Measurement of interoception and its role in autism and anxiety



Kiera Louise Adams

University College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2025

Measurement of interoception and its role in autism and anxiety

Kiera Louise Adams

University College, University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2025

Supervisor: Geoffrey Bird, PhD

Secondary supervisor: Polly Waite, PhD

Word count: 71,816

Acknowledgements

There are so many people without whom this thesis would not have been possible.

First and foremost, I would like to thank my amazing participants- I am so grateful for the time that each and every one of you has contributed to my studies! Without you, there would be no research. In particular, I would like to thank my sample of autistic adolescents from Chapter Six. Interviewing you was the highlight of my DPhil and I am in awe of the bravery you showed in speaking so candidly about such sensitive topics at a young age.

Thank you to my funders: the Clarendon Trust, University College, and the Experimental Psychology Department. Your generosity has allowed me to follow my passion and for that I am so grateful. Thank you also to Cecil Anthony-Mike for funding my attendance at INSAR, an experience that was so valuable to me!

Geoff Bird- I cannot imagine my rollercoaster of a PhD journey with any other supervisor. I am so grateful for all of our intense brain-workouts, for you giving me the freedom to explore the things I am most interested in, and for your constant advocacy of me. You never doubted in my ability as a student, even when I was doubting it myself. Thank you for *caring* and allowing me to develop both as a student and a person.

Polly Waite- your support when I was going through difficult times means more to me than I can put into words. Thank you for always balancing professionalism and warmth and sparking my passion for qualitative research. It has been an absolute pleasure.

Caroline Catmur- I am constantly in awe of your ability to drill down into all the small details that I tend to miss when I'm distracted by the big picture. I have learnt so much from you in our lab meetings and I feel privileged to have worked with you- I am undoubtedly a better scientist because of it!

To all my collaborators, it has been a joy to work alongside each and every one of you. Jenny Murphy- you have been a role model for me since the beginning of my DPhil. Thank you for letting my bounce ideas off you and inspiring me with your passion. And thank you for all the times you have let me rant to you when academia was getting too much. Will Mandy- it has been an absolute pleasure to work with you. Your knowledge of psychology is so endlessly impressive to me. I hope we get the chance to work together more in the future! Jonathan Bird- thank you for all your hard work and guiding me through my many queries about r, every meeting we have had has been an absolute joy.

To everyone in the Bird/Catmur Lab- my PhD would not have been the same without you geniuses! Our lab meetings were so challenging but taught me so much as a researcher! I have enjoyed our camaraderie over the years, and I am yet to encounter a logistical problem we could not all solve together over a pack of Percy Pigs.

Joanna Gibbs- you are an absolute inspiration to me! Thank you for all of your guidance, helping me to network, and aiding me with my recruitment.

Thank you to my support networks both in and outside of Oxford. To the beautiful OUPole family- competing and training alongside all of you has kept me (sort of) sane throughout this PhD and given me a confidence that I take with me on and off the pole. To the gymnasts, whilst I can't say that my prowess on the vault was my proudest achievement in Oxford (#faceplantqueen), the connections I made through the club were so important to me! I will miss all of the unhinged late-night conversations on Jason's sofa. To the Babes Who Brunch- I love you all so deeply, thank you for supporting me through all my dramas and reminding me to always demand respect for myself. To the Fri-yay Homertonians, We Saved Fabric group and the BGS squad- thank you for all the years of friendship, you guys are the real day

ones! Having a laugh or a dance or a DMC with all of you has provided much needed respite from the stress of academia.

Oliver Carter- you helped me so massively with recruitment and also took me out and made me eat on the days when I genuinely didn't think I would be able to leave my bed. I love you, King! Pia Pallot- you have been the best friend I could ever ask for and without you I don't think I would be here today, never mind on my way to becoming a Doctor. Ditto Charlotte Collingwood, who has met my every dark period with an unflappable support, deep care, and a willingness to go for our fifth desi breakfast of the week. Georgie Windsor and Claudio Castelnovo, you have both added some much-needed sparkle to my PhD years- I love you both so much!

Finally, thank you to my amazing family, who have supported me through thick and thin. Special thanks have to go to my wonderful mum, Anne Brown. I owe absolutely everything to you- you are my hero.

Declaration

This thesis is submitted to University of Oxford in support of Kiera Adams' application for the degree of Doctor of Philosophy. It has been composed by Kiera Adams and has not been submitted in any previous application for any other degree.

This thesis includes partial copies of published papers and papers currently under review at peer-reviewed journals. Kiera Adams is the (or a) lead author on all of these papers and authored the figures and text reproduced here. This work was formatted specifically for inclusion in this thesis. The associated, published papers and authors' contributions are included at the beginning of the relevant chapter. All co-authors on any published papers have given approval for this material to be submitted as part of this dissertation.

As well as the introduction and discussion sections that are specific to each chapter, this thesis includes a general introduction, including a literature review, and a general discussion of the main findings, interpretations, limitations and outstanding questions of this thesis.

Abbreviations

Note. Each abbreviation will be defined when it is first used in each chapter. All abbreviations used throughout the thesis are listed here.

2AFC: two-alternative forced-choice

6AFC: six-alternative forced-choice

AD: anxiety disorder

AN: anorexia nervosa

ANCOVA: analysis of covariance

ANOVA: analysis of variance

AQ: autism quotient

ASC: autism spectrum condition

ASI: Anxiety Sensitivity Index

BAI: Beck Anxiety Inventory

BAQ: Body Awareness Questionnaire

BMI: body mass index

BN: bulimia nervosa

BPQ: Body Perception Questionnaire

CASI: Child Anxiety Sensitivity Index

CATI: Comprehensive Autistic Trait Inventory

CD: clinical diagnosis

COPD: chronic obstructive pulmonary disease

DALYs: disease-adjusted life years

DASS-A: Depression and Anxiety Stress Scale – Anxiety

DL: dyslipidaemia

DPD: depersonalisation derealisation disorder

ECG: electrocardiogram

EDI: Eating Disorder Inventory

EFA: exploratory factor analysis

ES: effect size

F:M: female to male

FDR: false discovery rate

fMRI: functional magnetic resonance imaging

FNE: Fear of Negative Evaluation

GAD: generalised anxiety disorder

HBP: high blood pressure

HC: healthy control

HCT: heartbeat counting task

HDT: heartbeat detection task

HEP: heartbeat evoked potentials

IAS: Interoceptive Accuracy Scale

IATS: Interoceptive Attention Scale

IBI: inter-beat interval

IP: infrequent panic

ISAQ: Interoceptive Sensitivity and Attention Questionnaire

M: mean

MAIA: Multidimensional Assessment of Interoceptive Awareness

MCS: method of constant stimulation

MMD: major depressive disorder

MUS: medically unexplained symptoms

NIES: Negative Interoceptive Evaluation Scale

NR: not reported

NRMSE: normalized root mean square error

NS-SEC: National Statistics Socio-economic Classification

OA: Osteoarthritis

OCD: obsessive compulsive disorder

PAT: phase adjustment task

PCA: principal component analysis

PD: panic disorder

PES: pooled effect size

PFC: proportion of false classification

PH: primary hypertension

PHA: pathological health anxiety

PPI: public and patient involvement

PRISMA: Preferred Reporting Items for Systematic Review and Meta-Analyses

PST: phase selection task

Q: questionnaire

RCADS: Revised Children's Anxiety and Depression Scale

SAD: social anxiety disorder

SASC-R: Social Anxiety Scale for Children- Revised

SCARED: Screen for Child Anxiety Related Disorders

SoD: somatic disorder

SD: standard deviation

SeAD: separation anxiety disorder

SHAI: Short Health Anxiety Inventory

SM: selective mutism

SP: specific phobias

SPRAS: Sheehan Patient-Rated Anxiety Scale

STAI-S: State-Trait Anxiety Inventory- State

STAI-T: State-Trait Anxiety Inventory- Trait

TA: thematic analysis

TD: tic disorder

TDIS-Q: Three Domain Interoceptive Sensations Questionnaire

WI: Whitely Index

Contents

Acknowledgements ... 3
Declaration ... 6
Abbreviations ... 7
List of figures ... 15
List of tables ... 17
Abstract ... 20
Chapter One: Introduction ... 22
Defining interoception ... 22
Dimensions of interoception ... 22
Interoceptive domains ... 25
Level of definition ... 26
Why research interoception? ... 29
Anxiety and interoception ... 30
Autism and interoception ... 38
Alexithymia ... 45
Aims of this thesis ... 47
Research questions ... 49
Chapter Two: Measuring interoception- Methodological issues ... 51
Abstract ... 53
Critique of the most commonly used methods to measure interoception ... 54
Introduction ... 54
General critique ... 54
Self-report measures ... 54
Objective measures of interoceptive accuracy ... 58
Situational approach ... 67
Accuracy and bias ... 67
Confounders- do they matter? ... 68
Conclusion ... 69
Empirical evaluation of the 2AFC-HDT ... 70
Introduction ... 70
Methods ... 75
Participants ... 75
The PAT ... 81
Scoring ... 82
Analysis ... 84
Results ... 91

Analysis approach 1 ... *91*
Analysis approach 2 ... *92*
Discussion ... **94**
Conclusion ... **99**
Chapter Three: The Phase Selection Task Pilot ... **101**
Abstract ... **102**
Introduction ... **103**
Methods ... **104**
Participants ... *104*
The App ... *104*
Procedure ... *106*
Analysis ... *109*
Results ... **110**
Discussion ... **118**
Chapter Four: Validation of the Negative Interoceptive Evaluation Scale ... **122**
Abstract ... **123**
Introduction ... **124**
Methods ... **127**
Development of the Negative Interoceptive Evaluation Scale (NIES) ... *127*
Participants ... *128*
Measures ... *129*
Procedure ... *130*
Analysis ... *131*
Results ... **135**
Sample One ... *135*
Sample Two ... *138*
Item breakdown ... *140*
Convergent and divergent validity ... *145*
Test-retest reliability ... *147*
Group comparisons ... *147*
Questionnaire feedback ... *150*
EFA ... *152*
CFA ... *162*
Factor correlations ... *165*
Internal consistency ... *167*
Discussion ... **168**
Limitations and future directions ... *173*

