

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

Ana Neves<sup>1</sup>, Magdalena Pfaff<sup>1,2</sup>, Robyn Scharte<sup>1</sup>, Raquel Nogueira-Arjona<sup>1,3</sup>, Giulia L. Poerio<sup>1,4</sup>, and Dominique Makowski<sup>1,4</sup>

<sup>1</sup>School of Psychology, University of Sussex

<sup>2</sup>Brighton and Sussex Medical School, University of Sussex

<sup>3</sup>Sussex Addictions Research and Intervention Centre, University of Sussex

<sup>4</sup>Sussex Centre for Consciousness Science, University of Sussex

Recent reviews demonstrate that historic questionnaires designed to assess interoceptive abilities show low convergence and likely measure distinct constructs. Motivated by the theoretical distinction between interoceptive accuracy versus attention, the Interoceptive Accuracy Scale (IAS) is a recently developed, promising, and increasingly popular measure, but its psychometric properties and dimensional structure remain underexplored. Across two studies, we used a mega-analytic approach combining data from 17 datasets ( $N > 33,000$ ) to provide the most comprehensive evaluation to date of the IAS, including both its structure (Study 1) and dispositional correlates (Study 2). Study 1 triangulated the IAS structure across multiple dimensionality-exploration algorithms, identifying a stable core of 14-items, composed of 7 correlated but distinct pairs of items tied to specific bodily sensations (e.g., Hunger–Thirst, Urinate–Defecate, Cough–Sneeze). Study 2 examined associations between the IAS and measures of mood, anxiety, psychopathological and neurodevelopmental traits, personality, and belief-related constructs. Results revealed a nuanced pattern of associations, broadly reflecting positive links with adaptive affective and personality characteristics, and negative links with maladaptive traits. The 14-item IAS performed comparably to the original 21-item version, but item-pairing improved interpretability and highlighted domain-specific patterns. In particular, the Hungry–Thirsty and Bruise–Blood–sugar pairs displayed the strongest and most consistent associations with relevant outcomes. Taken together, these findings provide support for the idea that self-perceived interoceptive abilities is a differentiated construct, varying across bodily systems.

*Keywords:* Interoception, interoceptive accuracy, interoceptive sensibility, measurement, exploratory factor analysis, network analysis

## Introduction

1 Interception refers to the process of sensing, interpret-  
2 ing and integrating information of internal bodily stimuli  
3 by the nervous system, including both internal organs (e.g.,  
4 heart, lungs, gut) but also more broadly physiological tis-  
5 sues, providing information about the body’s current state  
6 (Khalsa, Adolphs, Cameron, Critchley, Davenport, Fein-  
7 stein, Feusner, Garfinkel, Lane, Mehling, & others, 2018).  
8 Although research emphasizes interoception’s fundamental  
9 role in numerous processes, such as emotion recognition  
10 (Terasawa et al., 2021) and regulation (Zamariola et al.,  
11 2019), decision making (Pollatos et al., 2023), learning  
12 (Joshi et al., 2023), body-ownership (Raimo et al., 2021)  
13 and outcomes including physical (Harrison & Pink, 2024)  
14 and psychological health (Nord et al., 2021) along with well-  
15 being (Ferentzi et al., 2019), the field remains hampered by  
16 concerns regarding both the conceptualization and measure-

18 ment of interoception (Desmedt et al., 2022, 2025; Murphy,  
19 2024).

## The Interoceptive Assessment Puzzle

21 Given the central nature of interoceptive processing to  
22 brain function, many measures could be seen as “indirectly”  
23 assessing (whether intentionally or not) interoceptive pro-  
24 cesses. This includes indices that quantify the coupling be-  
25 tween physiological signals and cognitive or behavioural pro-  
26 cesses (e.g., Heart Rate Variability), or constructs that the-  
27 oretically depend on interoceptive information (e.g., alex-  
28 ithymia). Critically, various *direct* measures of interocep-  
29 tion have been developed (see Figure 1), forming a combina-  
30 tion of “objective” and “subjective” assessments (i.e., phys-  
31 iological tasks such as the heart beat counting or tracking  
32 vs. questionnaires and subjective scales involving metacog-  
33 nitive judgments), “explicit” and “implicit” paradigms (i.e.,

34 directing participants' awareness and attention to interocep-  
 35 tive processes vs. measuring interoception covertly), vari-  
 36 ous interoceptive modalities (e.g., cardioception, respirocep-  
 37 tion, gastroception) and theoretical dimensions (e.g., accu-  
 38 racy, sensitivity, awareness). The diversity in measurement  
 39 approaches illustrates that rather than there being a "gold-  
 40 standard" measure of interoception in the field, each mea-  
 41 sure likely has strengths and limitations depending on the  
 42 research context and dimension of interoception being as-  
 43 sessed (Desmedt et al., 2023; Jahedi & Méndez, 2014).

#### 44 Self-reported Interoception



45 Although using self-report questionnaires to measure  
 46 deeply embodied functions might seem paradoxical, recent  
 47 redefinitions of interoception emphasize the importance of  
 48 high-level and metacognitive elaboration of interoceptive in-  
 49 formation, and support the idea that some facets of interocep-  
 50 tion, in particular participants' metacognitive beliefs, can be  
 51 assessed subjectively (Khalsa, Adolphs, Cameron, Critchley,  
 52 Davenport, Feinstein, Feusner, Garfinkel, Lane, Mehling, &  
 53 others, 2018; Suksasilp & Garfinkel, 2022). The notion that  
 54 self-reported interoception reflects a fundamentally different  
 55 aspect of conscious interoceptive processing compared to ob-  
 56 jective task measures is thought to be central to understand-  
 57 ing the reported lack of convergence between objective and  
 58 subjective measures (Desmedt et al., 2022). For instance,  
 59 there are typically weak or no correlation between question-

60 naires and objective measures (HCT, Schandry, 1981) and  
 61 the Heartbeat Detection Task (HDT, Kleckner et al., 2015),  
 62 even when both measures are intended to capture the same  
 63 theoretical interoceptive dimension (Arslanova et al., 2022;  
 64 Brand et al., 2023; e.g., task-based accuracy vs. self-reported  
 65 accuracy, Murphy et al., 2019). Perhaps more surprisingly,  
 66 there is also a reported lack of convergence amongst objec-  
 67 tive or subjective measures purporting to measure the same  
 68 interoceptive dimension. For instance, task-based measures  
 69 targeting accuracy typically show no or weak correlation  
 70 amongst themselves (respectively, Brand et al., 2023; Hick-  
 71 man et al., 2020), and subjective scores of interoceptive ac-  
 72 curacy also yield weaker than expected relationships (Vig  
 73 et al., 2022). As well as posing substantial validity con-  
 74 cerns, these findings raise the possibility that even within  
 75 broad measurement domain, different measures are target-  
 76 ing fundamentally different facets of interoception, under-  
 77 mining replicability and hindering theoretical and empirical  
 78 progress in the field.

79 One striking example concerns the assessment of in-  
 80 teroceptive sensibility, which is defined as the self-  
 81 reported tendency to focus on (with or without cor-  
 82 rectly detecting) internal sensations (Garfinkel et al., 2015;  
 83 Khalsa, Adolphs, Cameron, Critchley, Davenport, Feinstein,  
 84 Feusner, Garfinkel, Lane, Mehling, & others, 2018). A  
 85 recent systematic review suggested that various question-  
 86 naires designed to assess interoceptive sensibility may, in  
 87 fact, measure distinct constructs, with the risk of researchers  
 88 treating them as equivalent despite overall low convergence  
 89 (Desmedt et al., 2022). Several widely used questionnaires  
 90 were included in that review, such as the Body Perception  
 91 Questionnaire (BPQ, Porges, 1993), the Multidimensional  
 92 Assessment of Interoceptive Awareness (MAIA, Mehling et al.,  
 93 2012; MAIA-2, Mehling et al., 2018), the Body Aware-  
 94 ness Questionnaire (BAQ, Shields et al., 1989), the Private  
 95 subscale of the Body Consciousness Questionnaire (PBCS,  
 96 Miller et al., 1981), and the Self-Awareness Questionnaire  
 97 (SAQ, Longarzo et al., 2015) which comprised 86% of all  
 98 citations. The low to moderate correlations among ques-  
 99 tionnaires measures suggests that rather than measuring a  
 100 common construct they likely capture different aspects of  
 101 interoception, emphasizing the need for greater conceptual  
 102 clarity regarding what each measure captures, how they re-  
 103 late to different dimensions of interoception, and their poten-  
 104 tial overlaps with other constructs, such as alexithymia and  
 105 body awareness. One avenue for addressing these issues is  
 106 the careful investigation of newer measures that were not in-  
 107 cluded in prior reviews.

