads (N/C) or do not specify the lead used (N/M). None denotes studies which reconstructed ECG from EEG. (C) Locations of ECG electrodes that are placed to measure leads I, II, and III. Abbreviations: RC – right clavicle, LC – left clavicle, RS – right shoulder, LCh – left side of the chest, LAb – lower left part of the abdomen, RW – right wrist, LW – left wrist, LL – left leg/ankle, N/M – not mentioned, N/C – not classified.

3.5. Dealing with cardiac artifacts

Since HERs are time-locked to the cardiac cycle, it is essential to consider the confounding influence of cardio-vascular artifacts on HER effects. The two major factors that contribute to these artifacts are the electric field generated by the heart (Dirlich et al., 1997) and pulse-related movement of tissue or EEG electrodes (Kern et al., 2013). As mentioned above, researchers have predominantly employed one of three overarching strategies to mitigate those effects: avoiding HER space-time points where the influence of artifacts is potentially present, cleaning the data from artifacts, or including the ECG in the statistical analyses. This pattern also held in the reviewed sample of studies, with 89% (117/132) applying at least one

380 of those strategies. However, the applied strategies and their effect on HER might differ
381 substantially.

382 In terms of artifact removal, ICA is used by 52% of all studies (69/132) to remove cardiac
383 components. This trend is corroborated by research that demonstrates the effectiveness of
384 ICA in attenuating the influence of cardiac artifacts on the measured HERs (Buot et al., 2021).
385 However, there is no standard way of performing ICA, and approaches vary across studies.
386 Most frequently, ICA is applied to continuous data (76%, 74/97 studies that applied ICA). In
387 contrast, around 24% of studies (23/97) choose to perform it after data epoching, likely to
388 specifically target artifacts occurring around the R-peak and T-peak (Dirlich et al., 1997).

389 While cardiac-related activity often displays relatively consistent spatial patterns across
390 participants (Dirlich et al., 1997), ICA may not always accurately differentiate those from
391 genuine brain signals. Variability of the ICA decomposition across participants can introduce
392 additional variance in the data (Pontifex et al., 2017), potentially influencing HER findings.
393 Given that 86% of the studies that remove cardiac ICs (59/69) also report the approach used
394 for identifying these components, the field is moving towards transparent reporting, facilitating
395 the reproducible removal of cardiac artifacts.

396 Among the reports, we find that manual identification of CFA components (37/69, 54%) is used
397 more often than automatic (13/69, 19%) or semi-automatic (9/69, 13%) detection approaches
398 (Fig. 5A). The CFA-related components are commonly identified based on the time course
399 (32/69, 46%), topography (31/69, 45%), and power spectrum (13/69, 19%; Fig. 5B). The fully
400 automatic approaches often rely on algorithms such as CORRMAP (Campos Viola et al.,
401 2009), ICLabel (Pion-Tonachini et al., 2019), and SASICA (Chaumon et al., 2015). Semi-
402 automatic cardiac artifact removal is often guided by measures such as phase consistency
403 (7/69, 10%) or correlation (2/69, 2.9%) between ICs and ECG.

404 Only 29% of studies (20/69) that removed CFA-related ICs reported the number of removed
405 components, which ranged between 1 and 3 with a median of 1.7 (Fig. 5C). Since many

factors, such as the SNR or the M/EEG sensor configuration, influence the number of cardiac ICs obtained by the ICA, it could also be informative to report the combined explained variance of removed cardiac ICs. The explained variance might be a more generalizable measure across studies, allowing for better comparison of CFA removal efficiency. Thus, we suggest considering it for reporting in HER studies (Box 1).

Besides ICA, the reviewed studies employ several other techniques to suppress CFA (Fig. 5D). These include current source density analysis (CSD, 11/132, 8.3%), subtracting or regressing out the ECG artifact from the M/EEG data (9/132, 6.8%), principal component analysis (PCA, 2/132, 1.5%), and subtracting the average resting state HERs from task HERs (1/132, 0.8%). However, it remains an open question as to how these different methods compare with one another.

In addition, 21% of studies (28/132) aim to avoid the influence of cardiac artifacts by restricting the time range of interest used in their analyses. Some studies pursue this by excluding early time periods within the cardiac cycle (0-200 ms after the R-peak) from their analyses (Callara et al., 2023; Cambi et al., 2024; Engelen, Buot, et al., 2023; Giusti et al., 2024; Immanuel et al., 2014; Melo et al., 2022). Another common strategy is to choose a time window of approximately 455-595 ms post-R-peak (Gray et al., 2007; Lutz et al., 2019; Müller et al., 2015; Rapp et al., 2023; Schmitz et al., 2020, 2021; Schulz et al., 2013), following Dirlich et al. (1997), who demonstrated the minimal influence of the heart's electric field during this period. However, given the large variability across studies, there does not appear to be a consensus on a time range free from cardiac artifacts (Fig. 5F).

If HER trials are close to the surrounding heartbeats, it is reasonable to remove such trials to avoid contamination with surrounding HERs or heartbeat-related artifacts. Overall, 13% of studies (17/132) reported the removal of HER trials based on a minimal RR interval, the shortest time allowed between two consecutive R-peaks. We observe that the minimal allowed RR interval varies from 350 to 750 ms across these studies (Fig. 5E). Such large variability

could be attributed to diverse participant populations and experimental designs, warranting study-specific definitions of the cutoff. Note that a high minimal RR interval threshold may impose unintended imbalance across the numbers of epochs among conditions or groups if there are underlying differences in heart rate. On top of that, such differences may signify distinct arousal states and distort the interpretation of HER findings. Therefore, it is advisable to report the number of rejected epochs per group and condition (Box 1).

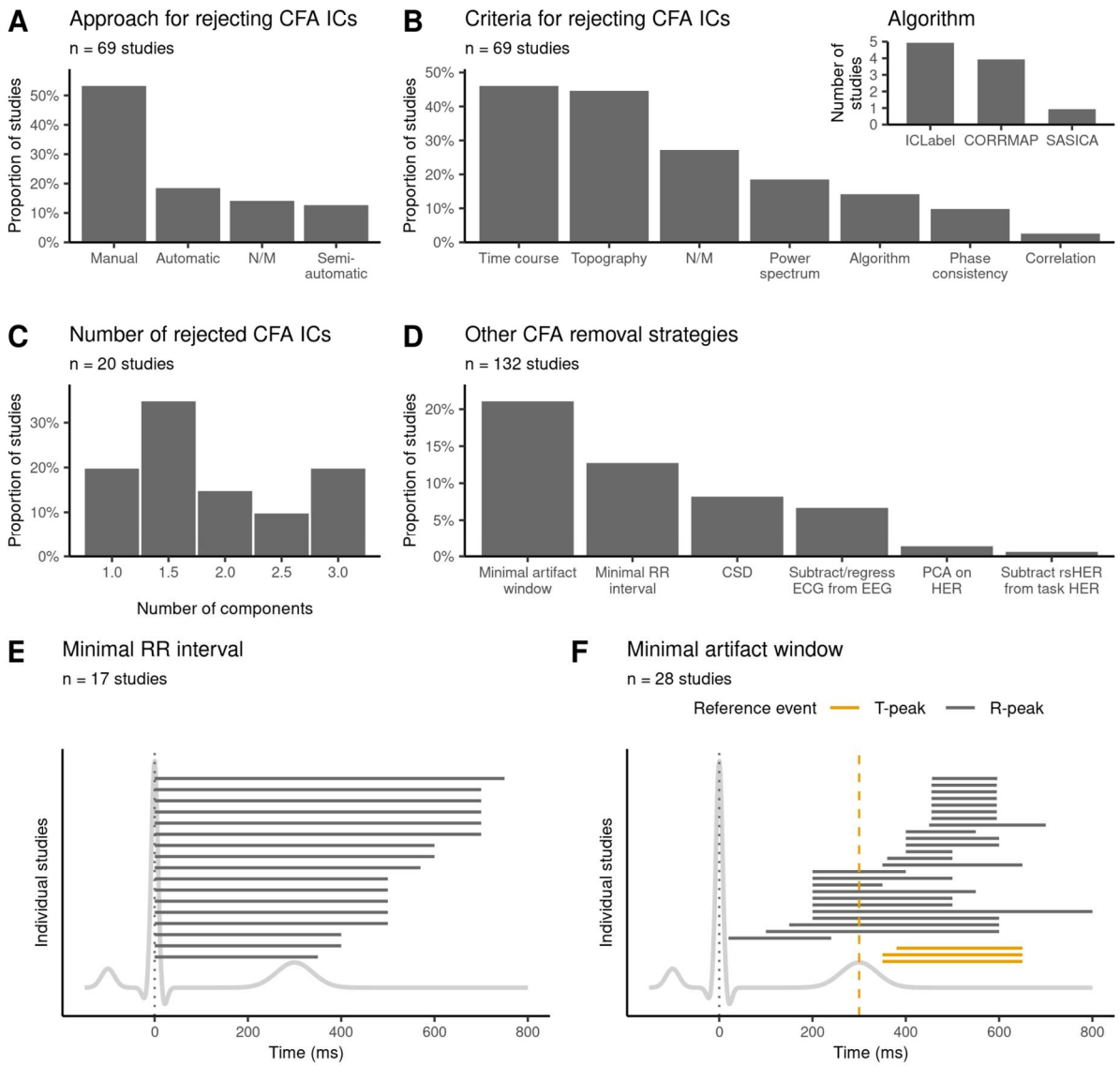

Figure 5. Approaches to address cardiac field artifacts. (A) CFA-related independent components (ICs) are rejected manually, based on automatic approaches or a combination of those (semi-automatic). (B) Decision criteria that are used to define CFA-related ICs. The inset plot displays the number of studies using different algorithms for CFA IC removal. (C)

443 The number of rejected CFA-related ICs per participant ranges from one to three across
444 studies. (D) Additional approaches to address cardiac artifacts. (E) Overlap with artifacts is
445 reduced by excluding trials with short RR intervals, or (F) by using specific time periods of
446 interest for analysis. Simulated ECG traces are shown in gray for reference in panels E and
447 F. T-peak-based time windows (in orange) are shifted by 300 ms relative to the R-peak for
448 illustrative purposes. Abbreviations: ICA – independent component analysis, ICs –
449 independent components, CFA – cardiac field artifact, CSD – current source density, PCA –
450 principal component analysis, rsHER – resting state HER, N/M – not mentioned.

451 3.6. Epoching and baseline correction

452 Critical decisions in HER research include the choice of the cardiac reference event for
453 epoching (e.g. R-peak vs. T-peak) and the spatiotemporal window used for statistical analysis.
454 These choices can directly influence the detection of HER effects and may define susceptibility
455 to the influence of artifacts. In this review, we find that the vast majority of reviewed studies
456 use the R-peak as reference (123/132, 93.2%). In comparison, only 4.5% (6/132) used the T-
457 peak, and 2.3% (3/132) performed analyses using both peaks within a single study (Fig. 6A).

458 After epoching, baseline correction is typically performed to remove vertical amplitude offsets
459 and slow drifts by subtracting the average pre-stimulus amplitude from the whole epoch.
460 Although it is a standard procedure in conventional ERP analyses (Luck, 2014), it is
461 controversial in HER research, because the heartbeat has a short inter-beat interval (~1s),
462 potentially allowing activity from one cycle to affect the pre-stimulus time window of the next
463 cardiac cycle (Azzalini et al., 2019).

464 Overall, only 61% of studies (81/132) clearly state whether baseline correction was performed.
465 Out of these, 91% (74/81) include baseline correction and 8.6% (7/81) state that they do not
466 use it (Fig. 6B). The remaining 39% (51/132) do not mention it in their manuscript, leaving it
467 unclear if they performed baseline correction or just omitted details on this step. Of note, none
468 of the reviewed studies has directly compared baseline-corrected to uncorrected values. Such

469 a comparison could be of value to ensure the observed effects are not due to baseline
470 differences (Box 1).

471 While 58% of the studies using the R-peak as a reference event (73/126) clearly state the use
472 of baseline correction, only two out of the nine studies using the T-peak (22%) did so. In both
473 studies, the baseline interval was set to -100 and 0 ms relative to the T-peak. A possible
474 explanation for the lower usage of baseline correction, here, could be that the time period
475 preceding the T-peak is close to the cardiac contraction, and hence more prone to
476 contamination by cardiac artifacts.

477 In R-peak locked analyses, baseline windows show considerable variability, ranging from -300
478 (min) to 50 (max) ms relative to the R-peak. Notably, 23% of these studies (17/73) define the
479 end of the baseline period at least 25 ms before the R-peak, potentially to avoid including
480 cardiac field artifacts (Coll et al., 2021). One study used a regression-based baseline
481 correction approach to avoid introducing baseline differences in the HER time window of
482 interest (Zaccaro et al., 2024). In future studies, it would be valuable to investigate the
483 effectiveness of baseline-correction methods and contrast them with the caveats associated
484 with not performing any baseline correction.

485 3.7. Predefined spatiotemporal HER regions and resulting findings

486 When it comes to quantifying the HER signal, amplitude is by far the most prevalent outcome
487 measure (128/132, 97% of studies). Alternative measures are rare, with only three studies
488 using HER latency (2.3%) and two studies examining global field power (1.5%). Note that
489 studies investigating heartbeat-locked neural activity using time-frequency analysis (e.g., Kim
490 & Jeong, 2019) were not considered in this review.

491 When it comes to choosing regions of interest, we observe that, overall, analyses locked to T-
492 peak cover a smaller time range (from -10 to 400 ms after T-peak). In comparison, R-peak
493 locked analyses start as early as -250 ms before the R-peak and range until 1000 ms

thereafter. Since studies often use averaging over a space-time window of interest or cluster-based permutation tests to identify HER effects (Fig. 6C, studies using only averaging: 47/115, 41%; only clustering: 30/115, 26%; both approaches: 38/115, 33%), we analyzed the data separately for each of these two approaches (Fig. 6E-F).

As expected, we see that all channels are typically included in the cluster-based analysis (Fig. 6F). However, averaging is performed more often only over frontocentral channels (Fig. 6E), often motivated by the results of previous studies (Billeci et al., 2021; Elkommos et al., 2023; Schulz et al., 2020). Similarly, clustering approaches use wider time windows more often than averaging approaches do. Nonetheless, there is a clear trend across both categories, averaging as well as clustering, to limit analysis to a time range between 200-600 ms relative to the R-peak (Fig. 6E-F).

Additionally, we also extracted information about channels and time points that show statistically significant test results and pooled them across all reviewed studies. Overall, significant HER effects were found between -10 and 1000 ms for R-peak locked analyses and between 50 and 389 ms for T-peak locked analyses. Interestingly, there is a difference between averaging- and cluster-based analysis in terms of where in time effects are most often detected, as illustrated in the cumulative sum plots in Fig. 6E-F. Specifically for the R-peak locked analyses, comparison of cumulative sum scores indicated that cluster-based approaches identified the most common latencies significantly earlier than averaging (mode = 250 ms vs. 490 ms; permutation test on the time difference between peaks, $p = 0.021$; 1000 permutations). Although the reason for this discrepancy is unclear, one possibility is that cluster-based permutation tests capture effects that are outside of narrow averaging windows by scanning a wider time range. Another possibility is that shorter effects are diluted when averaged across wider time ranges. In contrast, T-peak locked analyses showed no significant difference between clustering (mode = 307 ms) and averaging (mode = 234 ms; permutation test on time difference between peaks, $p = 0.744$; 1000 permutations), likely related to the overall smaller time range of interest in T-peak locked analyses. However, it is important to

note that permutation clustering does not allow for any precise inference about the spatiotemporal localization of effects (Sassenhagen & Draschkow, 2019), which means that the comparison to averaging-based analysis should be interpreted with caution.