Conclusion ... *174*
Chapter Five: The association between anxiety and interoception ... **175**
Abstract ... **176**
Introduction ... **177**
Method ... **180**
Search strategy ... *180*
Study selection and extraction ... *181*
Measures used: interoceptive accuracy ... *184*
Measures used: anxiety questionnaires ... *185*
Heart rate manipulation ... *199*
Clinical diagnoses ... *199*
Analysis strategy ... *199*
Results ... **201**
HCT and variants with trait anxiety ... *201*
HCT and variants with state anxiety ... *203*
HDT and variants with trait anxiety ... *205*
Moderation analysis ... *206*
Pre-registered moderators ... *208*
Exploratory moderators ... *209*
Discussion ... **214**
Conclusion ... *218*
Chapter Six: Interoception and anxiety in autistic adolescents ... **219**
Abstract ... **220**
Introduction ... **221**
Methods ... **226**
Study design ... *226*
Positioning statement ... *226*
Participants and sampling ... *227*
Public and patient involvement (PPI) ... *231*
Topic guide ... *233*
Procedure ... *234*
Analysis ... *236*
Assuring validity ... *237*
Results ... **238**
Theme 1: Interoceptive awareness ... *240*
Theme 2: Interoceptive Evaluation ... *243*
Theme 3: Situational Interoception ... *246*

Discussion ... **253**

Chapter Seven: The associations between self-reported dimensions of interoception and anxiety and autism ... **259**

Abstract ... **260**

Introduction ... **261**

Autism and interoception ... ***261***

Anxiety and interoception ... ***265***

Anxiety, autism and interoception ... ***266***

Covariates ... ***267***

The current study ... ***268***

Hypotheses ... ***268***

Methods ... **269**

Participants ... ***269***

Materials ... ***270***

Procedure ... ***272***

Data cleaning ... ***273***

Analysis ... ***274***

Results ... **278**

Trait level analyses ... ***278***

Autistic versus non-autistic ... ***282***

Anxiety disorder versus no anxiety disorder ... ***288***

Distributions across groups ... ***295***

Does autism moderate associations between interoception and anxiety? ... ***303***

Exploratory analysis ... ***307***

Discussion ... **313**

Chapter Eight: General Discussion ... **321**

Conclusion one: The need to employ caution when interpreting previous interoception research ... **322**

Conclusion Two: Interoceptive accuracy may not be as important as previously thought ... **324**

Conclusion Three: Importance of interoceptive evaluation ... **325**

Conclusion Four: Relationship between anxiety and interoception ... **328**

Conclusion Five: Relationship between autism and interoception ... **331**

General implications ... **335**

Interventions ... ***335***

Psychoeducation ... ***337***

Strengths of the work reported in this thesis ... **338**

Limitations ... **341**

Future directions ... **343**

Conclusion **345**
References **346**
Appendix **415**
Chapter Two **415**
Chapter Three **418**
Chapter Four **419**
Chapter Six **429**

List of figures

Figure 2.1 *Categorising delay choices as synchronous, asynchronous, or between*……...88

Figure 2.2 *Distribution of final IBIs across trials*……90

Figure 2.3 *Histogram of participants' median delay choices*……...92

Figure 2.4 *Delay choices within the cardiac cycle across all trials*…….93

Figure 2.5 *Delay choices within the cardiac cycle on trials where the IBI was between 650 to 750ms*……...94

Figure 4.1 *Scree Plot for initial EFA*……..153

Figure 4.2 *Scree plot for EFA with Question 7 removed*……....158

Figure 4.3 *Path diagram of Model 5*……..165

Figure 5.1 *Flow diagram of studies included in the meta-analysis*……...183

Figure 5.2 *Forest plot: trait anxiety and HCT variants*…….203

Figure 5.3 *Funnel plot: trait anxiety and HCT variants*……204

Figure 5.4 *Forest plot: state anxiety and HCT variants*…….....205

Figure 5.5 *Funnel plot: state anxiety and HCT variants*……206

Figure 5.6 *Forest plot: trait anxiety and HDT variants*…….207

Figure 5.7 *Funnel plot: trait anxiety and HDT variants*……...…….207

Figure 6.1 *The interview set up*……..236

Figure 6.2 *Themes and subthemes*……..240

Figure 6.3 *Pathways by which interoception may affect anxiety*……...254

Figure 7.1 *Interoception score distributions in autistic and non-autistic participants*……..297

Figure 7.2 *Interoception score distributions in participants with and without anxiety disorders*…….301

Figure 7.3 *A graph to show the effect of the interaction between interoceptive attention and evaluation on anxious traits*……...……310

List of tables

Table 2.1 *Inclusion and exclusion criteria*……...77

Table 2.2 *Demographic information for the interoceptive participants across all studies*…..84

Table 2.3 *Descriptive Statistics of the Different IBIs*……...85

Table 3.1 *Debriefing questions*……...108

Table 3.2 *Demographic information*……...111

Table 3.3 *Participant responses to the debriefing questions*……..……113

Table 3.4 *I found it possible to order the shapes and sounds in order of time after my actual heartbeat*……...…..116

Table 3.5 *I can usually feel my heartbeat, even if not using the app*..……117

Table 3.6 *I found that I could feel my heartbeat whilst using the app*……...……117

Table 4.1 *Questionnaire results for Sample One*……136

Table 4.2 *Questionnaire results for Sample Two*……...…….139

Table 4.3 *Item breakdown Sample One*……..140

Table 4.4 *Item breakdown Sample Two*……..143

Table 4.5 *Convergent and divergent validity across both* samples……145

Table 4.6 *Demographic information for the feedback sample*……...150

Table 4.7 *Number of factors suggested by each method*……..…..153

Table 4.8 *Variance explained*…….154

Table 4.9 *Communalities and factor loadings of all items*…….154

Table 4.10 *Factor correlation matrix*…….155

Table 4.11 *Three factor model suggested by EFA*……..156

Table 4.12 *Number of factors suggested for each method*…….158

Table 4.13 *Scree plot for EFA with Question 7 removed*……..……..159

Table 4.14 *Communalities and factor loadings of all items*……...159

Table 4.15 *Factor correlation matrix*……160

Table 4.16 *Three factor model suggested by EFA (with Question 7 removed)*……...161

Table 4.17 *CFA model fit statistics*……163

Table 4.18 *Model fit statistics when including correlated residuals*……..164

Table 4.19 *Correlations between factors and the ASI*……166

Table 4.20 *Internal consistency of the three factors*……...168

Table 5.1 *Anxiety questionnaires*……....187

Table 5.2 *HCT studies*……..……...192

Table 5.3 *HDT studies*……..…...198

Table 5.4 *Moderation analysis*……...208

Table 6.1 *Demographic information for interviewees*……230

Table 6.2 *Demographic information for the adolescent PPI group*…….……...233

Table 6.3 *Demographic information for the parent PPI group*…….……233

Table 7.1 *Other mental health and neurodevelopmental conditions reported by the whole sample*…….……280

Table 7.2 *Demographic information for the whole sample*……..……..281

Table 7.3 *Other mental health and neurodevelopmental conditions reported by the autistic group*……...284

Table 7.4 *Descriptive statistics for the autistic group*……285

Table 7.5 *Other mental health and neurodevelopmental conditions reported by the non-autistic group*……..……286

Table 7.6 *Descriptive information for the non-autistic group*……....287

Table 7.7 *Other mental health and neurodevelopmental conditions reported by the anxiety disorder group*……..……..290

Table 7.8 *Descriptive information for the anxiety disorder group*……291

Table 7.9 *Other mental health and neurodevelopmental conditions reported by the non-anxiety disorder group*……292

Table 7.10 *Descriptive information for the non-anxiety disorder group*……293

Table 7.11 *Demographic information for the anxiety disorder group for the logistic regression*……...306

Table 7.12 *Demographic information for the non-anxiety disorder group for the logistic regression*……...307

Abstract

Interoception, defined as the detection and interpretation of internal sensory information, is said to underlie a number of psychological processes. Despite considerable research into interoception, our understanding remains limited. This thesis had two primary objectives. First, to critically evaluate widely used measures of interoception and to highlight possible alternatives to more problematic measures. Second, to deepen understanding of the role of interoception in anxiety, autism, and the overlap between the two.

Chapter One provides an introduction to the thesis, outlining relevant empirical and theoretical work in a literature review. I compare various definitions of interoception, before outlining the definition that will be employed throughout this thesis. I also provide an overview of current knowledge of the role of interoception in anxiety and autism.

Chapter Two explores the limitations of widely used measures of the most researched aspects of interoception: self-reported interoceptive "awareness" and objective measures of interoceptive accuracy in the cardiac, respiratory and gastric domains. In this chapter, I also present an analysis of secondary data which empirically demonstrates that the heartbeat detection task, a widely used measure of cardiac interoceptive accuracy, may result in false negative results.

Chapter Three presents a pilot study of the Phase Selection Task (PST) a novel measure of cardiac interoceptive accuracy which avoids the limitations of the heartbeat detection task outlined in Chapter Two. The study confirmed that the PST can be used to assess cardiac interoceptive accuracy remotely.

Chapter Four focuses on a neglected construct of interoception: interoceptive evaluation. I introduce the Negative Interoceptive Evaluation Scale (NIES), a questionnaire that can be used to assess how negatively an individual evaluates interoceptive sensations

across seven domains. Test-retest reliability, group comparisons and convergent and divergent validity checks of the questionnaire suggest that the NIES is a valid measure, although exploratory and confirmatory factor analyses did not identify a consistent factor structure.

Chapter Five presents a meta-analysis of the relationship between cardiac interoceptive accuracy and anxiety. Findings showed a lack of any relationship, even when accounting for a number of potential moderators.

Chapter Six explores the role of interoception in the experiences of anxiety in a group of autistic adolescents. I conducted interviews which were then analysed using a reflexive thematic analysis approach. Results emphasised the importance of negative interoceptive evaluation, and that both hyper and hypo awareness of interoceptive sensations may lead to anxiety.

Chapter Seven presents a large-scale questionnaire-based study. The associations between self-reported interoceptive accuracy, attention and evaluation and anxiety and autism were explored. Findings revealed that greater interoceptive attention was associated with higher trait anxiety and autistic traits, as well as anxiety disorders and autism diagnoses. More negative interoceptive evaluation was associated with higher trait anxiety, higher autistic traits and diagnoses of autism. Lower interoceptive accuracy was associated with higher autistic traits.