#### 108 The Interoceptive Accuracy Scale (IAS)

109 Focusing on the self-perceived *accuracy* dimension of  
 110 interoceptive sensibility, a recently developed scale with a  
 111 rapidly growing popularity is the Interoceptive Accuracy

 Ana Neves  
 Magdalena Pfaff  
 Robyn Scharte  
 Raquel Nogueira-Arjona  
 Giulia L. Poerio  
 Dominique Makowski

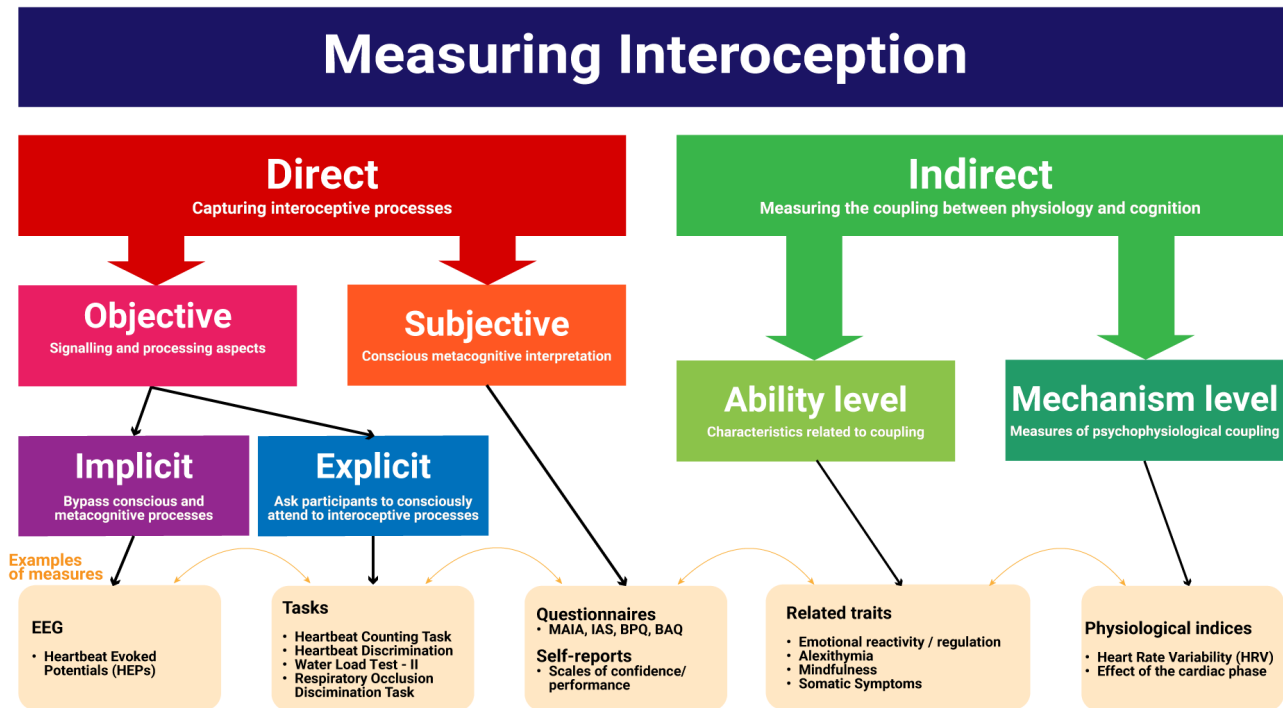


Author roles were classified using the Contributor Role Taxonomy (CRediT; <https://credit.niso.org/>) as follows: Ana Neves: Data curation, Formal Analysis, Investigation, Visualization, Writing – original draft, Writing – review & editing; Magdalena Pfaff: Data curation, Writing – original draft; Robyn Scharte: Writing – review & editing; Raquel Nogueira-Arjona: Writing – review & editing; Giulia L. Poerio: Writing – review & editing; Dominique Makowski: Project administration, Data curation, Formal Analysis, Investigation, Visualization, Writing – original draft, Writing – review & editing

Correspondence concerning this article should be addressed to Dominique Makowski, Email: [D.Makowski@sussex.ac.uk](mailto:D.Makowski@sussex.ac.uk)

**Figure 1**

*Schematic Representation of 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, or dependent on, interoceptive processes (e.g., physiological indices such as heart rate variability or effect of cardiac phase) or abilities (e.g., self-report measures of emotion regulation or identification abilities).*



112 Scale (IAS, Murphy et al., 2019). The authors devel-  
 113 oped the IAS to clarify not only how interoception is meas-  
 114 ured, but also what is being measured. Whereas many  
 115 existing interoceptive models (e.g., Garfinkel et al., 2015;  
 116 Khalsa, Adolphs, Cameron, Critchley, Davenport, Feinstein,  
 117 Feusner, Garfinkel, Lane, Mehling, & others, 2018) conflate  
 118 interoceptive accuracy and attention, Murphy and colleagues  
 119 conceptualise these as distinct dimensions. As a result, indi-  
 120 viduals may report attending to internal bodily signals while  
 121 simultaneously recognising that their perception of these sig-  
 122 nals is inaccurate. This distinction represents a conceptual  
 123 advance in the measurement of interoceptive sensibility and  
 124 positions the IAS as the only tool specifically designed to  
 125 assess subjective interoceptive accuracy.

126 The IAS comprises 21 Likert-scale items that ask how ac-  
 127 curately an individual can perceive different bodily signals,  
 128 with one item per physiological modality such as respira-  
 129 tion (“I can always accurately perceive when I am breath-  
 130 ing fast”), heart (“I can always accurately perceive when

131 my heart is beating fast”), skin (“I can always accurately  
 132 perceive when something is going to be ticklish”), arousal  
 133 or bodily functions like coughing (“I can always accurately  
 134 perceive when I am going to cough”) or urinating (“I can al-  
 135 ways accurately perceive when I need to urinate”). Appeal-  
 136 ingly, the IAS’ items are about specific interoceptive behav-  
 137 iors and “objective” experiences (albeit subjective reports of  
 138 them), a distinct difference compared to other popular intero-  
 139 ception questionnaires, such as the MAIA-2, which contains  
 140 more general and metacognitive items (e.g., “I trust my body  
 141 sensations”, “I can notice an unpleasant body sensation  
 142 without worrying about it”), as well as dimensions related  
 143 to attention and emotion regulation (e.g., Not-distracting and  
 144 Not-worrying subscales respectively).

145 The original IAS validation study suggested a two-factor  
 146 structure: one reflecting the perception of interoceptive sig-  
 147 nals (urinate, hungry, defecate, thirsty, pain, heart, taste,  
 148 breathing, temperature, muscles, affective touch, vomit, sex-  
 149 ual 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), with several subsequent studies identifying different optimal solutions and interpretations. For instance, Brand et al. (2023) reported a 1-factor solution to be the best fit, whereas Lin et al. (2023) and Campos et al. (2021) found bifactor solutions (i.e., one general factor above a set of lower-level factors, Rodriguez et al., 2016) 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) 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) 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) 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). Additionally, Campos et al. (2021) suggested that the IAS is predominately a unidimensional scale with all 20 items (except for “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; Mehling et al., 2018) except for the Not-Distracting and Not-Worrying subscales (Brand et al., 2023), which have been previously highlighted as related to non-interoceptive abilities (Ferentzi et al., 2021). 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; Gaggero et al., 2021; Koike & Nomura, 2023), while others find small negative correlations (Lin et al., 2023), a quadratic positive relationship (Campos et al., 2021) or no correlation at all (Murphy et al., 2019). Positive correlations have been observed for four of the five subscales from the Five Facet Mindfulness Questionnaire (FFMQ, Baer et al., 2006; Brand et al., 2023; Koike & No-

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

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; Campos et al., 2021; Koike & Nomura, 2023; Lin et al., 2023; Murphy et al., 2019), somatic symptoms (Brand et al., 2023; Koike & Nomura, 2023; Lin et al., 2023), depressive symptoms (Brand et al., 2023; Koike & Nomura, 2023; Lin et al., 2023), anxiety (Brand et al., 2023), neuroticism (Brand et al., 2023) and self-esteem (Murphy et al., 2019).

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. Accordingly, while the IAS shows promise, the findings highlight the need for further refinement, something which the present paper seeks to address with two aims.

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 (study 1), to clarify the structure of the IAS using a mega-analytic approach which leverages existing data by aggregating existing datasets and re-analysing 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 - EGA). By triangulating our results across different samples and 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. The second aim (study 2) is 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, psychopathological and neurodevelopmental traits 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; Brand et al., 2022; Brand et al., 2023; Campos et al., 2021; Gaggero et al., 2021; Lin et al., 2023; Murphy et al., 2019; Petzke et al., 2024; Poerio et al., 2024; Todd et al., 2022; Von Mohr et al., 2023). Combining studies has the advantage of providing more robust and generalizable understanding of the IAS' factor structure. However, given that studies differ in 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) 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 item scores. Datasets shared by the authors were also included. 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 =  $47.96 \pm 13.1$ , 71.3% Female).

### 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), 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 sam-

ple and each dataset separately (full details are available in the analysis folder at <https://github.com/RealityBending/InterceptionIAS>). 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) 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; H. F. Golino & Epskamp, 2017). EGA is considered a suitable but still comparable and interpretable alternative to traditional factor analysis that addresses several of its limitations, including the assumption of a "latent" source of variability and the potential bias in estimating the optimal number of factors as a function of sample size (Christensen & Golino, 2021b). Additionally, when compared to other factor retention methods such as Kaiser criterion and parallel analysis, EGA performs better in estimating number of factors in complex population structures, such as bifactor solutions (Jiménez et al., 2023). 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). We used the EGAnet package (Christensen & Golino, 2021a) 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 capacity to show potential discrepancies depending on the analysis methods used, we also include traditional EFA/CFA here for continuity with past literature. Unlike EGA, factor analysis assumes that a latent source of variability (i.e., a common latent variable) underlies the observed set of manifest variables (Cosermans et al., 2022). A critical step in factor analysis is determining the optimal number of factors, for which we used the Method Agreement Procedure (Lüdtke et al., 2021), a consensus-based decision method based on multiple factor estimation methods applied concurrently.

Finally, we also applied Hierarchical Clustering Analysis (HCA, Murtagh & Legendre, 2014), 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 (e.g., removing items with low stability, weak associations

**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/IlusionGameReliability
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. Samples using analogue scales were rescaled from a 0–1 range to a 1–5 scale to ensure comparability with the other samples.

with other items, low reliability, or high cross-loadings),  
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). As of note,  
datasets with missing items were excluded from analyses requiring  
the complete set of scale items.

