# Mega-analysis of the Interoceptive Accuracy Scale (IAS) Structure and its Dispositional Correlates

Ana Neves1, Magdalena Pfaff1, Robyn Scharte1, Raquel Nogueira Arjona1,2, Giulia Poerio3, and Dominique Makowski1,4

1School of Psychology, University of Sussex

2Sussex Addictions Research and Intervention Centre, University of Sussex

3

4Sussex Centre for Consciousness Science, University of Sussex

# Author Note

Ana Neves  <https://orcid.org/0009-0006-0020-7599>

Magdalena Pfaff  <https://orcid.org/0009-0006-2386-7936>

Robyn Scharte  <https://orcid.org/blabla>

Raquel Nogueira Arjona  <https://orcid.org/0000-0002-5998-701X>

Giulia Poerio  <https://orcid.org/0000-0002-2343-5109>

Dominique Makowski  <https://orcid.org/0000-0001-5375-9967>

Author roles were classified using the Contributor Role Taxonomy (CRediT; https://credit.niso.org/) as follows: *Ana Neves***:** data curation, formal analysis, investigation, visualization, writing – original draft, and writing – review & editing. *Magdalena Pfaff***:** data curation and writing – original draft. *Robyn Scharte***:** writing – review & editing. *Raquel Nogueira Arjona***:** writing – review & editing. *Giulia Poerio***:** writing – review & editing. *Dominique Makowski***:** project administration, data curation, formal analysis, investigation, visualization, writing – original draft, and writing – review & editing

Correspondence concerning this article should be addressed to Dominique Makowski, School of Psychology, University of Sussex, Brighton BN1 9RH, UK, Email: [D.Makowski@sussex.ac.uk](mailto:D.Makowski@sussex.ac.uk)

# Abstract

Interoception, interoception accuracy, interoception sensibility, measurement, exploratory factor analysis, network analysis

# Mega-analysis of the Interoceptive Accuracy Scale (IAS) Structure and its Dispositional Correlates

# Introduction

Interoception refers to the process of sensing, interpreting and integrating information of internal bodily stimuli by the nervous system, including both internal organs (e.g., the heart, the lungs or the gut) but also more broadly physiological tissues that inform the central nervous system about the body’s current state ([Khalsa, Adolphs, Cameron, Critchley, Davenport, Feinstein, Feusner, Garfinkel, Lane, Mehling, & others, 2018](#ref-khalsa2018interoception)). Although research emphasizes interoception’s fundamental role in numerous processes, such as emotion recognition ([Terasawa et al., 2021](#ref-terasawa2021effects)) and regulation ([Zamariola et al., 2019](#ref-zamariola2019relationship)), decision making ([Pollatos et al., 2023](#ref-pollatos2023interoceptive)), learning ([Joshi et al., 2023](#ref-joshi2023interoception)), body-ownership ([Raimo et al., 2021](#ref-raimo2021body)) and outcomes including physical ([Harrison & Pink, 2024](#ref-harrison2024interoception)) and psychological health ([Nord et al., 2021](#ref-nord2021disrupted)) along with well-being ([Ferentzi et al., 2019](#ref-ferentzi2019body)), the field remains hampered by concerns regarding both the conceptualization and measurement of interoception ([Desmedt et al., 2022](#ref-desmedt2022measures), [2025](#ref-desmedt2025discrepancies); [Murphy, 2024](#ref-murphy2024interoception)).

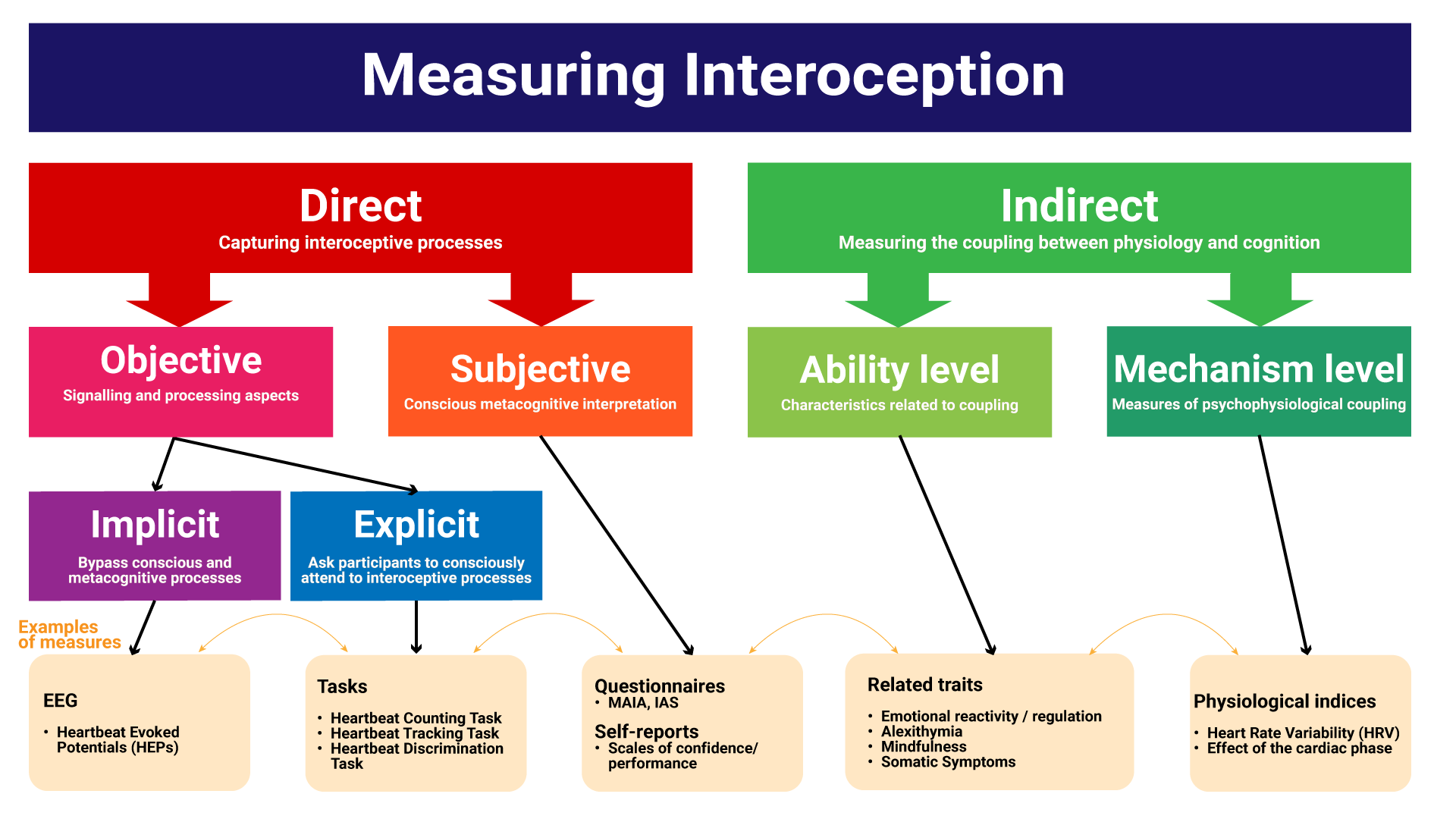
## The Interoceptive Assessment Puzzle

Numerous methods of assessing interoception have been developed (see [Figure 1](#fig-measures)), which can be broadly distinguished into “direct” and “indirect” approaches. Indirect approaches assess interoceptive processes by quantified by the coupling between physiological signals and cognitive or behavioural processes. At the “ability level”, such measures capture individual differences in traits related to this coupling (e.g., alexithymia) whereas at the “mechanism” level, they index psychophysiological interactions more directly (e.g., heart rate variability).

Although these approaches do not directly measure interoceptive perception, they rely on interoceptive mechanisms and therefore provide assessment of interoceptive functioning ([Lischke et al., 2021](#ref-lischke2021heart); [Wareing et al., 2024](#ref-wareing2024utility)). In contrast, direct measures try to capture interoceptive processes more explicitly and encompass both “objective” and “subjective” assessments (i.e., behavioral tests of heart beat counting or tracking vs. trait questionnaires and self-reported metacognitive judgments), as well as “explicit” and “implicit” paradigms (i.e., directing participants’ awareness and attention to interoceptive processes vs. measuring interoception covertly). Measures also differ on the interoceptive modalities (e.g., cardioception, respiroception, gastroception) and theoretical dimension (e.g., accuracy, sensitivity, awareness) assessed. The diversity in measurement approaches illustrates that rather than there being a ‘gold-standard’ measure of interoception in the field, each measure likely has strengths and limitations depending on the research context and dimension of interoception being assessed ([Desmedt et al., 2023](#ref-Desmedt2023accuracy); [Jahedi & Méndez, 2014](#ref-jahedi2014)).

Figure 1

The Interoceptive Assessment Puzzle. The different modalities of interoception (e.g., cardioception) can be assessed directly or indirectly. Direct assessments can be subjective or objective, depending on whether they use conscious metacognitive appraisals or performance-based indices. Objective interoceptive tasks can be explicit (participants must consciously attend to interoceptive signals; e.g., the heartbeat counting task) or implicit (measurements of interoception are covert; e.g., heartbeat evoked potentials measured during rest). Indirect assessments evaluate constructs typically related, and ideally dependent on, interoceptive processes (e.g., physiological indices such as HRV or effect of cardiac phase) or abilities (e.g., self-report measures of emotion regulation or identification abilities.



Although using self-report questionnaires to measure deeply embodied functions might seem paradoxical, recent redefinitions of interoception emphasize the importance of high-level and metacognitive elaboration of interoceptive information, and supportthe idea that some facets of interoception, including participants’ metacognitive beliefs, can be subjectively assessed ([Khalsa, Adolphs, Cameron, Critchley, Davenport, Feinstein, Feusner, Garfinkel, Lane, Mehling, & others, 2018](#ref-khalsa2018interoception); [Suksasilp & Garfinkel, 2022](#ref-suksasilp2022towards))..

Indeed, the fact that self-reported interoception reflects a fundamentally different aspect of conscious interoceptive processing compared to objective task measures, is thought to be central to understanding the reported lack of convergence between objective and subjective measures ([Desmedt et al., 2022](#ref-desmedt2022measures)). For instance, there are typically weak or no correlation between questionnaires and objective measures (HCT, [Schandry, 1981](#ref-schandry1981heart)) and the Heartbeat Detection Task (HDT, [Kleckner et al., 2015](#ref-kleckner2015methodological)), even when both measures are intended to capture the same theoretical interoceptive dimension ([Arslanova et al., 2022](#ref-arslanova2022); [Brand et al., 2023](#ref-brand2023); e.g., task-based accuracy vs. self-reported accuracy, [Murphy et al., 2019](#ref-murphy2019)).

Perhaps more surprisingly, there is also a reported lack of convergence across both objective and subjective measures: objective measures purporting to measure the same interoceptive dimension, such as accuracy, show no or weak correlation (respectively, [Brand et al., 2023](#ref-brand2023); [Hickman et al., 2020](#ref-hickman2020relationship)), and subjective questionnaire measures are only weakly correlated with each other ([Vig et al., 2022](#ref-vig2022questionnaires)). As well as posing substantial validity concerns, these findings raise the possibility that even within broad measurement domain, different measures are targeting fundamentally different facets of interoception, undermining replicability and hindering theoretical and empirical progress in the field.

The lack of convergence among self-reported trait measures of interoception, also called “interoceptive sensibility” [i.e., the tendency to focus on and/or accurately detect interoceptive signals; Garfinkel et al. ([2015](#ref-garfinkel2015knowing)); Khalsa, Adolphs, Cameron, Critchley, Davenport, Feinstein, Feusner, Garfinkel, Lane, Mehling, and others ([2018](#ref-khalsa2018interoception))], is particularly striking and problematic. A recent systematic review suggested that various questionnaires designed to assess interoceptive sensibility may, in fact, measure distinct constructs, with the risk of researchers treating them as equivalent despite overall low convergence ([Desmedt et al., 2022](#ref-desmedt2022measures); [Vig et al., 2022](#ref-vig2022questionnaires)).