In Chapter Eight, I recontextualise this work, discussing implications of findings and suggesting avenues for future research. This body of work sets the groundwork for more valid and reliable research into interoception, by illustrating the limitations of commonly used measures and providing some alternatives. It also allows for a better understanding of the role of interoceptive in autism and anxiety.

Chapter One: Introduction

Defining interoception

Sherrington (1906) first coined the term “interoceptive” to describe signals arising from the internal surface of the body, in contrast to exteroceptive (external) and proprioceptive (muscle) signals. Interoception received a renewed focus almost a century later, with Craig defining interoception as the processing of internal, bodily signals pertaining to homeostasis (Craig, 2002). Since then, interoception has become an increasingly popular topic of research (Khalsa et al., 2018). However, despite the prolific research interest, precisely defining interoception has proved contentious, with much debate as to the specific signals and pathways involved. This section will provide a brief overview of the different ways interoception has been defined in the literature, before outlining how interoception will be defined within the context of this thesis.

Dimensions of interoception

Interoception is known to be a multidimensional construct, but the number, nature, and name of these dimensions vary (see Desmedt et al., 2023; Khalsa et al., 2018; Murphy et al., 2020; Suksasilp & Garfinkel, 2022). The terminology that has most commonly been used in the literature splits interoception into three dimensions: interoceptive sensitivity (the objectively measurable ability to detect physiological activity), interoceptive sensibility (subjective awareness of physiological activity, measured by self-report) and metacognitive interoceptive awareness (the extent to which sensitivity matches sensibility) (Garfinkel & Critchley, 2013). However, recent concerns about the conflation of subjective accuracy and attention within the construct of “sensibility" have led to a refined model that this thesis will draw from. This model splits interoceptive accuracy (whether an individual can correctly

identify if, when, or the degree to which a particular interoceptive event has occurred) from interoceptive attention (the degree to which an individual's attention is focused on interoceptive signals, encompassing the amount of time they are aware, and how salient these signals are; (Murphy, Catmur, et al., 2018)), both of which can be measured objectively and subjectively. The difference between the two can be illustrated with an example. An individual might constantly feel hungry, even when they have just eaten. They have a high degree of interoceptive attention, as the proportion of the time they spend aware of the interoceptive signal of hunger is high. However, their interoceptive accuracy is low, as that hunger is not reflective of their body's need for food. In addition to interoceptive accuracy and attention, this thesis will also focus on interoceptive evaluation, which refers to how interoceptive signals are interpreted, including aspects such as affective experience and the meaning individuals take from certain signals (Herbert, 2020). These three dimensions do not reflect an exhaustive list of all possible interoceptive dimensions but were deemed the three dimensions where the most opportunity existed to develop existing methodologies, and where the strongest relationships to mental health are expected to exist.

Preliminary research supports the independence of interoceptive attention, accuracy, and evaluation. For example, dissociations have been observed between interoceptive accuracy and attention (Garfinkel et al., 2015; Murphy et al., 2020; Pollatos et al., 2016; Yang et al., 2023), interoceptive accuracy and evaluation (Herbert et al., 2013; Pollatos et al., 2016; Yang et al., 2023), and interoceptive attention and evaluation (Desmedt, Heeren, et al., 2022). This independence means it is important to study each dimension of interoception individually, especially when considering their effects on psychological outcomes.

These independent dimensions of interoception likely interact with one another. For instance, if one has extremely negatively biased evaluations of interoceptive signals, one may pay more attention to interoceptive signals that are perceived as a sign of a threat. An

example would be if someone who interprets their heart beating fast as a sign that they are going to have a panic attack subsequently attends more to their heartbeat. There appears to be some indirect support for this, as research finds that in a negative affective context, dyspnea (shortness of breath) elicits more respiratory symptom reporting than in a positive affective context (von Leupoldt & Dahme, 2013). Another association can be illustrated by the example of an individual with poor interoceptive accuracy, who may attend to signals less as they are seen as untrustworthy, or more, as they are seen as unpredictable, which captures attention (Itti & Baldi, 2009). Due to this potential for interplay across dimensions, it is important not to study the dimensions of interoception in isolation but to explore them within the context of one another (Suksasilp & Garfinkel, 2022).

One consideration that must be made when defining interoception is whether the varying dimensions are best conceptualised as state or trait features. Whilst a trait indicates a feature of an individual that persists over time and situations, states are short-lived and changeable. Although there is a degree of stability over time in interoceptive accuracy and attention (for example, self-reported interoceptive accuracy and attention both show good test-retest reliability over thirty days, Murphy et al., 2020) preliminary work seems to suggest that there is also state-like fluctuation in these dimensions (Murphy et al., 2017). Logically, one might expect that if one is in a situation where interoceptive signals are made stronger (for example, if one is in a situation that increases one's blood pressure), the signal will be easier to detect, and therefore accuracy and attention may increase. Indeed, evidence suggests stress, physical exercise and food deprivation can all modify individual cardiac interoceptive accuracy and interoceptive evaluation (Durlik et al., 2014; Herbert et al., 2010, 2012; Jones & Hollandsworth, 1981; Schulz et al., 2013; Schulz & Vögele, 2015). Therefore, within this thesis, each dimension of interoception will be assumed to be a trait, with state-like fluctuations.

Interoceptive domains

Interoception can be further split into domains (with each domain reflecting a specific physiological channel), although there is significant debate about which exact signals are deemed interoceptive (Desmedt, Luminet, Maurage, et al., 2023). The most agreed upon (and therefore those that are predominantly focused on within this thesis) are cardiac, respiratory, gastric, fluid regulation, and internal pain (Nord & Garfinkel, 2022), whilst more contentious signals that are detected on the surface of the skin rather than inside the body, such as temperature and external pain, are not discussed within this thesis. Evidence from neurological studies suggests the domains of interoception are independent at lower levels but combined at higher levels of processing. For example, Paton and colleagues (2000) provide evidence of the convergence of signals from distinct organs and viscero receptors in the nucleus of the solitary tract. Another study demonstrated that both the heart and stomach have overlapping cortical representations in the insula, a region of the brain which is commonly associated with interoception (Avery et al., 2015). Smith and Lane (2015) found evidence for separate mappings in the somatosensory cortices, posterior insula, nucleus of the tractus solitarii, hypothalamus, and parabrachial nucleus, but whole-body mappings in the mid and anterior insula. More recently, a study found evidence that asking participants to focus on either cardiac or gastric sensations resulted in differing patterns of brain activation within the left dorsal middle insula, but similar patterns in the posterior insula (Haruki & Ogawa, 2023).

Substantiating neurological findings of the independence of interoceptive domains, research shows that the domains of interoception also appear to be independent within specific dimensions. Although some evidence suggests that there is a degree of coherence in the accuracy of perception across domains (e.g., between cardiac and gastric interoception;

Herbert et al., 2012; Whitehead & Drescher, 1980), other evidence suggests a dissociation - for example between objective accuracy in detecting cardiac and respiratory signals (Garfinkel, Tiley, et al., 2016; Pollatos et al., 2016) and across cardiac, gastric, taste, and vestibular perception (Ferentzi et al., 2018), although this dissociation may be due to differences in task demands. One study, which aimed to validate a novel measure of interoceptive "perception," which is defined as the noticing and feeling of interoceptive signals, found a distinction between cardiorespiratory and gastroesophageal sensations in both exploratory and confirmatory factor analyses (Vlemincx et al., 2023), suggesting there may be some distinction between domains concerning interoceptive attention. However, there is a lack of other research comparing interoceptive attention across different domains, and to our knowledge, no research doing this for interoceptive evaluation.

Level of definition

Some researchers elect to define interoception in terms of the physiological pathways involved. One example comes from Craig (2002), who limits interoceptive signals to those involving homeostatic afferent pathways which involve the lamina I spinothalamic afferents, whereby visceral signals are relayed by the spinal cord, brainstem and nucleus of the solitary tract to the insular cortex. Critchley and colleagues (2004) also define interoception from a physiological standpoint, limiting their definition to visceroreceptors relaying information to the nucleus of the solitary tract, or thermoreceptors and nociceptors relaying information to the lamina I neurons; both of which send information to the thalamus, amygdala, hypothalamus, anterior cingulate cortex and insula.

However, Desmedt and colleagues warn of the limitations of purely physiological definitions, highlighting the near impossibility of empirically testing for this, as well as the fact that the processing of internal signals often relies upon a range of different pathways,

some of which are non-homeostatic (2023). The latter point can be illustrated with the example of cardiac interoception. Cardiac perception may occur either from interoceptive (baroreceptors and cardiac afferent neurons) or non-interoceptive (the somatosensory system) sources (Khalsa et al., 2009). It is difficult to uncover which of these is being used at a particular moment in a specific individual. Despite full denervation of the heart and aortic arch, some heart transplant recipients, or those with extracorporeal ventricular assist devices, may accurately feel their heartbeats, meaning one may interpret interoceptive information in the absence of interoceptive pathways (Barsky et al., 1998; Couto et al., 2014). Khalsa and colleagues further illustrated that a patient with extensive bilateral damage to interoception-related brain areas (the insula and anterior cingulate cortex) showed preserved interoceptive accuracy, which was only impaired following anaesthetising of the chest, suggesting the role of somatosensory information in "interoception" (Khalsa et al., 2009). Somatosensory cortices have also been shown to be activated in healthy participants when focusing on heartbeats (Critchley et al., 2004; Pollatos et al., 2007), and in other studies transcranial magnetic stimulation of the somatosensory cortex also influenced heartbeat perception (Pollatos et al., 2016). Moreover, the application of vibrotactile stimulation on the sternum, masking somatosensory signals, results in worse cardiac interoceptive accuracy (Knapp-Kline et al., 2021). Taken together, these results suggest that the detection and interpretation of cardiac information relies on physiological systems that would be considered interoceptive and exteroceptive. Similarly, both respiratory and gastric interoception have been demonstrated to rely on several different receptor systems, both interoceptive and exteroceptive (Camilleri et al., 2001; Harver et al., 1988).