## Results

The distribution of the items across samples suggested a  
consistent modal value (Figure 2, 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 difficulties in inferring  
and reporting them. Lower mode re-

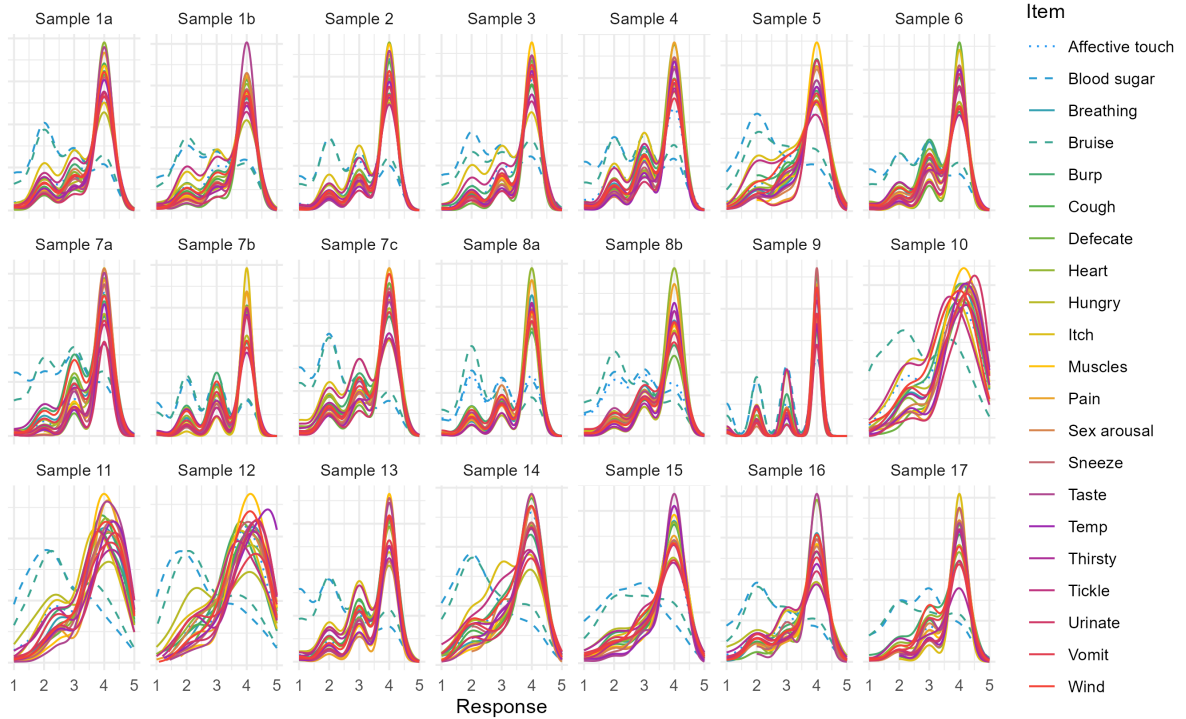
sponses were also observed for "affective touch" in samples  
using the Chinese versions, indicating translation or cultural  
discrepancies (Sorokowska et al., 2021). More generally, the  
response spread (i.e., scale coverage) was narrow (which can  
limit psychometric sensitivity), with a particularly 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)  
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).

UVA flagged two strongly redundant variables, "itch" and

**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-Cough).

**Item Distribution**



**Correlation Matrix**

N = 33526

Breathing	.17	.18	.25	.21	.33	.27	.32	.33	.33	.37	.33	.26	.28	.30	.31	.30	.32	.36	.36	.47	
Heart	.19	.20	.24	.24	.25	.22	.23	.24	.25	.29	.24	.23	.25	.25	.25	.26	.26	.26	.22		.47
Thirsty	.16	.16	.21	.19	.29	.22	.40	.34	.33	.30	.33	.25	.25	.26	.25	.25	.27	.49		.22	.36
Hungry	.17	.15	.21	.21	.30	.22	.35	.32	.33	.28	.33	.24	.23	.25	.24	.24	.24		.49	.26	.36
Cough	.27	.29	.32	.24	.33	.27	.31	.34	.41	.39	.34	.44	.49	.34	.43	.60		.24	.27	.26	.32
Sneeze	.27	.28	.29	.21	.34	.28	.32	.36	.38	.36	.33	.40	.43	.34	.48		.60	.24	.25	.26	.30
Vomit	.22	.23	.29	.22	.36	.30	.33	.38	.35	.34	.33	.34	.37	.35		.48	.43	.24	.25	.25	.31
Taste	.24	.24	.26	.21	.32	.30	.31	.33	.35	.36	.33	.31	.31		.35	.34	.34	.25	.26	.25	.30
Burp	.26	.27	.35	.24	.34	.26	.29	.32	.34	.40	.31	.62		.31	.37	.43	.49	.23	.25	.25	.28
Wind	.25	.25	.31	.23	.35	.25	.31	.34	.34	.37	.30		.62	.31	.34	.40	.44	.24	.25	.23	.26
Pain	.24	.24	.38	.24	.35	.33	.35	.34	.44	.48		.30	.31	.33	.33	.33	.34	.33	.33	.24	.33
Muscles	.25	.26	.37	.26	.37	.32	.33	.35	.42		.48	.37	.40	.36	.34	.36	.39	.28	.30	.29	.37
Temp	.23	.24	.28	.21	.40	.32	.38	.38		.42	.44	.34	.34	.35	.35	.38	.41	.33	.33	.25	.33
Defecate	.20	.19	.22	.17	.36	.28	.59		.38	.35	.34	.34	.32	.33	.38	.36	.34	.32	.34	.24	.33
Urinate	.18	.16	.22	.16	.34	.27		.59	.38	.33	.35	.31	.29	.31	.33	.32	.31	.35	.40	.23	.32
Affective touch	.37	.31	.28	.25	.36		.27	.28	.32	.32	.33	.25	.26	.30	.30	.28	.27	.22	.22	.22	.27
Sex arousal	.21	.18	.22	.18		.36	.34	.36	.40	.37	.35	.35	.34	.32	.36	.34	.33	.30	.29	.25	.33
Blood Sugar	.24	.27	.38		.18	.25	.16	.17	.21	.26	.24	.23	.24	.21	.22	.21	.24	.21	.19	.24	.21
Bruise	.29	.34		.38	.22	.28	.22	.22	.28	.37	.38	.31	.35	.26	.29	.29	.32	.21	.21	.24	.25
Itch	.61		.34	.27	.18	.31	.16	.19	.24	.26	.24	.25	.27	.24	.23	.28	.29	.15	.16	.20	.18
Tickle		.61	.29	.24	.21	.37	.18	.20	.23	.25	.24	.25	.26	.24	.22	.27	.27	.17	.16	.19	.17

395 “tickle”, suggesting the removal of the latter. We therefore  
 396 removed “tickle” from further analysis due to its high redun-  
 397 dancy (and because it is absent from some datasets due to  
 398 translation issues). Several more item pairs were flagged  
 399 as moderately redundant (“wind” and “burp”; “urinate” and  
 400 “defecate”) and mildly redundant (“sneeze” and “cough”;  
 401 “heart” and “breathing”; “hungry” and “thirsty”). These pat-  
 402 terns were observed consistently across most individual sam-  
 403 ples.

#### 404 Structure Exploration

405 HCA highlighted item pairs and triplets that were con-  
 406 sistent grouped together across samples, such as “wind”  
 407 and “burp”, “sneeze” and “cough”, “itch” and “bruise”, “uri-  
 408 nate” and “defecate”, and “pain”, “muscles”, and “tempera-  
 409 ture”. This pattern was largely replicated with EGA, with  
 410 the additional presence of a unique cluster comprising “sex-  
 411 ual arousal”, “affective touch”, “temperature”, “pain”, “mus-  
 412 cles”, and “taste”. EFA suggested an optimal 3-factor so-  
 413 lution with one factor comprising expulsion-related items  
 414 (“burp”, “wind”, “cough”, “sneeze”, and “vomit”), a second  
 415 factor comprising viscerosensitive items (“heart”, “breath-  
 416 ing”, “hungry”, “thirsty”, “urinate”, and “defecate”), and a  
 417 third factor comprising items that are not directly accessi-  
 418 ble via internal bodily sensations alone (“bruise” and “blood  
 419 sugar”). Results concerning the structure of the full scale  
 420 (excluding the “tickle” item) are available via the GitHub  
 421 repository referenced above.

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

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

#### 444 Confirmatory Factor Analysis (CFA)

445 We fitted and compared various candidate structures  
 446 emerging from the previous analyses, including a 1-factor  
 447 model (the G-model), a 3-factor model (EFA), a 3+1 model  
 448 (EFA + general factor), a 5-factor model (HCA), a 5+1 model  
 449 (HCA + general factor), a 7-factor model (EGA), and a 7+1  
 450 model (EGA + general factor), using CFA. The EGA model  
 451 with 7 factors of item pairs provided the best fit to the data  
 452 with the lowest RMSEA (0.035), lowest  $\chi^2$  (2,334.112), and  
 453 highest CFI (0.984). This was followed by the EGA + gen-  
 454 eral factor model (RMSEA = 0.054,  $\chi^2$  = 6,876.395, CFI  
 455 = 0.952), and then the HCA model with 5 factors (RMSEA  
 456 = 0.078,  $\chi^2$  = 13,441.878, CFI = 0.906). All other mod-  
 457 els performed poorly, with RMSEA > 0.08 and CFI < 0.90.  
 458 The EGA model yielded the lowest BIC (which favours parsim-  
 459 ony), followed by the EGA plus a general factor model. All  
 460 other models displayed significantly lower evidence (BIC-  
 461 based Bayes Factor < 1/100) for their structures compared to  
 462 the EGA model.

463 Across individual datasets, the EGA model provided the  
 464 best fit to the data in 8 out of 13 samples for which the  
 465 CFA converged. In these samples, the 7-factor EGA model  
 466 showed excellent performance, with low RMSEA values ( $\approx$   
 467 0.025–0.057), high CFI values ( $\geq$  0.95), and substantially  
 468 lower  $\chi^2$  values compared to alternative models. Adding a  
 469 general factor improved model fit in three samples, for which  
 470 the 7-factor EGA + general factor model showed acceptable  
 471 RMSEA ( $\approx$  0.048–0.061) and CFI ( $\approx$  0.929–0.957) values.  
 472 In two samples, the 5-factor plus general factor model pro-  
 473 vided the best fit, with acceptable RMSEA values (0.049  
 474 and 0.057) and CFI values (0.955 and 0.919). Bayes fac-  
 475 tor comparisons aligned with BIC-based conclusions; taken  
 476 together, these results suggest that although the EGA model  
 477 was preferred in most samples, bifactor models (i.e., EGA  
 478 and HCA models with a general factor) provided evidence  
 479 for a common IAS factor in some samples.