Several widely used questionnaires were included in the review, such as the Body Perception Questionnaire (BPQ, [Porges, 1993](#ref-porges1993body)), the Multidimensional Assessment of Interoceptive Awareness (MAIA, [Mehling et al., 2012](#ref-mehling2012); MAIA-2, [Mehling et al., 2018](#ref-mehling2018multidimensional)), the Body Awareness Questionnaire (BAQ, [Shields et al., 1989](#ref-shields1989body)), the Private subscale of the Body Consciousness Questionnaire (PBCS, [Miller et al., 1981](#ref-miller1981consciousness)), and the Self-Awareness Questionnaire (SAQ, [Longarzo et al., 2015](#ref-longarzo2015self)) which comprised 86% of all citations. The lack of correlations to moderate correlations among questionnaires measures suggests that rather than measuring a common construct they likely capture different aspects of interoception, emphasizing the need for greater conceptual clarity regarding what each measure captures, how they relate to different dimensions of interoception, and their potential overlaps with other constructs, such as alexithymia and body awareness. One avenue for addressing these issues is the careful investigation of newer measures that were not included in prior reviews.

## The Interoceptive Accuracy Scale (IAS)

Focusing on the accuracy dimension of interoceptive sensibility, a recently developed scale with a rapidly growing popularity is the Interoceptive Accuracy Scale (IAS, [Murphy et al., 2019](#ref-murphy2019)). The authors developed the IAS to clarify not only how interoception is measured, but also what is being measured. Whereas many existing interoceptive models (e.g., [Garfinkel et al., 2015](#ref-garfinkel2015knowing); [Khalsa, Adolphs, Cameron, Critchley, Davenport, Feinstein, Feusner, Garfinkel, Lane, Mehling, & others, 2018](#ref-khalsa2018interoception)) conflate interoceptive accuracy and attention, Murphy and colleagues conceptualise these as distinct dimensions. As a result, individuals may report attending to internal bodily signals while simultaneously recognising that their perception of these signals is inaccurate. This distinction represents a conceptual advance in the measurement of interoceptive sensibility and positions the IAS as the only tool specifically designed to assess subjective interoceptive accuracy.

The IAS comprises 21 Likert-scale items that query how accurately one can perceive different bodily signals, with one item per physiological modality such as respiration (“I can always accurately perceive when I am breathing fast”), heart (“I can always accurately perceive when my heart is beating fast”), skin (“I can always accurately perceive when something is going to be ticklish”), arousal or bodily functions like coughing (“I can always accurately perceive when I am going to cough”) or urinating (“I can always accurately perceive when I need to urinate”). Appealingly, the IAS’ items are about specific interoceptive behaviors and “objective” experiences (albeit subjective reports of them), a distinct difference compared to other popular interoception questionnaires, such as the MAIA-2, which contains more general and metacognitive items (e.g., “I trust my body sensations”, “I can notice an unpleasant body sensation without worrying about it”), as well as dimensions related to attention and emotion regulation (e.g., Not-distracting and Not-worrying subscales respectively).

The original IAS validation study suggested a two-factor structure: one reflecting the perception of interoceptive signals (urinate, hungry, defecate, thirsty, pain, heart, taste, breathing, temperature, muscles, affective touch, vomit, sexual arousal), and the other comprising signals that are either difficult to perceive solely through interoceptive information or reflect socially unacceptable bodily functions (itch, tickle, cough, burp, bruise, blood sugar, sneeze, wind). The authors highlight the imperfect fit of the 2-factor structure ([Murphy et al., 2019, p. 127](#ref-murphy2019)), with several subsequent studies identifying different optimal solutions and interpretations. Brand et al. ([2023](#ref-brand2023)) reported a 1-factor solution to be the best fit, whereas Lin et al. ([2023](#ref-lin2023)) and Campos et al. ([2021](#ref-campos2021)) found bifactor solutions (i.e., one general factor above a set of lower-level factors, [Rodriguez et al., 2016](#ref-rodriguez2016evaluating)) to be the best fit. Using Exploratory Factor Analysis (EFA), and constraining the solution to two factors to remain consistent with the original validation, Koike and Nomura ([2023](#ref-koike2023)) identified a different 2-factor structure that differentiated cutaneous (itching, tickling, coughing, burping, affective touch, bruising, passing gas, sneezing, muscle sensations, sexual arousal, and taste) and visceral sensations (urination, defecation, hunger, thirst, pain, breathing, fatigue/blood sugar, temperature, vomiting, and heartbeat).

As well as differences in the optimal factor structure of the IAS, there has also been discussion regarding the specificity of individual items. For instance, Murphy et al. ([2019](#ref-murphy2019)) note that while some items measure direct interoceptive signals such as cardioception, others likely capture phenomena that are difficult to detect solely through interoceptive signals (e.g., “bruising” might be detected through skin discoloration; p. 119). Lin et al. ([2023](#ref-lin2023)) highlight several psychometric limitations, including five locally dependent pairs and three items (touch, blood sugar, bruise) with exceptionally high difficulty and low discrimination, suggesting further refinement of IAS items to improve precision. Localization issues also arose, with both “itch” and “tickle” corresponding to the same Chinese character, leading to their collapse into a single item ([Lin et al., 2023](#ref-lin2023)). Additionally, Campos et al. ([2021](#ref-campos2021)) suggested that the IAS is predominately a unidimensional scale with all 20 items (with the exception of “tickle”) reflecting a general factor.

Regarding its validity, the IAS has been compared to other interoception-related measures. It shows positive correlations with most MAIA facets ([Gaggero et al., 2021](#ref-gaggero2021); [Mehling et al., 2018](#ref-mehling2018multidimensional)) except for the Not-Distracting and Not-Worrying subscales ([Brand et al., 2023](#ref-brand2023)), subscales previously highlighted as related to non-interoceptive abilities ([Ferentzi et al., 2021](#ref-ferentzi2021examining)). Comparisons between the IAS and the body awareness dimension of the BPQ, have been mixed: some studies report small positive correlations ([Brand et al., 2023](#ref-brand2023); [Gaggero et al., 2021](#ref-gaggero2021); [Koike & Nomura, 2023](#ref-koike2023)), while others find small negative correlations ([Lin et al., 2023](#ref-lin2023)), a quadratic positive relationship ([Campos et al., 2021](#ref-campos2021)) or no correlation at all ([Murphy et al., 2019](#ref-murphy2019)). Positive correlations have been observed for four of the five subscales from the Five Facet Mindfulness Questionnaire (FFMQ, [Baer et al., 2006](#ref-baer2006using); [Brand et al., 2023](#ref-brand2023); [Koike & Nomura, 2023](#ref-koike2023)).

Additionally, the IAS is positively correlated with the interoceptive awareness subscale of the Eating Disorder Inventory (EDI-IA, [Lin et al., 2023](#ref-lin2023)) and a negatively correlated with the Interoceptive Confusion Questionnaire ([Brand et al., 2023](#ref-brand2023); ICQ, [Brewer et al., 2016](#ref-brewer2016alexithymia); [Murphy et al., 2019](#ref-murphy2019)). Finally, supporting the distinction between interoceptive accuracy and attention, IAS’ correlation with the Interoceptive Attention Scale (IATS, [Gabriele et al., 2022](#ref-gabriele2022dissociations)) is typically small ([Koike & Nomura, 2023](#ref-koike2023); [Lin et al., 2023](#ref-lin2023)).

Although assessing the predictive validity of an interoception scale may be theoretically challenging, expected negative associations have been consistently observed between the IAS and alexithymia ([Brand et al., 2023](#ref-brand2023); [Campos et al., 2021](#ref-campos2021); [Koike & Nomura, 2023](#ref-koike2023); [Lin et al., 2023](#ref-lin2023); [Murphy et al., 2019](#ref-murphy2019)), somatic symptoms ([Brand et al., 2023](#ref-brand2023); [Koike & Nomura, 2023](#ref-koike2023); [Lin et al., 2023](#ref-lin2023)), depressive symptoms ([Brand et al., 2023](#ref-brand2023); [Koike & Nomura, 2023](#ref-koike2023); [Lin et al., 2023](#ref-lin2023)), anxiety ([Brand et al., 2023](#ref-brand2023)), neuroticism ([Brand et al., 2023](#ref-brand2023)) and self-esteem ([Murphy et al., 2019](#ref-murphy2019)).

Taken together, these findings support the IAS as measuring an adaptive aspect of interoception, although its pattern of associations with other interoception (or interoception-related) questionnaires indicate overlap across various theoretical dimensions. This further highlights limitations of strictly orthogonal models of interoception and the possibility of its faithful capture by questionnaires. Accordignly, while the IAS shows promise, the findings underscore the need for further refinement which the present paper seeks to address.

Given increasing interest in interoception across psychology and neuroscience, and the growing use of the IAS, careful scrutiny of this novel self-report measure of interoceptive accuracy is warranted. Therefore, the current study has two aims. First, to clarify the structure of the IAS using a mega-analytic approach which leverages existing data by aggregating existing datasets and re-analyzing them at the raw data level. This aggregated data will then be used to contrast the traditional CFA/SEM factor-based analyses and Hierarchical Clustering Analysis (HCA) with newer network-based ones (Exploratory Graph Analysis). By utilizing conceptually distinct analytical frameworks, the current paper aims to provide a more nuanced and method-independent view of the IAS factor structure, allowing the identification of both convergent and potentially overlooked structural features that may not be apparent when relying on a single analytic approach. Second, to provide an overview of construct validity and the dispositional correlates of the IAS, clarifying its general pattern of associations with interoception-related constructs, mood, psychopathology and other relevant measures and thereby situating interoception questionnaires within a broader context.

# Study 1

Study 1 re-analyses and assesses the factor structure of the IAS by re-analyzing the raw data across 17 open-access datasets ([Arslanova et al., 2022](#ref-arslanova2022); [Brand et al., 2022](#ref-brand2022); [Brand et al., 2023](#ref-brand2023); [Campos et al., 2021](#ref-campos2021); [Gaggero et al., 2021](#ref-gaggero2021); [Lin et al., 2023](#ref-lin2023); [Murphy et al., 2019](#ref-murphy2019); [Petzke et al., 2024](#ref-petzke2024somatic); [Poerio et al., 2024](#ref-poerio2024interoceptive); [Todd et al., 2022](#ref-todd2022); [Von Mohr et al., 2023](#ref-von2023)). Combining these studies has the advantage of providing more robust and generalizable understanding of the IAS’ factor structure. However, given that studies differ in their sample size, demographic characteristics, language, and procedure, we also provide an individual study analysis (i.e., applying the same method on all samples separately) to add nuance to the general picture.

## Methods

### Datasets

Our search focused on studies citing the original IAS validation paper ([Murphy et al., 2019](#ref-murphy2019)) which identified 136 papers (as of 01/05/2024). To qualify for inclusion, papers needed to provide open-access to raw data and report individual IAS items scores. A total of 17 datasets were used, including data from five unpublished studies (see **Table 1**). The total number of participants across these datasets was 33,526 participants (Mean = 47.96 13.1, 71.3% Female).

Table 1

Description of the samples used in Study 1 mega-analysis.

| Sample | Reference | Language | N | Difference | Age (Mean ± SD) | Range | Female % | Availability |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Murphy et al., (2020) |  |  |  |  |  |  | osf.io/3m5nh |
| Sample 1a |  | English | 451 |  | 25.8 ± 8.4 | 18-69 | 69.4% |  |
| Sample 1b |  | English | 375 |  | 35.3 ± 16.9 | 18-91 | 70.1% |  |
| Sample 2 | Gaggero et al., (2021) | English and Italian | 814 |  | 24.9 ± 5.3 | 18-58 | 60.3% | osf.io/5x9sg |
| Sample 3 | Campos et al., (2022) | Portuguese | 515 |  | 30.7 ± 10.5 | 18-72 | 59.6% | osf.io/j6ef3 |
| Sample 4 | Todd et al., (2022) | English | 802 |  | 48.6.6 ± 14.1\* | 18-92\* | 50%\* | osf.io/ms354 |
| Sample 5 | Arslanova et al., (2022) | English | 143 |  | 28.5 ± 7.6 | 18-73 | 46.8% | osf.io/mp3cy |
| Sample 6 | Brand et al., (2022) | German | 619 |  | 43.9 ± 14.5 | 18-78 | 78.7% | osf.io/xwz6g |
|  | Brand et al., (2023) |  |  |  |  |  |  | osf.io/3f2h6 |
| Sample 7a |  | German | 522 |  | 23.4 ± 6.7 | 18-79 | 79.5% |  |
| Sample 7b |  | German | 1993 |  | 32.0 ± 12.6 | 16-81 | 77.7% |  |
| Sample 7c |  | German | 802 |  | 27.3 ± 9.3 | 18-72 | 68.9% |  |
|  | Lin et al., (2023) |  |  |  |  |  |  | osf.io/3eztd |
| Sample 8a |  | Chinese | 1166 | Collapsed "Itch" and "Tickling" | 32.5 ± 8.4 | 16-60 | 57.0% |  |
| Sample 8b |  | Chinese | 500 | Collapsed "Itch" and "Tickling" | 37.4 ± 7.4 | 20-60 | 56.2% |  |
| Sample 9 | VonMohr et al., (2023) | English | 21843 |  | 56.5 ± 14.4 | 18-93 | 73.2% | osf.io/7p9u5 |
| Sample 10 | Petzke et al., (2024) | German | 254 |  | 31.5 ± 10.7 | 18-73 | 68.5% | osf.io/seru4 |
| Sample 11 | Makowski et al., (2023a) | English | 485 | Analog scales. No Temperature, Blood sugar and Cough items | 30.1 ± 10.1 | 17-76 | 50.3% | github.com/RealityBending/IllusionGameReliability |
| Sample 12 | Makowski et al., (2023b) | English | 836 | Analog scales | 25.1 ± 11.3 | 18-50 | 53.0% | github.com/DominiqueMakowski/PHQ4R |
| Sample 13 | Makowski et al., (2023c) | English | 146 | Analog scales | 21.1 ± 4.3 | 17, 87 | 76% | github.com/RealityBending/InteroceptionPrimals |
| Sample 14 | Makowski et al., (2024) | English | 737 |  | 36.8 ± 14.9 | 18-57 | 57.3% | github.com/RealityBending/InteroceptionScale |
| Sample 15 | Poerio et al., (2024) | English | 107 |  | 26.8 ± 9.2 | 18-60 | 74.8% | osf.io/49wbv |
| Sample 16 | Poerio et al., unpublished | English | 131 |  | 30.9 ± 12.0 | 18-79 | 75.9% |  |
| Sample 17 | Arjona et al., unpublished | English | 279 |  | 26.4 ± 13.2 | 22-69 | 67.7% |  |
| Sample 18 | Total |  | 33526 |  | 47.96 ± 13.1 | 17-93 | 71.3% |  |

\*Information taken from the sample description of relevant paper rather than recomputed.