The specific pathways involved in interoception may not matter from a practical standpoint, if one chooses to define interoception from a phenomenological perspective. For example, Chen and colleagues (2021) define interoception as the representation of the

internal states of an organism. The type of signal is irrelevant, so long as it provides information that one can use in their representation of the body state. Similarly, Ceunen and colleagues define interoception as multimodal integration of information about the subjective body state, not restricted to any sensory channel or mere sensation (Ceunen et al., 2016). In line with these definitions, it is increasingly accepted that the exact route of processing and whether it is purely interoceptive in a strict sense is not as important as the way the signals are interpreted, and whether it provides information about internal conditions of the body. As such, Desmedt and colleagues (2023) define interoception as "the top-down and bottom-up processes by which an organism senses, interprets, and integrates signals from within itself and below the skin, across conscious and non-conscious levels." They work back from this definition, characterising the physiological mechanisms involved as detection by chemoreceptors, humoral receptors, mechanoreceptors and nociceptors, which are converted to electrical, hormonal or non-neural signals and processed in subcortical structures of the brain before being projected to higher brain regions such as the hypothalamus, insula, anterior cingulate cortex, and somatosensory cortex.

This thesis will draw from Desmedt's definition (Desmedt et al., 2023), defining interoception as the processing of internal signals, regardless of the specific physiological pathways involved. As discussed, "internal signals" will be limited to those signals that have been most unanimously agreed upon as interoceptive in the literature: cardiac, respiratory, gastric, fluid regulatory and internal pain-related sensations. Although this processing may occur at both conscious and unconscious levels, the work in this thesis will focus on that which is consciously processed. Given the phenomenological, rather than biological, level at which interoception is henceforth defined, the work discussed in this thesis will be primarily behavioural rather than neurological, although reference to relevant neurological studies will be made where appropriate.

Why research interoception?

Interoception has been implicated in many psychological outcomes. If one subscribes to the Somatic Marker Hypothesis (Damasio, 1994), or indeed its precursor, the James Lange theory of emotion (Lange & James, 1967), then it is assumed that and individual's detection and perception of their internal physiological state gives rise to the subjective experience of emotions. Therefore, interoception should influence one's affective state (Domschke et al., 2010) and is in turn said to influence decision-making (Dunn, Galton, et al., 2010). Moreover, interoception has been linked to social cognition and judgement (Quattrocki & Friston, 2014; Feldman et al., 2023), agency (Koreki et al., 2022), action understanding (Zhou et al., 2022) and even consciousness (Damasio & Damasio, 2022). There is also an intuitive link to physical health, due to the role interoception plays in homeostasis, and physical health is also known to have an influence on one's mental health (Hernandez et al., 2018).

Given the wide array of ways interoception influences our psychology, it is unsurprising that interoception has been implicated in the aetiology of numerous mental health conditions (Brewer et al., 2021). These widespread theories and associations have led some to argue that interoception could be the p-factor proposed by Caspi and colleagues (Caspi et al., 2014) that is said to underly the high co-occurrence of several mental health conditions. Research into interoception allows us to establish whether the p-factor account could be true, and if it is, it would also provide us with a better understanding of one mechanism underlying the emergence and maintenance of mental health conditions.

Another reason that interoception is an appealing topic for research is that evidence has shown that interoceptive function is amenable to change over time, or through treatment (Farb et al., 2015; Khalsa et al., 2018), meaning that interoceptive styles that have been

demonstrated to be problematic for mental health have the potential to be addressed. Indeed, interventions have been trialled that seek to train interoceptive accuracy (e.g.: Quadt et al., 2021) and interoceptive evaluation (Boettcher et al., 2016).

As illustrated in this chapter, defining interoception is far from straightforward, which makes collating research and forming theories exceedingly difficult. Moreover, measuring interoception has proven difficult, with many of the most commonly used measures of interoception showing significant validity issues (a comprehensive discussion of this will be given in Chapter Two). However, given the exponential rate at which research on interoception is being published and the enormous potential for practical application, it is of the utmost importance that interoception research is high quality and based on clearly defined terms.

Anxiety and interoception

Anxiety, defined by excessive fear or worry (American Psychiatric Association, 2013), can be conceptualised dimensionally, and forms discrete disorders at the extreme end of the spectrum. When measured dimensionally, one can either focus on state anxiety, which is how one feels in that exact moment, or trait anxiety, which refers to a persisting characteristic of an individual (Spielberger et al., 1983). Those who experience high levels of trait anxiety may be diagnosed as having an anxiety disorder. Anxiety disorders are a subset of mental health conditions defined by excessive fear or worry, including panic disorder (PD), generalised anxiety disorder (GAD), social anxiety disorder (SAD), specific phobias (SP), agoraphobia, separation anxiety disorder (SeAD) and selective mutism (SM) (American Psychiatric Association, 2013). Anxiety disorders have been estimated to affect around 4% of the global population at a given time, resulting in a global disease burden of around 28.68 million disease-adjusted life years (DALYs; Yang et al., 2021). Moreover, around one in

three individuals will experience an anxiety disorder at some point in their life (Bandelow & Michaelis, 2015). If untreated, anxiety disorders can result in pervasive detrimental outcomes, ranging from a reduced likelihood of employment and poorer education attainment (Auerbach et al., 2016; Bruffaerts et al., 2018; Mojtabai et al., 2015), to persistent and recurrent physical and mental health problems (Kaufman & Charney, 2000; Scott et al., 2007). To minimise these outcomes through efficacious interventions, a comprehensive knowledge of the mechanisms underlying anxiety is required.

Of all conditions, anxiety is most often thought of as the prototypical disorder of interoception (Khalsa & Lapidus, 2016). Indeed, physiological symptoms, such as a racing heart rate and rapid breathing, are included in the diagnostic criteria for PD specifically (American Psychiatric Association, 2013; World Health Organization, 2022). One seminal theory linked interoception to PD (Ehlers & Breuer, 1992), with accompanying research focusing mainly on cardiac interoceptive accuracy. This theory outlined that those with a better perception of their bodily signals would be more aware of fluctuations in their heart rate, allowing for potentially catastrophic misinterpretations of somatic cues. For example, a slight increase in one's heart rate may be interpreted as a sign that something is wrong, resulting in heightened anxiety that further increases one's heart rate, and so on (Clark, 1986). This positive feedback loop was hypothesised to cause panic attacks. There is empirical support for this theory, with case-control studies demonstrating more accurate heartbeat perception in those with PD than healthy controls (Ehlers & Breuer, 1992; Willem Van der Does et al., 2000). In a prospective, naturalistic 1-year follow-up study of patients with current PD, remitted panic patients, infrequent panickers, and patients with SP, accurate heartbeat perception was found to be predictive of poor treatment outcome and recurrence of panic attacks after initial remission (Ehlers, 1995).

The theory above was later extended to include other anxiety disorders, with findings replicated in those with GAD, SAD and SPs, as well as within healthy participants with increased self-reported anxious traits (Anderson & Hope, 2009; De Berardis et al., 2007; Ludewig et al., 2005; Olatunji et al., 2007). A review by Willem van der Does and colleagues illustrated that participants with PD displayed a more accurate heartbeat perception than healthy controls, patients with palpitations, patients with only infrequent panic attacks or depressed patients, while patients with PD did not differ from patients with SAD, GAD or SPs (2000).

The results of empirical studies investigating the association between cardiac interoceptive accuracy and anxiety have been equivocal. Some studies have found enhanced interoceptive accuracy to be over-represented amongst anxiety patients, or positively correlated with anxious traits (Dunn, Stefanovitch, et al., 2010; Palser, Fotopoulou, et al., 2018; Pollatos et al., 2007, 2009). There is also some evidence linking higher cardiac interoceptive accuracy to greater *state* anxiety, before and after a public speaking task (Werner et al., 2009). However, other studies show no relationship between interoceptive accuracy and anxiety (Antony et al., 1995; Barsky et al., 1994; Yoris et al., 2015) or the reverse relationship, with higher anxiety in those with poorer interoceptive accuracy (De Pascalis et al., 1984; Ewing et al., 2017; Garfinkel, Tiley, et al., 2016; Smith, Feinstein, et al., 2021).

In attempts to draw together these disparate findings, meta-analyses have been conducted. One meta-analysis provided support for the association between anxiety and cardiac interoceptive accuracy. Those with higher trait anxiety had slightly better cardiac interoceptive accuracy than those with low trait anxiety, however, this effect size was small (d = 0.37) and findings were heterogeneous. Those with panic attacks also had better cardiac interoceptive accuracy than those without, with a medium effect size (d = 0.52) (Domschke et

al., 2010). However, only a small group of participants with PD could be qualified as “good heart rate perceivers,” meaning that heightened cardiac interoceptive accuracy was not ubiquitous across all individuals with PD. A more recent meta-analysis of studies using the heartbeat counting task (HCT) found no association between trait anxiety and cardiac interoceptive accuracy, in both clinical and non-clinical samples (Desmedt, Van Den Houte, et al., 2022). This has been substantiated by studies using novel measures of cardiac interoceptive accuracy, which have also found no associations with anxiety (Körmendi et al., 2023; Plans et al., 2021). When assessing self-reported, subjective interoceptive accuracy, one recent study found a negative correlation with anxious traits (Brand et al., 2023), which is the opposite of what one would expect given Ehler's theory. These studies, and potential reasons for these discrepancies in findings, will be further explored in Chapter Five of this thesis. Moreover, potential methodological issues that may lead to inconsistent findings will be discussed in Chapter Two.

As research into interoception developed and clearer distinctions were made between dimensions, more specific theories linking these dimensions to anxiety emerged. One such theory applies a predictive coding model to interoception within the context of anxiety. A predictive coding account of interoception suggests that the perception of one’s internal state is constructed through the comparison of bottom-up sensory input and top-down cognitions and expectations (Owens et al., 2018; Seth, 2013). In response to a large discrepancy between the two (prediction error), one will either adjust their expectations or re-assess the sensory information (Adams et al., 2015). Paulus and Stein suggest that individuals who are more prone to anxiety likely have an “augmented” discrepancy between their observed and expected body states than those who are less anxious (Paulus & Stein, 2006; Schmitz et al., 2012). This means that anxious individuals are more reactive when somatic, sensory information does not match with top-down, cognitive expectations about what they should

feel. Paulus and Stein suggest that this prediction error causes the engagement of cognitive resources to resolve this discrepancy, which heightens anxiety. Although this mechanism is not directly explained, one might speculate that the distress of being unable to predict, understand, and regulate the body may lead to anxiety, especially in combination with the diffuse arousal caused by the engagement of cognitive resources.