#### 480 Discussion

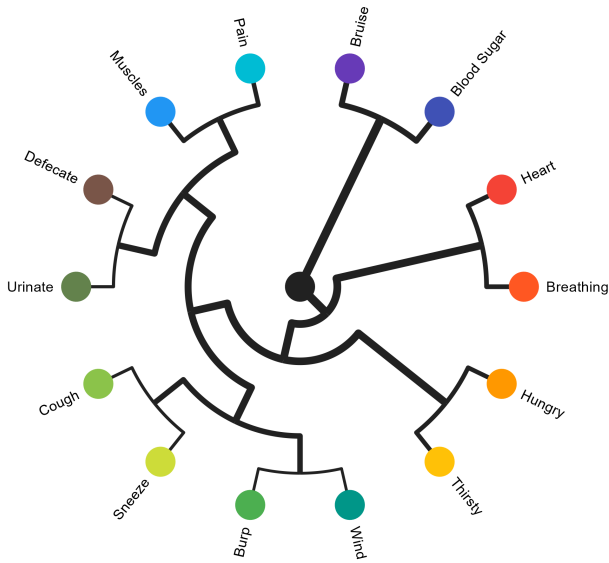
481 Study 1 aimed to systematically evaluate the structure of  
 482 the IAS using a mega-analytic approach with both factor-  
 483 analytic and network-based methods. Previous research  
 484 across different datasets has highlighted inconsistency in op-  
 485 timal factor structure (e.g., unidimensional, bifactor, and two  
 486 factor solutions have all been reported Brand et al., 2023;  
 487 Campos et al., 2021; Koike & Nomura, 2023; Lin et al.,  
 488 2023; Murphy et al., 2019) and raised concerns about spec-  
 489 ific items (e.g., low discrimination, local dependencies, and  
 490 items reflecting phenomena less accessible to interoception).  
 491 To address these concerns and inconsistencies we combined  
 492 17 datasets with over 33 thousand participants using the IAS,  
 493 providing the most comprehensive and robust analysis of  
 494 the IAS structure to date. We conducted analyses across

**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.

**Hierarchical Clustering Analysis (HCA)**

Method = Correlation



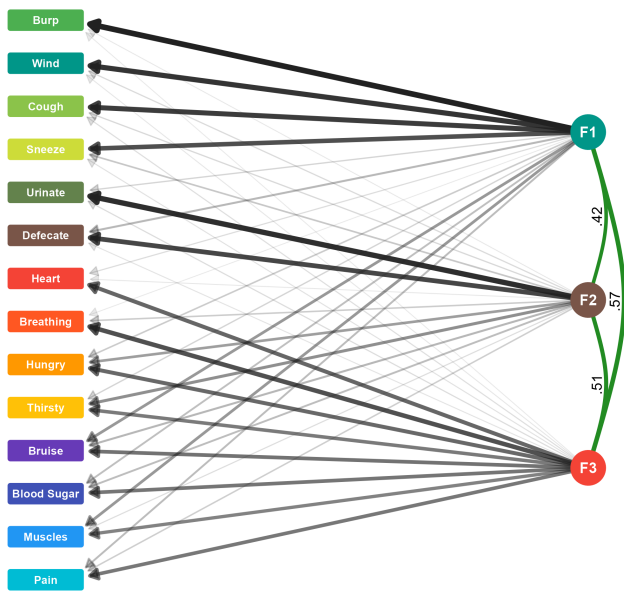
**Exploratory Graph Analysis (EGA)**

Method = Leiden



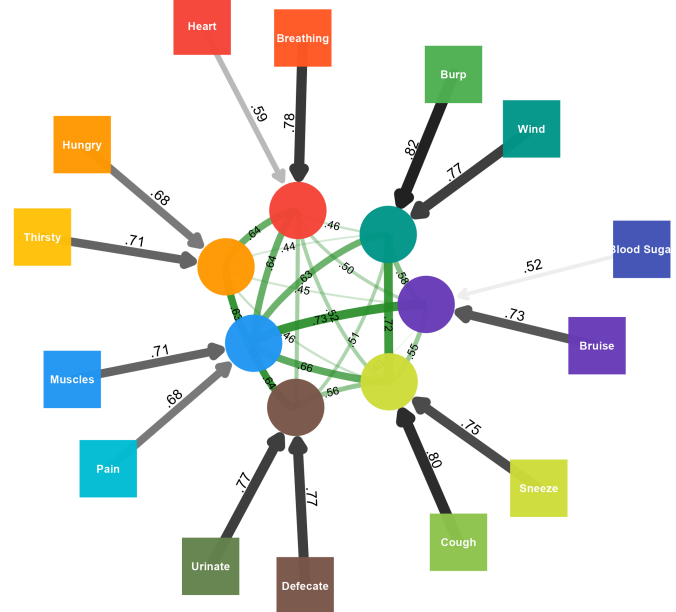
**Exploratory Factor Analysis (EFA)**

Method = Oblimin



**Confirmatory Factor Analysis (CFA)**

Method = Maximum Likelihood



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. Notably, while  $\chi^2$  and BIC can be sensitive to very large sample sizes, the conclusions are supported not only by model fit but also by the replication of the structure across analytic approaches and individual datasets.

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 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; Lin et al., 2023).

Study 1 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, meaning that context-specific phrasing may be necessary to more accu-

rately capture interoceptive beliefs and reduce measurement noise in questionnaire measures (Makowski, Neves, et al., 2025; e.g., Vlemincx et al., 2023).

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; Lin et al., 2023; Murphy et al., 2019; Todd et al., 2022). 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, more interpretable, and more reliable external associations by comparing associations to dispositional variables with the original 21-item unidimensional solution, 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, psychopathological and neurodevelopmental traits, 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 Measures.** Interoception-related measures included the Multidimensional Assessment of Interoceptive Awareness Version-2 (MAIA-2, Mehling et al., 2018), the Body Perception Questionnaire Short Form (BPQ-SF, Cabrera et al., 2018), and the Interoceptive Confusion Questionnaire (ICQ, Brewer et al., 2016). We also included the Bermond-Vorst Alexithymia Questionnaire (BVAQ, Vorst & Bermond, 2001), and the Toronto Alexithymia Scale (TAS-20, Bagby et al., 1994), due to the close relationship between interoception deficits and alexithymia (e.g., Brewer et al., 2016; Gaggero et al., 2021).

**Mood and Anxiety Related Measures.** 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); namely the Patient Health Questionnaire-4 (PHQ-4, Kroenke et al., 2009), the Patient Health Questionnaire-15 (PHQ-15, Kroenke et al., 2002), the Patient Health Questionnaire-9 (PHQ-9, Kroenke et al., 2001), the Short Mood and Feelings Questionnaire (MFQ, Angold et al., 1995), the Beck Depression Inventory-II (BDI-II, Beck et al., 1996), the State-Trait Anxiety Inventory Trait Version (STAI-T, Spielberger, 1970; and STAIT-T, Zsido et al., 2020), the Generalized Anxiety Disorder Scale (GAD-7, Spitzer et al., 2006), and General Anxiety Disorder-2 (GAD-2, Kroenke et al., 2007), obtained from the PHQ-4).

**Psychopathological and Neurodevelopmental Traits.** Dimensional psychopathological and neurodevelopmental-related trait scores were included from the Personality Inventory for DSM-5 Short Form (PID-5-SF, Thimm et al., 2016), the Autism Spectrum Quotient Short Form (ASQ-S, Hoekstra et al., 2011), the Schizotypal Personality Questionnaire – Brief Revised (SPQ-BR, Davidson et al., 2016), and the McLean Screening Instrument for Borderline Personality Disorder (MSI-BPD, Zanarini, 2003).

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

**Beliefs and Misbeliefs.** Beliefs- and misbeliefs-related measures included the Lie Scale (LIE, Makowski et al., 2023), the Generic Conspiracist Beliefs Scale (GCBS, Brotherton et al., 2013), and the Primal World Beliefs Inventory (PI-99, Clifton et al., 2019) and its short form (PI-18, Clifton & Yaden, 2021). 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.

### 636 Data Analysis

637 All variables from the 17 open-access datasets were  
638 extracted and standardized (z-scored) for the aggregated  
639 dataset. Scores were harmonized and combined where mean-  
640 ingful (e.g., scores from the short 18-item form of the PI and  
641 scores from its longer 99-item version). Missing data was ex-  
642 cluded on a per-model basis, to maximize statistical power.

643 Akin to a meta-analytic framework, datasets were treated  
644 as clusters (random groups), and pooled estimates of the as-  
645 sociations that account for both within- and between-sample  
646 variability were computed. This approach allowed for the es-  
647 timation of multiple associations while still accounting for

648 study heterogeneity and intra-study uncertainty (i.e., mea-  
649 surement uncertainty within the studies). Specifically, as-  
650 sociations were estimated using mixed models implemented  
651 in the *glmmTMB* package (McGillycuddy et al., 2025), and  
652 postprocessed with the easystats R ecosystem (Ben-Shachar  
653 et al., 2020; Lüdtke et al., 2020; Makowski, Ben-Shachar,  
654 et al., 2025). Each predictor was entered separately with the  
655 seven IAS item pairs identified in Study 1 (Hungry-Thirsty,  
656 Muscles-Pain, Wind-Burp, Urinate-Defecate, Breathing-  
657 Heart, Bruise-Blood sugar, Cough-Sneeze) and the origi-  
658 nal IAS “total” score serving as outcomes. Models included  
659 dataset as a random intercept and the predictor as a random  
660 slope. Note that given the large sample size and iterative  
661 hypothesis testing, we focus on effect sizes and confidence  
662 intervals rather than *p*-values which are not relevant for in-  
663 ference in this context.

664 From each fitted model, we extracted standardized re-  
665 gression coefficients (interpretable as a correlation index),  
666 and 95% confidence intervals (additional details on *p*-values,  
667 convergence status, and sample sizes are available in the  
668 GitHub repository). Results are summarized in Figure 4 and  
669 we report the mean coefficient in the text when consistent  
670 across measures.

## 671 Results

### 672 Demographics

673 Age showed consistent positive standardized associations  
674 across IAS correlates ( $\beta = .13$ ), including both the IAS total  
675 scores and the scores for the 7-pairs, with all effects signifi-  
676 cant. Gender displayed near-zero effects ( $\beta = -.06$ ) and was  
677 only significant for the Hungry-Thirsty, Wind-Burp, and  
678 Urinate-Defecate pairs.