### Data Analysis

Psychometrically good items should exhibit various qualities, such as validity, reliability and discrimination, to which one of the contributing factors is the amount of inter-individual variability captured by an item. IAS Items to which most participants’ answers are concentrated around one option - i.e., exhibiting a narrow distribution - were flagged as potentially problematic.

After examining the distributions and correlations of IAS items, we tested for “redundant” items (e.g., due to multicollinearity or local dependency) using Unique Variable Analysis (UVA, [Christensen et al., 2023](#ref-christensen2023unique)), a novel method derived from network psychometrics to identify and merge items that share substantial variance (which can distort the structure estimation). We used a conservative threshold of 0.30 that detects “large” to “very large” overlap between items and only suggest scale modifications if strongly justified.

Following the analysis of items, we analyzed the factor structure of the IAS using three different approaches, each with particular trade-offs and assumptions, to provide a multi-verse picture of likely solutions. Specifically, we applied traditional exploratory and confirmatory Factor Analysis (EFA/CFA), hierarchical clustering (HCA), and Exploratory Graph Analysis (EGA), to both the whole sample and each dataset separately (details being available in the analysis document at https://github.com/RealityBending/InteroceptionIAS). Decisions and conclusions consider both overall and dataset specific results.

EGA is a novel framework that combines network analysis with psychometric methods ([H. F. Golino & Epskamp, 2017](#ref-golino2017exploratory)) to jointly estimate the number of dimensions (i.e., groups of items) as well as the structure stability of a scale ([H. Golino et al., 2020](#ref-golino2020investigating); [H. F. Golino & Epskamp, 2017](#ref-golino2017exploratory)). EGA is a considered a suitable but still comparable and interpretable alternative to traditional factor analysis that addresses some of its limitations, including the assumption of a “latent” source of variability, possible biasing in optimal factor number estimation by sample size ([Christensen & Golino, 2021b](#ref-christensen2021equivalency)). Additionally, when compared to other factor retention methods such as Kaiser criterion and parallel analysis, EGA performed better in estimating number of factors in complex population structures, such as bi-factor solutions ([Jiménez et al., 2023](#ref-jimenez2023dimensionality)). At a fundamental level, EGA conceptualizes items as nodes in a network, with connections (edges) between nodes reflecting associations between items. Clustering of nodes reveals distinct communities of related items, in practice akin to traditional latent factors albeit without explicitly assuming their presence ([Christensen & Golino, 2021b](#ref-christensen2021equivalency)). We used the EGAnet package ([Christensen & Golino, 2021a](#ref-christensen2021estimating)) to fit a hierarchical EGA with the Leiden community detection algorithm.

Although EGA is an attractive and robust alternative to traditional factor analysis, factor analysis remains a widely used method for dimensionality assessment. Since our goal was to provide a general, yet nuanced picture of IAS factor structure, with capability to show potential discrepancies depending on analysis methods used, we also include traditional EFA/CFA here. Unlike EGA, factor analysis assumes that a latent source of variability (i.e., a common latent variable) underlies the observed set of manifest variables ([Cosemans et al., 2022](#ref-cosemans2022exploratory)). A critical step in factor analysis is determining the optimal number of factors, for which we used the Method Agreement Procedure ([Lüdecke et al., 2021](#ref-ludecke2021performance)), a consensus-based decision method based on multiple factor estimation methods applied concurrently.

Finally, we also applied Hierarchical Clustering Analysis (HCA, [Murtagh & Legendre, 2014](#ref-murtagh2014ward)), which unlike factor analysis, does not assume any latent source of variability, but instead iteratively groups items based on their similarity (e.g., correlation) into a hierarchy of clusters. The benefits of HCA include its interpretability and ability to capture complex relationships among items without relying on strict assumptions about data distribution or latent variables.

In a typical 2-step fashion, initial analyses were conducted, followed by structure refinement via item selection, with the final item pool tested for a second time. Additionally, various solutions (e.g., adding general factors) were statistically compared using Confirmatory Factor Analysis (CFA).

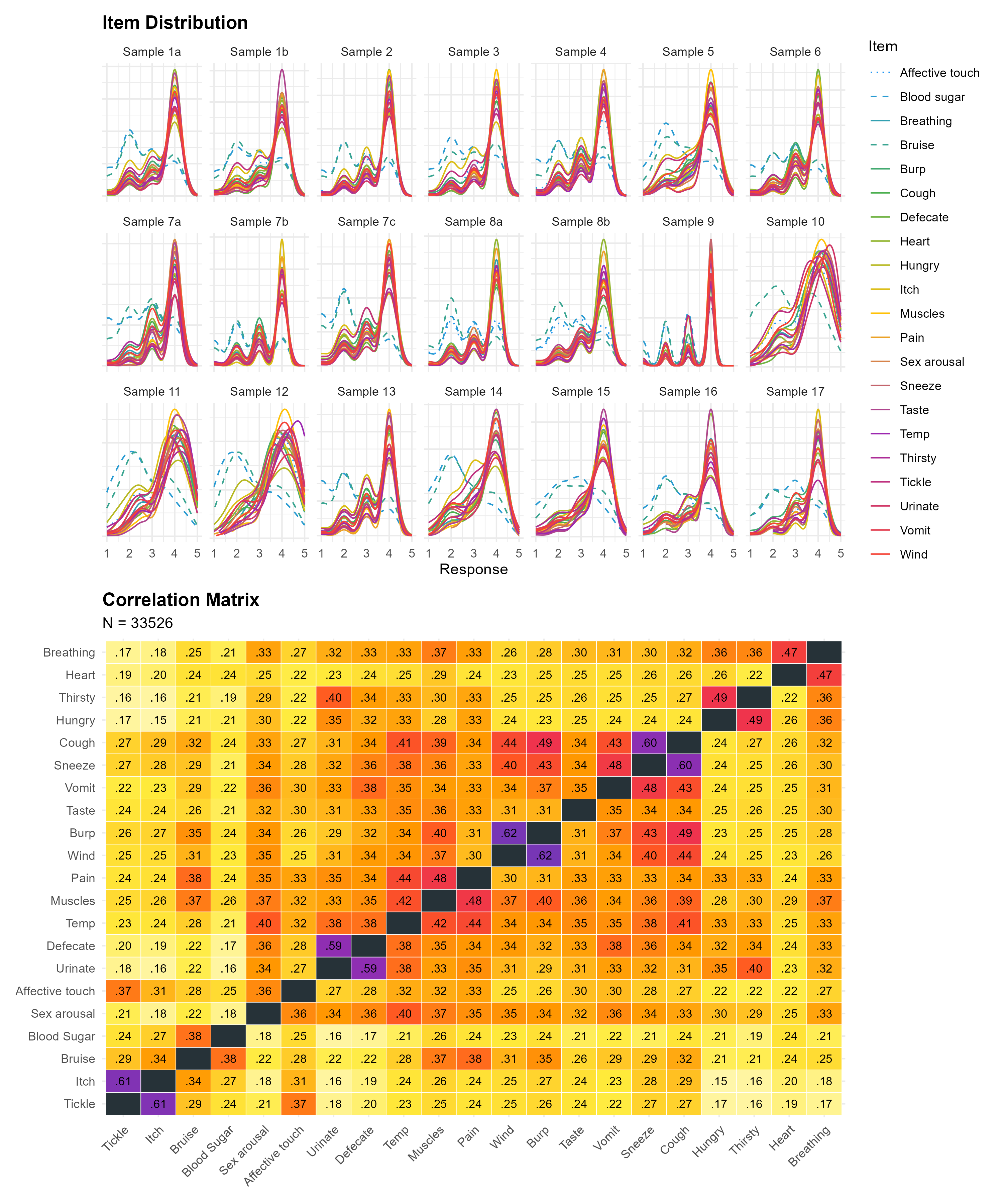
## Results

The distribution of the items across samples suggested a consistent modal value ([Figure 2](#fig-distributions), top panel), such that participants are most likely to answer 4/5 (i.e., agree) on all IAS items except “blood sugar” and “bruise”, which show a different distributional pattern with a lower mode (~2/5).

While this difference is not problematic per se, the contrasting distributions might be indicative of items with non-homogeneous psychometric “difficulty”, potentially due to difficulty in inferring and reporting them. **Consistently with Sorokowska et al. (**[**2021**](#ref-sorokowska2021affective)**),** lower mode responses were also observed for “affective touch” in samples the Chinese versions, indicating translation or cultural discrepancies. More generally, the response spread (i.e., scale coverage) was narrow (which can limit psychometric sensitivity), with a particularly low occurrence of extreme values (1 and 5), and a strong mode with most responses clustering around 4 (assuming the IAS is implemented as a 5-point Likert scale following its validation). Samples using an analogue response scale (samples 10, 11 and 12 in [Figure 2](#fig-distributions)) displayed a more continuous spread of answers, seemingly improving the interindividual variability, although potentially displaying a secondary lower mode at ~2 suggesting the existence of potential clusters of participants. The correlation matrix between all items (see Figure 2, bottom panel) shows an overall positive correlation pattern, with highly correlated item pairs (e.g., Tickle-Itch, Urinate-Defecate, Pain-Wind, Hungry-Thirsty) and triplets (e.g., Vomit-Sneeze-Cough, Temperature-Muscles-Pain).

Figure 2

Top: Distribution of responses across datasets reveals a consistent modal value, typically 4 or 5 (indicating agreement), except for ‘blood sugar’ and ‘bruise’ (dashed lines) and ‘affective touch’ (dotted lines) in the Chinese validation sample, which have lower modes. Most responses cluster around middle values, with few extreme scores (1 and 5). Samples using an analogue scale (10a-10c in Table 1 but 10-12 below) show a more continuous distribution and increased interindividual variability. Bottom: The correlation matrix between all items shows an overall positive correlation pattern, with correlated item pairs (e.g., Wind, Burp) and triplets (e.g., Vomit, Sneeze and Cough).



UVA flagged two strongly redundant variables, “itch” and “tickle” indicating to remove the latter. We removed “tickle” from further analysis due to its high redundancy (and because it is absent from some datasets due to translation issues). Several more item pairs were flagged as moderately redundant (“wind” and “burp”; “urinate” and “defecate”) and mildly redundant (“sneeze” and “cough”; “heart” and “breathing”; “hungry” and “thirsty”). These patterns were observed consistently across most individual samples.