Garfinkel and colleagues provide a related theory, proposing that anxiety may be associated with a greater discrepancy between interoceptive accuracy, and what participants subjectively perceive regarding the degree of their attention and confidence in their accuracy (Garfinkel, Tiley, et al., 2016), labelling this interoception trait prediction error. Although not explicitly stated by the authors, it could be presumed that for interoceptive information to be used optimally, self-perceived accuracy should match actual accuracy. If this is not the case then as described under predictive coding theories, individuals will be unable to effectively predict or regulate interoceptive states, leading to anxiety. Moreover, an over-attendance to inaccurately perceived interoceptive signals may heighten these effects.

It has also been speculated that anxiety disorders may be related to a bias in the evaluation of interoceptive signals, resulting in more perceived danger and catastrophising interpretative styles (which is where one magnifies the likelihood of the worst possible outcome, such as interpreting an increase in heart rate as a heart attack or panic attack). Evidence appears to support this; however, research is focused mainly on the evaluation of anxiety-related signals (mainly heart rate and breathing), rather than interoceptive signals more broadly (Barlow, 1988; Beck et al., 1985; Clark, 1986; Palser, Palmer, et al., 2018; Paulus, 2013; Paulus & Stein, 2010; Taylor et al., 1991). Theories of overly negative interoceptive evaluation in anxiety have provided a basis for a lot of modern-day cognitive behavioural therapy (CBT; Barlow et al., 2000; Boswell et al., 2013; Clark et al., 1994; Craske et al., 1997). Relatedly, interoceptive exposure therapy that aims to reduce overly

negative interoceptive evaluation has been used clinically to reduce negative emotions associated with interoceptive cues in those with anxiety and PD (Boettcher et al., 2016; Boswell et al., 2013). This logic underpins the concept of "anxiety sensitivity," which is defined as a fear of anxiety-related sensations and symptoms based on beliefs that these sensations and symptoms have harmful consequences (Reiss, 1991). This is slightly different to interoceptive evaluation as a whole, as it focuses specifically on those signals most tightly bound to anxiety. Research has suggested that anxiety sensitivity underlies several anxiety disorders (Olatunji & Wolitzky-Taylor, 2009).

Another theory is that those with anxiety may have a higher degree of interoceptive attention, reflecting hypervigilance to somatic signals (Brewer et al., 2021). This has so far received some empirical support (Anderson & Hope, 2009; Noda et al., 2022; Palser, Palmer, et al., 2018; Pickard et al., 2020), although it is critical to note that these studies tend to rely on self-report measures that often conflate attention and accuracy. In a similar vein, some have suggested that internal cues may be stronger in those with anxiety (e.g., higher arousal; Hoehn-Saric & McLeod, 2000; Paulus & Stein, 2010), which may lead to them being consciously perceived more often and to a greater degree.

It is important for research to consider how each of the different dimensions of interoception relate to anxiety. One study that makes an important step in this direction is Slotta and colleagues' (2021) network analysis, which explores the interrelationships between trait anxiety, cardiac interoceptive accuracy, and eight aspects of self-reported interoceptive attention and evaluation (attention regulation, self-regulation, body-listening, trusting, emotional awareness, noticing, not worrying and not distracting). They found that cardiac interoceptive accuracy was unrelated to the other variables. Except for not distracting and not worrying, the self-report measures of interoception were tightly related to one another and not worrying and trusting exhibited negative associations with trait anxiety.

Despite most research focusing on cardiac interoception, there is some limited research implicating respiratory interoception in anxiety (Paulus, 2013). However, the picture appears to be similarly complex and inconclusive to the cardiac interoception literature. Some studies have found reduced respiratory interoceptive accuracy associated with anxiety disorders (Bogaerts et al., 2005; Van Den Bergh et al., 2004) and anxious symptoms in healthy participants (Garfinkel, Manassei, et al., 2016). However, Harrison and colleagues found that individuals with moderate levels of trait anxiety were *more* accurate at perceiving respiratory signals than those with low trait anxiety (2021). Moderately anxious participants also demonstrated more breathing-related catastrophising, more negatively minded general interoceptive awareness (meaning a greater awareness of signals that could be interpreted as a sign something was wrong), higher overall reported body awareness, and reduced confidence in respiratory interoceptive accuracy than those with low-level trait anxiety.

There is a lack of research on other domains of interoception, although one study showed a lack of an association between anxiety sensitivity and gastric interoception, using a novel vibrating pill task (Smith, Mayeli, et al., 2021), and another showed that in anorexia nervosa patients, activation of the dorsal mid-insula during stomach interoception was correlated with measures of anxiety (Kerr et al., 2016).

Further evidence of an association between interoception and anxiety comes from neurological studies. Areas associated with interoceptive function (such as the insula and anterior cingulate cortex) are noted to differ in activation when attending to interoceptive information between those with and without anxiety disorders (Etkin & Wager, 2007; Holzschneider & Mulert, 2011; Nagai et al., 2007; Paulus & Stein, 2006) and those with high versus low trait anxiety (Harrison et al., 2021; Rosso et al., 2010). Another study found that activation of the anterior insula when participants were instructed to focus on cardiac sensations was positively correlated with state and trait anxiety levels and mid-insular activity

related to participants' interoceptive accuracy, state and trait anxiety (Tan et al., 2018). However, it is possible that those with higher anxiety tried harder on the task, and it was this that caused the greater activation within this group. Other methodologies focus on heartbeat-evoked potentials (HEPs), which have been interpreted as indicators of the cortical processing of cardiac signals (Banellis & Cruse, 2020; Mai et al., 2018; Schandry et al., 1986), and also appear to be heightened in those with anxiety (Pang et al., 2019), although Coll and colleagues warn of the limitations of this research, highlighting the heterogenous methodologies used and varying locations of the HEPs in the brain across studies (Coll et al., 2021).

Evidence supporting an association between interoception and anxiety also comes from intervention studies. Heim and colleagues review thirty-one randomised control trials of interoception interventions in adults. 64.5% of these studies demonstrated improved interoception in the experimental group compared to controls. Of the three studies that measured changes in anxiety as an outcome measure, all three found decreases in trait anxiety immediately following the intervention. However, there was significant heterogeneity in the exact methods involved in each of these interventions and how "improved interoception" was defined (Heim et al., 2023). Sugawara and colleagues (2020) found that an interoceptive accuracy intervention improved state, but not trait anxiety, whilst Feinstein and colleagues reported that their mindfulness-based intervention that attempted to reduce exteroceptive sensory input to the nervous system led to more adaptive interoceptive interpretations and attention levels, and reduced anxiety (Feinstein et al., 2018).

It is worth noting that any relationship that exists between anxiety and interoception might be bidirectional. Although many of the theories discuss certain interoceptive profiles causing anxiety, it is also just as likely that anxiety may impact interoception. For example, if an individual has anxious cognitions, this may cause them to have more negative

interoceptive evaluation or pay more attention to interoceptive signals as they are consistently monitoring them. This could easily cause a positive feedback loop between the two, for instance, if a small increase in anxiety causes an increased heart rate, which prompts an individual to pay more attention to their heartrate, which causes a greater increase in anxiety, which causes a greater increase in cardiac interoceptive attention, and so on.

In summary, there are several theories linking interoception to PD, as well as anxiety more generally. Research on the role of interoceptive accuracy is extremely mixed. Research on negative interoceptive evaluation is more consistent, however, research focuses on the evaluation of anxiety-related signals only rather than general interoceptive evaluation. There is a lack of research on the role of interoceptive attention in anxiety that does not conflate attention with other aspects of interoception. Further research is required to provide clarity about the exact interoceptive mechanisms (if any) involved in anxiety.

Autism and interoception

Autism is a diagnosable, life-long neurodevelopmental profile characterised by social and communication difficulties and restricted and repetitive patterns of behaviour and interests (American Psychiatric Association, 2013). Estimates of the prevalence of autism range from 1 in 100 children around the world (Zeidan et al., 2022) to one in 36 children in the United States (Maenner et al., 2023). Autism is highly heterogeneous, meaning the exact presentation and support needs can vary largely from person to person. In line with the preferences of many in the autistic community, this paper will use identity-first language (i.e., “autistic person” rather than “person with autism”; Kenny et al., 2016).

With the rise of neurodivergence and associated movements, the field is experiencing a shift away from conceptualising autism as a disorder and seeing it instead as a different, but not lesser, cognitive profile (Fletcher-Watson & Happé, 2019). As such, this thesis aims to

avoid a pathologizing perspective on autism and does not focus on whether interoception is a "cause" of autism. Instead, the approach of this thesis follows guidance from the autistic community. In their survey, Cusack and Sterry (2016) found that those in the autistic community highlighted mental health interventions and anxiety as, respectively, the first and fourth most pressing concerns for future research. Interoception is an important lens through which we may better be able to understand some of the associations between autism and mental health conditions.

There are numerous different ways autism has been linked to interoception theoretically. The link was perhaps most notably proposed in a seminal paper which suggests that an "aberrant" oxytocin system directly underlies reduced interoceptive functioning, which in turn causes autism (Quattrocki & Friston, 2014). Specifically, the authors claim that oxytocin allows for encoding of the saliency or precision of interoceptive signals through a predictive coding framework and that an aberrant oxytocin system in autistic individuals has a range of downstream effects on things such as sensory and social processing that form the core features of autism. The authors cite evidence that oxytocin receptor distribution in the brain aligns with areas associated with interoception (Lawson et al., 2012; Riem et al., 2011; Striepens et al., 2012) and that similarly, oxytocin signalling is linked to interoceptive information such as thirst (James et al., 1995)(however, Quattrocki and Friston refer to other domains that may not necessarily be interoceptive, such as sexual arousal and thermoregulation). They also cite evidence for oxytocin peptide and plasma level abnormalities (Green et al., 2001; Modahl et al., 1998), receptor genetic variant associations (Campbell et al., 2011; Liu et al., 2010; Skuse et al., 2014; Wermter et al., 2010; Wu et al., 2005), and epigenetic alterations (Gregory et al., 2009; Gurrieri & Neri, 2009; Jacob et al., 2007; Kumsta et al., 2013) related to oxytocin in autistic populations. It should be noted,

however, that research linking oxytocin to autism is not conclusive (see review in Adams et al., 2024).