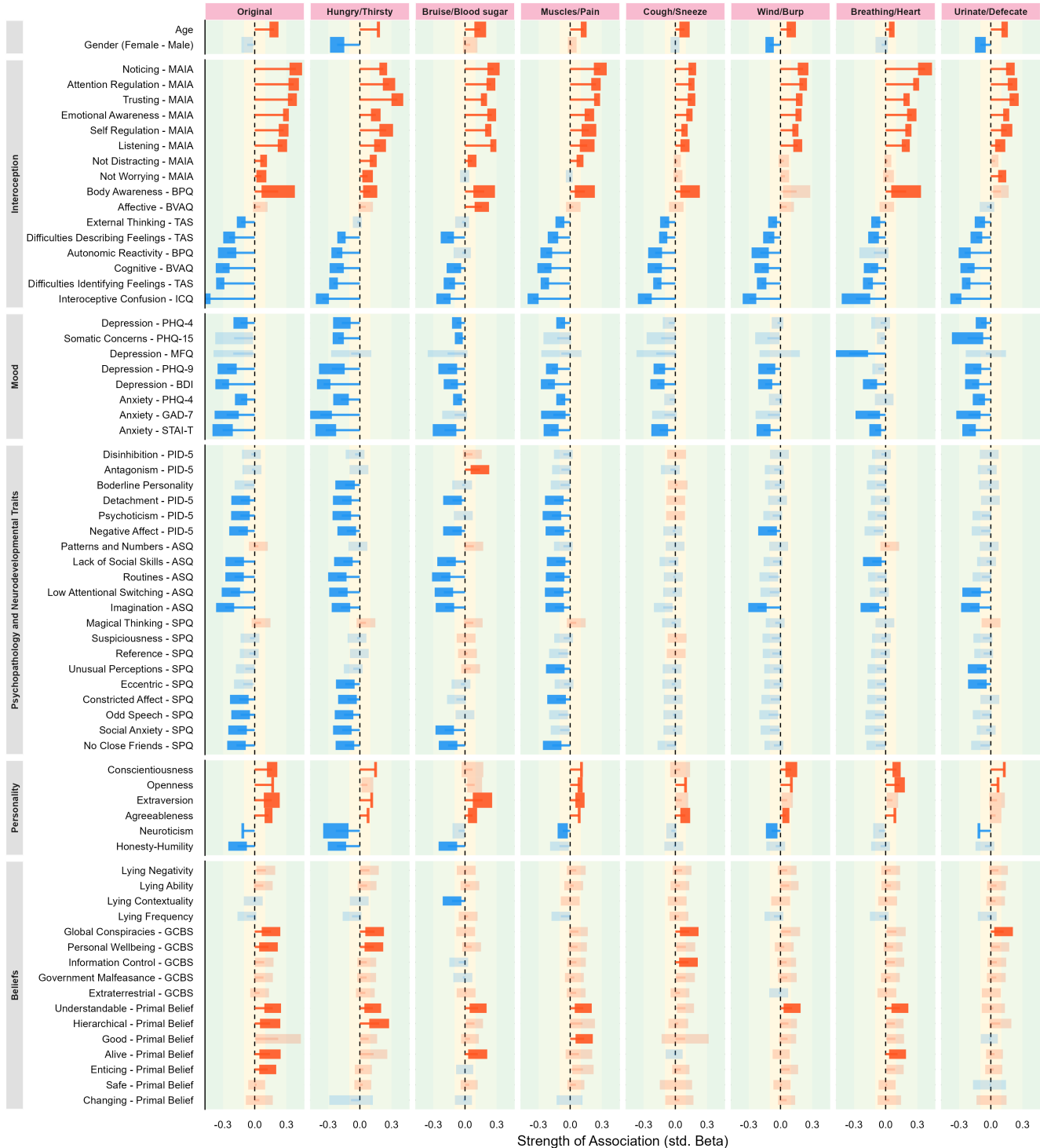
### 679 Interoception and Interoception Related Measures

680 MAIA subscales showed robust positive standardized as-  
681 sociations with the IAS. Noticing had the strongest associ-  
682 ation ( $\beta = .26$ ), followed by Attention Regulation ( $\beta = .25$ )  
683 with other subscales (Body Listening, Trusting, Emotional  
684 Awareness and Self-Regulation) in the .19-.24 range, all sig-  
685 nificant. Effects were generally strongest for the full IAS, ex-  
686 cept for Body Listening, which peaked for the Bruise-Blood  
687 sugar pair. Not Distracting was significant only for the full  
688 IAS, Hungry-Thirsty, Bruise-Blood sugar and the Muscles-  
689 Pain pairs, albeit weakly ( $\beta \approx -.02$  to  $.06$ ). Not Worrying  
690 was significant only for the full IAS and the Hungry-Thirsty  
691 and Urinate-Defecate pairs ( $\beta \approx -.01$  to  $.04$ ).

692 The BPQ’s Autonomic Reactivity subscale was reliably  
693 negatively associated across the seven item pairs ( $\beta = -.19$ ),  
694 reaching significance for all pairs except Bruise-Blood sugar  
695 and Breathing-Heart. In contrast, the Body Awareness sub-  
696 scale of the BPQ was positive and significant across most

**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 psychopathological and neurodevelopmental traits 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.*



697 pairs ( $\beta = .15$ ; range = .10 to .22) except for Wind–Burp and  
698 Urinate–Defecate.

699 Interoceptive deficits showed consistent negative associ-  
700 ations: Alexithymia measures (TAS, BVAQ) were robustly  
701 negative ( $\beta \approx -.13$  to  $-.2$ ), except for the BVAQ Affective  
702 subscale, which was positive for Bruise–Blood sugar ( $\beta =$   
703  $.16$ ). Interoceptive confusion (ICQ) was also negative across  
704 all IAS measures ( $\beta = -.32$ ; range =  $-.49$  to  $-.20$ ).

#### 705 *Mood and Anxiety Related Measures*

706 Depression ( $\beta = -.20$ ) and anxiety ( $\beta = -.20$ ) showed con-  
707 sistent negative associations across the IAS and all its pairs,  
708 with the strongest effects observed for the Hungry–Thirsty  
709 dimension. However, looking at specific measures revealed  
710 some differences.

711 The MFQ was largely nonsignificant except for a nega-  
712 tive association with Breathing–Heart, while somatic con-  
713 cerns (PHQ-15) were unrelated to the IAS total scale but  
714 had weak negative associations with the Hungry–Thirsty,  
715 Bruise–Blood sugar, and Urinate–Defecate pairs (from  $-.22$   
716 to  $-.05$ ).

717 Other mood indices (PHQ-9, PHQ-4, GAD-7) produced  
718 generally weaker but often reliable negative coefficients (typ-  
719 ically  $-.10$  to  $-.15$ ), though effects sometimes appeared only  
720 for specific pairs (e.g., PHQ-4 was significant for all pairs ex-  
721 cept Wind–Burp and Breathing–Heart). Across anxiety and  
722 depression measures, the Hungry–Thirsty pair consistently  
723 had the strongest standardized effects.

#### 724 *Psychopathological and Neurodevelopmental Traits*

725 Psychopathological traits were generally negatively as-  
726 sociated with the IAS scores. Detachment, Psychoticism,  
727 and Negative Affect demonstrated the most reliable, albeit  
728 weak, negative associations with IAS scores (from  $-.08$   
729 to  $-.10$ ), again primarily driven by the Hungry–Thirsty,  
730 Bruise–Blood sugar, and Muscles–Pain pairs.

731 Other maladaptive traits were unrelated to IAS scores.  
732 Notably, Antagonism showed a significant positive asso-  
733 ciation, limited to the Bruise–Blood sugar pair ( $\beta = .14$ ),  
734 whereas Borderline Personality was uniquely negatively as-  
735 sociated with Hungry–Thirsty ( $\beta = -.14$ ).

736 A comparable trend was observed for autistic traits  
737 (ASQ), with the strongest effects driven by the Hungry–  
738 Thirsty, Bruise–Blood sugar, and Muscles–Pain pairs, which  
739 yielded the largest negative coefficients. Among ASQ sub-  
740 scales, Imagination ( $\beta = -.18$ ) exhibited the most consis-  
741 tent pattern of negative associations across pairs except for  
742 Cough–Sneeze.

743 Schizotypal traits showed a similarly negative pattern of  
744 associations to IAS scores. Social Anxiety ( $\beta = -.11$ ) and  
745 No Close Friends ( $\beta = -.12$ ) displayed the most consistent  
746 effects. Although some facets were unrelated to the total IAS  
747 score, they were associated with specific item pairs: Unusual

748 Perceptions with Muscles–Pain and Urinate–Defecate, and  
749 Eccentric with Hungry–Thirsty and Urinate–Defecate.

#### 750 *Personality*

751 Among the Big Five traits, Openness ( $\beta = .10$ ), Agree-  
752 ableness ( $\beta = .08$ ), and Conscientiousness ( $\beta = .11$ ) exhib-  
753 ited small yet consistent positive associations, most clearly  
754 for the total IAS score. Extraversion showed a weak but posi-  
755 tive association, driven primarily by the Bruise–Blood sugar  
756 pair ( $\beta = .17$ ). In contrast, Honesty–Humility ( $\beta = -.10$ )  
757 had a small negative effect, most clearly for the Hungry–  
758 Thirsty and Bruise–Blood sugar pairs. Finally, Neuroticism  
759 ( $\beta = -.23$ ) was negatively associated with the IAS, with the  
760 strongest effect observed for the Hungry–Thirsty pair ( $\beta =$   
761  $-.22$ ).

#### 762 *Beliefs and Misbeliefs*

763 Within the Lying Profile, only Lying Contextuality  
764 showed a modest negative coefficient for Bruise–Blood  
765 sugar ( $\beta = -.12$ ). Global Conspiracies (GCBS) beliefs were  
766 associated with the total IAS and three pairs ( $\beta = .10$ ):  
767 Hungry–Thirsty, Cough–Sneeze and Urinate–Defecate. Pri-  
768 mal world beliefs were largely null, with small positive asso-  
769 ciations for appraising the world as being “Understandable”,  
770 “Alive”, and “Hierarchical”, with Hungry–Thirsty ( $\beta = .09$ ).  
771 Appraising the world as “Good” was uniquely positively re-  
772 lated with Muscles–Pain ( $\beta = .09$ ).