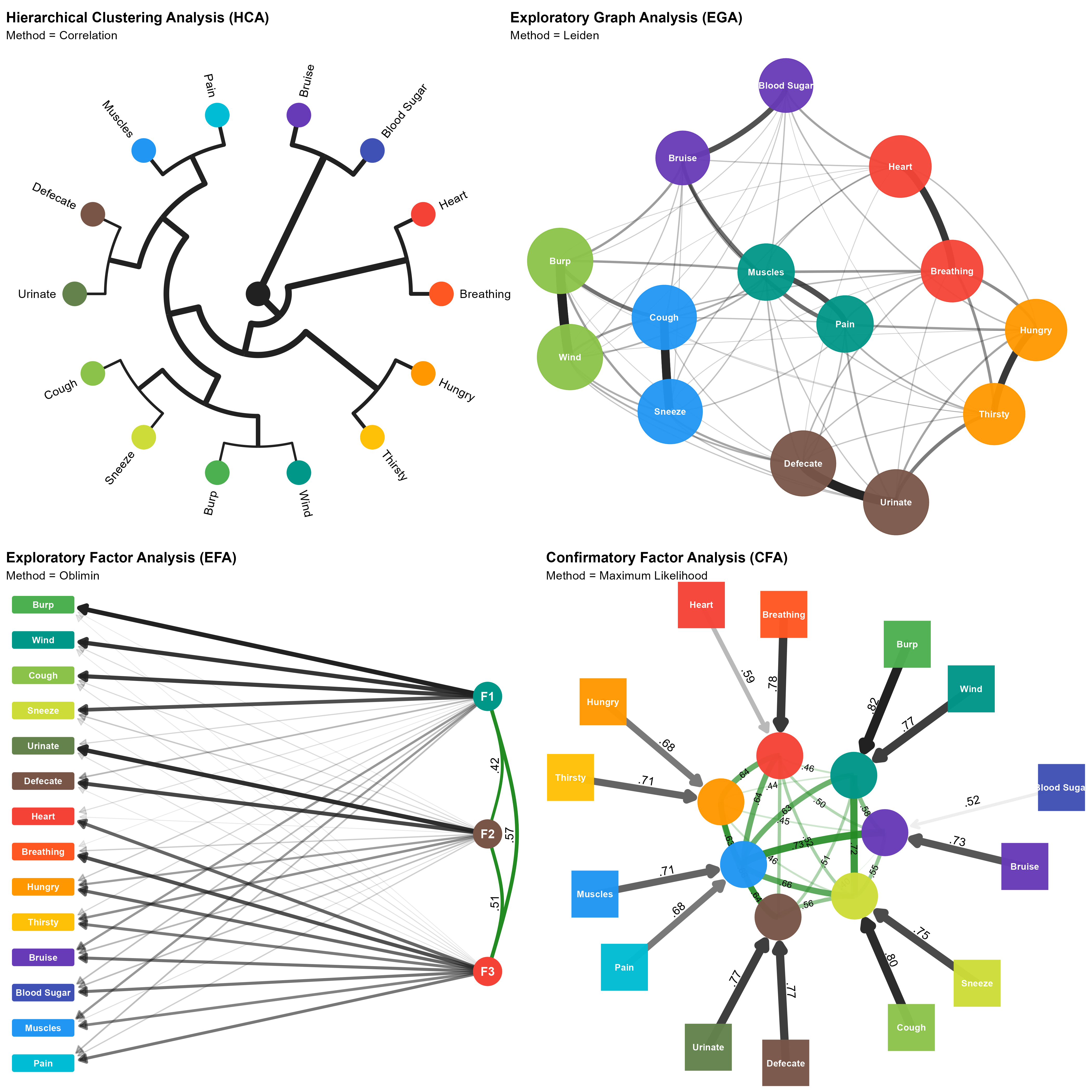
### Step 1

HCA highlighted item pairs and triplets that were consistently grouped together across samples, such as “wind” and “burp”, “sneeze” and “cough”, “itch” and “bruise”, “urinate” and “defecate”, and “pain”, “muscles”, and “temperature”. This pattern was largely replicated with EGA, with the additional presence of a unique cluster comprising “sex arousal”, “affective touch”, “temperature”, “pain”, “muscles”, and “taste”. EFA suggested an optimal 3-factor solution with one factor comprising expulsion-related items (“burp”, “wind”, “cough”, “sneeze”, and “vomit”), a second factor comprising visceroceptive items (“heart”, “breathing”, “hungry”, “thirsty”, “urinate”, and “defecate”), and a third factor comprising items that are not directly accessible via internal bodily sensations alone (“bruise” and “blood sugar”). Results on the structure of the full scale (excluding “tickle”) are presented in the **Supplementary Materials.**

Importantly, these initial structural analyses highlighted six problematic items in addition to “tickle” identified from UVA. “Taste” typically displayed a lone or unstable pattern of associations, “affective touch” exhibited cross-loadings and instability, “vomit” was weakly associated with other items, and “itch” did not form a consistent cluster. Finally, “temperature” and “sexual arousal” showed redundant patterns of associations but were less reliable. These 6 items as well as “tickle” were removed, and a second run of structural analyses were performed on the remaining 14 items (summarized in [Figure 3](#fig-structure)).

Figure 3

Four structure analysis methods (HCA, EGA, EFA, CFA) were applied and converged on a consistent optimal solution of 14 items formed seven pairs: Hungry-Thirsty, Bruise-Blood sugar, Urinate-Defecate, Muscles-Pain, Breathing-Heart, Cough-Sneeze, Wind-Burp. While EFA suggested 3 factors, CFA confirmed the superiority of the 7-factor model over alternative structures.



HCA and EGA yielded highly consistent results, emphasizing seven item pairs: Hungry-Thirsty, Bruise-Blood sugar, Urinate-Defecate, Muscles-Pain, Breathing-Heart, Cough-Sneeze, Wind-Burp. HCA also meaningfully grouped the Urinate-Defecate and Muscles-Pain pairs, and expulsion items pairs (Wind-Burp and Cough-Sneeze). EFA again suggested an optimal 3-factor solution with the first factor comprising expulsion-related items (“burp”, “wind”, “cough”, “sneeze”), the second factor comprising the Urinate-Defecate pair, and the third factor comprising the remaining items.

### Step 2

We fitted and compared various candidate structures emerging from the previous analyses, including a 1-factor model (the G-model), a 3-factor model (EFA), a 3+1 model (EFA + general factor), a 5-factor model (HCA), a 5+1 model (HCA + general factor), a 7-factor model (EGA), and a 7+1 model (EGA + general factor), using CFA. The EGA model with 7 factors of item pairs provided the best fit to the data with the lowest RMSEA (0.035), lowest (2,334.112), and highest CFI (0.984). This was followed by the EGA + general factor model (RMSEA = 0.054, = 6,876.395, CFI = 0.952), and then the HCA model with 5 factors (RMSEA = 0.078, = 13,441.878, CFI = 0.906). All other models performed poorly, with RMSEA > 0.08 and CFI < 0.90. The EGA model yielded the lowest BIC (which considers the number of factors), followed by the EGA plus a general factor model. All other models displayed significantly (BIC-based Bayes Factor < 1/100) lower evidence for their structures compared to the EGA model.

Across individual datasets, the EGA model provided the best fit to the data in 8 out of 13 samples for which the CFA converged. In these samples, the 7-factor EGA model showed excellent performance, with low RMSEA values (≈ 0.025–0.057), high CFI values (≥ 0.95), and substantially lower values compared to alternative models. Adding a general factor improved model fit in three samples, for which the 7-factor EGA + general factor model showed acceptable RMSEA (≈ 0.048–0.061) and CFI (≈ 0.929–0.957) values. In two samples, the 5-factor plus general factor model provided the best fit, with acceptable RMSEA values (0.049 and 0.057) and CFI values (0.955 and 0.919). Bayes factor comparisons aligned with BIC-based conclusions; taken together, these results suggest that although the EGA model was preferred in most samples, bifactor models (i.e., EGA and HCA models with a general factor) provided evidence for a common IAS factor in some samples.

## Discussion

Study 1 aimed to systematically evaluate the structure of the IAS using a mega-analytic approach with both factor-analytic and network-based methods. Previous research across different datasets has highlighted inconsistency in optimal factor structure (e.g., unidimensional, bifactor, and two factor solutions have all been reported [Brand et al., 2023](#ref-brand2023); [Campos et al., 2021](#ref-campos2021); [Koike & Nomura, 2023](#ref-koike2023); [Lin et al., 2023](#ref-lin2023); [Murphy et al., 2019](#ref-murphy2019)) and raised concerns about specific items (e.g., low discrimination, local dependencies, and items reflecting phenomena less accessible to interoception). To address these concerns and inconsistencies we combined 17 datasets with over 33 thousand participants using the IAS, providing the most comprehensive and robust analysis of the IAS structure to date. We conducted analyses across both combined and individual datasets allowing an examination of overall and dataset specific results and how they compared across three different structural analyses methods (EFA, HCA, EGA).

Analyses revealed that seven item pairs (Hungry-Thirsty, Bruise-Blood sugar, Urinate-Defecate, Muscles-Pain, Breathing-Heart, Cough-Sneeze, Wind-Burp) provided a robust and replicable structure, leading to a refined 14-item IAS. These seven pairs emerged consistently across datasets and had superior fit indices compared to alternative models for most samples. Results suggest that rather than measuring the broad latent construct of “interoceptive accuracy” the IAS reflects clusters of tightly coupled items tied to specific bodily sensations, , including both “directly” felt sensations and inferred or anticipatory physical processes in the case of the Bruise-Blood Sugar pairing.

Alternative mechanisms may also contribute to the emergence of these item pairings. Clustering may reflect shared temporal or functional characteristics of bodily events. For example, the Cough-Sneeze and Wind-Burp pairings may involve rapid reflexive responses and frequent experiential co-occurrence, whereas the Hungry-Thirsty and the Urinate-Defecate pairings may be shaped by basic homeostatic regulation and recurrent need states. Such shared characteristics may elicit similar response strategies even when the underlying physiological processes are distinct. Importantly, these alternative explanations do not exclude the interpretation of the pairings as reflecting intuitive bodily domains (e.g., visceral, expulsion, musculoskeletal). Temporal and functional similarities are themselves grounded in bodily organisation and lived bodily experiences, suggesting multiple and overlapping sources of structure may shape responses to the IAS items. This convergence may contribute to the intuitive coherence of the item pairs while also increasing redundancy and instability when weakly discriminating items are retained, thereby complicating factor-analytic solutions and likely contributing to inconsistencies observed in prior findings ([Campos et al., 2021](#ref-campos2021); [Lin et al., 2023](#ref-lin2023)).

These findings highlight the need for targeted scale refinement and motivate future development of context-sensitive, multidimensional interoception measures. For example, many bodily sensations are inherently ambiguous without contextual information (e.g., a racing heart may signal physical exertion, anxiety, or excitement) and without specific context (e.g., physiological vs. affective triggers), it is unclear which state participants reference when responding. Such ambiguity likely constrains ecological validity and contributes to variability across studies, therefore adopting context-specific phrasing may therefore be necessary to more accurately capture interoceptive beliefs and reduce measurement noise ([Makowski, Neves, et al., 2025](#ref-makowski2025mint); e.g., [Vlemincx et al., 2023](#ref-vlemincx2023novel)).

Overall, Study 1 suggests that, after removal of problematic items the IAS is best conceptualized as a set of item clusters rather than a unidimensional scale. This structural insight provides the foundation for Study 2, where dispositional correlates are examined to further evaluate the IAS’s construct validity.

# Study 2

Although the IAS has been widely adopted, evidence regarding its associations with personality traits, affective dispositions, and related psychological constructs is scattered and inconsistent ([Brand et al., 2023](#ref-brand2023); [Lin et al., 2023](#ref-lin2023); [Murphy et al., 2019](#ref-murphy2019); [Todd et al., 2022](#ref-todd2022)). In Study 2 we aimed to provide a robust overview of the dispositional correlates of the IAS by leveraging and combining a large number of available datasets containing both the IAS, other measures of interoception, and dispositional correlates. We also sought to assess whether structural improvements of the 14-item IAS translate into clearer and more reliable external associations by comparing associations to dispositional variables with the original 21-item unidimensional solution, thereby potentially offering a more efficient alternative to the longer version for future research.

## Methods

### Materials

We selected measures that appeared at least once in the 17 datasets from Study 1 and/or were relevant given the scope of the study, i.e., constructs related to physiology, mood, personality, psychopathology, and beliefs and misbeliefs. We merged scores of regular and abridged versions of scales for conciseness where applicable (e.g., the 2-item Generalized Anxiety Disorder and Patient Health Questionnaire measures were combined into the Patient Health Questionnaire-4).

#### Interoception and Interoception Related.

Interoception-related measures included the Multidimensional Assessment of Interoceptive Awareness Version-2 (MAIA-2, [Mehling et al., 2018](#ref-mehling2018multidimensional)), the Body Perception Questionnaire Short Form (BPQ-SF, [Cabrera et al., 2018](#ref-cabrera2018assessing)), and the Interoceptive Confusion Questionnaire (ICQ, [Brewer et al., 2016](#ref-brewer2016alexithymia)). We also included the Bermond–Vorst Alexithymia Questionnaire (BVAQ, [Vorst & Bermond, 2001](#ref-vorst2001validity)), and the Toronto Alexithymia Scale (TAS-20, [Bagby et al., 1994](#ref-bagby1994twenty)), due to the close relationship between interoception deficits and alexithymia (e.g., [Brewer et al., 2016](#ref-brewer2016alexithymia); [Gaggero et al., 2021](#ref-gaggero2021)).