An alternate perspective suggests the opposite direction of effect - that the specific cognitive profile and core features of autism lead to a reduced ability to integrate sensory information to form a cohesive percept of the internal state of the body (Hatfield et al., 2019). This theory is based on the weak central coherence theory of autism, whereby autism is thought to reflect a bias towards local rather than global processing (Happé & Frith, 2006). This results in a focus on details, or lower-level sensory information, rather than the global picture, or the picture provided by the integration of multiple sources of sensory information. Given that the interpretation of interoceptive signals relies upon the integration of multiple sources of information, this may be impaired in those with a cognitive style that avoids the formation of a global percept from multiple sources of information. Others have suggested that an inability to identify and communicate somatic experiences may result in a greater need for external control, as well as more intolerance of uncertainty (Maisel et al., 2016), although at present there is no direct research on this. Despite the difference in the direction of effect, each of these theories suggests an association between autism and *reduced* interoceptive accuracy.

An exception to this is an account which links interoception to more general models of sensory processing in autism (Proff et al., 2022). It is known that many autistic individuals experience atypical sensory processing, although the exact presentation is highly heterogeneous (Crane et al., 2009). Whilst some experience sensory hypersensitivity, others experience hyposensitivity. Under the framework proposed by Proff and colleagues, both heightened or reduced interoceptive accuracy and attention could be associated with autism, due to difficulties integrating and interpreting sensory information. The authors explain this

by linking autism to both weak neural synchronization and atypical perceptual inference (see Chapter Eight for a more in-depth discussion of this theory).

Another theory of autism that may be relevant to interoception is the theory of monotropism, although it is yet to be explicitly linked to interoception. Monotropism refers to a cognitive style whereby one focuses intensely on a small number of things at a given time, rather than having a broader or more diffuse spotlight of attention. It has been proposed to be a core component of autism (Murray et al., 2005). If interoceptive signals are attended to, this could result in a higher-than-average degree of interoceptive attention or accuracy, whilst if they are not, it could result in the opposite. Similarly to sensory models, the monotropism theory of autism is consistent with the potential for interoceptive attention and accuracy to be both heightened and reduced in autistic individuals compared to non-autistic individuals.

Despite first-hand accounts suggesting interoceptive difficulties in autism (Trevisan, Parker, et al., 2021), research exploring interoception in autism has been just as equivocal as research into anxiety. Williams and colleagues (2023) recently conducted a meta-analysis on interoception in autism. They found that autistic participants showed significantly reduced cardiac interoceptive accuracy when the HCT was used, but this was not replicated with findings using the heartbeat discrimination task. They also found autistic individuals had higher confidence in their heartbeat counting abilities than non-autistic individuals, but groups were equivalent on self-reported interoceptive attention (although there was substantial heterogeneity across interoceptive attention studies).

When considering cardiac interoceptive accuracy specifically, whilst some studies have found autistic adults (Garfinkel, Tiley, et al., 2016; Mul et al., 2018) and children (Nicholson et al., 2019; Palser, Fotopoulou, et al., 2018) perform worse when compared to non-autistics, others have found either no relationship in adults (Failla et al., 2020; Mash et al., 2017; Nicholson et al., 2018, 2019) and adolescents (Pickard et al., 2020), or the opposite

pattern in autistic children (but only when trials were over longer time intervals; Schauder et al., 2015). To our knowledge, only one study has looked at respiratory interoceptive accuracy in autistic and non-autistic adults and it found no difference in performance (Nicholson et al., 2019).

Of the self-report studies, some found reduced self-reported interoceptive "sensibility" in autistic adults (Fiene & Brownlow, 2015; Mul et al., 2018) and adolescents (Pickard et al., 2020) compared to non-autistic control groups. Similarly, a negative association between autistic traits and interoceptive sensibility has also been reported in a general population sample (Yang et al., 2023). Other studies have found interoceptive sensibility to be heightened in autistic adults (Garfinkel, Tiley, et al., 2016), or for there to be no difference between autistic and non-autistic adults (Ben Hassen et al., 2023) and children (Butera et al., 2023; Palser, Fotopoulou, et al., 2018). One study found autistic adults reported greater autonomic reactivity than neurotypical controls but were no different in self-reported "awareness" (Zdankiewicz-Ścigała et al., 2021). Trevisan and colleagues distinguish between adaptive and non-adaptive interoceptive attention in their commentary, using secondary data to illustrate that autistic individuals may show reduced adaptive attention to interoceptive signals, but heightened maladaptive and anxious attention (Trevisan, Mehling, et al., 2021). In terms of self-reported accuracy, one study found lower accuracy in autistic adults compared to neurotypical controls (Fiene et al., 2018).

Evidence from neuroimaging studies would appear to suggest there is an association between autism and interoception, however, results are difficult to interpret. A few studies have found hypoconnectivity of interoception-related brain regions such as the anterior insula, thalamus, and superior frontal gyrus in autistic participants compared to neurotypical controls (Ebisch et al., 2011; Francis et al., 2019; Martínez et al., 2020). Moreover, connectivity of the lateral ventral anterior insula and anterior cingulate cortex has been found

to correlate with autistic traits, as well as interoceptive accuracy and sensibility, in healthy controls (Yang et al., 2023).

One explanation for the mixed findings obtained so far may be the heterogeneity of autism. Firstly, the presence of both hypo and hyper-awareness within autistic populations could make the interpretation of mean scores difficult. It may also be that heightened or diminished interoception is only observed in some sub-groups or may be driven by the presence of co-occurring conditions such as alexithymia (Proff et al., 2022; Shah et al., 2016). It could also be that the trajectories of interoception development differ in autistic versus non-autistic individuals, resulting in different relationships at different ages. There is some support for this, as one study found that interoception worsens with age from childhood to adulthood in autistic participants, whilst it remained stable in non-autistic individuals (Mash et al., 2017).

The high co-occurrence rates between autism and numerous mental health conditions have been well-reported in the literature (Lai et al., 2019). Of particular concern in the context of this thesis is the co-occurrence between autism and anxiety, given that interoception has been linked to both. Autistic individuals face a higher risk of developing anxiety disorders (Postorino et al., 2017; van Steensel et al., 2011), with an estimated 47% of autistic adults experiencing “severe” anxiety (National Autistic Society, 2021). Autistic individuals also show higher levels of trait anxiety (Jolliffe et al., 2023; van Steensel & Heeman, 2017), than their neurotypical counterparts. This is concerning given that autistic individuals also benefit less from conventional interventions for anxiety, such as CBT, than non-autistic individuals (van Steensel & Bögels, 2015).

There are likely multiple, interacting mechanisms that underlie the association between anxiety and autism, including the fact that social environments often are not accessible or accommodating to autistic individuals and autistic individuals may be less able

to access social support networks than non-autistic individuals (Bishop-Fitzpatrick et al., 2018). The association can be conceptualized within a wider developmental psychopathology framework, in which development is both active and dynamic, with both environmental- and individual-level factors interacting to determine mental health outcomes throughout one's lifespan (Rutter & Sroufe, 2000). Interoception is one of the many potential factors that might contribute to the observed co-occurrences between autism and mental health conditions within this framework, interacting with other factors in complex ways. This thesis, however, will focus specifically on the role of interoception in anxiety and autism.

When exploring associations between autism, anxiety and interoception, many interesting questions are raised. Firstly, whether interoception can explain some of the association between autism and anxiety. The mechanism by which interoception is related to anxiety may be the same in both autistic in neurotypical individuals, but it could exert more of an effect within autistic populations who show greater difficulties with interoception. It could also be that the relationship between interoception and anxiety is different between autistic and non-autistic populations, or even that autism and anxiety interact to influence interoception.

There is reason to believe that for autistic individuals specifically, interoception may be an important factor in anxiety. However, as with anxiety, there is still considerable ambiguity as to the nature of associations between interoception and autism. A small-scale study found that autistic traits, subjectively reported interoceptive awareness and the difference between subjectively reported interoception and objectively measured interoceptive accuracy all independently contributed to state anxiety, in a sample of both autistic and non-autistic adults (Garfinkel, Tiley, et al., 2016). This was replicated in a sample of autistic and non-autistic children (Palser, Fotopoulou, et al., 2018). However, these studies did not check for interaction effects between any of the dimensions of interoception and

autism in predicting anxiety. Similarly, one study focusing on only autistic young people found that interoceptive "sensibility" was associated with social anxiety (Wood et al., 2022). These studies are all relatively small scale, and a lot remains unknown about the role of interoception in the association between autism and anxiety.

Alexithymia

Alexithymia is critical to consider when exploring the associations between interoception, autism and anxiety. Alexithymia is defined as difficulty identifying and describing emotions (Nemiah et al., 1976). Given the assumed role of somatic sensations in emotions, some consider alexithymia to be the result of dysfunctional interoception (Brewer et al., 2016; Van Der Cruijsen et al., 2019). Others employ a slightly less extreme stance, suggesting that aberrant interoception may be one route by which alexithymia, or certain types of alexithymia, develop, but that there may also be other causes of alexithymia (Jakobson & Rigby, 2021). Similar brain areas have also been implicated in both alexithymia and interoception (Ernst et al., 2014; Hogeveen et al., 2016; Meza-Concha et al., 2017).

There is significant evidence to support a relationship between alexithymia and interoception, although confusingly alexithymia has been linked to both heightened and reduced interoceptive attention and accuracy. Intuitively, one may expect that those who struggle to notice interoceptive signals may be less able to detect certain emotions. Indeed, several studies have found evidence of negative associations between alexithymia and self-reported interoceptive awareness (Brewer et al., 2016; Gaggero et al., 2021), and objective tasks of interoceptive accuracy (Bornemann & Singer, 2017; Murphy, Brewer, et al., 2018; Murphy, Catmur, et al., 2018). However, other studies have shown a positive association (Nakao et al., 2002; Betka et al., 2018; Longarzo et al., 2015; Scarpazza et al., 2022), or no association at all (Nicholson et al., 2018), between self-reported interoceptive awareness and

alexithymia, and a positive association between interoceptive accuracy and alexithymia (Nakao et al., 2002; Scarpazza et al., 2022). One explanation for this inconsistency is that alexithymia could arise from hyper as well as hypoawareness of interoceptive signals. If one is hyperaware of interoceptive signals, it may become difficult to disentangle noise from signals that inform about one's emotional state, and therefore recognising distinct emotions may become more difficult.