773 Overall, the original IAS total score had the greatest number  
774 of significant standardized effects (28), though many were  
775 driven by specific item pairs rather than uniform scale-wide  
776 associations. Hungry–Thirsty (17) and Bruise–Blood sugar  
777 (10) consistently yielded the largest standardized coeffi-  
778 cients, especially for interoception, mood, and psychopatho-  
779 logical and neurodevelopmental trait measures. Other item  
780 pairs (Cough–Sneeze, Wind–Burp, Urinate–Defecate) con-  
781 tributed little unique variance and rarely had meaningful  
782 standardized effects.

#### 783 *Discussion*

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

789 The IAS had positive associations with all MAIA facets  
790 except for the Not-Distracting and Not-Worrying subscales,  
791 replicating prior findings (Brand et al., 2023; Todd et al.,  
792 2022). Similarly, associations between the IAS and BPQ  
793 Body Awareness subscale were positive, consistent with pre-  
794 vious research (Brand et al., 2023; Gaggero et al., 2021;  
795 Koike & Nomura, 2023). Indices of autonomic reactivi-  
796 ty, 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). 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; Koike & Nomura, 2023; Lin et al., 2023; Murphy et al., 2019).

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), negative (Lin et al., 2023), and even quadratic relationships (e.g., Campos et al., 2021). 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; e.g., Gabriele et al., 2022). 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 findings (Koike & Nomura, 2023; Murphy et al., 2019). In contrast, associations with psychopathological and neurodevelopmental traits 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 reflecting contributions from specific pairs. 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 partially align with prior work reporting a negative association between Neuroticism and the IAS (Brand et al., 2023), but associations involving other personality traits are less well established and require replication to situate the present findings within the broader literature. Overall, however, results suggest that associations between interoceptive self-report and personality dimensions may depend on the bodily domain represented.

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 di-

mensions, 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.

Notably, the primary focus of this study was to capture general patterns across the aggregated datasets rather than systematic differences between samples; accordingly, dataset-specific effects were not examined in detail. While variability across datasets is already reflected in the uncertainty of the estimates, a promising direction for future research would be to investigate sample-specific associations within the individual open-data datasets and investigate the cause of potential divergences across samples.

In summary, the 14-item IAS performed similarly to the original 21-item version, showing comparable patterns of association with mood, psychopathological and neurodevelopmental traits, 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 psychopathological and neurodevelopmental traits and beliefs. The homeostatic salience and universality of the Hungry–Thirsty pair, and the contextual relevance and visual/psychological prominence of the Bruise–Blood sugar pair, might contribute to their stronger associations with dispositional traits.

Overall, our findings both support the validity of the original IAS general score, but also point at the refined 14-item scale as an alternative offering a more efficient, nuanced, and interpretation-specific measure of self-reported interoceptive accuracy abilities.

## 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; Murphy et al., 2019) 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–

902 Defecate, Muscles–Pain, Breathing–Heart, Cough–Sneeze,  
 903 Wind–Burp, and Bruise–Blood sugar), yielding a refined 14-  
 904 item structure with superior fit and cross-dataset stability  
 905 compared to previously proposed unidimensional or bifactor  
 906 models. The presence of redundant or unstable items  
 907 (e.g., Tickle, Taste, Affective Touch) might weaken the inter-  
 908 pretability of the original scale and contribute to the inconsis-  
 909 tencies observed in prior factor-analytic findings (Campos  
 910 et al., 2021; Lin et al., 2023). Unlike Lin et al. (2023),  
 911 whose 12-item solution supported unidimensionality, Study  
 912 1 findings suggest that local dependencies between items re-  
 913 flect meaningful theoretical coupling among functionally re-  
 914 lated signals (e.g., Heart–Breathing). Such clustering likely  
 915 mirrors the way bodily signals co-occur physiologically or  
 916 phenomenologically (e.g., hunger with thirst, coughing with  
 917 sneezing) and are therefore perceived and evaluated together.  
 918 Perceived accuracy in one bodily system may generalize to  
 919 conceptually or functionally linked sensations, yielding a  
 920 cluster rather than continuous organization of interoceptive  
 921 beliefs.

922 Consistent with this interpretation, self-reported intero-  
 923 ceptive abilities across related bodily axes (e.g., cardiac  
 924 and respiratory) show positive associations (Garfinkel et al.,  
 925 2016), whereas objective measures of interoceptive accuracy  
 926 across modalities typically show weak or absent correlations  
 927 (Banellis et al., 2026; Bruni, 2023; Ferentzi et al., 2018).  
 928 This dissociation indicates that objective interoceptive per-  
 929 formance may be largely domain-specific whereas subjective  
 930 interoceptive beliefs may generalize across functionally  
 931 related systems. We therefore interpret the present cluster-  
 932 ing as reflecting the conceptual organization of interoceptive  
 933 beliefs which are structured by perceived bodily interdepend-  
 934 ence, as opposed to a single general interoceptive capacity.

935 Study 2 extended these findings by examining the IAS’s  
 936 associations with interoceptive, affective, personality, and  
 937 belief-related traits. Consistent with previous research, the  
 938 IAS correlated positively with adaptive aspects of interocep-  
 939 tive sensibility (e.g., MAIA subscales of reflecting notic-  
 940 ing, attention regulation, and body trusting) and negatively  
 941 with interoceptive difficulties, namely alexithymia and intero-  
 942 ceptive confusion (e.g., Brand et al., 2023; Gaggero et al.,  
 943 2021; Garfinkel et al., 2015). Also consistent with previous  
 944 research, individuals reporting higher symptoms of depres-  
 945 sion and anxiety tended to perceive themselves as less ac-  
 946 curate in detecting bodily signals (e.g., Brand et al., 2023;  
 947 Khalsa, Adolphs, Cameron, Critchley, Davenport, Feinstein,  
 948 Feusner, Garfinkel, Lane, Mehling, Meuret, et al., 2018;  
 949 Nord & Garfinkel, 2022). Crucially, however, these associa-  
 950 tions were not uniform across bodily domains. The strongest,  
 951 and most robust, effects involved drive-states (e.g., hunger,  
 952 thirst), whereas regulatory systems, such as cardiorespira-  
 953 tory and expulsion domains showed weaker or inconsistent  
 954 relationships to anxiety and depression metrics. Such do-

955 main specificity implies either that beliefs about gastroin-  
 956 testinal accuracy are more psychologically and clinically rel-  
 957 evant than beliefs about reflexive or surface sensations, or  
 958 that the experience of these signals is more salient and eas-  
 959 ier to accurately appraise, leading to meaningful variation  
 960 across individuals.

961 Patterns of domain-specific links to anxiety and depres-  
 962 sion may explain inconsistencies in the literature linking in-  
 963 teroception and emotion. Reviews and meta-analyses report  
 964 that performance-based measures of interoceptive accuracy,  
 965 usually cardioception, show weak or absent correlations with  
 966 depression and anxiety (Adams et al., 2022; Banellis et al.,  
 967 2025; Jenkinson et al., 2024) whereas self-reported intero-  
 968 ceptive measures, containing items relevant to multiple bod-  
 969 ily domains, typically show stronger, often negative, asso-  
 970 ciations with internalizing traits (e.g., Brand et al., 2022;  
 971 Clemente et al., 2024; Lin et al., 2023).

972 This discrepancy likely reflects both conceptual distinc-  
 973 tions between perceptual accuracy and metacognitive be-  
 974 liefs (Garfinkel et al., 2015; Khalsa, Adolphs, Cameron,  
 975 Critchley, Davenport, Feinstein, Feusner, Garfinkel, Lane,  
 976 Mehling, Meuret, et al., 2018; Suksasilp & Garfinkel, 2022)  
 977 as well as variation in the bodily domains assessed. Study  
 978 2 results extend this idea by demonstrating that associations  
 979 between interoceptive beliefs and affective traits are concen-  
 980 trated in psychologically salient visceral domains. Thus,  
 981 the field’s heavy reliance on cardioceptive tasks may have  
 982 obscured more meaningful relationships between affective  
 983 traits and other interoceptive modalities.

984 Beyond mood, similar albeit weaker patterns emerged for  
 985 autistic traits, schizotypy, and maladaptive personality di-  
 986 mensions. These traits tended to correlate negatively with  
 987 subjective interoceptive accuracy, particularly for visceral  
 988 and musculoskeletal domains but not other areas. This se-  
 989 lective pattern suggests that, rather than reflecting a general  
 990 interoceptive deficit, diminished accuracy beliefs in certain  
 991 bodily domains may contribute to specific socio-emotional  
 992 and personality features.. This idea fits well with previous  
 993 research showing that interoceptive self-beliefs are meaning-  
 994 fully related to emotional awareness, social connectedness,  
 995 and psychopathology (Brand et al., 2022; Torregrossa et al.,  
 996 2022; Williams et al., 2023). The present findings extend  
 997 this literature by showing that such associations are likely  
 998 domain-specific rather than global across all bodily systems.