#### Mood and Anxiety Related.

Scores from depression and anxiety-related measures were included due to their established associations with interoceptive processing ([Khalsa, Adolphs, Cameron, Critchley, Davenport, Feinstein, Feusner, Garfinkel, Lane, Mehling, Meuret, et al., 2018](#ref-khalsa2018)); namely the Patient Health Questionnaire-4 (PHQ-4, [Kroenke et al., 2009](#ref-kroenke2009ultra)), the Patient Health Questionnaire-15 (PHQ-15, [Kroenke et al., 2002](#ref-kroenke2002phq)), the Patient Health Questionnaire-9 (PHQ-9, [Kroenke et al., 2001](#ref-kroenke2001phq)), the Short Mood and Feelings Questionnaire (MFQ, [Angold et al., 1995](#ref-angold1995development)), the Beck Depression Inventory-II (BDI-II, [Beck et al., 1996](#ref-beck1996beck)), the State–Trait Anxiety Inventory Trait Version (STAI-T, [Spielberger, 1970](#ref-spielberger1970manual); and STAIT-T, [Zsido et al., 2020](#ref-zsido2020development)), and the Generalized Anxiety Disorder Scale (GAD-7, [Spitzer et al., 2006](#ref-spitzer2006brief)), and General Anxiety Disorder-2 (GAD-2, [Kroenke et al., 2007](#ref-kroenke2007anxiety), obtained from the PHQ-4).

#### Psychopathology and Neurodevelopmental Traits.

Dimensional psychopathology and neurodevelopmental-related scores were included from the Personality Inventory for DSM-5 Short Form (PID-5-SF, [Thimm et al., 2016](#ref-thimm2016personality)), the Autism Spectrum Quotient Short Form (ASQ-S, [Hoekstra et al., 2011](#ref-hoekstra2011construction)), the Schizotypal Personality Questionnaire – Brief Revised (SPQ-BR, [Davidson et al., 2016](#ref-davidson2016schizotypal)), and the McLean Screening Instrument for Borderline Personality Disorder (MSI-BPD, [Zanarini, 2003](#ref-zanarini2003zanarini)).

#### Personality.

Personality-related measures included the NEO Five-Factor Inventory, Neuroticism subscale (NEO-FFI, [Costa et al., 1992](#ref-costa1992personality)), the Mini International Personality Item Pool (MINI-IPIP6, [Sibley et al., 2011](#ref-sibley2011mini)), and the Big Five Inventory Short Form (BFI-S, [Rammstedt & John, 2007](#ref-rammstedt2007measuring)).

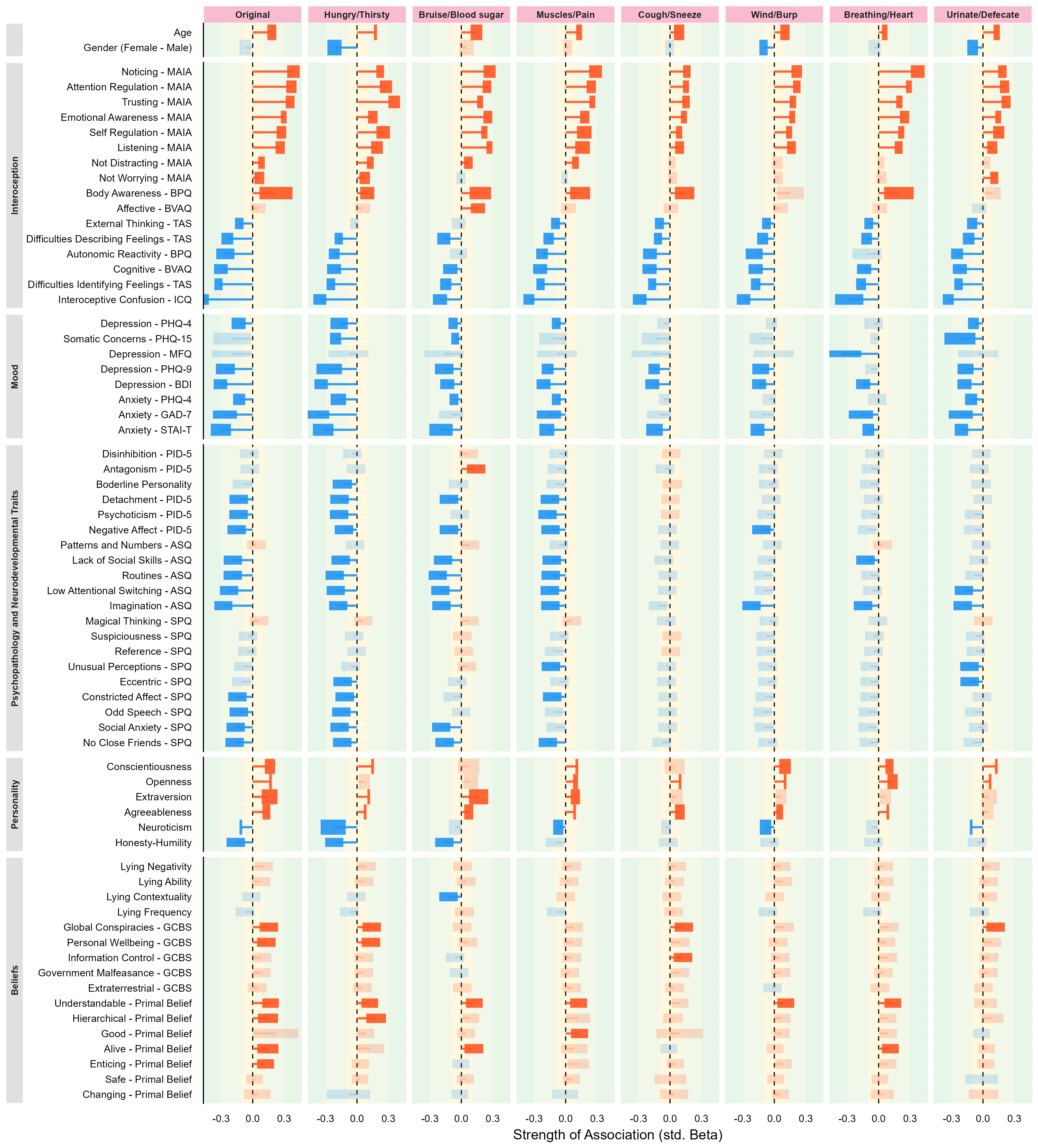
#### Beliefs and Misbeliefs.

Beliefs- and misbeliefs-related measures included the Lie Scale (LIE, [Makowski et al., 2023](#ref-makowski2023structure)), the Generic Conspiracist Beliefs Scale (GCBS, [Brotherton et al., 2013](#ref-brotherton2013measuring)), and the Primal World Beliefs Inventory (PI-99, [Clifton et al., 2019](#ref-clifton2019primal)) and its short form (PI-18, [Clifton & Yaden, 2021](#ref-clifton2021brief)). These instruments assess individual differences in conspiracist, world, and lying beliefs. This domain represents a novel avenue of research in the context of interoception, as such beliefs have not been previously assessed in relation to interoceptive processes.

### Data Analysis

Figure 4

Meta-analytic associations (95% CI of standardized coefficients) between the IAS (total score and seven item pairs) and dispositional measures across 17 open-access datasets. Positive associations are shown in red, negative associations in blue. Non-significant associations (CI crossing zero) are transparent. The IAS-21 showed positive associations with most MAIA facets and BPQ-Body Awareness, negative associations with interoceptive deficit measures (TAS, ICQ), negative associations with the BPQ-Autonomic Reactivity, moderate negative associations with mood measures, and weaker, more selective associations with psychopathology and belief measures. Certain item pairings, particularly Hungry–Thirsty and Bruise–Blood sugar, drove the strongest effects, highlighting domain-specific patterns in self-perceived interoceptive accuracy.



All variables from the 17 open-access datasets were extracted and standardized (z-scored). Scores were harmonized and combined where meaningful (e.g., scores from the short 18-item form of the PI and scores from its longer 99-item version). Missing data was excluded on a per-model basis, to maximize statistical power.

Akin to a meta-analytic framework, datasets were treated as clusters (random groups), and pooled estimates of the associations that account for both within- and between-sample variability were computed. This approach allowed the estimation of multiple associations while still accounting for study heterogeneity and intra-study uncertainty (i.e., measure uncertainty within the studies). Specifically, associations were estimated using mixed models implemented in the *glmmTMB* package ([McGillycuddy et al., 2025](#ref-glmTMB)), and postprocessed with the easystats R ecosystem ([Ben-Shachar et al., 2020](#ref-ben2020effectsize); [Lüdecke et al., 2020](#ref-ludecke2020extracting); [Makowski, Ben-Shachar, et al., 2025](#ref-makowski2025modelbased)). Each predictor was entered separately with the seven IAS item pairs identified in Study 1 (Hungry–Thirsty, Muscles–Pain, Wind–Burp, Urinate–Defecate, Breathing–Heart, Bruise–Blood, Cough–Sneeze) and the original IAS “total” score serving as outcomes. Models included dataset as a random intercept and the predictor as a random slope. Note that we focus on effect sizes and confidence intervals rather than p-values to allay concerns of multiple hypothesis testing.

From each fitted model, we extracted standardized regression coefficients (interpretable as a correlation index), and 95% confidence intervals (details on p-values, convergence status, and sample size are included in the Supplementary Materials). Results are summarized in [Figure 4](#fig-correlations) and we report the mean coefficient in the text when consistent across measures.

## Results

### Demographics

Age showed consistent positive standardized associations across IAS correlates (mean β = .13), including both the IAS total scores and the scores for the 7-pairs, with all effects significant. Gender displayed near-zero effects (mean β = −.06) and was only significant for the Hungry–Thirsty, Wind–Burp, and Urinate–Defecate pairs.

### Interoception and Interoception Related

MAIA subscales showed robust positive standardized associations with the IAS. Noticing had the strongest association (mean β = .26), followed by Attention Regulation (mean β = .25) with other subscales (Body Listening, Trusting, Emotional Awareness and Self-Regulation) in the .19–.24 range, all significant. Effects were generally strongest for the full IAS, except for Body Listening, which peaked for the Bruise–Blood pairing. Not Distracting was significant only for the full IAS, Hungry-Thirsty, Bruise-Blood sugar and the Muscles-Pain pairings, albeit weakly (β ≈ −.02 to .06). Not Worrying was significant only for the full IAS and the Hungry–Thirsty and Urinate–Defecate pairings (β ≈ −.01 to .04).

The BPQ’s Autonomic Reactivity subscale was reliably negatively associated across the seven item pairs (mean β = −.19), reaching significance for all pairs except Bruise–Blood sugar and Breathing–Heart. In contrast, the Body Awareness subscale of the BPQ was positive and significant across most pairings (mean β = .15; range = .10 to .22) except for Wind-Burp and Urinate-Defecate.

Interoceptive deficits showed consistent negative associations: Alexithymia measures (TAS, BVAQ) were robustly negative (β ≈ −.13 to −.2), except for the BVAQ Affective subscale, which was positive for Bruise–Blood (β = .16). Interoceptive confusion (ICQ) was also negative across all IAS measures (mean β = −.32; range = −.49 to −.20).

### Mood and Anxiety Related

Depression (mean β = −.20) and anxiety (mean β = −.20) showed consistent negative associations across the IAS and all its pairs, with the strongest effects observed for the Hungry–Thirsty dimension. However, looking at specific measures revealed some differences.

The MFQ was largely nonsignificant except for a negative association with Breathing–Heart, while somatic concerns (PHQ-15) were unrelated to the IAS total scale but had weak negative associations with the Hungry–Thirsty, Bruise–Blood sugar, and Urinate–Defecate pairs (from −.22 to −.05).

Other mood indices (PHQ-9, PHQ-4, GAD-7) produced generally weaker but often reliable negative coefficients (typically −.10 to −.15), though effects sometimes appeared only for specific pairings (e.g., PHQ-4 was significant for all pairs except Wind–Burp and Breathing–Heart). Across anxiety and depression measures, the Hungry–Thirsty pair consistently had the strongest standardized effects.