There also appears to be a high co-occurrence between autism and alexithymia, therefore meaning that alexithymia could be a confounder when considering associations between autism and interoception. Alexithymia frequently co-occurs with autism, with rates of alexithymia in autistic populations at around 50%, compared to approximately 5% in the general population (Berthoz & Hill, 2005; Kinnaird et al., 2019; Oakley et al., 2022). Alexithymia also correlates with autistic traits in healthy populations (Yang, Zhou et al., 2022). Historically, many of the symptoms of alexithymia have been conflated with autism, such as difficulty describing emotions, which numerous measures of autistic traits still include. Some have expressed concern that the high co-occurrence and conflation of autism and alexithymia artificially inflates a lot of reported associations between autism and other things that may actually be driven by alexithymia. This is known as the alexithymia hypothesis (Bird & Cook, 2013).

Atypical interoception is an example of something that may be misattributed to autism when it is in fact attributable to alexithymia. Indeed, some empirical work suggests that interoception is associated with alexithymia and not autism (Brewer et al., 2015; Shah et al., 2016). As a result, it's important to at the very least include alexithymia as a covariate when exploring relationships between autism and interoception. However, this is not common practice in the literature.

Further complicating the issue, there are increased rates of alexithymia in numerous clinical populations (Leweke. et al., 2012; Taylor & Bagby, 2004). In fact, alexithymia has been proposed to explain variance in mental health in the general population (Berardis et al., 2008; Fietz et al., 2018; Preece et al., 2022). Alexithymia also appears to be highly relevant to mental health within autistic populations specifically. One cross-sectional study demonstrated that self-reported alexithymic traits partially mediated associations between core autism traits and depression and anxiety in young adults (Maisel et al., 2016), with another showed that autistic adolescents with elevated alexithymia self-reported greater state anxiety than those without elevated alexithymia (Milosavljevic et al., 2016). This was substantiated by later longitudinal work that found that difficulty identifying feelings was predictive of anxiety in both the whole sample (including both autistic and non-autistic participants) and within autistic participants only (Oakley et al., 2022). In terms of research which explicitly measures interoception, one study found that the positive association between interoceptive sensibility and trait anxiety was partially mediated by alexithymia in a general population sample (Palser, Palmer, et al., 2018), further illustrating the importance of considering alexithymia when researching interoception.

To summarise, as alexithymia appears to be related to the three main constructs discussed in this thesis (anxiety, autism and interoception), it is something which will be considered throughout and controlled for where relevant.

Aims of this thesis

Despite the enormous research interest in interoception, our understanding of how exactly it affects mental health remains limited. To guide future research, it is paramount for us to first understand what interoception is and how it is best measured. This thesis takes an important step in addressing the limitations of commonly used measures of interoception and

makes proposals for methods that will lead to higher-quality research, leading to more valid and replicable conclusions.

It is important to gain a comprehensive understanding of the mechanisms by which interoception influences mental health to ensure any interoception-based interventions are as efficacious as possible and avoid potential iatrogenic effects. Take the example of cardiac interoceptive accuracy. Both greater and reduced cardiac interoceptive accuracy have been linked to anxiety. Therefore, whilst there is the potential that an intervention enhancing cardiac interoceptive accuracy might reduce anxiety, it could also be anxiogenic. Similarly, mindfulness is a highly popular technique for alleviating mental health symptoms that is recommended by mental healthcare professionals. However, if certain groups of people struggle because of overly high levels of interoceptive attention, then mindfulness may have harmful effects. Gaining a better understanding of the exact interoceptive profiles that are related to different mental health outcomes reduces the likelihood of this happening.

There is also a lack of certain kinds of research into interoception. Specific dimensions of interoception such as interoceptive evaluation have been neglected in the literature, despite numerous theories suggesting it may be especially important in mental health. Research into interoception also rarely takes a bottom-up approach, where theories are guided by individual experience. This may be especially useful given the complexity of interoception, and the importance of understanding the nuanced relationships between each dimension within certain contexts. Taking a bottom-up approach guided by qualitative methodology may provide useful guidance as to the specific aspects of interoception that are most important, thus guiding research towards the most fruitful avenues. This approach is also pertinent when considering research conducted on specific groups, such as autistic individuals, as it allows the communities involved to guide the direction of research.

The principal aim of this thesis is to advance research into interoception and better understand its role in mental health. As such, the sub-aims of this thesis are to illustrate the limitations of current measures of interoception, develop and refine alternative measures, and then apply some of these measures when researching the role of interoception in anxiety and autism. It is hoped that this knowledge may inform the development of specialised therapies to reduce symptom loads in these populations.

Chapter Two provides an outline and evaluation of the most used measures of interoception and novel variants of these, focusing specifically on self-report measures and objective tests of cardiac, respiratory, and gastric interoceptive accuracy. The latter half of Chapter Two empirically substantiates the principal problem with one commonly used measure of cardiac interoceptive accuracy- the Two Alternative Forced Choice (2AFC) version of the Heartbeat Detection Task (HDT)- using secondary data. Following on from these critiques, the third and fourth chapters of this thesis outline two novel measures of interoception. Chapter Three involves the pilot study of the Phase Selection Task (PST) which measures cardiac interoceptive accuracy. Chapter Four covers the development and validation of the Negative Interoceptive Evaluation Scale (NIES), a self-report measure of interoceptive evaluation. Chapter Five of this thesis is a meta-analysis of the association between cardiac interoceptive accuracy and state and trait anxiety. Chapter Six is a qualitative study employing a reflexive thematic analysis approach to the role of interoception in the experiences of anxiety in autistic adolescents. Chapter Seven quantitatively extends the findings of the previous chapter, using questionnaire data to explore associations between self-reported interoceptive accuracy, attention, and evaluation and autism and anxiety in adults. Finally, Chapter Eight collates these findings and outline a roadmap for future research.

Research questions

The following research questions are addressed:

1. What are the limitations of the most-used measures of interoception?
2. Is the PST (a novel measure of cardiac interoceptive accuracy) of practical utility?
3. Is the NIES a valid and reliable measure of interoceptive evaluation?
4. Based on a meta-analysis of existing studies, is there a relationship between cardiac interoceptive accuracy and state and trait anxiety?
5. What role does interoception play in experiences of anxiety in autistic adolescents?
6. Do self-reported interoceptive accuracy, attention and evaluation relate to both the presence of anxiety disorders and self-reported trait anxiety?
7. Do self-reported interoceptive accuracy, attention and evaluation relate to both the presence of an autism diagnosis, and self-reported autistic traits?
8. Does self-reported interoceptive accuracy, attention or evaluation contribute to the association between anxiety and autism?

Chapter Two: Measuring interoception- Methodological issues

The first section of this chapter is an edited version of the below published paper. Kiera Adams and Geoffrey Bird developed the idea for the paper. Kiera Adams conducted the literature review and drafted the initial manuscript, which was edited based on feedback from Geoffrey Bird, Jennifer Murphy, and Caroline Catmur.

Content adapted from: **Adams, K. L.,** Murphy, J., Catmur, C., & Bird, G. (2022). The role of interoception in the overlap between eating disorders and autism: Methodological considerations. *European eating disorders review: the journal of the Eating Disorders Association, 30*(5), 501–509. https://doi.org/10.1002/erv.2905

The second section of this chapter is an edited version of the below paper which is in preparation. Kiera Adams, Geoffrey Bird, and Jennifer Murphy designed the study. David Plans, Adam Cunningham, Tom Piercy, and Davide Morelli developed the app for the PAT. Ria Spooner, Rhea Clemente, Nerea Irigoras Izagirre, Elisa Fernandez Fueyo, Dawn Watling, Rebecca Brewer, Gemma Budworth, Dorina Cocirla, Mateo Leganes-Fonteneau, Jennifer Todd, Jane E. Aspell, and Jennifer Murphy collected the data. Jonathan Bird and Kiera Adams developed the analysis scripts and conducted analysis. Kiera Adams conducted the literature review, interpreted the results, and drafted the initial manuscript, which was edited based on feedback from co-authors.

Content adapted from: **Adams, K. L.*,** Bird, J.*, Plans, D., Cunningham, A., Morelli, D., Piercy, T., Spooner, R., Clemente, R., Irigoras Izagirre, N., Fernandez Fueyo, E., Watling, D., Brewer, R., Budworth, G., Cocirla, D., Leganes-Fonteneau, M., Todd, J., Aspell, J.E., Bird, G.† & Murphy, J.† (in preparation). Individual differences in heartbeat-tone

synchronicity judgements suggest the two-alternative heartbeat detection task is a poor test of interoceptive accuracy and insight.

* Co-first authors
† Co-senior authors

Abstract

Interoception research is burgeoning. In the rush to publish novel findings, there has been little emphasis on achieving cohesion across the field, ensuring the validity of measures of interoception, and maintaining the reliability of results. This chapter first presents a critical analysis of the most used methods for assessing the most researched facets of interoception: self-report measures of interoceptive "awareness" and objective measures of interoceptive accuracy in the cardiac, respiratory, and gastric domains. Second, this chapter presents an empirical study which illustrates the main limitation of the two-alternative forced-choice variant of the heartbeat detection task (2AFC-HDT) that is frequently used to measure cardiac interoceptive accuracy. Consideration of the logic of this task suggests that the use of pre-defined delays may result in false negatives (individuals judged to be not interoceptive when they are), either if participants perceive their heartbeats as occurring closer to tones occurring at a delay predefined as asynchronous rather than synchronous with their heartbeats, or if they perceive their heartbeats at a delay equidistant between the "synchronously" and "asynchronously" presented tones. Currently, there is little data on the likelihood of such potential false negatives. This chapter addresses this lack of knowledge using the Phase Adjustment Task (PAT). The PAT allows us to determine if a participant is interoceptive and if so, when they perceive their heartbeat in the cardiac cycle. A conservative analysis suggests that between 7 and 9% of interoceptive participants would be misclassified as not interoceptive by the 2AFC-HDT due to the use of predefined points in the cardiac cycle at which participants are expected to perceive their heartbeat. These issues suggest that the 2AFC-HDT should not be used to measure interoceptive accuracy or to assess other relevant dimensions of interoception, including interoceptive insight (also known as awareness or meta-cognition).

Critique of the most commonly used methods to measure interoception

Introduction

Research into interoception has increased exponentially within the last couple of decades (Khalsa & Lapidus, 2016). This rush to publish findings on interoception has resulted in a lack of standardised definitions and methodologies that lack validity. This makes gaining a cohesive understanding across the field very difficult. Given that research into interoception has enormous potential to guide interventions that could benefit mental health, research must be of the highest possible quality.