999 Taken together, findings from Study 1 and 2 suggest  
 1000 that self-perceived interoceptive accuracy is a differentiated  
 1001 construct, varying systematically across distinct bodily sys-  
 1002 tems. Beliefs about gastrointestinal accuracy, in particu-  
 1003 lar, showed the most consistent associations with affective  
 1004 and dispositional features. Although modest in magnitude,  
 1005 these effects emphasize the value of considering domain-  
 1006 specificity when linking interoceptive beliefs to emotion  
 1007 and mental health. Clinically, the domain-specific structure

of the IAS highlights that interoceptive disruptions in affective, psychopathological and neurodevelopmental 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 indicate use of the refined 14-item IAS which balances parsimony and predictive validity, illustrating how network-informed scale refinement (H. F. Golino & Epskamp, 2017) 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) 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)' 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 generalizes 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; Ferentzi et al., 2021; Rogowska et al., 2023; Vig et al., 2023) 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; e.g.,

cardiac accuracy vs subjective accuracy, Brand et al., 2023; Petzke et al., 2024), 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 related to intuitive bodily domains might be better suited to capture interpretable and ecologically valid domain-specific interoceptive beliefs. 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., Ehmsen, J. F., Courtin, A. S., Vejlo, M., Tyrer, A., Böhme, R. A., Fardo, F., & Allen, M. G. (2026). Interoceptive ability is uncorrelated across respiratory and cardiac axes in a large scale psychophysical study. *Communications Psychology*.
- Banellis, L., Nikolova, N., Fischer Ehmsen, J., Courtin, A. S., Vejlo, M., Tyrer, A., Bohme, R., Bavato, F., Hooger-

- 1113 vorst, K., Fardo, F., et al. (2025). Interoceptive ability  
1114 is unrelated to mental health symptoms: Evidence from  
1115 a large scale multi-domain psychophysical investigation.  
1116 *medRxiv*, 2025–2008.
- 1117 Beck, A. T., Steer, R. A., & Brown, G. (1996). Beck depres-  
1118 sion inventory–II. *Psychological Assessment*.
- 1119 Ben-Shachar, M. S., Lüdtke, D., & Makowski, D. (2020).  
1120 Effectsize: Estimation of effect size indices and stan-  
1121 dardized parameters. *Journal of Open Source Software*,  
1122 5(56), 2815.
- 1123 Brand, S., Meis, A. C., Tünte, M. R., Murphy, J., Woller,  
1124 J. P., Jungmann, S. M., Witthöft, M., Hoehl, S., Wey-  
1125 mar, M., Hermann, C., & Ventura-Bort, C. (2023). A  
1126 multi-site German validation of the Interoceptive Accu-  
1127 racy Scale and its relation to psychopathological symp-  
1128 tom burden. *Communications Psychology*, 1(1). <https://doi.org/10.1038/s44271-023-00016-x>
- 1130 Brand, S., Petzke, T. M., & Witthöft, M. (2022). The  
1131 differential relationship between self-reported intero-  
1132 ceptive accuracy and attention with psychopathology.  
1133 *Zeitschrift für Klinische Psychologie Und Psychothera-  
1134 pie*.
- 1135 Brewer, R., Cook, R., & Bird, G. (2016). Alexithymia: A  
1136 general deficit of interoception. *Royal Society Open Sci-  
1137 ence*, 3(10), 150664.
- 1138 Brotherton, R., French, C. C., & Pickering, A. D. (2013).  
1139 Measuring belief in conspiracy theories: The generic  
1140 conspiracist beliefs scale. *Frontiers in Psychology*, 4,  
1141 279.
- 1142 Bruni, P. (2023). Interoception across domains: Assessing  
1143 the role of age. *Innovation in Aging*, 7, 755.
- 1144 Cabrera, A., Kolacz, J., Pailhez, G., Bulbena-Cabre, A.,  
1145 Bulbena, A., & Porges, S. W. (2018). Assessing body  
1146 awareness and autonomic reactivity: Factor structure  
1147 and psychometric properties of the body perception  
1148 questionnaire-short form (BPQ-SF). *International Jour-  
1149 nal of Methods in Psychiatric Research*, 27(2), e1596.
- 1150 Campos, C., Rocha, N. B., & Barbosa, F. (2021). *Untan-  
1151 gling self-reported interoceptive attention and accuracy:  
1152 Evidence from the european portuguese validation of the  
1153 body perception questionnaire and the interoceptive ac-  
1154 curacy scale*. <http://dx.doi.org/10.31234/osf.io/a7wdj>
- 1155 Christensen, A. P., Garrido, L. E., & Golino, H. (2023).  
1156 Unique variable analysis: A network psychometrics  
1157 method to detect local dependence. *Multivariate Behav-  
1158 ior Research*, 58(6), 1165–1182.
- 1159 Christensen, A. P., & Golino, H. (2021a). Estimating the  
1160 stability of psychological dimensions via bootstrap ex-  
1161 ploratory graph analysis: A monte carlo simulation and  
1162 tutorial. *Psych*, 3(3), 479–500.
- 1163 Christensen, A. P., & Golino, H. (2021b). On the equiva-  
1164 lency of factor and network loadings. *Behavior Research  
1165 Methods*, 53(4), 1563–1580.
- 1166 Clemente, R., Murphy, A., & Murphy, J. (2024). The rela-  
1167 tionship between self-reported interoception and anxiety:  
1168 A systematic review and meta-analysis. *Neuroscience &  
1169 Biobehavioral Reviews*, 167, 105923.
- 1170 Clifton, J. D., Baker, J. D., Park, C. L., Yaden, D. B., Clifton,  
1171 A. B., Terni, P., Miller, J. L., Zeng, G., Giorgi, S.,  
1172 Schwartz, H. A., et al. (2019). Primal world beliefs. *Psy-  
1173 chological Assessment*, 31(1), 82.
- 1174 Clifton, J. D., & Yaden, D. B. (2021). Brief measures of the  
1175 four highest-order primal world beliefs. *Psychological  
1176 Assessment*, 33(12), 1267.
- 1177 Cosemans, T., Rosseel, Y., & Gelper, S. (2022). Exploratory  
1178 graph analysis for factor retention: Simulation results for  
1179 continuous and binary data. *Educational and Psycholog-  
1180 ical Measurement*, 82(5), 880–910.
- 1181 Costa, P., McCrae, R. R., & Revised, N. (1992). Personal-  
1182 ity inventory (NEO-PI-r) and NEO five-factor inventory  
1183 (NEO-FFI): Professional manual. *Psychological Assess-  
1184 ment Resources*, Odessa, FL.
- 1185 Davidson, C. A., Hoffman, L., & Spaulding, W. D. (2016).  
1186 Schizotypal personality questionnaire–brief revised (up-  
1187 dated): An update of norms, factor structure, and item  
1188 content in a large non-clinical young adult sample. *Psy-  
1189 chiatry Research*, 238, 345–355.
- 1190 Desmedt, O., Heeren, A., Corneille, O., & Luminet, O.  
1191 (2022). What do measures of self-report interoception  
1192 measure? Insights from a systematic review, latent factor  
1193 analysis, and network approach. *Biological Psychology*,  
1194 169, 108289.
- 1195 Desmedt, O., Luminet, O., Maurage, P., & Corneille, O.  
1196 (2025). Discrepancies in the definition and measurement  
1197 of human interoception: A comprehensive discussion  
1198 and suggested ways forward. *Perspectives on Psycholog-  
1199 ical Science*, 20(1), 76–98.
- 1200 Desmedt, O., Luminet, O., Walentynowicz, M., &  
1201 Corneille, O. (2023). The new measures of in-  
1202 teroceptive accuracy: A systematic review and as-  
1203 sessment. *Neuroscience & Biobehavioral Reviews*,  
1204 153, 105388. [https://doi.org/https://doi.org/10.1016/j.  
1205 neubiorev.2023.105388](https://doi.org/https://doi.org/10.1016/j.neubiorev.2023.105388)
- 1206 Ferentzi, E., Drew, R., Tihanyi, B. T., & Köteles, F. (2018).  
1207 Interoceptive accuracy and body awareness–temporal  
1208 and longitudinal associations in a non-clinical sample.  
1209 *Physiology & Behavior*, 184, 100–107.
- 1210 Ferentzi, E., Horváth, Á., & Köteles, F. (2019). Do body-  
1211 related sensations make feel us better? Subjective well-  
1212 being is associated only with the subjective aspect of in-  
1213 teroception. *Psychophysiology*, 56(4), e13319.
- 1214 Ferentzi, E., Oлару, G., Geiger, M., Vig, L., Köteles, F.,  
1215 & Wilhelm, O. (2021). Examining the factor structure  
1216 and validity of the multidimensional assessment of in-  
1217 teroceptive awareness. *Journal of Personality Assessment*,  
1218 103(5), 675–684.