When performing analysis on high-dimensional data, it is advisable to a priori define spatiotemporal regions of interest to reduce the multiple comparison problem and increase statistical sensitivity (Luck, 2014). Our review indicates that most studies follow this procedure and have predefined analysis windows (88/132, 67%, referred to as primary in Fig. 6D). However, we also observe that in 33% of the studies (44/132), researchers perform a two-step procedure, where the selected space-time window is based on the analyzed data within the same study (referred to as secondary in Fig. 6D). While this approach can be useful to identify the electrode sites and time windows where an effect appears in the studied population (Luck & Gaspelin, 2017), it is not always trivial to prevent the analyses from becoming circular (Kriegeskorte, Lindquist, et al., 2010; Sassenhagen & Draschkow, 2019).

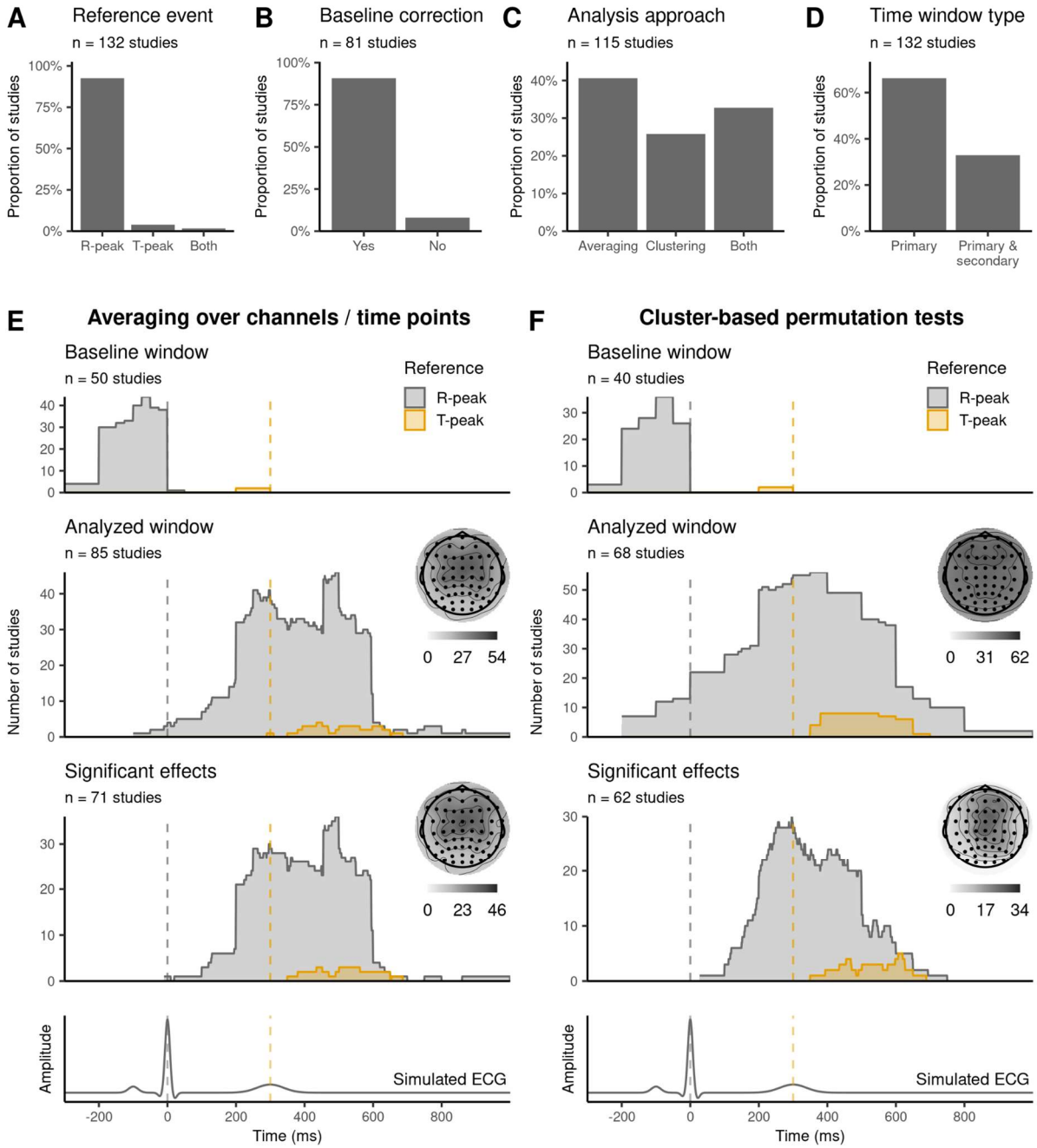

Figure 6. Predefined spatiotemporal regions and detected HER effects. Proportion of studies using: (A) R- vs. T-peak as reference event, (B) baseline correction (applied vs. not applied, unreported cases omitted), (C) averaging approaches vs. cluster-based permutation tests (or both, if performed alongside each other), (D) only primary or both primary and secondary analyses. For (E) averaging approaches or (F) cluster-based permutation tests, baseline windows (top row), time ranges and regions of interest (second row), as well as spatiotemporal distributions of significant effects (third row) are summarized. Panels E and F use cumulative

sum plots to visualize the different HER window types. Results are split by heartbeat reference R-peak (gray) or T-peak (orange) and aligned to the respective time range of the simulated ECG trace in the bottom row (gray / orange dashed line). T-peak based time windows are shifted by 300 ms relative to the R-peak for illustrative purposes. Colorbars below topographies indicate the number of studies.

3.8. Statistical analysis and power considerations

Neuroscientific studies are often underpowered, resulting in low research reliability and a potential waste of resources (Button et al., 2013). To assess power in the growing field of HER research, we estimate the minimal effect sizes that the reviewed pipelines can detect with a power of 80% at a 5% significance level, based on their sample size, number of groups and conditions, as well as the type of chosen statistical test.

Across studies, the number of included participants lies between 1 and 552, with a median of 32 (Fig. 7A), the number of studied groups lies between 1 and 6 (Fig. 7B), and the number of conditions lies between 1 and 20 (Fig. 7C). Many pipelines compare groups or conditions (301/410, 73%), or perform a correlational/regression analysis to study a linear association (109/410, 27%). As mentioned earlier, and where possible (Section 2.3), we approximate the number of heartbeat-related epochs used to obtain HERs for the many studies (73/132, 55%) that did not report this information. The estimated and reported numbers of heartbeat-locked epochs range from 22 to 5517 with a median of 300 averaged epochs per participant and condition (Fig. 7D). As shown in Fig. 7E, the t-test family (one-sample, two-sample) is the predominant choice in statistical analysis (74/222 pipelines, 33%), closely followed by variations of correlation (62/222, 28%) and ANOVA (52/222, 23%).

The resulting minimal detectable effect sizes range from 0.18 to 2.37, with a median of 0.86 Hedges' g. Fig. 7F shows their distribution across pipelines that used averaged HER values in one- or two-sample t-tests, correlations, or ANOVAs (estimates including clustering and regression pipelines, which require more assumptions, are shown in Fig. S1). In general, the

568 smaller the minimal detectable effect size, the broader the range of effect sizes that the
569 pipeline can robustly detect. When comparing the estimated detectable effect sizes to the four
570 meta-level effect sizes (specific dimensions in parentheses below) reported in a previous HER
571 review by Coll et al., 2021, we observe that only 0.7% (1/134) of pipelines had enough power
572 to detect Hedges' g of 0.35 (interoceptive performance meta-estimate), 1.5% (2/134) of
573 pipelines could reliably detect Hedges' g of 0.37 (attention to heart meta-estimate), 8.2%
574 (11/134) of pipelines could reliably detect Hedges' g of 0.49 (clinical vs. control groups meta-
575 estimate), and 33% (44/134) of pipelines could reliably detect Hedges' g of 0.72 (arousal meta-
576 estimate). These observations suggest that 67% (90/134) to 99% (133/134) of pipelines may
577 have yielded underpowered and potentially unreliable findings. Such evidence draws attention
578 to potential reproducibility and reliability problems of current findings within the HER field.
579 Studies with larger sample sizes could test the robustness of previous findings, and future
580 studies may consider adjusting their sampling strategy to detect smaller effect sizes more
581 reliably.

582 It is essential to note three key points. First, the review by Coll et al. (2021) focuses on a
583 subset of 45 HER studies that report interoception-related HER effects. We did not assess the
584 actual effect sizes among the studies we reviewed, as they varied in topics and experimental
585 designs. Second, the Hedges' g estimates presented here are based on approximations and
586 therefore should be interpreted with caution. Third, as mentioned previously, HERs are
587 susceptible to noise contamination due to their relatively low amplitudes of 0.5-2 μV (Pollatos
588 & Schandry, 2004; Schandry & Weitkunat, 1990). It is unclear how many trials, or heartbeat-
589 locked epochs, are required to obtain a stable evoked response per participant. Previous ERP
590 studies note that there is no straightforward answer, but it is clear that smaller effects (as
591 typically observed in HER amplitudes) generally require larger samples and more epochs to
592 be reliably detected (Boudewyn et al., 2018; Gibney et al., 2020). However, Hedges' g
593 estimates discussed in the previous paragraphs do not account for the levels of within-
594 participant noise across HER epochs. To examine this, we simulated the possible dynamics

of minimal detectable effect sizes, relative to the number of averaged HER epochs, by varying the degrees of intra-individual variability in HER values (Fig. S2). By construction, minimal detectable effect sizes are increasing after adjusting for within-participant variations in HER across epochs. Therefore, in Fig. 7F, we demonstrate optimistic estimates, unadjusted for possible measurement noise. Future methodological studies should examine the levels of HER variability across trials and recommend recording settings. In practice, pilot data can be used to assess variability and adjust the number of epochs. At the very least, we recommend reporting the number of averaged HER epochs per participant and condition (Box 1).

A lack of transparency related to sample size calculation is characteristic of evoked response studies in neuroscience (Larson & Carbine, 2017). To confront this, we recommend preregistering the planned HER analysis alongside power calculations (Box 1). Preregistration is a declaration of planned analysis in a public registry, and it reduces the risk of bias (adjusting methods after knowing outcomes), as well as increases transparency by allowing other researchers to judge the reliability of outcomes (Hardwicke & Wagenmakers, 2023). Amongst the reviewed studies, only 3.8% (5/132) mentioned having published a preregistration. Further, there were no registered reports, advanced forms of preregistration in which the methods are reviewed before data collection and a publication is guaranteed regardless of the outcome (Chambers & Tzavella, 2022). It is possible that the high variability in methods (Sections 3.2-3.6) and findings (Section 3.7) prevents HER researchers from preregistering their hypotheses. This problem can be best remedied if, firstly, most research groups working in the HER field preregister their studies by default, and secondly, more research is devoted to the advancement of robust HER analysis methods.

Another potential benefit of preregistrations is overcoming publication bias, the tendency for publications with significant findings to be published more frequently than those with null findings (Thornton & Lee, 2000). There is evidence for publication bias in the HER field: we observe that only 6.8% of studies (9/132) were published with no significant HER results. We suggest that a combination of preregistrations with formal null testing could reduce publication

bias and make null findings more visible and reliable (Box 1). Such results would contribute to the structured accumulation of knowledge in the field, as future studies could use this information to refine their approach. Statistical assessment of “absent” effects is possible using Bayesian statistics (Wagenmakers et al., 2018) and frequentist equivalence testing (Lakens et al., 2018). Few studies in the reviewed sample have used Bayesian statistics to assert the absence of effects, and we did not observe any use of equivalence testing or, importantly, any definitions of the smallest effect sizes of interest. Determining the smallest effect size of interest (Lakens et al., 2018) or the region of practical equivalence (Kruschke, 2021) is not only useful for establishing the absence of effects, but also for making sure that a study does not focus on significant effects too small to be reliably replicated (Box 1).

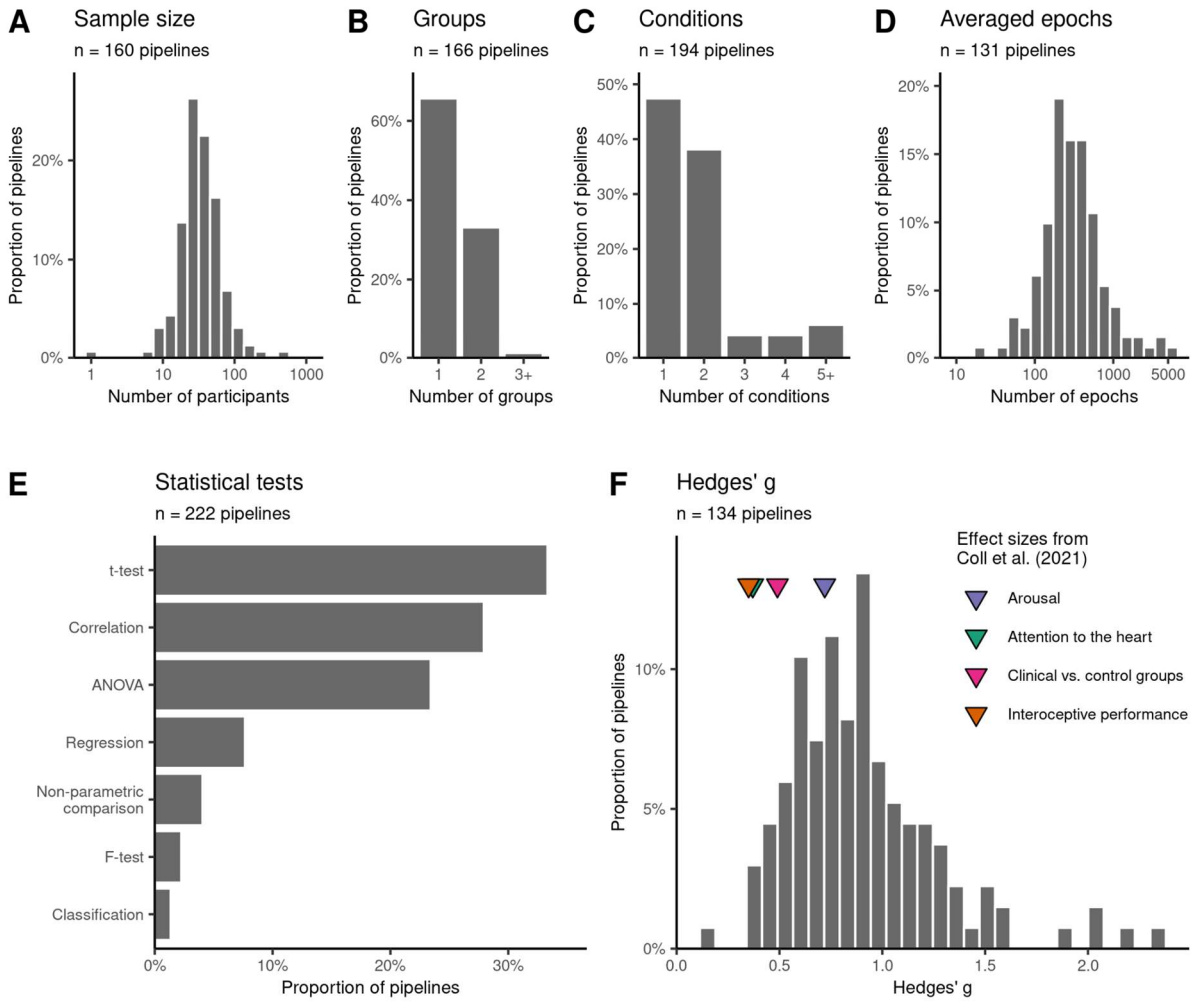

Figure 7. Statistical analysis and minimal detectable effect sizes. The settings used in experimental design and statistical analysis across studies are as follows: (A) sample size (x-

axis is log-scaled); (B) number of groups; (C) number of conditions; and (D) estimated number of epochs averaged to obtain HERs per participant and condition (x-axis is log-scaled). (E) The statistical tests that were used in the reviewed studies. (F) Minimal detectable effect sizes with 80% power at 5% significance level, and four HER effect sizes from the meta-analysis by Coll et al. (2021), visualized by colored triangles. Estimated minimal effect sizes are based on the type of statistical test, sample size, number of participants, groups, and conditions, and are only shown for pipelines that use averaged HER values for t-tests, correlations, and ANOVAs.