### Psychopathology and Neurodevelopmental Traits.

Maladaptive personality traits were generally negatively associated with the IAS scores. Detachment, Psychoticism, and Negative Affect demonstrated the most reliable, albeit weak, negative associations with IAS scores (from −.08 to −.10), again primarily driven by the Hungry–Thirsty, Bruise–Blood sugar, and Muscles–Pain pairings.

Other maladaptive traits were unrelated to IAS scores. Notably, Antagonism showed a significant positive association, limited to the Bruise–Blood sugar pair (β = .14), whereas Borderline Personality was uniquely negatively associated with Hungry–Thirsty (mean β = −.14).

A comparable trend was observed for autistic traits (ASQ), with the strongest effects driven by the Hungry–Thirsty, Bruise–Blood sugar, and Muscles–Pain pairs, which yielded the largest negative coefficients. Among ASQ subscales, Imagination (mean β = −.18) exhibited the most consistent pattern of negative associations across pairs except for Cough–Sneeze.

Schizotypal traits showed a similarly negative pattern of associations to IAS scores. Social Anxiety (mean β = −.11) and No Close Friends (mean β = −.12) displayed the most consistent effects. Although some facets were unrelated to the total IAS score, they were associated with specific item pairs: Unusual Perceptions with Muscles–Pain and Urinate–Defecate, and Eccentric with Hungry–Thirsty and Urinate–Defecate.

### Personality

Among the Big Five traits, Openness (mean β = .10), Agreeableness (mean β = .08), and Conscientiousness (mean β = .11) exhibited small yet consistent positive associations, most clearly for the total IAS score. Extraversion showed a weak but positive association, driven primarily by the Bruise–Blood pair (β = .17). In contrast, Honesty–Humility (mean β = −.10) had a small negative effect, most clearly for the Hungry–Thirsty and Bruise–Blood sugar pairs. Finally, Neuroticism (mean β = −.23) was negatively associated with the IAS, with the strongest effect observed for the Hungry–Thirsty pair (β = −.22).

### Beliefs and Misbeliefs

Within the Lying Profile, only Lying Contextuality showed a modest negative coefficient for Bruise–Blood (β = −.12). Global Conspiracies (GCBS) beliefs were associated with the total IAS and three pairings (mean β = .10): the Hungry–Thirsty, Cough-Sneeze and the Urinate-Defecate. Primal world beliefs were largely null, with small positive associations for appraising the world as being “Understandable”, “Alive”, and “Hierarchical”, with Hungry–Thirsty (β = .09). Appraising the world as “Good” was uniquely positively related with Muscles–Pain (β = .09).

Overall, the original IAS total score had the greatest number of significant standardized effects (28), though many were driven by specific item pairs rather than uniform scale-wide associations. Hungry–Thirsty (17) and Bruise–Blood sugar (10) consistently yielded the largest standardized coefficients, especially for interoception, mood, and psychopathology measures. Other item pairs (Cough–Sneeze, Wind–Burp, Urinate–Defecate) contributed little unique variance and rarely had meaningful standardized effects.

## Discussion

Using a meta-analytical approach with data across 17 datasets, Study 2 examined the dispositional correlates of the IAS and explored whether structural improvements of the 14-item IAS translated into clearer and more reliable associations compared to the original 21-item scale.

The IAS had positive associations with all MAIA facets except for the Not-Distracting and Not-Worrying subscales, replicating prior findings ([Brand et al., 2023](#ref-brand2023); [Todd et al., 2022](#ref-todd2022)). Similarly, associations between the IAS and BPQ Body Awareness subscale were positive, consistent with previous research ([Brand et al., 2023](#ref-brand2023); [Gaggero et al., 2021](#ref-gaggero2021); [Koike & Nomura, 2023](#ref-koike2023)). Indices of autonomic reactivity, however, were more strongly and negatively associated with IAS scores, supporting earlier observations that reactivity may reflect aspects of bodily signal processing captured by self-reported interoceptive accuracy ([Todd et al., 2022](#ref-todd2022)). In contrast, measures of interoceptive difficulties, including alexithymia and interoceptive confusion, were negatively associated with IAS scores, consistent with previous research ([Brand et al., 2023](#ref-brand2023); [Koike & Nomura, 2023](#ref-koike2023); [Lin et al., 2023](#ref-lin2023); [Murphy et al., 2019](#ref-murphy2019)).

Although associations between the IAS and the BPQ Body Awareness subscale replicated several previous findings, other studies have found non-existent (e.g., [Murphy et al., 2019](#ref-murphy2019)), negative ([Lin et al., 2023](#ref-lin2023)), and even quadratic relationships (e.g., [Campos et al., 2021](#ref-campos2021)). Mixed findings may stem from inconsistent interpretations of the items in the BPQ subscales, specifically, whether they capture attentional focus on bodily sensations or perceived accuracy in detecting them ([Campos et al., 2021](#ref-campos2021); e.g., [Gabriele et al., 2022](#ref-gabriele2022dissociations)). Thus, although we observed a positive association between the BPQ Body Awareness and the IAS, this result should be interpreted with caution, and future studies should assess participants’ understanding of BPQ items to clarify the nature of this relationship.

Across mood, most item pairs displayed small-to-moderate standardized coefficients, indicating shared variance between interoceptive self-reports and affective tendencies, aligning with previous finding ([Koike & Nomura, 2023](#ref-koike2023); [Murphy et al., 2019](#ref-murphy2019)). In contrast, associations with psychopathology were more selective: negative associations were primarily observed for visceral pairings such as Hungry–Thirsty and Bruise–Blood sugar, whereas other domains had weaker or nonsignificant effects. This pattern suggests that self-reported interoceptive accuracy for visceral sensations may be more strongly associated with affective and clinical traits compared to accuracy for somatic or respiratory reflexes.

Associations with personality traits followed a similar domain-specific structure, with effects for the full IAS largely reflected contributions from specific pairings. Extraversion was most strongly associated with Bruise-Blood sugar, Neuroticism and Honesty–Humility with Hungry–Thirsty, and Openness, Conscientiousness and Agreeableness with the full scale. These results indicate that associations between interoceptive self-report and personality dimensions may depend on the bodily domain represented. These results partially align with prior work reporting a negative association between Neuroticism and the IAS ([Brand et al., 2023](#ref-brand2023)). However, associations involving other personality traits are less well established and require replication to situate the present findings within the broader literature.

Associations with belief systems were uniformly small. Weak positive associations emerged between the IAS and certain primal world beliefs (e.g., perceiving the world as comprehensible or alive) and selected conspiracy belief dimensions, particularly Global Conspiracies and Personal Wellbeing. These effects were limited to specific item pairs, most often Hungry–Thirsty and Bruise–Blood sugar, indicating that links between interoceptive self-report and broader cognitive or worldview constructs are likely modest and domain-specific

In summary, the 14-item IAS performed similarly to the original 21-item version, showing comparable patterns of association with mood, psychopathology, and interoceptive measures. Although associations between the IAS and external constructs were typically captured by the original score, they were often not uniform across bodily domains and were instead driven by specific item pairs. In particular, the Hungry–Thirsty and Bruise–Blood sugar items pairs had the strongest and most consistent associations to interoceptive, affective, and personality measures. Other pairs contributed little unique information, particularly in measures of psychopathologies and beliefs. The homeostatic salience and universality of the Hungry-Thirsty pair, and the contextual relevance and visual/psychological prominence of the Bruise-Blood pair, might contribute to their stronger associations with dispositional traits. Overall, findings suggest that the refined 14-item scale captured domain-specific patterns, potentially offering a more efficient, nuanced, and interpretation-precise measure of self-reported interoceptive accuracy.

# General Discussion

Across two studies, we used a mega-analytic approach combining data from 17 datasets with over 33 thousand participants to provide the most comprehensive evaluation to date of the IAS including both its structure (Study 1) and dispositional correlates (Study 2). These studies offer new insights into the conceptual interpretation, and practical applications of the IAS, with suggestions for how researchers may use the IAS in future research.

Study 1 challenges interpretations of the IAS as a single latent construct (e.g., [Brand et al., 2023](#ref-brand2023); [Murphy et al., 2019](#ref-murphy2019)) and instead suggests that the IAS is best described as capturing multiple, interrelated but domain-specific, clusters of bodily sensations. Across samples, seven stable item pairs emerged (Hungry–Thirsty, Urinate–Defecate, Muscles–Pain, Breathing–Heart, Cough–Sneeze, Wind–Burp, and Bruise–Blood sugar), yielding a refined 14-item structure with superior fit and cross-dataset stability compared to previously proposed unidimensional or bifactor models.

Unlike Lin et al. ([2023](#ref-lin2023)), whose 12-item solution supported unidimensionality, Study 1 findings suggest that local dependencies between items reflect meaningful theoretical coupling among functionally related signals (e.g., Heart–Breathing). Such clustering likely mirrors the way bodily signals co-occur physiologically or phenomenologically (e.g., hunger with thirst, coughing with sneezing) and are therefore perceived and evaluated together. Perceived accuracy in one bodily system may generalize to conceptually or functionally linked sensations, yielding a cluster rather than continuous organization of interoceptive beliefs.

Consistent with this interpretation, self-reported interoceptive sensibility across related bodily axes (e.g., cardiac and respiratory) show positive associations ([Garfinkel et al., 2016](#ref-garfinkel2016interoceptive)), whereas objective measures of interoceptive accuracy across modalities typically show weak or absent correlations ([Bruni, 2023](#ref-bruni2023interoception); [Ferentzi et al., 2018](#ref-ferentzi2018interoceptive)). This dissociation indicates that objective interoceptive performance may be largely domain-specific whereas subjective interoceptive beliefs may generalize across functionally related systems. We therefore interpret the present clustering as reflecting the conceptual organization of interoceptive beliefs which are structured by perceived bodily interdependence, as opposed to a single general interoceptive capacity.

Study 2 extended these findings by examining the IAS’s associations with interoceptive, affective, personality, and belief-related traits. Consistent with previous research, the IAS correlated positively with adaptive aspects of interoceptive sensibility (e.g., MAIA subscales of reflecting noticing, attention regulation, and body trusting) and negatively with interoceptive difficulties, namely alexithymia and interoceptive confusion (e.g., [Brand et al., 2023](#ref-brand2023); [Gaggero et al., 2021](#ref-gaggero2021); [Garfinkel et al., 2015](#ref-garfinkel2015knowing)).

Also consistent with previous research, individuals reporting higher symptoms of depression and anxiety tended to perceive themselves as less accurate in detecting bodily signals (e.g., [Brand et al., 2023](#ref-brand2023); [Khalsa, Adolphs, Cameron, Critchley, Davenport, Feinstein, Feusner, Garfinkel, Lane, Mehling, Meuret, et al., 2018](#ref-khalsa2018); [Nord & Garfinkel, 2022](#ref-nord2022interoceptive)). Crucially, however, these associations were not uniform across bodily domains. The strongest, and most robust, effects involved drive-states (e.g., hunger, thirst), whereas regulatory systems, such as cardiorespiratory and expulsion domains showed weaker or inconsistent relationships to anxiety and depression metrics. Such domain specificity implies either that beliefs about gastrointestinal accuracy are more psychologically and clinically relevant than beliefs about reflexive or surface sensations, or that the experience of these signals is more salient and easier to accurately appraise, leading to meaningful variation across individuals.

Patterns of domain-specific links to anxiety and depression may explain inconsistencies in the literature linking interoception and emotion. Reviews and meta-analyses report that performance-based measures of interoceptive accuracy, usually cardioception, show weak or absent correlations with depression and anxiety ([Adams et al., 2022](#ref-adams2022association); [Banellis et al., 2025](#ref-banellis2025interoceptive); [Jenkinson et al., 2024](#ref-jenkinson2024interoception)) whereas self-reported interoceptive measures, containing items relevant to multiple bodily domains, typically show stronger, often negative, associations with internalizing traits (e.g., [Brand et al., 2022](#ref-brand2022); [Clemente et al., 2024](#ref-clemente2024relationship); [Lin et al., 2023](#ref-lin2023)).

This discrepancy likely reflects both conceptual distinctions between perceptual accuracy and metacognitive beliefs ([Garfinkel et al., 2015](#ref-garfinkel2015knowing); [Khalsa, Adolphs, Cameron, Critchley, Davenport, Feinstein, Feusner, Garfinkel, Lane, Mehling, Meuret, et al., 2018](#ref-khalsa2018); [Suksasilp & Garfinkel, 2022](#ref-suksasilp2022towards)) as well as variation in the bodily domains assessed. Study 2 results extend this idea by demonstrating that associations between interoceptive beliefs and affective traits are concentrated in psychologically salient visceral domains. Thus, the field’s heavy reliance on cardioceptive tasks may have obscured more meaningful relationships between affective traits and other interoceptive modalities.