- 1219 Gabriele, E., Spooner, R., Brewer, R., & Murphy, J. (2022).  
1220 Dissociations between self-reported interoceptive accu-  
1221 racy and attention: Evidence from the interoceptive at-  
1222 tention scale. *Biological Psychology*, *168*, 108243.
- 1223 Gaggero, G., Bizzego, A., Dellantonio, S., Pastore, L., Lim,  
1224 M., & Esposito, G. (2021). Clarifying the relationship  
1225 between alexithymia and subjective interoception. *PLoS*  
1226 *One*, *16*(12), e0261126.
- 1227 Garfinkel, S. N., Manassei, M. F., Hamilton-Fletcher, G., In  
1228 den Bosch, Y., Critchley, H. D., & Engels, M. (2016).  
1229 Interoceptive dimensions across cardiac and respiratory  
1230 axes. *Philosophical Transactions of the Royal Society B:*  
1231 *Biological Sciences*, *371*(1708), 20160014.
- 1232 Garfinkel, S. N., Seth, A. K., Barrett, A. B., Suzuki, K.,  
1233 & Critchley, H. D. (2015). Knowing your own heart:  
1234 Distinguishing interoceptive accuracy from interocep-  
1235 tive awareness. *Biological Psychology*, *104*, 65–74.
- 1236 Golino, H. F., & Epskamp, S. (2017). Exploratory graph  
1237 analysis: A new approach for estimating the number of  
1238 dimensions in psychological research. *PloS One*, *12*(6),  
1239 e0174035.
- 1240 Golino, H., Shi, D., Christensen, A. P., Garrido, L. E., Ni-  
1241 to, M. D., Sadana, R., Thiyagarajan, J. A., & Martinez-  
1242 Molina, A. (2020). Investigating the performance of  
1243 exploratory graph analysis and traditional techniques to  
1244 identify the number of latent factors: A simulation and  
1245 tutorial. *Psychological Methods*, *25*(3), 292.
- 1246 Harrison, O. K., & Pink, A. (2024). Interoception and physi-  
1247 cal health. In *Interoception: A comprehensive guide* (pp.  
1248 227–264). Springer.
- 1249 Hickman, L., Seyedsalehi, A., Cook, J. L., Bird, G., & Mur-  
1250 phy, J. (2020). The relationship between heartbeat count-  
1251 ing and heartbeat discrimination: A meta-analysis. *Bio-*  
1252 *logical Psychology*, *156*, 107949.
- 1253 Hoekstra, R. A., Vinkhuyzen, A. A., Wheelwright, S., Bar-  
1254 tels, M., Boomsma, D. I., Baron-Cohen, S., Posthuma,  
1255 D., & Van Der Sluis, S. (2011). The construction and  
1256 validation of an abridged version of the autism-spectrum  
1257 quotient (AQ-short). *Journal of Autism and Develop-*  
1258 *mental Disorders*, *41*, 589–596.
- 1259 Jahedi, S., & Méndez, F. (2014). On the advantages and  
1260 disadvantages of subjective measures. *Journal of Eco-*  
1261 *nomic Behavior & Organization*, *98*, 97–114. <https://doi.org/10.1016/j.jebo.2013.12.016>
- 1262  
1263 Jenkinson, P. M., Fotopoulou, A., Ibañez, A., & Rossell, S.  
1264 (2024). Interoception in anxiety, depression, and psy-  
1265 chosis: A review. *EClinicalMedicine*, *73*.
- 1266 Jiménez, M., Abad, F. J., Garcia-Garzon, E., Golino, H.,  
1267 Christensen, A. P., & Garrido, L. E. (2023). Dimension-  
1268 ality assessment in bifactor structures with multiple gen-  
1269 eral factors: A network psychometrics approach. *Psy-*  
1270 *chological Methods*.
- 1271 Joshi, S. A., Aupperle, R. L., & Khalsa, S. S. (2023). Intero-  
1272 ception in fear learning and posttraumatic stress disorder.  
1273 *Focus*, *21*(3), 266–277.
- 1274 Khalsa, S. S., Adolphs, R., Cameron, O. G., Critchley, H.  
1275 D., Davenport, P. W., Feinstein, J. S., Feusner, J. D.,  
1276 Garfinkel, S. N., Lane, R. D., Mehling, W. E., et al.  
1277 (2018). Interoception and mental health: A roadmap.  
1278 *Biological Psychiatry: Cognitive Neuroscience and Neu-*  
1279 *roimaging*, *3*(6), 501–513.
- 1280 Khalsa, S. S., Adolphs, R., Cameron, O. G., Critchley, H.  
1281 D., Davenport, P. W., Feinstein, J. S., Feusner, J. D.,  
1282 Garfinkel, S. N., Lane, R. D., Mehling, W. E., Meuret,  
1283 A. E., Nemeroff, C. B., Oppenheimer, S., Petzschner, F.  
1284 H., Pollatos, O., Rhudy, J. L., Schramm, L. P., Simmons,  
1285 W. K., Stein, M. B., ... Zucker, N. (2018). Interoception  
1286 and Mental Health: A Roadmap. *Biological Psychiatry:*  
1287 *Cognitive Neuroscience and Neuroimaging*, *3*(6), 501–  
1288 513. <https://doi.org/10.1016/j.bpsc.2017.12.004>
- 1289 Kleckner, I. R., Wormwood, J. B., Simmons, W. K., Bar-  
1290 rett, L. F., & Quigley, K. S. (2015). Methodological rec-  
1291 ommendations for a heartbeat detection-based measure  
1292 of interoceptive sensitivity. *Psychophysiology*, *52*(11),  
1293 1432–1440.
- 1294 Koike, H., & Nomura, M. (2023). Development and val-  
1295 idation of japanese versions of the interoceptive accu-  
1296 racy scale and interoceptive attention scale. *SAGE Open*,  
1297 *13*(4), 21582440231214639.
- 1298 Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The  
1299 PHQ-9: Validity of a brief depression severity measure.  
1300 *Journal of General Internal Medicine*, *16*(9), 606–613.
- 1301 Kroenke, K., Spitzer, R. L., & Williams, J. B. (2002). The  
1302 PHQ-15: Validity of a new measure for evaluating the  
1303 severity of somatic symptoms. *Psychosomatic Medicine*,  
1304 *64*(2), 258–266.
- 1305 Kroenke, K., Spitzer, R. L., Williams, J. B., & Löwe, B.  
1306 (2009). An ultra-brief screening scale for anxiety and  
1307 depression: The PHQ-4. *Psychosomatics*, *50*(6), 613–  
1308 621.
- 1309 Kroenke, K., Spitzer, R. L., Williams, J. B., Monahan, P. O.,  
1310 & Löwe, B. (2007). Anxiety disorders in primary care:  
1311 Prevalence, impairment, comorbidity, and detection. *An-*  
1312 *nals of Internal Medicine*, *146*(5), 317–325.
- 1313 Lin, X.-X., Shen, H.-R., Lin, J.-X., Zhang, Y.-H., Murphy,  
1314 J., Wang, Y.-Z., Sun, Y.-B., Wang, N., Wang, J.-Y., Wei,  
1315 G.-X., & Luo, F. (2023). Psychometric validation and  
1316 refinement of the Chinese Interoceptive Accuracy Scale  
1317 (IAS) in general population and patients with chronic  
1318 pain. *Journal of Psychosomatic Research*, *175*, 111541.  
1319 <https://doi.org/10.1016/j.jpsychores.2023.111541>
- 1320 Longarzo, M., D'Olimpio, F., Chiavazzo, A., Santangelo,  
1321 G., Trojano, L., & Grossi, D. (2015). Self-awareness  
1322 questionnaire. *Frontiers in Psychology*.
- 1323 Lüdecke, D., Ben-Shachar, M. S., Patil, I., & Makowski, D.  
1324 (2020). Extracting, computing and exploring the parame-

- ters 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., Frackowiak, T., Karowski, 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*).

- 1431 Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. 1464  
1432 (2006). A brief measure for assessing generalized anxiety 1465  
1433 disorder: The GAD-7. *Archives of Internal Medicine*, 1466  
1434 166(10), 1092–1097. 1467
- 1435 Suksasilp, C., & Garfinkel, S. N. (2022). Towards a 1468  
1436 comprehensive assessment of interoception in a multi- 1469  
1437 dimensional framework. *Biological Psychology*, 168, 1470  
1438 108262. 1471
- 1439 Terasawa, Y., Motomura, K., Natsume, A., Iijima, K., 1472  
1440 Chalise, L., Sugiura, J., Yamamoto, H., Koyama, K., 1473  
1441 Wakabayashi, T., & Umeda, S. (2021). Effects of insular 1474  
1442 resection on interactions between cardiac interoception 1475  
1443 and emotion recognition. *Cortex*, 137, 271–281. 1476
- 1444 Thimm, J. C., Jordan, S., & Bach, B. (2016). The personal- 1477  
1445 ity inventory for DSM-5 short form (PID-5-SF): Psycho- 1478  
1446 metric properties and association with big five traits and 1479  
1447 pathological beliefs in a norwegian population. *BMC* 1480  
1448 *Psychology*, 4, 1–11. 1481
- 1449 Todd, J., Swami, V., Aspell, J. E., Furnham, A., Horne, G., 1482  
1450 & Stieger, S. (2022). Are some interoceptive sensibility 1483  
1451 components more central than others? Using item pool 1484  
1452 visualisation to understand the psychometric representa- 1485  
1453 tion of interoception. *Plos One*, 17(12), e0277894. 1486
- 1454 Torregrossa, L. J., Amedy, A., Roig, J., Prada, A., & Park, S. 1487  
1455 (2022). Interoceptive functioning in schizophrenia and 1488  
1456 schizotypy. *Schizophrenia Research*, 239, 151–159. 1489
- 1457 Vig, L., Ferentzi, E., & Köteles, F. (2023). Self-reported in- 1490  
1458 teroception, worries and protective behaviors during the 1491  
1459 COVID-19 pandemic: A longitudinal study. *Psicologia:* 1492  
1460 *Reflexão e Crítica*, 36(1), 23. 1493
- 1461 Vig, L., Köteles, F., & Ferentzi, E. (2022). Questionnaires 1494  
1462 of interoception do not assess the same construct. *Plos* 1495  
1463 *One*, 17(8), e0273299.
- 1464 Vlemincx, E., Walentynowicz, M., Zamariola, G., Van 1465  
1466 Oudenhove, L., & Luminet, O. (2023). A novel self- 1467  
1468 report scale of interoception: The three-domain intero- 1469  
1470 ceptive sensations questionnaire (THISQ). *Psychology* 1471  
1472 *& Health*, 38(9), 1234–1253. 1473
- 1469 Von Mohr, M., Silva, P. C., Vagnoni, E., Bracher, A., 1470  
1471 Bertoni, T., Serino, A., Banissy, M. J., Jenkinson, P. M., 1472  
1473 & Fotopoulou, A. (2023). My social comfort zone: At- 1474  
1475 tachment anxiety shapes peripersonal and interpersonal 1476  
1477 space. *Iscience*, 26(2). 1478
- 1474 Vorst, H. C., & Bermond, B. (2001). Validity and reliability 1475  
1476 of the bermond–vorst alexithymia questionnaire. *Person- 1477  
1478 ality and Individual Differences*, 30(3), 413–434. 1479
- 1477 Williams, Z. J., Suzman, E., Bordman, S. L., Markfeld, J. E., 1480  
1478 Kaiser, S. M., Dunham, K. A., Zoltowski, A. R., Failla, 1481  
1479 M. D., Cascio, C. J., & Woynaroski, T. G. (2023). Char- 1482  
1480 acterizing interoceptive differences in autism: A system- 1483  
1481 atic review and meta-analysis of case–control studies. 1484  
1482 *Journal of Autism and Developmental Disorders*, 53(3), 1485  
1483 947–962. 1486
- 1484 Zamariola, G., Frost, N., Van Oost, A., Corneille, O., & 1487  
1485 Luminet, O. (2019). Relationship between interoception 1488  
1486 and emotion regulation: New evidence from mixed meth- 1489  
1487 ods. *Journal of Affective Disorders*, 246, 480–485. 1490
- 1488 Zanarini, M. C. (2003). Zanarini rating scale for borderline 1491  
1489 personality disorder (ZAN-BPD): A continuous measure 1492  
1490 of DSM-IV borderline psychopathology. *Journal of Per- 1493  
1491 sonality Disorders*, 17(3), 233–242. 1494
- 1492 Zsido, A. N., Teleki, S. A., Csokasi, K., Rozsa, S., & Bandi, 1495  
1493 S. A. (2020). Development of the short version of the 1494  
1494 spielberger state—trait anxiety inventory. *Psychiatry Re- 1495  
1495 search*, 291, 113223.