3.9. Control analysis

While small effect sizes could pose serious problems for the reproducibility of HER findings, even the presence of strongly significant, “large” effects does not guarantee that the observed HER effect reflects genuine heartbeat-evoked neural activity (Park & Blanke, 2019; Steinfath et al., 2025). Although many strategies exist to avoid or remove cardiac artifacts (discussed in Section 3.5), it is unclear how effective these are. To ensure that HER effects do not simply reflect cardiac artifacts, 65% of the studies (86/132) include ECG and other heartbeat-related controls in their statistical analysis (Fig. 8A). Specifically, it is common to repeat the analysis on ECG data alone or include it as a regressor of no interest in analyses (combined: 59/132, 45% of studies; Fig. 8B) to demonstrate that effects found at the level of M/EEG sensors do not only reflect cardiac field artifacts. Furthermore, many studies control for the length of the RR interval (45/132, 34%), for instance, by including it as a regressor of no interest in their analyses or by comparing the RR intervals of the studied groups. In addition, 26% of studies (34/132) control for heart rate variability, 6.1% (8/132) for respiration, and 3% (4/132) for blood pressure (Fig. 8A).

Apart from confounding factors that are directly related to cardiac activity, various heartbeat-independent factors can also affect HER. For instance, if HERs are analyzed that were recorded on-task, the concurrent activity evoked by external stimuli can overlap with HERs,

leading to spurious results. Similarly, group-related differences in SNR, for example, due to an imbalanced number of trials or systematic differences in artifacts, can lead to contaminated HER results. One solution to these problems is surrogate heartbeat control analysis (Park et al., 2014; Steinfath et al., 2025). To verify that observed effects are truly locked to the heartbeat, the onset timings of the cardiac markers within each condition, and the same analyses are repeated multiple times (e.g., >100 iterations). The outcome statistic of the original analysis is then compared to the distribution of statistics obtained from these permutations. Hence, it can only be concluded that the original effect is truly related to the heartbeat if it exceeds most of the permuted effects. In our review, we find that 19% of the studies (25/132) perform surrogate heartbeat control analyses or other variations of data shuffling. Currently, robust control methods for HER research are still being developed (those that exist are listed in the quality control checklist; Box 1). For a full overview of control variables that appeared in the reviewed studies, see Fig. S3.

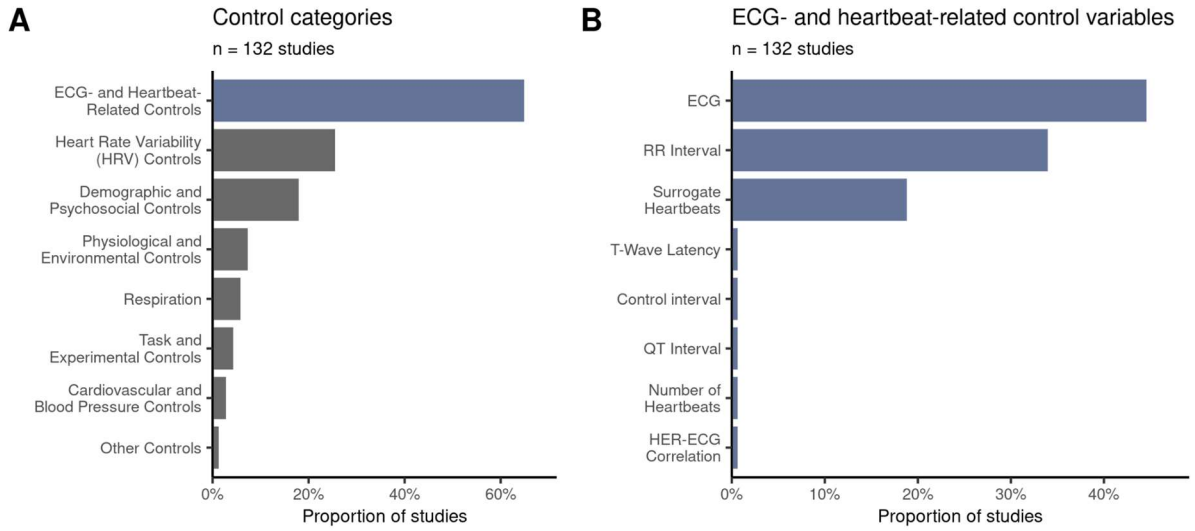

Figure 8. Variables and approaches commonly used for control analysis sorted by the frequency of occurrence in descending order. (A) Control variables grouped by category. (B) Detailed breakdown of ECG- and heartbeat-related control variables. Blue color indicates that items shown in Panel B compose the category ECG and Heartbeat Related Controls from Panel A.

680 *Box 1*. Quality-control checklist for HER analysis and reporting. We suggest key reporting
681 points and recommendations for HER studies based on the extracted information from the
682 reviewed studies and prior experience analyzing HERs.

Step 1 – Designing an experiment

A. Preregistration. Supports the reproducibility and reliability of findings, as well as the publication of null results. May incorporate several control analysis pipelines (e.g., with and without baseline correction).
B. Relationship between heartbeats and experimental design. HERs can overlap with task-related activity, and they might co-vary with differences in background activity across groups. Avoid overlap if possible, or prepare a strategy to account for it afterwards.
C. ECG electrode leads, ground, and location. Affect ECG-based control analyses. Report lead configuration used for analysis. Consider recording several leads to account for the three-dimensional field distribution.
D. Recording of physiological activity. Breathing, eye movements, cardiac and other parameters can confound HER measures.

Step 2 – Data preprocessing

A. Filtering. High- and low-pass filter settings can affect the HER shape.
B. Reference electrode. The choice of online and offline references can affect HER results and artifact removal.
C. R-peak or T-peak. Or other points of reference. Can be compared to each other.
D. Minimal IBI. Prevents contamination by overlapping heartbeats. Larger thresholds may induce imbalance in the number of epochs if the heart rate is different between conditions or groups, potentially signifying differences in arousal. Report the number of removed epochs.
E. General cleaning. Report approaches used for noise removal. For ICA, include the description of the algorithm, types of identified components (e.g., ocular, muscle), and rejection criteria.

Step 3 – Cardiac Field Artifact Removal (CFA)

A. Rejection of cardiac ICs (if applicable). State whether ICA was applied to epoched or continuous data. Describe rejection criteria and approach, number of removed cardiac-related components, and explained variance.
B. Other CFA removal methods. Detailed description of other approaches used to remove CFA. For example, regressing out ECG from M/EEG data.
C. Limiting analysis time window (optional). Avoid time ranges that are prone to cardiac artifact contamination.

Step 4 – HER computation

A. ROI in time and space. Describe start/end times & preselected sensors as well as the motivation for the selection.
B. Baseline correction. Clearly state if omitted or performed. If performed, include information on the method and time window.
C. Number of epochs. Describe how many epochs were averaged or otherwise analyzed to obtain HER per participant, group, and condition. This affects the signal-to-noise ratio.
D. Visualization. Show figures with HER results as evoked responses and as topographies.

Step 5 – Statistical analysis

A. Statistical power. Ensure that the analysis has sufficient statistical power to detect expected results reliably.
B. Regions of practical equivalence. For both Bayesian and frequentist statistics, define the smallest effect of interest to test if observed effects are negligibly small (i.e., "absent") and to avoid highlighting practically irreproducible (too small) effects.

Step 6 – Controls

A. Surrogate heartbeat analysis. Confirms that observed effects are genuinely locked to the heartbeat. Describe how it was performed and how it addressed potential confounding effects.
B. Task overlap controls. For example, pseudotrial correction can help isolate heartbeat-driven effects from task-related activity. Describe how it was performed and how it addressed potential confounding effects.
C. ECG control analysis. Assesses whether differences in ECG exist that might confound the HERs via cardiac artifacts, e.g. by repeating statistical analysis on the ECG channel. Use multiple leads if possible.
D. Effect of baseline. It is advisable to show the HER results without baseline correction.
E. Effect of CFA removal approach. It is advisable to show the HER results without CFA removal or with alternative cleaning approaches.
F. Other cardiac parameters. Can be included in statistical analysis. Examples: heart rate, heart rate variability, blood pressure.

Step 7 – After publication

A. Data sharing. Publicly shared data enable other researchers to reproduce current findings, develop new research questions, and, importantly, further HER methods. This facilitates more reliable, faster, cost-effective research.
B. Code sharing. Allows for the highest level of transparency in understanding the analysis steps.

Abbreviations: HERs – heartbeat-evoked responses, ECG – electrocardiography, IBI – inter-beat interval, M/EEG – magneto- or electroencephalography, ICA – independent component

analysis, ICs – independent components, ROI – region of interest

4. Discussion

This systematic review aimed to comprehensively characterize the diverse set of methods used in non-invasive heartbeat-evoked responses research using M/EEG. Our analysis of 132 publications revealed substantial heterogeneity across many processing steps, including data acquisition, data preprocessing, and statistical analyses. While some predominant methodological patterns emerge, such as a higher percentage of EEG compared to MEG studies and the frequent use of the R-peak as a cardiac reference marker, most other steps show considerable variability, as indicated by large entropy scores. For the majority of preprocessing steps, studies typically use only one of the many possible options.

Apart from documenting large variability in methodological choices, this review also reveals a large proportion (40-80%) of unreported decisions for processing steps that are critical for HER analysis such as the number of rejected CFA-related independent components (ICs), the number of epochs averaged to obtain HERs per participant and condition, or the baseline window approach used. Taken together with a worrying estimation that 67-99% of studies might be underpowered for detecting meta-level HER effects, it is critical that the community takes more decisive steps towards transparency and reliability of findings.

To alleviate reporting ambiguity and improve the robustness of experimental designs and analyses, we compiled a checklist containing key points that should be considered when conducting and reporting HER studies (Box 1). Since HERs are a subclass of evoked responses, many potential problems are general to the field of M/EEG research and are already summarized elsewhere (e.g. Keil et al., 2014). However, some important considerations are unique to HER analysis and require special attention to ensure the quality and reproducibility of results. While previous reviews already highlight some HER-specific best reporting practices (Azzalini et al., 2019; Park & Blanke, 2019), a comprehensive list of

707 reporting guidelines could be useful for experienced and new HER researchers, as well as
708 reviewers of HER studies. Our checklist provides a broad overview of existing approaches for
709 tackling artifacts (mainly, the CFA) and other confounders (e.g., overlap with task-evoked
710 activity).

711 Despite the overall lack of consensus in methodological approaches, our results highlight the
712 most common choices for each processing step. However, it is important to note that the
713 correctness and suitability of each choice cannot be judged solely based on its occurrence
714 frequency, especially when most studies only use one choice for each processing step. To
715 better understand the impact of the various methodological choices on the observed effects,
716 it is crucial to either perform simulation-based studies with known ground truth derived from
717 single cell recordings (De Falco et al., 2024), intracranial measurements (Kern et al., 2013),
718 or to perform multiverse analyses on real data (Steegen et al., 2016).

719 We provide an overview of the currently used methods that could serve as a starting point for
720 novel experimental designs or systematic method comparisons. In addition, we highlight
721 several processing steps which are likely to affect HERs, but their influence has not been
722 systematically studied yet (Box 2). Below, we reiterate the motivation for investigating these
723 steps in more detail.

724 *Box 2.* Outstanding methodological questions in HER research.

To further HER research, the development of methodology is required. Effects of the following processing steps, as well as their interactions, on HER results require further systematic evaluation:

1. High- and low-pass filter parameters
2. ECG recording and subsequent ECG-based CFA controls
3. CFA removal with ICA or other spatial filtering techniques with varying settings
4. Influence of individual differences in cardiac parameters on HER
5. Ocular, muscular, and respiratory artifact influence and cleaning
6. Baseline correction
7. Within-participant variability of HER epochs and required number of trials

The high likelihood of being confounded by cardiac artifacts is a major factor that distinguishes HERs from other ERPs. Importantly, the complex three-dimensional nature of cardiac electric fields, with dynamic changes in scalp potential patterns across the heart cycle, poses a critical challenge (Dirlich et al., 1997; Pérez et al., 2005). Decomposition methods, such as ICA, which aim to represent the data in lower-dimensional components, often struggle with time-dependent artifacts. Although the majority of studies use ICA to remove artifacts from the data, there is no consensus on the best approach for identifying CFA-related ICs. While different approaches lead to varying results in artifact suppression, no direct comparison between methods has been conducted to date. On one side, simulations are needed to evaluate the influence and removal efficacy of cardiac artifacts on the measured signal. On the other side, high-density ECG recordings with sophisticated forward models offer a promising tool for a better separation of CFA and brain activity in real data. The complex nature of the CFA is also affecting ECG-based control analyses, which aim to confirm that HER effects reflect differences in brain activity, rather than volume-conducted CFA. Although low-density ECG derivations are generally sufficient for identifying R- or T-peaks as reference events, they only capture a portion of the complex cardiac electrical field. Since experimental conditions can induce changes in ECG that might be detectable at some recording sites but not at others (Gray et al. 2007), several ECG electrodes covering all three dimensions of the ECG are necessary to show that HER effects are not volume-conducted CFA.

Even though the removal of cardiac artifacts plays an important role in HER research, artifact correction also carries the risk of removing genuine parts from the HER. Hence, many researchers analyze predefined time ranges for their analysis, which show low influence of the CFA (Dirlich et al., 1997). While early HER studies aimed to identify spatiotemporal points where the influence of cardiac-related artifacts is minimal (Dirlich et al., 1997), our review finds little progress in recent years toward identifying and validating these optimal time ranges. Therefore, it remains essential to ensure that changes in HER are not accompanied by changes in cardiac parameters.

Our review supports the idea of considerable temporal variability with predominantly frontocentral HER effects across M/EEG studies. We observe a wide range of chosen time windows, with a preference for limiting analyses to a time range of approximately 200 to 600 ms after the R-peak. This time range is in agreement with intracranial recordings, revealing heartbeat-related brain activity to emerge primarily between 200 and 400 ms after the R-peak (De Falco et al., 2024; Kern et al., 2013). However, even in these invasive studies, there is considerable variability regarding the onset, duration, and anatomical location of cardiac-related neural responses. This spatiotemporal variability in observed effects suggests that there is no single manifestation of HER. Instead, those pieces of evidence point to a wide range of cardiac-locked modulations of neuronal activity depending on the specific task, cognitive state, experimental condition, or pathology.

Spatiotemporal variability of the HER affects the suitable analysis options. Ideally, spatiotemporal regions of interest are defined a priori based on previous studies or independent analyses. However, many researchers use hypothesis-free cluster-based permutation testing, likely due to inconsistent HER findings in the field. In addition, we observe that many studies use a two-step procedure, where the region of interest and time range are first defined based on primary cluster-based permutation analysis. The resulting significant cluster is then used in further secondary analyses. While this approach is convenient to circumvent the variable nature of the HER, it is important to keep in mind that cluster-based permutation tests do not allow for exact spatial or temporal interpretation of effects (Maris & Oostenveld, 2007; Sassenhagen & Draschkow, 2019). Furthermore, if the analysis used for data selection is not independent of the secondary analysis, circularity may become a concern (Kriegeskorte, Simmons, et al., 2010). While this does not mean that the results are incorrect, it underscores the need for independent replication and careful evaluation of the statistical validity of such two-step analyses.