Beyond mood, similar albeit weaker patterns emerged for autistic traits, schizotypy, and maladaptive personality dimensions. These traits tended to correlate negatively with subjective interoceptive accuracy, particularly for visceral and musculoskeletal domains but not other areas. This selective pattern suggests that, rather than reflecting a general interoceptive deficit, diminished beliefs in certain bodily domains may contribute to specific socio-emotional and personality features.. This idea fits well with previous research showing that interoceptive self-beliefs are meaningfully related to emotional awareness, social connectedness, and psychopathology ([Brand et al., 2022](#ref-brand2022); [Torregrossa et al., 2022](#ref-torregrossa2022interoceptive); [Williams et al., 2023](#ref-williams2023characterizing)). The present findings extend this literature by showing that such associations are likely domain-specific rather than global across all bodily systems.

Taken together, findings across Study 1 and 2 suggest that self-perceived interoceptive accuracy is a differentiated construct, varying systematically across distinct bodily systems. Beliefs about gastrointestinal accuracy, in particular, showed the most consistent associations with affective and dispositional features. Although modest in magnitude, these effects emphasize the value of considering domain-specificity when linking interoceptive beliefs to emotion and mental health. Clinically, the domain-specific structure of the IAS highlights that interoceptive disruptions in affective and psychopathological conditions may center primarily on beliefs about visceral sensations rather than global interoceptive deficits. This perspective suggests new routes for targeted assessment and intervention focusing on specific bodily domains.

Our results also suggest use of the refined 14-item IAS which balances parsimony and predictive validity, illustrating how network-informed scale refinement ([H. F. Golino & Epskamp, 2017](#ref-golino2017exploratory)) can enhance construct clarity. Methodologically, combining traditional psychometrics with network-based approaches was valuable with EGA complementing, and often surpassing, factor analysis by identifying meaningful inter-item dependencies to produce a more interpretable and stable structure.

Nevertheless, use of the 14-item IAS should be carefully considered. Although item pairs improved interpretability and outperformed the overall IAS in associations with affective and dispositional traits, dimensions with only two items (one per bodily domain) remain suboptimal. This is particularly true for constructs that are broad or conceptually ambiguous ([Allen et al., 2022](#ref-allen2022single)) such as beliefs about interoceptive accuracy across multiple bodily systems.

Importantly, because perceived accuracy in one domain can generalize to conceptually or functionally related sensations, two-item clusters may overemphasize these links whilst simultaneously failing to capture variability across unrelated domains. However, use of overall scores risks obscuring meaningful relationships between domain-specific items entirely, and rests on the assumption that interoceptive abilities generalize across bodily domains. Building on Desmedt et al. ([2022](#ref-desmedt2022measures))’ observation of a “jingle fallacy” in using self-reported interoceptive measures, the present results also suggest that researchers should be cautious in assuming self-reported accuracy generalises across domains and can be captured with a single construct.

Although our study focused exclusively on the IAS future research should re-evaluate the structure and dispositional correlates of other commonly used interoception questionnaires, such as the BPQ or MAIA, which have documented shortcomings ([Campos et al., 2021](#ref-campos2021); [Ferentzi et al., 2021](#ref-ferentzi2021examining); [Rogowska et al., 2023](#ref-rogowska2023validation); [Vig et al., 2023](#ref-vig2023self)) ideally with a mega-analytic approach.

Future research might also profitably incorporate objective measures of interoceptive accuracy across specific interoceptive domains (e.g., cardiac, respiratory, and gastric performance-based measures). This would enable an investigation of how beliefs about interoceptive accuracy relate to actual performance, clarifying whether domain-specific self-reports reflect true perceptual ability or instead primarily capture metacognitive and conceptual beliefs. Although previous research reports weak or absent correlations between subjective and objective measures within the same domain ([Arslanova et al., 2022](#ref-arslanova2022); e.g., cardiac accuracy vs subjective accuracy, [Brand et al., 2023](#ref-brand2023); [Petzke et al., 2024](#ref-petzke2024somatic)), examining additional modalities, such as gastric perception, could reveal whether these observations generalize to other interoceptive systems.

In summary, our research demonstrates that the IAS is a solid measure of self-reported interoceptive accuracy with typical patterns of associations with external measures. However, it also challenges its optimal structure and scoring approach, instead suggesting that a refined 14-item version with seven item pairs might better capture domain-specific interoceptive beliefs and how they relate to dispositional traits. More broadly, our results emphasize the importance and potential value of conceptualizing interoceptive processes as domain-specific rather than unidimensional and highlight the need for targeted scale refinement as well as the development of novel, context-sensitive, and multi-dimensional, interoception measures.

# References

Adams, K. L., Edwards, A., Peart, C., Ellett, L., Mendes, I., Bird, G., & Murphy, J. (2022). The association between anxiety and cardiac interoceptive accuracy: A systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews*, *140*, 104754.

Allen, M. S., Iliescu, D., & Greiff, S. (2022). Single item measures in psychological science. In *European journal of psychological assessment*. Hogrefe Publishing.

Angold, A., Costello, E. J., Messer, S. C., & Pickles, A. (1995). Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *International Journal of Methods in Psychiatric Research*.

Arslanova, I., Galvez-Pol, A., Kilner, J., Finotti, G., & Tsakiris, M. (2022). Seeing through each other’s hearts: Inferring others’ heart rate as a function of own heart rate perception and perceived social intelligence. *Affective Science*, *3*(4), 862–877.

Baer, R. A., Smith, G. T., Hopkins, J., Krietemeyer, J., & Toney, L. (2006). Using self-report assessment methods to explore facets of mindfulness. *Assessment*, *13*(1), 27–45.

Bagby, R. M., Parker, J. D., & Taylor, G. J. (1994). The twenty-item toronto alexithymia scale—i. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research*, *38*(1), 23–32.

Banellis, L., Nikolova, N., Fischer Ehmsen, J., Courtin, A. S., Vejlo, M., Tyrer, A., Bohme, R., Bavato, F., Hoogervorst, K., Fardo, F., et al. (2025). Interoceptive ability is unrelated to mental health symptoms: Evidence from a large scale multi-domain psychophysical investigation. *medRxiv*, 2025–2008.

Beck, A. T., Steer, R. A., & Brown, G. (1996). Beck depression inventory–II. *Psychological Assessment*.

Ben-Shachar, M. S., Lüdecke, D., & Makowski, D. (2020). Effectsize: Estimation of effect size indices and standardized parameters. *Journal of Open Source Software*, *5*(56), 2815.

Brand, S., Meis, A. C., Tünte, M. R., Murphy, J., Woller, J. P., Jungmann, S. M., Witthöft, M., Hoehl, S., Weymar, M., Hermann, C., & Ventura-Bort, C. (2023). A multi-site German validation of the Interoceptive Accuracy Scale and its relation to psychopathological symptom burden. *Communications Psychology*, *1*(1). <https://doi.org/10.1038/s44271-023-00016-x>

Brand, S., Petzke, T. M., & Witthöft, M. (2022). The differential relationship between self-reported interoceptive accuracy and attention with psychopathology. *Zeitschrift für Klinische Psychologie Und Psychotherapie*.

Brewer, R., Cook, R., & Bird, G. (2016). Alexithymia: A general deficit of interoception. *Royal Society Open Science*, *3*(10), 150664.

Brotherton, R., French, C. C., & Pickering, A. D. (2013). Measuring belief in conspiracy theories: The generic conspiracist beliefs scale. *Frontiers in Psychology*, *4*, 279.

Bruni, P. (2023). Interoception across domains: Assessing the role of age. *Innovation in Aging*, *7*, 755.

Cabrera, A., Kolacz, J., Pailhez, G., Bulbena-Cabre, A., Bulbena, A., & Porges, S. W. (2018). Assessing body awareness and autonomic reactivity: Factor structure and psychometric properties of the body perception questionnaire-short form (BPQ-SF). *International Journal of Methods in Psychiatric Research*, *27*(2), e1596.

Campos, C., Rocha, N. B., & Barbosa, F. (2021). *Untangling self-reported interoceptive attention and accuracy: Evidence from the european portuguese validation of the body perception questionnaire and the interoceptive accuracy scale*. <http://dx.doi.org/10.31234/osf.io/a7wdj>

Christensen, A. P., Garrido, L. E., & Golino, H. (2023). Unique variable analysis: A network psychometrics method to detect local dependence. *Multivariate Behavioral Research*, *58*(6), 1165–1182.

Christensen, A. P., & Golino, H. (2021a). Estimating the stability of psychological dimensions via bootstrap exploratory graph analysis: A monte carlo simulation and tutorial. *Psych*, *3*(3), 479–500.

Christensen, A. P., & Golino, H. (2021b). On the equivalency of factor and network loadings. *Behavior Research Methods*, *53*(4), 1563–1580.

Clemente, R., Murphy, A., & Murphy, J. (2024). The relationship between self-reported interoception and anxiety: A systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews*, *167*, 105923.

Clifton, J. D., Baker, J. D., Park, C. L., Yaden, D. B., Clifton, A. B., Terni, P., Miller, J. L., Zeng, G., Giorgi, S., Schwartz, H. A., et al. (2019). Primal world beliefs. *Psychological Assessment*, *31*(1), 82.

Clifton, J. D., & Yaden, D. B. (2021). Brief measures of the four highest-order primal world beliefs. *Psychological Assessment*, *33*(12), 1267.

Cosemans, T., Rosseel, Y., & Gelper, S. (2022). Exploratory graph analysis for factor retention: Simulation results for continuous and binary data. *Educational and Psychological Measurement*, *82*(5), 880–910.

Costa, P., McCrae, R. R., & Revised, N. (1992). Personality inventory (NEO-PI-r) and NEO five-factor inventory (NEO-FFI): Professional manual. *Psychological Assessment Resources, Odessa, FL*.

Davidson, C. A., Hoffman, L., & Spaulding, W. D. (2016). Schizotypal personality questionnaire–brief revised (updated): An update of norms, factor structure, and item content in a large non-clinical young adult sample. *Psychiatry Research*, *238*, 345–355.

Desmedt, O., Heeren, A., Corneille, O., & Luminet, O. (2022). What do measures of self-report interoception measure? Insights from a systematic review, latent factor analysis, and network approach. *Biological Psychology*, *169*, 108289.

Desmedt, O., Luminet, O., Maurage, P., & Corneille, O. (2025). Discrepancies in the definition and measurement of human interoception: A comprehensive discussion and suggested ways forward. *Perspectives on Psychological Science*, *20*(1), 76–98.

Desmedt, O., Luminet, O., Walentynowicz, M., & Corneille, O. (2023). The new measures of interoceptive accuracy: A systematic review and assessment. *Neuroscience & Biobehavioral Reviews*, *153*, 105388. https://doi.org/<https://doi.org/10.1016/j.neubiorev.2023.105388>

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

Ferentzi, E., Horváth, Á., & Köteles, F. (2019). Do body-related sensations make feel us better? Subjective well-being is associated only with the subjective aspect of interoception. *Psychophysiology*, *56*(4), e13319.

Ferentzi, E., Olaru, G., Geiger, M., Vig, L., Köteles, F., & Wilhelm, O. (2021). Examining the factor structure and validity of the multidimensional assessment of interoceptive awareness. *Journal of Personality Assessment*, *103*(5), 675–684.

Gabriele, E., Spooner, R., Brewer, R., & Murphy, J. (2022). Dissociations between self-reported interoceptive accuracy and attention: Evidence from the interoceptive attention scale. *Biological Psychology*, *168*, 108243.

Gaggero, G., Bizzego, A., Dellantonio, S., Pastore, L., Lim, M., & Esposito, G. (2021). Clarifying the relationship between alexithymia and subjective interoception. *PLoS One*, *16*(12), e0261126.

Garfinkel, S. N., Manassei, M. F., Hamilton-Fletcher, G., In den Bosch, Y., Critchley, H. D., & Engels, M. (2016). Interoceptive dimensions across cardiac and respiratory axes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1708), 20160014.

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

Golino, H. F., & Epskamp, S. (2017). Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PloS One*, *12*(6), e0174035.

Golino, H., Shi, D., Christensen, A. P., Garrido, L. E., Nieto, M. D., Sadana, R., Thiyagarajan, J. A., & Martinez-Molina, A. (2020). Investigating the performance of exploratory graph analysis and traditional techniques to identify the number of latent factors: A simulation and tutorial. *Psychological Methods*, *25*(3), 292.

Harrison, O. K., & Pink, A. (2024). Interoception and physical health. In *Interoception: A comprehensive guide* (pp. 227–264). Springer.