Careful experimental design and control analyses are essential to isolate genuine HER activity from other sources of variability. When HERs are studied in task settings, as is the case in

most of the reviewed studies, task-related evoked responses may overlap with internal responses like the HER. This opens additional challenges. On the one hand, researchers have to make sure that potential HER effects are not a result of temporally overlapping task-related differences in neural activity. Currently, we find that there are two approaches used in the literature to deal with this situation: surrogate heartbeat analyses (Park et al., 2014) and pseudotrial correction (Steinfath et al., 2025). On the other hand, if effects in task-evoked activity are observed, researchers have to make sure that these are not a result of overlapping cardiac-related activity, given that heartbeats may align to task triggers (e.g., Kunzendorf et al., 2019). Overall, disentangling overlapping effects might also be approached by regression-based ERP modelling (Ehinger & Dimigen, 2019). However, the studies included in our review indicate that, so far, it has not gained wider use in the HER research community. Further work is therefore required to show in which scenarios the different methods work best and where the limitations are.

Adding to these challenges, the heart does not operate in isolation and is coupled to several other physiological systems (Hall & Hall, 2021). This coupling can complicate the delineation of genuine HER effects from other correlated sources of variability, such as eye movements (Ohl et al., 2016) or respiration (Yasuma & Hayano, 2004). Consequently, careful experimental design and control analyses are required to delineate heartbeat-related effects from other influences. Importantly, our review revealed that, across HER studies, a broad range of different control variables has been included in the analyses. While eye movements or respiration are not among the most frequent choices, some experimental designs may be more prone to these specific confounds, which warrant explicit care.

The current review has several limitations. Since only information about methods was extracted, we cannot compare and recommend any methods solely based on how often they are used. Furthermore, methods used by particular HER research laboratories might be overrepresented in the overall results. In fact, we observed that 20% of studies in the reviewed sample (27/132) have the same three last authors. In addition, we did not extract the effect

sizes per reviewed study or relate them to the employed methods due to the substantial heterogeneity in research questions. Therefore, an objective comparison of methodological choices for HER analysis on the same dataset is a crucial next step for the advancement of HER methodology. Additionally, we did not extensively extract all non-significant control analyses, which might also bias the representativeness of some of the results. It is critical to remember that any review is subject to publication bias when it comes to the included studies (Thornton & Lee, 2000). Overall, we find that a stunning 93% of reviewed studies reported statistically significant findings. Given that we also observe that a large proportion of studies might be underpowered, the small number of studies with non-significant findings is another evidence of a possible publication bias (Button et al., 2013).

The problem of publication bias, however, can be remedied in the future of the HER field by pursuing more transparent practices such as preregistrations and registered reports, and reporting the significance of null findings through Bayesian approaches (Wagenmakers et al., 2018) or frequentist equivalence testing (Lakens et al., 2018). We suggest that, just as for other research domains, preregistrations and especially registered reports can increase the quality of HER research and are a key solution to many of the issues outlined in this review (Chambers & Tzavella, 2022; Hardwicke & Wagenmakers, 2023; Soderberg et al., 2021). In such a growing field, it might be difficult to navigate the multiplicity of methodological choices across the many processing and analysis steps, and prepare a robust research plan in advance. This is further complicated by variability in reported HER findings. Characterizing the temporal, spatial, and spectral components of HER across contexts is the key direction of future research in the field, but this, first and foremost, warrants the advancement of understanding how to analyze HERs in the most robust way and heavily relies on transparent and reliable findings. This review and future, much-needed, methodology-oriented HER studies will shed light on these decisions and help researchers in preparing their HER studies. Ultimately, accumulation of robust evidence through reliable research with transparent

methodology is the way to understanding what HER is, and thereby, how the brain and body interact.

Our review provides a set of reasonable candidate methods to consider for the preprocessing and analysis of HERs, and the extracted information is published with the article. To make the results more accessible, we also provide an easy-to-use interactive interface for inspecting the dataset. We hope that this review will serve as a foundation for improving the methodological consistency across HER studies.

Statement on gender composition within HER research and citation diversity

Recent evidence suggests that publications with female names at prominent positions in the author list are cited less often in the field of neuroscience than should be expected (Dworkin et al., 2020; Hefter et al., 2025). In light of recent discussions by journals and publishers (e.g., "Citation Diversity Statement", 2023; Fulvio et al., 2021; Sweet, 2021), we analyzed gender approximations across the HER literature included in this review to guide the benchmarking of future citation practices.

As done by Dworkin et al. (2020), gender estimates (F = female; M = male) were extracted for the names of the first and last authors across all references (n = 132) through an automated service (https://gender-api.com) and the cleanBib toolbox (D. Zhou et al., 2022). In case of low assignment probability (<.80), publicly available pronouns were manually searched for. Following the exclusion of those for which this failed (n = 2), the author pairs were classified as 18% (n = 24) FF (first author/last author), 10% (n = 13) MF, 31% (n = 40) FM, and 41% (n = 53) MM. There were no publications by singular authors. Four publications categorized as FM were authored by our lab.

A median split by publication year showed that gender balance improved over time (since October 2020 [Δ before]: 23% FF [+9.2%], 14% MF [+7.7%], 28% FM [-6.2%], 35% MM [-

11%]). The observed pattern resembles the gender distribution (18% FF, 14% MF, 29% FM, 39% MM) that resulted from large-scale analyses by Hefter et al. (2025). For more detailed analysis see Section D of the Supplementary Material.

To note, the applied procedure falls short of some aspects: 1) shared first and/or last authorship were not considered, 2) gender estimates and openly accessible pronouns may not have been, in full, reflective of the respective individuals' gender identities, and 3) gender was taken as dichotomous variable only, and thus, the procedure does not allow for monitoring gender representation in its full variety. Lastly, some clustering of HER publications (using a threshold of ≥5) was observed, with 27 studies (20%) last authored by the same three names (2 F, 1 M), and 15 studies (11%) first authored by another set of three names (1 F, 2 M).
Lastly, we examined our own citation behavior in the main text (n = 102). After excluding four references with undetermined author gender, the remaining were classified as 16% (n = 16) FF, 13% (n = 13) MF, 28% (n = 27) FM, and 43% (n = 42) MM. These proportions include five single-authored papers (1 FF, 4 MM) and four from our lab (1 MM; 3 FM, with two amongst the reviewed studies). In total, 42 of the 132 studies included in the review (31.8%) were also cited in the main text.

Studies included in the review

(Al et al., 2020, 2021, 2023; Alkhachroum et al., 2024; Azzalini et al., 2021; Babo-Rebelo et al., 2019; Babo-Rebelo, Richter, et al., 2016; Babo-Rebelo, Wolpert, et al., 2016; Banellis & Cruse, 2020; Baranauskas et al., 2017; Billeci et al., 2021; Birba et al., 2022; Bogdány et al., 2022; Buot et al., 2021; Callara et al., 2023; Cambi et al., 2024; Candia-Rivera, Catrambone, et al., 2021; Canales-Johnson et al., 2015; Candia-Rivera, Annen, et al., 2021; Candia-Rivera et al., 2023; Candia-Rivera & Machado, 2023a, 2023b; Couto et al., 2014, 2015; de la Fuente et al., 2019; Di Bernardi Luft & Bhattacharya, 2015; Dirlich et al., 1997, 1998; Elkommos et al., 2023; Engelen, Buot, et al., 2023; Fittipaldi et al., 2020; Flasbeck et al., 2020, 2021, 2024; Fouragnan et al., 2024; Fukushima et al., 2011; Gao et al., 2023; García-Cordero et al., 2017,

2016; Gentsch et al., 2019; Giusti et al., 2024; Gray et al., 2007; Hermann et al., 2024; Hodossy et al., 2021; Immanuel et al., 2014; Ito et al., 2019; Judah et al., 2018; Kamp et al., 2021, 2023; Katkin et al., 1991; Kato et al., 2020; Khoshnoud et al., 2022, 2024; Kim et al., 2019; Koreki, 2024; Kritzman et al., 2022; Kumral et al., 2022; Lechinger et al., 2015; Legaz et al., 2022; Leopold & Schandry, 2001; Liu et al., 2022, 2023; Liuzzi et al., 2024; Lutz et al., 2019; MacKinnon et al., 2013; Mai et al., 2018; Maister et al., 2017; Majeed et al., 2022; Marshall et al., 2017, 2018, 2019, 2020, 2022; McCraty et al., 2004; Melo et al., 2022; Migeot et al., 2023; Montoya et al., 1993; Müller et al., 2015; Ono et al., 2024; Pang et al., 2019; Park et al., 2014, 2016; Pelentritou et al., 2024; Pérez et al., 2005; Perogamvros et al., 2019; Petzschner et al., 2019; Pollatos et al., 2005, 2016; Pollatos & Schandry, 2004; Poppa et al., 2022; Rapp et al., 2023; Ren et al., 2022; Richter et al., 2021; Richter & Ibáñez, 2021; Salamone et al., 2018, 2020, 2021; Schandry et al., 1986; Schandry & Montoya, 1996; Schandry & Weitkunat, 1990; Schmitz et al., 2020, 2021; Schulz et al., 2013, 2018, 2020; Schulz, Ferreira de Sá, et al., 2015; Schulz, Köster, et al., 2015; Sel et al., 2016; Shao et al., 2011; Simor et al., 2021; Solcà et al., 2020; Stoupi et al., 2024; Tanaka et al., 2023; Terhaar et al., 2012; Todd et al., 2021; Ventura-Bort & Weymar, 2024; Verdonk et al., 2021, 2024; Villena-González et al., 2017, 2023; Wei et al., 2016; Weijs et al., 2023; A. Yoris et al., 2017, 2018; A. E. Yoris et al., 2024; Yuan et al., 2007; Zaccaro et al., 2022, 2024; Zhang et al., 2023; H. Zhou et al., 2022; M. Zhou et al., 2024; Zwienenberg et al., 2023)

CRediT

P.S.; M.A.; N.K.: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing – original draft, Writing – review & editing, Visualization, Project administration. T. L.: Data Curation, Writing – review & editing. L.E.: Gender Citation Diversity Statement, Writing – review & editing. V.N.; A.V.: Writing – review & editing, Supervision, Funding acquisition.

Data and code availability statement

The full data table containing all relevant information extracted from the literature that is used in this systematic review can be accessed at https://osf.io/znrbm. In addition, we provide an interactive web application that allows browsing the dataset: https://paulsteinfath.shinyapps.io/her-systematic-review/. All analysis scripts are available at https://github.com/PaulSteinfath/systematic-hep-review. Citation diversity code along with gender approximation datasets can be accessed at https://github.com/liobaenk/citation-diversity-tracking

Funding sources

This project was supported by the Max Planck Society. T.L. received funding from the Ministry of Health and Social Security of the Government of Luxembourg.

Declaration of generative AI in scientific writing

During the preparation of this work the authors used OpenAI chatGPT (GPT-4o), Anthropic Claude 4, and Grammarly in order to proofread the text, and assist during analysis, code generation and documentation. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Declaration of competing interest

The authors declare no conflicts of interest.

References

Al, E., Iliopoulos, F., Forschack, N., Nierhaus, T., Grund, M., Motyka, P., Gaebler, M., Nikulin, V. V., & Villringer, A. (2020). Heart-brain interactions shape somatosensory perception

and evoked potentials. Proceedings of the National Academy of Sciences of the United States of America, 117(19), 10575–10584. https://doi.org/10.1073/pnas.1915629117

Al, E., Iliopoulos, F., Nikulin, V. V., & Villringer, A. (2021). Heartbeat and somatosensory perception. NeuroImage, 238(February), 118247. https://doi.org/10.1016/j.neuroimage.2021.118247

Al, E., Stephani, T., Engelhardt, M., Haegens, S., Villringer, A., & Nikulin, V. V. (2023). Cardiac activity impacts cortical motor excitability. PLOS Biology, 21(11), e3002393. https://doi.org/10.1371/journal.pbio.3002393

Alkhachroum, A., Fló, E., Manolovitz, B., Cohan, H., Shammassian, B., Bass, D., Aklepi, G., Monexe, E., Ghamasaee, P., Sobczak, E., Samano, D., Saavedra, A. B., Massad, N., Kottapally, M., Merenda, A., Cordeiro, J. G., Jagid, J., Kanner, A. M., Rundek, T., … Sitt, J. D. (2024). Resting-State EEG Signature of Early Consciousness Recovery in Comatose Patients with Traumatic Brain Injury. Neurocritical Care, 41(3), 855–865. https://doi.org/10.1007/s12028-024-02005-2

Azzalini, D., Buot, A., Palminteri, S., & Tallon-Baudry, C. (2021). Responses to heartbeats in ventromedial prefrontal cortex contribute to subjective preference-based decisions. Journal of Neuroscience, 41(23), 5102–5114. https://doi.org/10.1523/JNEUROSCI.1932-20.2021

Azzalini, D., Rebollo, I., & Tallon-Baudry, C. (2019). Visceral Signals Shape Brain Dynamics and Cognition. Trends in Cognitive Sciences, 23(6), 488–509. https://doi.org/10.1016/j.tics.2019.03.007

Babo-Rebelo, M., Buot, A., & Tallon-Baudry, C. (2019). Neural responses to heartbeats distinguish self from other during imagination. NeuroImage, 191, 10–20. https://doi.org/10.1016/j.neuroimage.2019.02.012

Babo-Rebelo, M., Richter, C. G., & Tallon-Baudry, C. (2016). Neural responses to heartbeats in the default network encode the self in spontaneous thoughts. Journal of Neuroscience, 36(30), 7829–7840. https://doi.org/10.1523/JNEUROSCI.0262-16.2016

Babo-Rebelo, M., Wolpert, N., Adam, C., Hasboun, D., & Tallon-Baudry, C. (2016). Is the cardiac monitoring function related to the self in both the default network and right anterior insula? Philosophical Transactions of the Royal Society B: Biological Sciences, 371(1708). https://doi.org/10.1098/rstb.2016.0004

Banellis, L., & Cruse, D. (2020). Skipping a beat: Heartbeat-evoked potentials reflect predictions during interoceptive-exteroceptive integration. Cerebral Cortex Communications, 1–14. https://doi.org/10.1093/texcom/tgaa060

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 and Behavior, 180, 1–7. https://doi.org/10.1016/j.physbeh.2017.07.032

Billeci, L., Faraguna, U., Santarcangelo, E. L., d'Ascanio, P., Varanini, M., & Sebastiani, L. (2021). Heartbeat-Evoked Cortical Potential during Sleep and Interoceptive Sensitivity: A Matter of Hypnotizability. Brain Sciences, 11(8), 1089. https://doi.org/10.3390/brainsci11081089

Birba, A., Santamaría-García, H., Prado, P., Cruzat, J., Ballesteros, A. S., Legaz, A., Fittipaldi, S., Duran-Aniotz, C., Slachevsky, A., Santibañez, R., Sigman, M., García, A. M., Whelan, R., Moguilner, S., & Ibáñez, A. (2022). Allostatic-Interoceptive Overload in Frontotemporal Dementia. Biological Psychiatry, 92(1), 54–67. https://doi.org/10.1016/j.biopsych.2022.02.955

Bogdány, T., Perakakis, P., Bódizs, R., & Simor, P. (2022). The heartbeat evoked potential is a questionable biomarker in nightmare disorder: A replication study. NeuroImage: Clinical, 33. https://doi.org/10.1016/j.nicl.2021.102933

Botvinik-Nezer, R., & Wager, T. D. (2023). Reproducibility in Neuroimaging Analysis: Challenges and Solutions. Biological Psychiatry. Cognitive Neuroscience and Neuroimaging, 8(8), 780–788. https://doi.org/10.1016/j.bpsc.2022.12.006

Boudewyn, M. A., Luck, S. J., Farrens, J. L., & Kappenman, E. S. (2018). How many trials does it take to get a significant ERP effect? It depends. Psychophysiology, 55(6), e13049. https://doi.org/10.1111/psyp.13049

Buot, A., Azzalini, D., Chaumon, M., & Tallon-Baudry, C. (2021). Does stroke volume influence heartbeat evoked responses? Biological Psychology, 165, 108165. https://doi.org/10.1016/j.biopsycho.2021.108165

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. Nature Reviews. Neuroscience, 14(5), 365–376. https://doi.org/10.1038/nrn3475

Callara, A. L., Fontanelli, L., Belcari, I., Rho, G., Greco, A., Zelič, Ž., Sebastiani, L., & Santarcangelo, E. L. (2023). Modulation of the heartbeat evoked cortical potential by hypnotizability and hypnosis. Psychophysiology, 60(9), e14309. https://doi.org/10.1111/psyp.14309

Cambi, S., Solcà, M., Micali, N., & Berchio, C. (2024). Cardiac interoception in Anorexia Nervosa: A resting-state heartbeat-evoked potential study. European Eating Disorders Review, 32(3), 417–430. https://doi.org/10.1002/erv.3049

Campos Viola, F., Thorne, J., Edmonds, B., Schneider, T., Eichele, T., & Debener, S. (2009). Semi-automatic identification of independent components representing EEG artifact. Clinical Neurophysiology, 120(5), 868–877. https://doi.org/10.1016/j.clinph.2009.01.015

Canales-Johnson, A., Silva, C., Huepe, D., Rivera-Rei, Á., Noreika, V., Del Carmen Garcia, M., 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, 25(11), 4490–4503. https://doi.org/10.1093/cercor/bhv076

Candia-Rivera, D., Annen, J., Gosseries, O., Martial, C., Thibaut, A., Laureys, S., & Tallon-Baudry, C. (2021). Neural responses to heartbeats detect residual signs of consciousness during resting state in postcomatose patients. Journal of Neuroscience, 41(24), 5251–5262. https://doi.org/10.1523/JNEUROSCI.1740-20.2021

Candia-Rivera, D., Catrambone, V., & Valenza, G. (2021). The role of electroencephalography electrical reference in the assessment of functional brain–heart interplay: From methodology to user guidelines. Journal of Neuroscience Methods, 360, 109269. https://doi.org/10.1016/j.jneumeth.2021.109269

Candia-Rivera, D., & Machado, C. (2023a). Multidimensional assessment of heartbeat-evoked responses in disorders of consciousness. European Journal of Neuroscience, 58(4), 3098–3110. https://doi.org/10.1111/ejn.16079

Candia-Rivera, D., & Machado, C. (2023b). Reduced Heartbeat-Evoked Responses in a Near-Death Case Report. Journal of Clinical Neurology, 19(6), 581. https://doi.org/10.3988/jcn.2022.0415

Candia-Rivera, D., Raimondo, F., Pérez, P., Naccache, L., Tallon-Baudry, C., & Sitt, J. D. (2023). Conscious processing of global and local auditory irregularities causes differentiated heartbeat-evoked responses. eLife, 12, e75352. https://doi.org/10.7554/eLife.75352

Chambers, C. D., & Tzavella, L. (2022). The past, present and future of Registered Reports. Nature Human Behaviour, 6(1), 29–42. https://doi.org/10.1038/s41562-021-01193-7

Champely, S., Ekstrom, C., Dalgaard, P., Gill, J., Weibelzahl, S., Anandkumar, A., Ford, C., Volcic, R., & Rosario, H. D. (2020). pwr: Basic Functions for Power Analysis (Version 1.3-0) [Computer software]. https://cran.r-project.org/web/packages/pwr/index.html

Chaumon, M., Bishop, D. V. M., & Busch, N. A. (2015). A practical guide to the selection of independent components of the electroencephalogram for artifact correction. Journal of Neuroscience Methods, 250, 47–63. https://doi.org/10.1016/j.jneumeth.2015.02.025

Citation diversity statement. (2023). Nature Reviews Bioengineering, 1(4), 227–227. https://doi.org/10.1038/s44222-023-00061-5

Clayson, P. E., Baldwin, S. A., Rocha, H. A., & Larson, M. J. (2021). The data-processing multiverse of event-related potentials (ERPs): A roadmap for the optimization and standardization of ERP processing and reduction pipelines. NeuroImage, 245(November), 118712. https://doi.org/10.1016/j.neuroimage.2021.118712

Coll, M.-P., Hobson, H., Bird, G., & Murphy, J. (2021). Systematic review and meta-analysis of the relationship between the heartbeat-evoked potential and interoception. Neuroscience and Biobehavioral Reviews, 122, 190–200. https://doi.org/10.1016/j.neubiorev.2020.12.012

Couto, B., Adolfi, F., Velasquez, M., Mesow, M., Feinstein, J., Canales-Johnson, A., Mikulan, E., Martínez-Pernía, D., Bekinschtein, T., Sigman, M., Manes, F., & Ibanez, A. (2015). Heart evoked potential triggers brain responses to natural affective scenes: A preliminary study. Autonomic Neuroscience: Basic & Clinical, 193, 132–137. https://doi.org/10.1016/j.autneu.2015.06.006

Couto, B., Salles, A., Sedeño, L., Peradejordi, M., Barttfeld, P., Canales-Johnson, A., Dos Santos, Y. V., Huepe, D., Bekinschtein, T., Sigman, M., Favaloro, R., Manes, F., & Ibanez, A. (2014). The man who feels two hearts: The different pathways of interoception. Social Cognitive and Affective Neuroscience, 9(9), 1253–1260. https://doi.org/10.1093/scan/nst108

Criscuolo, A., Schwartze, M., & Kotz, S. A. (2022). Cognition through the lens of a body–brain dynamic system. Trends in Neurosciences, 45(9), 667–677. https://doi.org/10.1016/j.tins.2022.06.004

De Falco, E., Solcà, M., Bernasconi, F., Babo-Rebelo, M., Young, N., Sammartino, F., Tallon-Baudry, C., Navarro, V., Rezai, A. R., Krishna, V., & Blanke, O. (2024). Single neurons in the thalamus and subthalamic nucleus process cardiac and respiratory signals in humans. Proceedings of the National Academy of Sciences, 121(11), e2316365121. https://doi.org/10.1073/pnas.2316365121

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

Di Bernardi Luft, C., & Bhattacharya, J. (2015). Aroused with heart: Modulation of heartbeat evoked potential by arousal induction and its oscillatory correlates. Scientific Reports, 5(March), 1–11. https://doi.org/10.1038/srep15717

Dirlich, G., Dietl, T., Vogl, L., & Strian, F. (1998). Topography and morphology of heart action-related EEG potentials. Electroencephalography and Clinical Neurophysiology, 108(3), 299–305. https://doi.org/10.1016/s0168-5597(98)00003-3

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

Dworkin, J. D., Linn, K. A., Teich, E. G., Zurn, P., Shinohara, R. T., & Bassett, D. S. (2020). The extent and drivers of gender imbalance in neuroscience reference lists. Nature Neuroscience, 23(8), 918–926. https://doi.org/10.1038/s41593-020-0658-y

Ehinger, B. V., & Dimigen, O. (2019). Unfold: An integrated toolbox for overlap correction, non-linear modeling, and regression-based EEG analysis. PeerJ, 2019(10), 1–33. https://doi.org/10.7717/peerj.7838

Elkommos, S., Martin-Lopez, D., Koreki, A., Jolliffe, C., Kandasamy, R., Mula, M., Critchley, H. D., Edwards, M. J., Garfinkel, S., Richardson, M. P., & Yogarajah, M. (2023). Changes in the heartbeat-evoked potential are associated with functional seizures. Journal of Neurology, Neurosurgery & Psychiatry, 94(9), 769–775. https://doi.org/10.1136/jnnp-2022-330167

Elson, M. (2019). Examining Psychological Science Through Systematic Meta-Method Analysis: A Call for Research. Advances in Methods and Practices in Psychological Science, 2(4), 350–363. https://doi.org/10.1177/2515245919863296

Engelen, T., Buot, A., Grèzes, J., & Tallon-Baudry, C. (2023). Whose emotion is it? Perspective matters to understand brain-body interactions in emotions. NeuroImage, 268, 119867. https://doi.org/10.1016/j.neuroimage.2023.119867

Engelen, T., Solcà, M., & Tallon-Baudry, C. (2023). Interoceptive rhythms in the brain. Nature Neuroscience. https://doi.org/10.1038/s41593-023-01425-1

Feldman, M. J., Bliss-Moreau, E., & Lindquist, K. A. (2024). The neurobiology of interoception and affect. Trends in Cognitive Sciences, 28(7), 643–661. https://doi.org/10.1016/j.tics.2024.01.009

Fittipaldi, S., Abrevaya, S., Fuente, A. de la, Pascariello, G. O., Hesse, E., Birba, A., Salamone, P., Hildebrandt, M., Martí, S. A., Pautassi, R. M., Huepe, D., Martorell, M. M., Yoris, A., Roca, M., García, A. M., Sedeño, L., & Ibáñez, A. (2020). A multidimensional and multi-feature framework for cardiac interoception. NeuroImage, 212, 116677. https://doi.org/10.1016/j.neuroimage.2020.116677

Flasbeck, V., Bamberg, C., & Brüne, M. (2021). Short-Term Fasting and Ingestion of Caloric Drinks Affect Heartbeat-Evoked Potentials and Autonomic Nervous System Activity in Males. Frontiers in Neuroscience, 15. https://www.frontiersin.org/articles/10.3389/fnins.2021.622428

Flasbeck, V., Jungilligens, J., Lemke, I., Beckers, J., Öztürk, H., Wellmer, J., Seliger, C., Juckel, G., & Popkirov, S. (2024). Heartbeat evoked potentials and autonomic arousal during dissociative seizures: Insights from electrophysiology and neuroimaging. BMJ Neurology Open, 6(1), e000665. https://doi.org/10.1136/bmjno-2024-000665

Flasbeck, V., Popkirov, S., Ebert, A., & Brüne, M. (2020). Altered interoception in patients with borderline personality disorder: A study using heartbeat-evoked potentials. Borderline Personality Disorder and Emotion Dysregulation, 7(1), 24. https://doi.org/10.1186/s40479-020-00139-1

Fouragnan, E. F., Hosking, B., Cheung, Y., Prakash, B., Rushworth, M., & Sel, A. (2024). Timing along the cardiac cycle modulates neural signals of reward-based learning. Nature Communications, 15(1), 2976. https://doi.org/10.1038/s41467-024-46921-5

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

Fulvio, J. M., Akinnola, I., & Postle, B. R. (2021). Gender (Im)balance in Citation Practices in Cognitive Neuroscience. Journal of Cognitive Neuroscience, 33(1), 3–7. https://doi.org/10.1162/jocn_a_01643

Gao, J., Sun, R., Leung, H. K., Roberts, A., Wu, B. W. Y., Tsang, E. W., Tang, A. C. W., & Sik, H. H. (2023). Increased neurocardiological interplay after mindfulness meditation: A brain oscillation-based approach. Frontiers in Human Neuroscience, 17, 1008490. https://doi.org/10.3389/fnhum.2023.1008490

García-Cordero, I., Esteves, S., Mikulan, E. P., Hesse, E., Baglivo, F. H., Silva, W., García, M. del 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(JUL). https://doi.org/10.3389/fnins.2017.00411

García-Cordero, I., Sedeno, L., Slachevsky, A., Forno, G., Klein, F., Lillo, P., Ferrari, J., Rodriguez, C., Bustin, J., Torralva, T., Baez, S., Yoris, A., Esteves, S., Melloni, M., Salamone, P., Huepe, D., Manes, F., Garcıa, A. M., & Ibanez, A. (2016). Feeling, learning from and being aware of inner states: Interoceptive dimensions in neurodegeneration and stroke. 371.

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. The Journal of Neuroscience, 34(19), 6573–6582. https://doi.org/10.1523/JNEUROSCI.3507-13.2014

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

Gibney, K. D., Kypriotakis, G., Cinciripini, P. M., Robinson, J. D., Minnix, J. A., & Versace, F. (2020). Estimating statistical power for event-related potential studies using the late positive potential. Psychophysiology, 57(2), e13482. https://doi.org/10.1111/psyp.13482

Giusti, G., Zelič, Ž., Callara, A. L., Sebastiani, L., & Santarcangelo, E. L. (2024). Interoception as a function of hypnotizability during rest and a heartbeat counting task. Psychophysiology, 61(6), e14535. https://doi.org/10.1111/psyp.14535

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

Hall, J. E., & Hall, M. E. (2021). Guyton and Hall textbook of medical physiology (14th edition). Elsevier.

Hardwicke, T. E., & Wagenmakers, E.-J. (2023). Reducing bias, increasing transparency and calibrating confidence with preregistration. Nature Human Behaviour, 7(1), 15–26. https://doi.org/10.1038/s41562-022-01497-2

Hedges, L. V., & Olkin, I. (1985). Statistical Methods for Meta-Analysis. Academic Press.

Hefter, K., Teich, E. G., & Bassett, D. S. (2025). Gender imbalance in citations most pronounced in high-impact neuroscience journals (p. 2025.02.19.638905). bioRxiv. https://doi.org/10.1101/2025.02.19.638905

Hermann, B., Candia-Rivera, D., Sharshar, T., Gavaret, M., Diehl, J., Cariou, A., & Benghanem, S. (2024). Aberrant brain–heart coupling is associated with the severity of post cardiac arrest brain injury. Annals of Clinical and Translational Neurology, 11(4), 866–882. https://doi.org/10.1002/acn3.52000

Hodossy, L., Ainley, V., & Tsakiris, M. (2021). How do we relate to our heart? Neurobehavioral differences across three types of engagement with cardiac interoception. Biological Psychology, 165, 108198. https://doi.org/10.1016/j.biopsycho.2021.108198

Immanuel, S. A., Pamula, Y., Kohler, M., Martin, J., Kennedy, D., Nalivaiko, E., Saint, D. A., & Baumert, M. (2014). Heartbeat evoked potentials during sleep and daytime behavior in children with sleep-disordered breathing. American Journal of Respiratory and Critical Care Medicine, 190(10), 1149–1157. https://doi.org/10.1164/rccm.201405-0920OC

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

Jammal Salameh, L., Bitzenhofer, S. H., Hanganu-Opatz, I. L., Dutschmann, M., & Egger, V. (2024). Blood pressure pulsations modulate central neuronal activity via mechanosensitive ion channels. Science, 383(6682), eadk8511. https://doi.org/10.1126/science.adk8511

Judah, M. R., Shurkova, E. Y., Hager, N. M., White, E. J., Taylor, D. L., & Grant, D. M. M. (2018). The relationship between social anxiety and heartbeat evoked potential amplitude. Biological Psychology, 139(December 2017), 1–7. https://doi.org/10.1016/j.biopsycho.2018.09.013

Kamp, S.-M., Buntić, N., Amtmann, J., Scharpf, A., Schönen, A., Wagner, L., & Schulz, A. (2023). Reduced concentration performance and heartbeat-evoked potential in individuals with a history of a SARS-CoV-2 infection. Neuroscience Letters, 814, 137466. https://doi.org/10.1016/j.neulet.2023.137466

Kamp, S.-M., Schulz, A., Forester, G., & Domes, G. (2021). Older adults show a higher heartbeat-evoked potential than young adults and a negative association with everyday metacognition. Brain Research, 1752(January), 147238. https://doi.org/10.1016/j.brainres.2020.147238

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. https://doi.org/10.3109/00207459108990745

Kato, Y., Takei, Y., Umeda, S., Mimura, M., & Fukuda, M. (2020). Alterations of Heartbeat Evoked Magnetic Fields Induced by Sounds of Disgust. Frontiers in Psychiatry, 11, 683. https://doi.org/10.3389/fpsyt.2020.00683

Keil, A., Debener, S., Gratton, G., Junghöfer, M., Kappenman, E. S., Luck, S. J., Luu, P., Miller, G. A., & Yee, C. M. (2014). Committee report: Publication guidelines and recommendations for studies using electroencephalography and magnetoencephalography. Psychophysiology, 51(1), 1–21. https://doi.org/10.1111/psyp.12147

Kern, M., Aertsen, A., Schulze-Bonhage, A., & Ball, T. (2013). Heart cycle-related effects on event-related potentials, spectral power changes, and connectivity patterns in the human ECoG. NeuroImage, 81, 178–190. https://doi.org/10.1016/j.neuroimage.2013.05.042

Khoshnoud, S., Alvarez Igarzábal, F., & Wittmann, M. (2022). Brain–Heart Interaction and the Experience of Flow While Playing a Video Game. Frontiers in Human Neuroscience, 16, 819834. https://doi.org/10.3389/fnhum.2022.819834

Khoshnoud, S., Leitritz, D., Çinar Bozdağ, M., Alvarez Igarzábal, F., Noreika, V., & Wittmann, M. (2024). When the Heart Meets the Mind: Exploring the Brain–Heart Interaction during Time Perception. The Journal of Neuroscience, 44(34), e2039232024. https://doi.org/10.1523/JNEUROSCI.2039-23.2024

Kim, J., & Jeong, B. (2019). Heartbeat induces a cortical theta-synchronized network in the resting state. eNeuro, 6(4), 1–11. https://doi.org/10.1523/ENEURO.0200-19.2019

Kim, J., Park, H. D., Kim, K. W., Shin, D. W., Lim, S., Kwon, H., Kim, M. Y., Kim, K., & Jeong, B. (2019). Sad faces increase the heartbeat-associated interoceptive information flow within the salience network: A MEG study. Scientific Reports, 9(1), 1–15. https://doi.org/10.1038/s41598-018-36498-7

Koreki, A. (2024). Aberrant heartbeat-evoked potential in schizophrenia. Progress in Neuropsychopharmacology.