Hickman, L., Seyedsalehi, A., Cook, J. L., Bird, G., & Murphy, J. (2020). The relationship between heartbeat counting and heartbeat discrimination: A meta-analysis. *Biological Psychology*, *156*, 107949.

Hoekstra, R. A., Vinkhuyzen, A. A., Wheelwright, S., Bartels, M., Boomsma, D. I., Baron-Cohen, S., Posthuma, D., & Van Der Sluis, S. (2011). The construction and validation of an abridged version of the autism-spectrum quotient (AQ-short). *Journal of Autism and Developmental Disorders*, *41*, 589–596.

Jahedi, S., & Méndez, F. (2014). On the advantages and disadvantages of subjective measures. *Journal of Economic Behavior & Organization*, *98*, 97–114. <https://doi.org/10.1016/j.jebo.2013.12.016>

Jenkinson, P. M., Fotopoulou, A., Ibañez, A., & Rossell, S. (2024). Interoception in anxiety, depression, and psychosis: A review. *EClinicalMedicine*, *73*.

Jiménez, M., Abad, F. J., Garcia-Garzon, E., Golino, H., Christensen, A. P., & Garrido, L. E. (2023). Dimensionality assessment in bifactor structures with multiple general factors: A network psychometrics approach. *Psychological Methods*.

Joshi, S. A., Aupperle, R. L., & Khalsa, S. S. (2023). Interoception in fear learning and posttraumatic stress disorder. *Focus*, *21*(3), 266–277.

Khalsa, S. S., Adolphs, R., Cameron, O. G., Critchley, H. D., Davenport, P. W., Feinstein, J. S., Feusner, J. D., Garfinkel, S. N., Lane, R. D., Mehling, W. E., et al. (2018). Interoception and mental health: A roadmap. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *3*(6), 501–513.

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

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

Koike, H., & Nomura, M. (2023). Development and validation of japanese versions of the interoceptive accuracy scale and interoceptive attention scale. *SAGE Open*, *13*(4), 21582440231214639.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*(9), 606–613.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2002). The PHQ-15: Validity of a new measure for evaluating the severity of somatic symptoms. *Psychosomatic Medicine*, *64*(2), 258–266.

Kroenke, K., Spitzer, R. L., Williams, J. B., & Löwe, B. (2009). An ultra-brief screening scale for anxiety and depression: The PHQ–4. *Psychosomatics*, *50*(6), 613–621.

Kroenke, K., Spitzer, R. L., Williams, J. B., Monahan, P. O., & Löwe, B. (2007). Anxiety disorders in primary care: Prevalence, impairment, comorbidity, and detection. *Annals of Internal Medicine*, *146*(5), 317–325.

Lin, X.-X., Shen, H.-R., Lin, J.-X., Zhang, Y.-H., Murphy, J., Wang, Y.-Z., Sun, Y.-B., Wang, N., Wang, J.-Y., Wei, G.-X., & Luo, F. (2023). Psychometric validation and refinement of the Chinese Interoceptive Accuracy Scale (IAS) in general population and patients with chronic pain. *Journal of Psychosomatic Research*, *175*, 111541. <https://doi.org/10.1016/j.jpsychores.2023.111541>

Lischke, A., Pahnke, R., Mau-Moeller, A., & Weippert, M. (2021). Heart rate variability modulates interoceptive accuracy. *Frontiers in Neuroscience*, *14*, 612445.

Longarzo, M., D’Olimpio, F., Chiavazzo, A., Santangelo, G., Trojano, L., & Grossi, D. (2015). Self-awareness questionnaire. *Frontiers in Psychology*.

Lüdecke, D., Ben-Shachar, M. S., Patil, I., & Makowski, D. (2020). Extracting, computing and exploring the parameters of statistical models using r. *Journal of Open Source Software*, *5*(53), 2445.

Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). Performance: An r package for assessment, comparison and testing of statistical models. *Journal of Open Source Software*, *6*(60).

Makowski, D., Ben-Shachar, M. S., Wiernik, B. M., Patil, I., Thériault, R., & Lüdecke, D. (2025). Modelbased: An r package to make the most out of your statistical models through marginal means, marginal effects, and model predictions. *Journal of Open Source Software*, *10*(109), 7969.

Makowski, D., Neves, A., Benn, E., Bennett, M., & Poerio, G. (2025). *The mint scale: A fresh validation of the multimodal interoception questionnaire and comparison to the MAIA, BPQ and IAS*.

Makowski, D., Pham, T., Lau, Z. J., Raine, A., & Chen, S. A. (2023). The structure of deception: Validation of the lying profile questionnaire. *Current Psychology*, *42*(5), 4001–4016.

McGillycuddy, M., Warton, D. I., Popovic, G., & Bolker, B. M. (2025). Parsimoniously fitting large multivariate random effects in glmmTMB. *Journal of Statistical Software*, *112*(1), 1–19. <https://doi.org/10.18637/jss.v112.i01>

Mehling, W. E., Acree, M., Stewart, A., Silas, J., & Jones, A. (2018). The multidimensional assessment of interoceptive awareness, version 2 (MAIA-2). *PloS One*, *13*(12), e0208034.

Mehling, W. E., Price, C., Daubenmier, J. J., Acree, M., Bartmess, E., & Stewart, A. (2012). The Multidimensional Assessment of Interoceptive Awareness (MAIA). *PLoS ONE*, *7*(11), e48230. <https://doi.org/10.1371/journal.pone.0048230>

Miller, L. C., Murphy, R., & Buss, A. H. (1981). Consciousness of body: Private and public. *Journal of Personality and Social Psychology*, *41*(2), 397.

Murphy, J. (2024). Interoception: Where do we go from here? *Quarterly Journal of Experimental Psychology*, *77*(2), 223–229.

Murphy, J., Brewer, R., Plans, D., Khalsa, S. S., Catmur, C., & Bird, G. (2019). Testing the independence of self-reported interoceptive accuracy and attention. *Quarterly Journal of Experimental Psychology*, *73*(1), 115–133. <https://doi.org/10.1177/1747021819879826>

Murtagh, F., & Legendre, P. (2014). Ward’s hierarchical agglomerative clustering method: Which algorithms implement ward’s criterion? *Journal of Classification*, *31*(3), 274–295.

Nord, C. L., & Garfinkel, S. N. (2022). Interoceptive pathways to understand and treat mental health conditions. *Trends in Cognitive Sciences*, *26*(6), 499–513.

Nord, C. L., Lawson, R. P., & Dalgleish, T. (2021). Disrupted dorsal mid-insula activation during interoception across psychiatric disorders. *American Journal of Psychiatry*, *178*(8), 761–770.

Petzke, T. M., Köteles, F., Pohl, A., & Witthöft, M. (2024). Somatic symptom distress is not related to cardioceptive accuracy. *Journal of Psychosomatic Research*, *181*, 111655.

Poerio, G. L., Klabunde, M., Bird, G., & Murphy, J. (2024). Interoceptive attention and mood in daily life: An experience-sampling study. *Philosophical Transactions of the Royal Society B*, *379*(1908), 20230256.

Pollatos, O., Mönkemöller, K., Groppe, K., & Elsner, B. (2023). Interoceptive accuracy is associated with benefits in decision making in children. *Frontiers in Psychology*, *13*, 1070037.

Porges, S. (1993). Body perception questionnaire. *Laboratory of Developmental Assessment, University of Maryland*, *10*, s15327752jpa5304\_1.

Raimo, S., Boccia, M., Di Vita, A., Cropano, M., Guariglia, C., Grossi, D., & Palermo, L. (2021). The body across adulthood: On the relation between interoception and body representations. *Frontiers in Neuroscience*, *15*, 586684.

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the big five inventory in english and german. *Journal of Research in Personality*, *41*(1), 203–212.

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, *21*(2), 137.

Rogowska, A. M., Tataruch, R., & Klimowska, K. (2023). Validation of the shortened 24-item multidimensional assessment of interoceptive awareness, version 2 (brief MAIA-2). *Scientific Reports*, *13*(1), 21270.

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

Shields, S. A., Mallory, M. E., & Simon, A. (1989). The body awareness questionnaire: Reliability and validity. *Journal of Personality Assessment*, *53*(4), 802–815.

Sibley, C. G., Luyten, N., Purnomo, M., Mobberley, A., Wootton, L. W., Hammond, M. D., Sengupta, N., Perry, R., West-Newman, T., Wilson, M. S., et al. (2011). The mini-IPIP6: Validation and extension of a short measure of the big-six factors of personality in new zealand. *New Zealand Journal of Psychology*, *40*(3).

Sorokowska, A., Saluja, S., Sorokowski, P., Frąckowiak, T., Karwowski, M., Aavik, T., Akello, G., Alm, C., Amjad, N., Anjum, A., et al. (2021). Affective interpersonal touch in close relationships: A cross-cultural perspective. *Personality and Social Psychology Bulletin*, *47*(12), 1705–1721.

Spielberger, C. D. (1970). Manual for the state-trait anxiety inventory (self-evaluation questionnaire). *(No Title)*.

Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine*, *166*(10), 1092–1097.

Suksasilp, C., & Garfinkel, S. N. (2022). Towards a comprehensive assessment of interoception in a multi-dimensional framework. *Biological Psychology*, *168*, 108262.

Terasawa, Y., Motomura, K., Natsume, A., Iijima, K., Chalise, L., Sugiura, J., Yamamoto, H., Koyama, K., Wakabayashi, T., & Umeda, S. (2021). Effects of insular resection on interactions between cardiac interoception and emotion recognition. *Cortex*, *137*, 271–281.

Thimm, J. C., Jordan, S., & Bach, B. (2016). The personality inventory for DSM-5 short form (PID-5-SF): Psychometric properties and association with big five traits and pathological beliefs in a norwegian population. *BMC Psychology*, *4*, 1–11.

Todd, J., Swami, V., Aspell, J. E., Furnham, A., Horne, G., & Stieger, S. (2022). Are some interoceptive sensibility components more central than others? Using item pool visualisation to understand the psychometric representation of interoception. *Plos One*, *17*(12), e0277894.

Torregrossa, L. J., Amedy, A., Roig, J., Prada, A., & Park, S. (2022). Interoceptive functioning in schizophrenia and schizotypy. *Schizophrenia Research*, *239*, 151–159.

Vig, L., Ferentzi, E., & Köteles, F. (2023). Self-reported interoception, worries and protective behaviors during the COVID-19 pandemic: A longitudinal study. *Psicologia: Reflexão e Crı́tica*, *36*(1), 23.

Vig, L., Köteles, F., & Ferentzi, E. (2022). Questionnaires of interoception do not assess the same construct. *Plos One*, *17*(8), e0273299.

Vlemincx, E., Walentynowicz, M., Zamariola, G., Van Oudenhove, L., & Luminet, O. (2023). A novel self-report scale of interoception: The three-domain interoceptive sensations questionnaire (THISQ). *Psychology & Health*, *38*(9), 1234–1253.

Von Mohr, M., Silva, P. C., Vagnoni, E., Bracher, A., Bertoni, T., Serino, A., Banissy, M. J., Jenkinson, P. M., & Fotopoulou, A. (2023). My social comfort zone: Attachment anxiety shapes peripersonal and interpersonal space. *Iscience*, *26*(2).

Vorst, H. C., & Bermond, B. (2001). Validity and reliability of the bermond–vorst alexithymia questionnaire. *Personality and Individual Differences*, *30*(3), 413–434.

Wareing, L., Readman, M. R., Longo, M. R., Linkenauger, S. A., & Crawford, T. J. (2024). The utility of heartrate and heartrate variability biofeedback for the improvement of interoception across behavioural, physiological and neural outcome measures: A systematic review. *Brain Sciences*, *14*(6), 579.

Williams, Z. J., Suzman, E., Bordman, S. L., Markfeld, J. E., Kaiser, S. M., Dunham, K. A., Zoltowski, A. R., Failla, M. D., Cascio, C. J., & Woynaroski, T. G. (2023). Characterizing interoceptive differences in autism: A systematic review and meta-analysis of case–control studies. *Journal of Autism and Developmental Disorders*, *53*(3), 947–962.

Zamariola, G., Frost, N., Van Oost, A., Corneille, O., & Luminet, O. (2019). Relationship between interoception and emotion regulation: New evidence from mixed methods. *Journal of Affective Disorders*, *246*, 480–485.

Zanarini, M. C. (2003). Zanarini rating scale for borderline personality disorder (ZAN-BPD): A continuous measure of DSM-IV borderline psychopathology. *Journal of Personality Disorders*, *17*(3), 233–242.

Zsido, A. N., Teleki, S. A., Csokasi, K., Rozsa, S., & Bandi, S. A. (2020). Development of the short version of the spielberger state—trait anxiety inventory. *Psychiatry Research*, *291*, 113223.