Kriegeskorte, N., Lindquist, M. A., Nichols, T. E., Poldrack, R. A., & Vul, E. (2010). Everything You Never Wanted to Know about Circular Analysis, but Were Afraid to Ask. Journal of Cerebral Blood Flow & Metabolism, 30(9), 1551–1557. https://doi.org/10.1038/jcbfm.2010.86

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., & Baker, C. I. (2010). Circular inference in neuroscience: The dangers of double dipping. Journal of Vision, 8(6), 88–88. https://doi.org/10.1167/8.6.88

Kritzman, L., Eidelman-Rothman, M., Keil, A., Freche, D., Sheppes, G., & Levit-Binnun, N. (2022). Steady-state visual evoked potentials differentiate between internally and externally directed attention. NeuroImage, 254(March), 119133. https://doi.org/10.1016/j.neuroimage.2022.119133

Kruschke, J. K. (2021). Bayesian Analysis Reporting Guidelines. Nature Human Behaviour, 5(10), 1282–1291. https://doi.org/10.1038/s41562-021-01177-7

Kumral, D., Al, E., Cesnaite, E., Kornej, J., Sander, C., Hensch, T., Zeynalova, S., Tautenhahn, S., Hagendorf, A., Laufs, U., Wachter, R., Nikulin, V., & Villringer, A. (2022). Attenuation of the Heartbeat-Evoked Potential in Patients With Atrial Fibrillation. JACC: Clinical Electrophysiology, 8(10), 1219–1230. https://doi.org/10.1016/j.jacep.2022.06.019

Kunzendorf, S., Klotzsche, F., Akbal, M., Villringer, A., Ohl, S., & Gaebler, M. (2019). Active information sampling varies across the cardiac cycle. Psychophysiology, 56(5), e13322. https://doi.org/10.1111/psyp.13322

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence Testing for Psychological Research: A Tutorial. Advances in Methods and Practices in Psychological Science, 1(2), 259–269. https://doi.org/10.1177/2515245918770963

Larson, M. J., & Carbine, K. A. (2017). Sample size calculations in human electrophysiology (EEG and ERP) studies: A systematic review and recommendations for increased rigor. International Journal of Psychophysiology, 111, 33–41. https://doi.org/10.1016/j.ijpsycho.2016.06.015

Lechinger, J., Heib, D. P. J., Gruber, W., Schabus, M., & Klimesch, W. (2015). Heartbeat-related EEG amplitude and phase modulations from wakefulness to deep sleep: Interactions with sleep spindles and slow oscillations. Psychophysiology, 52(11), 1441–1450. https://doi.org/10.1111/psyp.12508

Legaz, A., Yoris, A., Sedeño, L., Abrevaya, S., Martorell, M., Alifano, F., García, A. M., & Ibañez, A. (2022). Heart–brain interactions during social and cognitive stress in hypertensive disease: A multidimensional approach. In European Journal of Neuroscience (Vol. 55, Issues 9–10). https://doi.org/10.1111/ejn.14979

Leopold, C., & Schandry, R. (2001). The heartbeat-evoked brain potential in patients suffering from diabetic neuropathy and in healthy control persons. 112, 674–682.

Liu, H., Liang, H., Yu, X., Wang, G., Han, Y., Yan, M., Li, S., & Wang, W. (2023). Enhanced external counterpulsation modulates the heartbeat evoked potential. Frontiers in Physiology, 14, 1144073. https://doi.org/10.3389/fphys.2023.1144073

Liu, H., Yu, X., Wang, G., Han, Y., & Wang, W. (2022). Effects of 24-h acute total sleep deprivation on physiological coupling in healthy young adults. Frontiers in Neuroscience, 16, 952329. https://doi.org/10.3389/fnins.2022.952329

Liuzzi, P., Cassioli, T., Secci, S., Hakiki, B., Scarpino, M., Burali, R., di Palma, A., Toci, T., Grippo, A., Cecchi, F., Frosini, A., & Mannini, A. (2024). A neurophysiological profiling of the heartbeat-evoked potential in severe acquired brain injuries: A focus on unconsciousness. The European Journal of Neuroscience, 60(3), 4201–4216. https://doi.org/10.1111/ejn.16394

Luck, S. J. (2014). An introduction to the event-related potential technique (Second edition). The MIT Press.

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

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, 1–12. https://doi.org/10.7554/eLife.25318

Majeed, U., Aftab, M. F., Baloch, D. M., Ahmed, S., Yusuf, I. M., Hasan, M. A., & Qureshi, M. S. (2022). Modulation of Heart and Brain Function by Surah Al-Rehman Recitation Among Distressed Diabetic Patients in Pakistan. Journal of Religion and Health, 61(5), 3852–3865. https://doi.org/10.1007/s10943-021-01431-2

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. Journal of Neuroscience Methods, 164(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Marshall, A. C., Gentsch, A., Blum, A. L., Broering, C., & Schütz-Bosbach, S. (2019). I feel what I do: Relating interoceptive processes and reward-related behavior. NeuroImage, 191(November 2018), 315–324. https://doi.org/10.1016/j.neuroimage.2019.02.032

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), 1–15. 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. (2020). Interoceptive cardiac expectations to emotional stimuli predict visual perception. Emotion, 20(7), 1113–1126. https://doi.org/10.1037/emo0000631

Marshall, A. C., Gentsch-Ebrahimzadeh, A., & Schütz-Bosbach, S. (2022). From the inside out: Interoceptive feedback facilitates the integration of visceral signals for efficient sensory processing: Interoceptive feedback facilitates sensory processing. NeuroImage, 251(July 2021), 119011. https://doi.org/10.1016/j.neuroimage.2022.119011

McCraty, R., Atkinson, M., & Bradley, R. T. (2004). Electrophysiological Evidence of Intuition: Part 2. A System-Wide Process? The Journal of Alternative and Complementary Medicine, 10(2), 325–336. https://doi.org/10.1089/107555304323062310

Melo, E., Fiel, J., Milhomens, R., Ribeiro, T., Navegantes, R., Gomes, F., Duarte Gomes, B., & Pereira, A. (2022). Dynamic coupling between the central and autonomic cardiac nervous systems in patients with refractory epilepsy: A pilot study. Frontiers in Neurology, 13, 904052. https://doi.org/10.3389/fneur.2022.904052

Migeot, J., Hesse, E., Fittipaldi, S., Mejía, J., Fraile, M., García, A. M., García, M. D. C., Ortega, R., Lawlor, B., Lopez, V., & Ibáñez, A. (2023). Allostatic-interoceptive anticipation of social rejection. NeuroImage, 276, 120200. https://doi.org/10.1016/j.neuroimage.2023.120200

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Systematic Reviews and Meta-Analyses: The PRISMA Statement. Annulas of Internal Medicine, 151(4), 264–269. https://doi.org/10.1371/journal.pmed1000097

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/ Evoked Potentials, 88(3), 163–172. https://doi.org/10.1016/0168-5597(93)90001-6

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

Niizeki, K., Kawahara, K., & Miyamoto, Y. (1993). Interaction among cardiac, respiratory, and locomotor rhythms during cardiolocomotor synchronization. Journal of Applied Physiology, 75(4), 1815–1821. https://doi.org/10.1152/jappl.1993.75.4.1815

Ohl, S., Wohltat, C., Kliegl, R., Pollatos, O., & Engbert, R. (2016). Microsaccades are coupled to heartbeat. Journal of Neuroscience, 36(4), 1237–1241. https://doi.org/10.1523/JNEUROSCI.2211-15.2016

Ono, K., Mizuochi, R., Yamamoto, K., Sasaoka, T., & Ymawaki, S. (2024). Exploring the neural underpinnings of chord prediction uncertainty: An electroencephalography (EEG) study. Scientific Reports, 14(1), 4586. https://doi.org/10.1038/s41598-024-55366-1

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(September), 1–9. 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. Journal of 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(April), 502–511. https://doi.org/10.1016/j.neuroimage.2019.04.081

Park, H.-D., Correia, S., Ducorps, A., & Tallon-Baudry, C. (2014). Spontaneous fluctuations in neural responses to heartbeats predict visual detection. Nature Neuroscience, 17(4), 612–618. https://doi.org/10.1038/nn.3671

Pelentritou, A., Pfeiffer, C., Schwartz, S., & De Lucia, M. (2024). Cardio-audio synchronization elicits neural and cardiac surprise responses in human wakefulness and sleep. Communications Biology, 7(1), 226. https://doi.org/10.1038/s42003-024-05895-2

Pérez, J. J., Guijarro, E., & Barcia, J. A. (2005). Suppression of the cardiac electric field artifact from the heart action evoked potential. Medical and Biological Engineering and Computing, 43(5), 572–581. https://doi.org/10.1007/BF02351030

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(December 2018), 101701. https://doi.org/10.1016/j.nicl.2019.101701

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(November 2018), 595–606. https://doi.org/10.1016/j.neuroimage.2018.11.037

Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. NeuroImage, 198(May), 181–197. https://doi.org/10.1016/j.neuroimage.2019.05.026

Pollatos, O., Herbert, B. M., Mai, S., & Kammer, T. (2016). Changes in interoceptive processes following brain stimulation. Philosophical Transactions of the Royal Society 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/10.1111/1469-8986.2004.00170.x

Pontifex, M. B., Gwizdala, K. L., Parks, A. C., Billinger, M., & Brunner, C. (2017). Variability of ICA decomposition may impact EEG signals when used to remove eyeblink artifacts. Psychophysiology, 54(3), 386–398. https://doi.org/10.1111/psyp.12804

Poppa, T., Benschop, L., Horczak, P., Vanderhasselt, M.-A., Carrette, E., Bechara, A., Baeken, C., & Vonck, K. (2022). Auricular transcutaneous vagus nerve stimulation modulates the heart-evoked potential. Brain Stimulation, 15(1), 260–269. https://doi.org/10.1016/j.brs.2021.12.004

R Core Team. (2025). R: A Language and Environment for Statistical Computing [Computer software]. R Foundation for Statistical Computing.

Rapp, L., Mai-Lippold, S. A., Georgiou, E., & Pollatos, O. (2023). Elevated EEG heartbeat-evoked potentials in adolescents with more ADHD symptoms. Biological Psychology, 184, 108698. https://doi.org/10.1016/j.biopsycho.2023.108698

Ren, Q., Marshall, A. C., Kaiser, J., & Schütz-Bosbach, S. (2022). Response inhibition is disrupted by interoceptive processing at cardiac systole. Biological Psychology, 170(August 2021), 108323. https://doi.org/10.1016/j.biopsycho.2022.108323

Richter, F., García, A. M., Rodriguez Arriagada, N., Yoris, A., Birba, A., Huepe, D., Zimmer, H., Ibáñez, A., & Sedeño, L. (2021). Behavioral and neurophysiological signatures of interoceptive enhancements following vagus nerve stimulation. Human Brain Mapping, 42(5), 1227–1242. https://doi.org/10.1002/hbm.25288

Richter, F., & Ibáñez, A. (2021). Time is body: Multimodal evidence of crosstalk between interoception and time estimation. Biological Psychology, 159(January). https://doi.org/10.1016/j.biopsycho.2021.108017

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

Salamone, P. C., Legaz, A., Sedeño, L., Moguilner, S., Fraile-Vazquez, M., Campo, C. G., Fittipaldi, S., Yoris, A., Miranda, M., Birba, A., Galiani, A., Abrevaya, S., Neely, A., Caro, M. M., Alifano, F., Villagra, R., Anunziata, F., Okada De Oliveira, M., Pautassi, R. M., … Ibañez, A. (2021). Interoception Primes Emotional Processing: Multimodal

Evidence from Neurodegeneration. The Journal of Neuroscience, 41(19), 4276–4292. https://doi.org/10.1523/JNEUROSCI.2578-20.2021

Salamone, P. C., Sedeño, L., Legaz, A., Bekinschtein, T., Martorell, M., Adolfi, F., Fraile-Vazquez, M., Rodríguez Arriagada, N., Favaloro, L., Peradejordi, M., Absi, D. O., García, A. M., Favaloro, R., & Ibáñez, A. (2020). Dynamic neurocognitive changes in interoception after heart transplant. Brain Communications, 2(2). https://doi.org/10.1093/braincomms/fcaa095

Sassenhagen, J., & Draschkow, D. (2019). Cluster-based permutation tests of MEG/EEG data do not establish significance of effect latency or location. Psychophysiology, 56(6), e13335. https://doi.org/10.1111/psyp.13335

Schandry, R., & Montoya, P. (1996). Event-related brain potentials and the processing of cardiac activity. Biological Psychology, 42(1–2), 75–85. https://doi.org/10.1016/0301-0511(95)05147-3

Schandry, R., Sparrer, B., & Weitkunat, R. (1986). From the heart to the brain: A study of heartbeat contingent scalp potentials. International Journal of Neuroscience, 30(4), 261–275. https://doi.org/10.3109/00207458608985677

Schandry & Weitkunat. (1990). Enhancement of heartbeat-related brain potentials through cardiac awareness training. International Journal of Neuroscience, 53(2–4), 243–253. https://doi.org/10.3109/00207459008986611

Schmitz, M., Müller, L. E., Schulz, A., Kleindienst, N., Herpertz, S. C., & Bertsch, K. (2020). Heart and brain: Cortical representation of cardiac signals is disturbed in borderline personality disorder, but unaffected by oxytocin administration. Journal of Affective Disorders, 264, 24–28. https://doi.org/10.1016/j.jad.2019.11.139

Schmitz, M., Müller, L. E., Seitz, K. I., Schulz, A., Steinmann, S., Herpertz, S. C., & Bertsch, K. (2021). Heartbeat evoked potentials in patients with post-traumatic stress disorder: An unaltered neurobiological regulation system? European Journal of Psychotraumatology, 12(1). https://doi.org/10.1080/20008198.2021.1987686

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. https://doi.org/10.1097/PSY.0000000000000195

Schulz, A., Rost, S., Flasinski, T., Dierolf, A. M., Lutz, A. P. C., Münch, E. E., Mertens, V.-C., Witthöft, M., & Vögele, C. (2020). Distinctive body perception mechanisms in high versus low symptom reporters: A neurophysiological model for medically-unexplained symptoms. Journal of Psychosomatic Research, 137, 110223. https://doi.org/10.1016/j.jpsychores.2020.110223

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. (2016). Heartfelt Self: Cardio-Visual Integration Affects Self-Face Recognition and Interoceptive Cortical Processing. Cerebral Cortex, cercor;bhw296v1. https://doi.org/10.1093/cercor/bhw296

Shannon, C. E. (1948). A mathematical theory of communicatio. The Bell System Technical Journal, 27(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x.

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

Sheppard, J. P., Barker, T. A., Ranasinghe, A. M., Clutton-Brock, T. H., Frenneaux, M. P., & Parkes, M. J. (2011). Does modifying electrode placement of the 12 lead ECG matter in healthy subjects? International Journal of Cardiology, 152(2), 184–191. https://doi.org/10.1016/j.ijcard.2010.07.013

Simor, P., Bogdány, T., Bódizs, R., & Perakakis, P. (2021). Cortical monitoring of cardiac activity during rapid eye movement sleep: The heartbeat evoked potential in phasic and tonic rapid-eye-movement microstates. Sleep, 44(9), 1–12. https://doi.org/10.1093/sleep/zsab100

Skora, L. I., Livermore, J. J. A., & Roelofs, K. (2022). The functional role of cardiac activity in perception and action. Neuroscience & Biobehavioral Reviews, 137, 104655. https://doi.org/10.1016/j.neubiorev.2022.104655

Soderberg, C. K., Errington, T. M., Schiavone, S. R., Bottesini, J., Thorn, F. S., Vazire, S., Esterling, K. M., & Nosek, B. A. (2021). Initial evidence of research quality of registered reports compared with the standard publishing model. Nature Human Behaviour, 5(8), 990–997. https://doi.org/10.1038/s41562-021-01142-4

Solcà, M., Park, H.-D., Bernasconi, F., & Blanke, O. (2020). Behavioral and neurophysiological evidence for altered interoceptive bodily processing in chronic pain. NeuroImage, 217(April). https://doi.org/10.1016/j.neuroimage.2020.116902

Steegen, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016). Increasing Transparency Through a Multiverse Analysis. Perspectives on Psychological Science, 11(5), 702–712. https://doi.org/10.1177/1745691616658637

Steinfath, P., Herzog, N., Fourcade, A., Sander, C., Nikulin, V., & Villringer, A. (2025). Validating genuine changes in Heartbeat Evoked Potentials using Pseudotrials and Surrogate Procedures. Imaging Neuroscience. https://doi.org/10.1162/IMAG.a.30

Stoupi, N. A., Weijs, M. L., Imbach, L., & Lenggenhager, B. (2024). Heartbeat-evoked potentials following voluntary hyperventilation in epilepsy patients: Respiratory influences on cardiac interoception. Frontiers in Neuroscience, 18, 1391437. https://doi.org/10.3389/fnins.2024.1391437

Sweet, D. J. (2021). New at Cell Press: The Inclusion and Diversity Statement. Cell, 184(1), 1–2. https://doi.org/10.1016/j.cell.2020.12.019

Tallon-Baudry, C., Campana, F., Park, H. D., & Babo-Rebelo, M. (2018). The neural monitoring of visceral inputs, rather than attention, accounts for first-person perspective in conscious vision. Cortex, 102, 139–149. https://doi.org/10.1016/j.cortex.2017.05.019

Tanaka, Y., Ito, Y., Terasawa, Y., & Umeda, S. (2023). Modulation of heartbeat-evoked potential and cardiac cycle effect by auditory stimuli. Biological Psychology, 182, 108637. https://doi.org/10.1016/j.biopsycho.2023.108637

Tanner, D., Morgan-Short, K., & Luck, S. J. (2015). How inappropriate high-pass filters can produce artifactual effects and incorrect conclusions in ERP studies of language and cognition. Psychophysiology, 52(8), 997–1009. https://doi.org/10.1111/psyp.12437

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

Thornton, A., & Lee, P. (2000). Publication bias in meta-analysis: Its causes and consequences. Journal of Clinical Epidemiology, 53(2), 207–216. https://doi.org/10.1016/S0895-4356(99)00161-4

Todd, J., Cardellicchio, P., Swami, V., Cardini, F., & Aspell, J. E. (2021). Weaker implicit interoception is associated with more negative body image: Evidence from gastric-alpha phase amplitude coupling and the heartbeat evoked potential. Cortex, 143(xxxx), 254–266. https://doi.org/10.1016/j.cortex.2021.07.006

Trübutschek, D., Yang, Y.-F., Gianelli, C., Cesnaite, E., Fischer, N. L., Vinding, M. C., Marshall, T. R., Algermissen, J., Pascarella, A., Puoliväli, T., Vitale, A., Busch, N. A.,

& Nilsonne, G. (2024). EEGManyPipelines: A Large-scale, Grassroots Multi-analyst Study of Electroencephalography Analysis Practices in the Wild. Journal of Cognitive Neuroscience, 36(2), 217–224. https://doi.org/10.1162/jocn_a_02087

Umeda, S., Tochizawa, S., Shibata, M., & Terasawa, Y. (2016). Prospective memory mediated by interoceptive accuracy: A psychophysiological approach. Philosophical Transactions of the Royal Society B: Biological Sciences, 371(1708), 20160005. https://doi.org/10.1098/rstb.2016.0005

Ventura-Bort, C., & Weymar, M. (2024). Transcutaneous auricular vagus nerve stimulation modulates the processing of interoceptive prediction error signals and their role in allostatic regulation. Human Brain Mapping, 45(3), e26613. https://doi.org/10.1002/hbm.26613

Verdonk, C., Teed, A. R., White, E. J., Ren, X., Stewart, J. L., Paulus, M. P., & Khalsa, S. S. (2024). Heartbeat-evoked neural response abnormalities in generalized anxiety disorder during peripheral adrenergic stimulation. Neuropsychopharmacology, 49(8), 1246–1254. https://doi.org/10.1038/s41386-024-01806-5

Verdonk, C., Trousselard, M., Di Bernardi Luft, C., Medani, T., Billaud, J. B., Ramdani, C., Canini, F., Claverie, D., Jaumard-Hakoun, A., & Vialatte, F. (2021). The heartbeat evoked potential does not support strong interoceptive sensibility in trait mindfulness. Psychophysiology, 58(10), 1–13. https://doi.org/10.1111/psyp.13891

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

Villena-González, M., Rojas-Thomas, F., Morales-Torres, R., & López, V. (2023). Autonomous sensory meridian response is associated with a larger heartbeat-evoked potential amplitude without differences in interoceptive awareness. Psychophysiology, 60(6), e14277. https://doi.org/10.1111/psyp.14277

Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., Selker, R., Gronau, Q. F., Šmíra, M., Epskamp, S., Matzke, D., Rouder, J. N., & Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. Psychonomic Bulletin & Review, 25(1), 35–57. https://doi.org/10.3758/s13423-017-1343-3

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

Weijs, M. L., Daum, M. M., & Lenggenhager, B. (2023). Cardiac interoception in infants: Behavioral and neurophysiological measures in various emotional and self-related contexts. Psychophysiology, 60(12), e14386. https://doi.org/10.1111/psyp.14386

Yasuma, F., & Hayano, J. (2004). Respiratory Sinus Arrhythmia. Chest, 125(2), 683–690. https://doi.org/10.1378/chest.125.2.683

Yoris, A., 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

Yoris, A. E., Cira, L. F., Luque-Casado, A., Salvotti, C., Tajadura-Jiménez, A., Avancini, C., Zarza-Rebollo, J. A., Sanabria, D., & Perakakis, P. (2024). Delving into the relationship between regular physical exercise and cardiac interoception in two cross-sectional studies. Neuropsychologia, 198, 108867. https://doi.org/10.1016/j.neuropsychologia.2024.108867

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

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

Zaccaro, A., Della Penna, F., Mussini, E., Parrotta, E., Perrucci, M. G., Costantini, M., & Ferri, F. (2024). Attention to cardiac sensations enhances the heartbeat-evoked potential during exhalation. iScience, 27(4), 109586. https://doi.org/10.1016/j.isci.2024.109586

Zaccaro, A., Perrucci, M. G., Parrotta, E., Costantini, M., & Ferri, F. (2022). Brain-heart interactions are modulated across the respiratory cycle via interoceptive attention. NeuroImage, 262, 119548. https://doi.org/10.1016/j.neuroimage.2022.119548

Zhang, Y., Zhang, J., Xie, M., Ding, N., Zhang, Y., & Qin, P. (2023). Dual interaction between heartbeat-evoked responses and stimuli. NeuroImage, 266(December 2022). https://doi.org/10.1016/j.neuroimage.2022.119817

Zhou, D., Stiso, J., Bertolero, M., Cornblath, E., Teich, E., Blevins, A. S., Murphy, K., Oudyk, K., Michael, C., Urai, A., Matelsky, J., Virtualmario, Camp, C., Castillo, R. A., Saxe, R., Dworkin, J., & Bassett, D. (2022). dalejn/cleanBib: V1.1.3 [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.7375250

Zhou, H., Zou, H., Dai, Z., Zhao, S., Hua, L., Xia, Y., Han, Y., Yan, R., Tang, H., Huang, Y., Du, Y., Wang, X., Yao, Z., & Lu, Q. (2022). Interoception Dysfunction Contributes to the Negative Emotional Bias in Major Depressive Disorder. Frontiers in Psychiatry, 13(April), 1–13. https://doi.org/10.3389/fpsyt.2022.874859

Zhou, M., Cheng, L., Zhou, Y., Zhu, S., Zhang, Y., Kendrick, K. M., & Yao, S. (2024). Intranasal Oxytocin Improves Interoceptive Accuracy and Heartbeat-Evoked Potentials During a Cardiac Interoceptive Task. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, S245190222400137X. https://doi.org/10.1016/j.bpsc.2024.05.004

Zwienenberg, L., van Dijk, H., Enriquez-Geppert, S., van der Vinne, N., Gevirtz, R., Gordon, E., Sack, A. T., & Arns, M. (2023). Heartbeat-Evoked Potential in Major Depressive

Disorder: A Biomarker for Differential Treatment Prediction between Venlafaxine and rTMS? Neuropsychobiology, 1–10. https://doi.org/10.1159/000529308

Supplementary material

Contents

A. Estimation of minimal detectable Hedges' g across reviewed pipelines 73

B. Simulation of minimal detectable Hedges' g effect sizes depending on the number of averaged epochs 75

C. Other supplementary figures 79

D. Details on evaluation of citation diversity 80

References 81

A. Estimation of minimal detectable Hedges' g across reviewed pipelines

Evaluation of the minimal detectable Hedges' g was performed under the following assumptions:

- We always considered a two-sided alternative hypothesis.
- We assumed equality of groups, so the sample size per group was equal to the sample size divided by the number of groups.
- We have encountered only a couple of uses of Bayesian statistics across all papers. For estimation purposes, we treated all tests as frequentist.
- In ANOVAs, we assumed that the number of cells (unique clusters of observations that we use for power estimation) was equal to the product of the number of groups and the number of conditions. We also assumed that each group performed all conditions.
- For clustering, we assumed it to be at least as powerful as a test on averaged HERs.
- For regression, we assumed a regression with one predictor.

Since the last two assumptions might not hold in practice, Hedges' g estimates for cluster-based tests and regression are only shown in Fig. S1 but not in Fig. 7F.

For all considered types of statistical tests, we first estimated Cohen's d:

- For t-tests, we obtained it directly from the power estimation function of pwr package in R.
- For ANOVAs and regressions, we obtained Cohen's f from the power estimation function and transformed it as follows (Cohen, 1988):

$$d = 2f$$

- For correlations, we obtained the correlation coefficient (denoted as *r* in the equation below) and converted it to Cohen's d (Cohen, 1988; Ruscio, 2008):

$$d = \frac{2r}{\sqrt{1 - r^2}}$$

The exact Hedges & Olkin (1985) correction was performed according to the following formula:

$$J(\upsilon) = \frac{\Gamma\left(\frac{\upsilon}{2}\right)}{\sqrt{\frac{\upsilon}{2}}\Gamma\left(\frac{\upsilon-1}{2}\right)},$$

where υ is the number of degrees of freedom. Hedges' g is then obtained by multiplying Cohen's d and $J(\upsilon)$. For further implementation details, see the code at: https://github.com/PaulSteinfath/systematic-hep-review/blob/main/functions/analysis/hedges_g.R

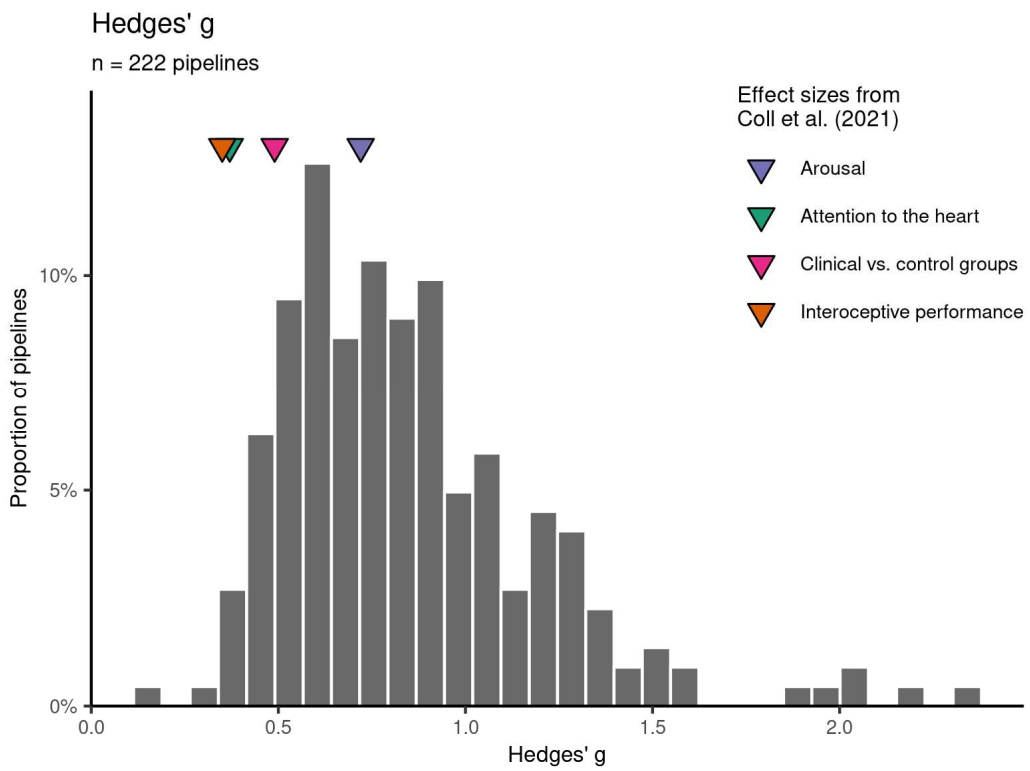

Figure S1. Minimal detectable effect sizes across reviewed pipelines. Estimates for cluster-based permutation tests and regression are also included. These additional estimates are not

reflected in Fig. 7F, which only includes estimates from one- or two-sample t-tests, ANOVAs, and correlations that were performed on averaged HER values. When comparing the estimated detectable effect sizes to the four effect sizes reported in a previous HER review by Coll et al., 2021 (marked with colored triangles), we observe that only 0.9% of pipelines (2/222) had enough power to detect Hedges' g of 0.35 (interoceptive performance meta-estimate), 2.3% of pipelines (5/222) could reliably detect Hedges' g of 0.37 (attention to heart meta-estimate), 9.9% of pipelines (22/222) could reliably detect Hedges' g of 0.49 (clinical vs. control groups meta-estimate), and 41% of pipelines (91/222) could reliably detect Hedges' g of 0.72 (arousal meta-estimate).

B. Simulation of minimal detectable Hedges' g effect sizes depending on the number of averaged epochs

The estimated effect sizes shown in Fig. 7F and S1 were obtained assuming that we have only one observation per participant. In practice, HERs have relatively low amplitude (0.5–2 μV in the case of EEG), so multiple trials per participant and condition are typically averaged to obtain a more stable response. To understand the influence of the number of trials on the minimal detectable effect size, we performed simulations with different levels of between- and within-subject variability.

Cohen'd is defined as (Cohen, 1988):

$$d = \frac{\mu_1 - \mu_2}{s}$$

$$s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

where $\mu_{1,2}$ and $s_{1,2}$ are the means and the standard deviations (SDs) of both groups, $n_{1,2}$ are group sizes, and s is the pooled SD. If we assume that the group sizes are equal and between-

participant variance is same for both groups, pooled SD for the case of no averaging over trials is equal to the between-subject SD σ_s:

$$s = \sigma_s$$

If we now consider the case when each observation is obtained by averaging k trials per participant, we need to take the standard error of the mean into account. Assuming that within-subject SD σ_t is the same for all subjects (independent and identically distributed), pooled variance is equal to the sum of between- and within-subject variance. Hence, pooled SD for the case of averaging over k trials is equal to:

$$s_k = \sqrt{\sigma_s^2 + \frac{\sigma_t^2}{k}} = \sigma_s\sqrt{1 + \frac{1}{r^2 k}} \text{ where } r = \frac{\sigma_s}{\sigma_t}$$

is the ratio of between- and within-subject variability (later referred to as between-within variability ratio). As a result, the estimated effect size with k trials per participant is equal to:

$$d_k = \frac{\mu_1 - \mu_2}{s_k} = \frac{\mu_1 - \mu_2}{\sigma_s\sqrt{1 + \frac{1}{r^2 k}}} = \frac{d}{\sqrt{1 + \frac{1}{r^2 k}}}$$

Which shows that the effect size adjusted for number of epochs is the true effect size multiplied by $\frac{1}{\sqrt{1+\frac{1}{r^2 k}}}$. Since the square root in the denominator is always at least 1, the estimated effect size d_k is always smaller than or equal to the true effect size d. Equality is reached only in case of infinite r (equivalent to zero within-participant variance) or k (equivalent to having a perfect estimate). As the estimated effects become smaller in size, the minimal detectable true effect size increases.

As the estimation approach for Hedges' g (Section A) didn't account for averaging over epochs, the values of the minimal detectable effect size need to be corrected. Let d_0 be the minimal detectable effect size in case of no within-subject variability (as obtained from the

power estimation function). If we now assume averaging over k epochs, the adjusted minimal detectable effect size is equal to:

$$d_{adj} = d_0 \sqrt{1 + \frac{1}{r^2 k}}$$

Since correction for Hedges' g depends only on the number of degrees of freedom, and applies equally to both d_{adj} and d_0, the equation for Hedges' g is very similar:

$$g_{adj} = g_0 \sqrt{1 + \frac{1}{r^2 k}}$$

The equations for d_{adj} and g_{adj} imply that the minimal detectable effect size becomes larger if only a limited number of epochs per participant is available. However, the extent of this effect depends on the between-within variability ratio r. We varied this ratio in simulations discussed below, considering values of 0.1, 0.25, 0.5, 1, and 2. In practice, the ratio can also be estimated from pilot data based on the SD of HER values within and between participants.

We used the estimated or, where available, reported number of epochs (per participant and condition) and adjusted the original minimal detectable values of Hedges' g (Fig. 7F) accordingly. As shown in Fig. S2-A, the effect of trials noticeably affects the resulting distribution only in the case of very high within-subject variability ($r = 0.1$). Fig. S2-B shows how, for a one-sample t-test, the minimal detectable effect size depends on the number of participants and the number of epochs in simulations.

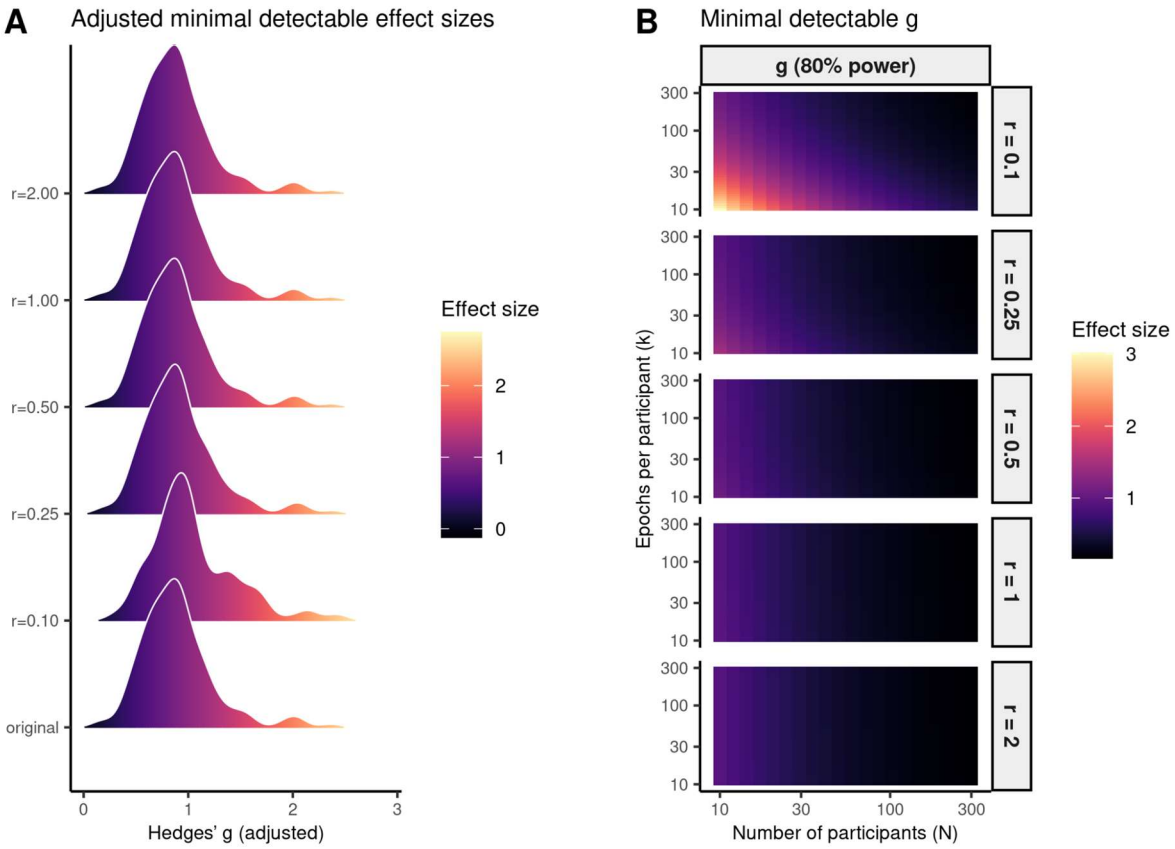

Figure S2. Minimal detectable Hedges' g adjusted for the levels of between- and within-subject variability. (A) The distribution of Hedges' g that was not adjusted for the number of epochs (Fig. 7F), labelled as 'original', compared to its adjusted counterparts for different values of between-within variability ratio. (B) The simulated distributions depending on between-within variability ratio, the number of participants, and the number of averaged epochs per participant, for a one-sample t-test. Note that both axes are log-scaled.

C. Other supplementary figures

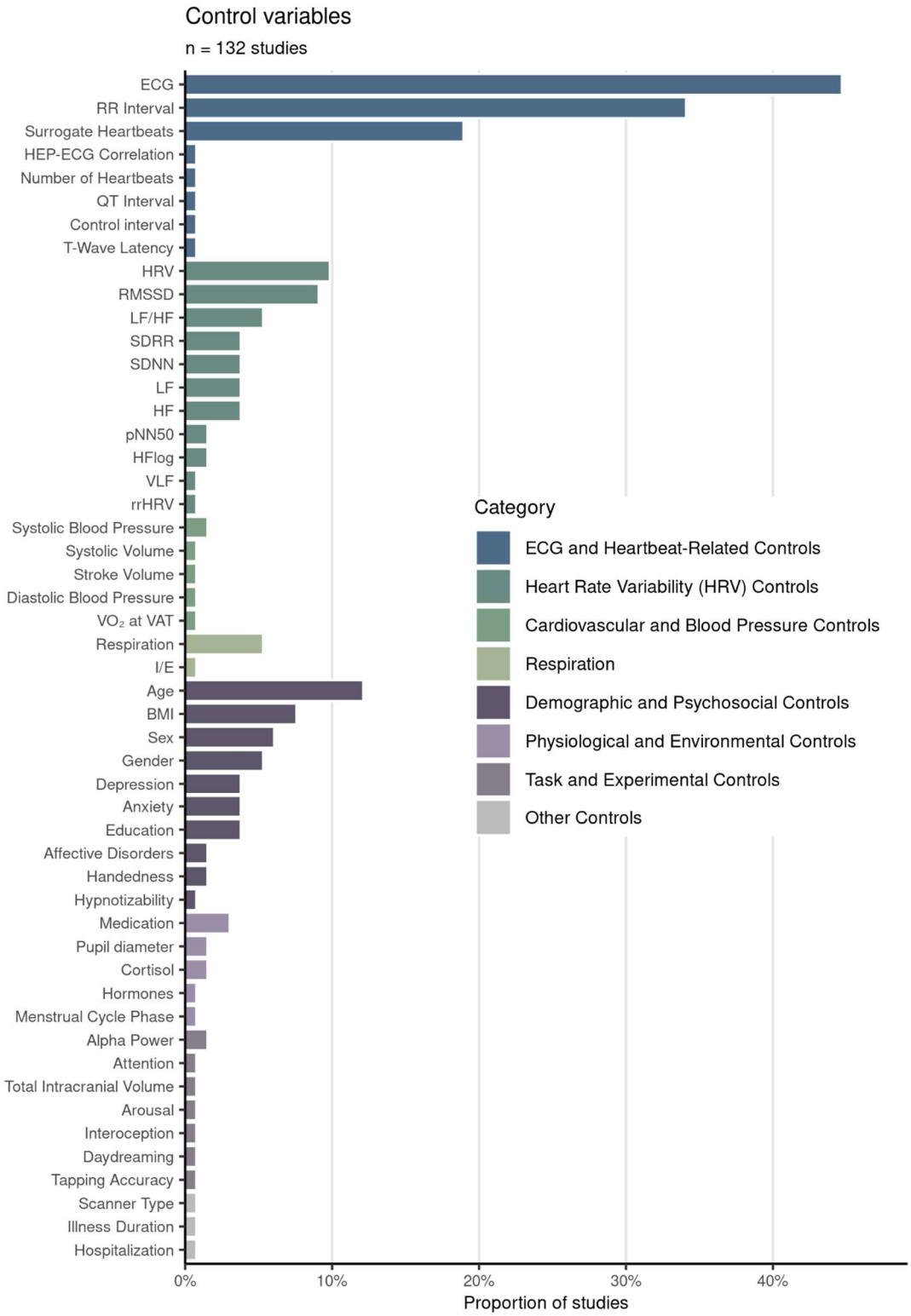

Figure S3. Control variables used across pipelines.

Diverse variables are considered to control for spurious findings. The majority reflect heartbeat and cardiovascular system related variables. HR: Heart Rate, SDRR: Standard Deviation of RR intervals, SDNN: Standard deviation of NN intervals, RMSSD: Root Mean Square of Successive Differences, pNN50: Percentage of NN Intervals > 50ms Differences, VLF: Very Low Frequency, LF: Low Frequency, HF: High Frequency, LF/HF: Low Frequency / High Frequency Ratio, HFLog: Log-transformed High Frequency, HRV: Heart Rate Variability, rrHRV: RR Interval-based Heart Rate Variability, VO_2 at VAT: Oxygen Consumption at Ventilatory Anaerobic Threshold, BMI: Body Mass Index, I/E: Inhalation/Exhalation ratio

D. Details on evaluation of citation diversity

To contextualize gender distribution in HER research within the broader neuroscience field, we subsampled journals according to Dworkin et al. (2020) and (Hefter et al., 2025), which had both ranked neuroscience journals by Eigenfactor scores (Bergstrom et al., 2008) from Thomson Reuters' Web of Science (WoS). Their selection in 2020 included the five highest-ranking neuroscience journals from 1995–2018 (Procedure 1) and in 2025 the 50 highest-ranking ones from 2023 (Procedure 2). Since few of the reviewed HER papers had been published in those journals (Procedure 1: n = 19; Procedure 2: n = 43), we also included the papers from journals with latest Eigenfactor scores (from the WoS' 2024 Journal Citation Report; Clarivate, 2024) above Hefter et al.'s lowest cutoff (Procedure 3: n = 77). This way we found that authors with female names were more prevalent in higher-ranked journals, decreasing as lower-ranked journals were included (Procedure 1: 84.21% > Procedure 2: 69.77%; Procedure 3: 66.23%; Procedure 3 [Δ remaining papers]: 19.48% [-2.50] FF, 10.39% [-0.96%] MF, 36.6% [-13.72%] FM, 33.77% [+17.18%] MM). Nonetheless, the observed trend more closely resembles the gender distribution (18% FF, 14% MF, 29% FM, 39% MM) that resulted from the large-scale analyses by Hefter et al. (2025).

References

Bergstrom, C. T., West, J. D., & Wiseman, M. A. (2008). The EigenfactorTM Metrics. Journal of Neuroscience, 28(45), 11433–11434. https://doi.org/10.1523/JNEUROSCI.0003-08.2008

Clarivate. (2024). Journal Citation Reports (JCR): Impact Factor 2024. PDF (Web of Science). https://doi.org/10.13140/RG.2.2.34797.60640

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

Coll, M.-P., Hobson, H., Bird, G., & Murphy, J. (2021). Systematic review and meta-analysis of the relationship between the heartbeat-evoked potential and interoception. Neuroscience and Biobehavioral Reviews, 122, 190–200. https://doi.org/10.1016/j.neubiorev.2020.12.012

Dworkin, J. D., Linn, K. A., Teich, E. G., Zurn, P., Shinohara, R. T., & Bassett, D. S. (2020). The extent and drivers of gender imbalance in neuroscience reference lists. Nature Neuroscience, 23(8), 918–926. https://doi.org/10.1038/s41593-020-0658-y

Hefter, K., Teich, E. G., & Bassett, D. S. (2025). Gender imbalance in citations most pronounced in high-impact neuroscience journals (p. 2025.02.19.638905). bioRxiv. https://doi.org/10.1101/2025.02.19.638905

Ruscio, J. (2008). A probability-based measure of effect size: Robustness to base rates and other factors. Psychological Methods, 13(1), 19–30. https://doi.org/10.1037/1082-989X.13.1.19