This chapter aims to outline and evaluate the *most used* measures of interoception, highlighting their chief limitations and how this may affect research and contribute to the observed inconsistency in findings so far. This is not intended to be a comprehensive overview of all measures, but rather to illustrate the most pervasive problems within the field. This sets the groundwork for the rest of the thesis, which will propose and then apply novel measures of interoception.

General critique

Self-report measures

The primary issue with self-report measures of interoception is that multiple questionnaires are intended to measure the same underlying construct but have low agreement with one another. Many of the most used questionnaires of interoception claim to measure "interoceptive awareness" or "sensibility," however, neither of these terms are clearly or consistently defined across studies (despite the efforts of researchers to create a unified nomenclature; Khalsa et al., 2018). As well as enabling a large degree of variability in what these questionnaires measure, thus resulting in inconsistent findings in the literature, these questionnaires typically conflate separable constructs such as interoceptive accuracy

and attention (Murphy, Catmur, et al., 2019) and adaptive and non-adaptive awareness (Trevisan, Mehling, et al., 2021).

One such questionnaire is the Body Awareness section of the Body Perception Questionnaire (BPQ; Porges, 1993). Substantiating concerns of a conflation of accuracy and attention, participants show a large degree of heterogeneity in whether they interpret the questionnaire as a measure of accuracy, attention, or the objective frequency of interoceptive signals (Gabriele et al., 2022). Moreover, the questionnaire is frequently misused. Whilst only the Body Awareness subscale pertains to interoception, many researchers either use a combined total score of this subscale *and* the Autonomic Symptoms subscale or do not clarify that they are only using the Body Awareness subscale (e.g.: Desmedt et al., 2022), resulting in an impure measure of interoception.

Another commonly used interoceptive “awareness” questionnaire is the Multidimensional Assessment of Interoceptive Awareness (MAIA; Mehling et al., 2012). As with the BPQ, although the MAIA was intended to measure multiple, separate dimensions of interoception within different subscales, and the authors advised against the use of a total score, studies continue to analyse using the total score (for example, Montoya-Hurtado et al., 2023). There is also some concern that the MAIA lacks construct validity, as it includes items that may better pertain to alexithymia (a sub-clinical condition characterised by difficulty identifying or describing one's own emotions; (Nemiah et al., 1976)) rather than pure interoception. For example: "I notice how my body changes when I feel happy/joyful."

Similarly, Vlemincx and colleagues raise concerns that several measures of interoception conflate the noticing of interoceptive signals with higher cognitions, including emotional evaluation. They cite both the MAIA and the Interoception Sensory Questionnaire (Fiene et al., 2018) as examples of questionnaires that do this. Vlemincx and colleagues argue

that this may result in artificially inflated associations between interoceptive “sensibility” (which also captures emotional factors) and emotion-based pathologies (Vlemincx et al., 2023).

A low agreement across measures of interoceptive sensibility or awareness has been demonstrated by Vig and colleagues (2022), who found the BPQ, the MAIA (including the eight sub-scales assessed separately) and the Body Awareness Questionnaire (BAQ; Shields et al., 1989) show only weak to moderate associations with one another. This has been further substantiated by a systematic review, latent factor analysis, and network analysis, which concluded that the five most cited questionnaires of interoceptive sensibility (the BPQ, MAIA, BAQ, the Private Subscale of the Body Consciousness Questionnaire (Miller et al., 1981) and the Self-Awareness Questionnaire (Longarzo et al., 2016)) appear to measure different constructs (Desmedt, Heeren, et al., 2022). This does not mean the questionnaires are inherently bad, however, it does mean that researchers should exercise caution when using them, ensuring that they define exactly what they are measuring, rather than citing them simply as measures of interoceptive awareness or sensibility.

When looking specifically at eating-disordered populations, a substantial number of studies have assessed interoception using the Interoceptive Accuracy Subscale of the Eating Disorder Inventory (Garner et al., 1983), where respondents report their ratings of confidence in identifying emotions and sensations of hunger. As with the MAIA, this conflates interoception with alexithymia. The same critiques may also be levelled at a novel measure of interoception designed for use within populations with eating disorders- the Eating Disorder-Specific Interoceptive Perception Questionnaire (Ortmann et al., 2024), which includes subscales of emotion, hunger, satiety and discrimination. Whilst both measures may be useful within eating-disordered populations specifically, their applicability to interoception more widely is limited by the constrained domains they consider.

Other self-reported measures of interoception have since been developed which benefit from being more clearly defined than previous measures and focus on more specific constructs than general “awareness” or “sensibility.” However, they have not yet been subject to sufficient validation or been widely adopted in interoception research. Some examples include the Interoceptive Accuracy Scale (IAS) and the Interoceptive Attention Scale (IATS), which research has shown are interpreted as pure measures of interoceptive accuracy and attention respectively (Gabriele et al., 2022; Murphy et al., 2020), the Three Domain Interoceptive Sensations Questionnaire (TDIS-Q; Vlemincx et al., 2023), which aims to assess noticing and feeling of neutral respiratory, cardiac, and gastroesophageal sensations, the Interoceptive Sensitivity and Attention Questionnaire (ISAQ), which measures sensitivity to neutral bodily sensations, attention to unpleasant bodily sensations, and difficulty disengaging from unpleasant bodily sensations (Bogaerts et al., 2022), and the Interoceptive Attention subscale of the Body Mind Connection Questionnaire (Van Bael et al., 2023). In their current instantiations, each of these requires further refinement and psychometric evaluations across larger and more diverse samples.

A further issue with each of these questionnaires, including the latest questionnaires, is there is little consistency in the exact domains included. For example, both the IAS and IATS consider pain and affective touch as interoceptive, the TDIS-Q includes questions about the temperature of food in the oesophagus, the ISAQ includes questions about thermoregulation, and the Body Mind Connection Questionnaire simply refers to “bodily sensations,” which does allow for a large degree of variability in terms of both which signals are considered by participants, and the valence of the signals. This reflects a wider issue within interoception research, which is the contention surrounding which signals should be considered interoceptive (Desmedt, Luminet, Maurage, et al., 2023; see Chapter One). Again,

researchers must exercise caution to ensure consistency in how they choose to define interoception and the measures they use.

In addition to the above measures, several self-report measures of interoceptive evaluation exist. These will be more comprehensively outlined and evaluated in Chapter Four of this thesis.

Objective measures of interoceptive accuracy

Cardiac. Problematically, most experimental work on cardiac interoceptive accuracy has used the heartbeat counting task (HCT, Schandry, 1981). In this task, participants report how many heartbeats they feel in a given window of time, and their reported number of heartbeats is compared to the objectively measured number of heartbeats. Critiques of this approach have been well documented (Desmedt et al., 2018, 2020; Murphy, Millgate, et al., 2018). In short, the most salient issue is that participants could achieve perfect performance on the task without being able to perceive their heartbeats at all if they have prior knowledge about what their heart rate should be (Ring et al., 2015; Ring & Brener, 1996). For example, if a participant knows that their heart rate is around 60 beats per minute, they need only count the number of seconds in the interval to make an accurate estimation of the number of times their heart has beat. This results in false positives (identifying participants as interoceptive who are not). The fact that participants likely do this has been evidenced by studies demonstrating that participants' reported heart rates remain consistent, even with changes in position, or when a pacemaker is used to increase their heart rate (Ring & Brener, 1996; Windmann et al., 1999). Moreover, if participants are explicitly instructed not to guess, and only to count a heartbeat if they feel it, performance has been shown to decrease (Desmedt et al., 2020).

Some studies employing the HCT have used a control task, where participants estimate the amount of time that has elapsed in varying intervals. Although this allows for the control of non-specific factors that may influence performance, such as motivation and attention, it is not sufficiently matched to the HCT, in that it does not control for differences in detection thresholds relating to decision parameters. A further issue with the HCT is that given people tend to under-report the number of heartbeats (one is unlikely to perceive a beat that is not there, and more likely to miss a beat; Zamariola et al., 2018), one is more likely to be accurate if one has a bias towards reporting a higher number of felt heartbeats (Larsson et al., 2021), allowing for the potential conflation of accuracy and bias. The HCT can also be influenced by a range of extraneous variables such as body mass index (BMI; Rouse et al., 1988) and cardiac variables (Knapp-Kline & Kline, 2005).

Adaptations of the task have aimed to reduce the role of working memory that may constrain HCT performance. One such adaptation is the heartbeat tapping task, where participants tap every time they feel a heartbeat, and the total number of taps is compared to the total number of heartbeats within a given trial (De Pascalis et al., 1984). In its more recent instantiation, this task forgoes tapping for finger movements, so the tactile response does not detract from heartbeat detection (Körmendi et al., 2022). This task requires that the finger movement be within a specific time window after the r-wave for it to be counted as valid, which authors argue reduces the likelihood of false positives obtained through participants guessing or moving their fingers without feeling anything. However, this assumes that everyone perceives their heartbeat at a similar time point in the cardiac cycle. The issues with this approach will be illustrated in the second section of this chapter.

Another modern variation also attempts to advance on the HCT. Larsson and colleagues propose assessing accuracy as how variations in reported heart rate across several short trials correlate with actual variation (Larsson et al., 2021), which is intended to

disambiguate between accuracy and biases in reporting, as well as reduce the likelihood of false positives. However, this task requires further validation. It may be that the limited variability in heart rate at rest precludes significant results from being obtained.

Another alternative to the HCT is the HDT (Whitehead et al., 1977), in which participants judge which set of auditory or visual tones is synchronous with their heartbeats. Most studies employing this task use the two-alternative forced choice (2AFC) variant, where two sets of pre-defined tones at set delays after the heart's r-wave are presented, and the participant selects between them (Hickman et al., 2020). Similarly to Körmendi's task, this relies on the—false (see the second section in this chapter)—assumption that there are no individual differences in the delay following cardiac contraction with which individuals perceive their heartbeat. This may result in false negatives (identifying participants as not interoceptive when they are).

There are variations of the HDT that aim to address this issue by presenting tones at multiple different delays after cardiac contraction and assessing the consistency of synchronicity judgements (e.g. the six-alternative forced-choice [6AFC] and method of constant stimulation [MCS] methods; Brener et al., 1993; Clemens, 1984; Yates et al., 1985). These HDT variants—given they are coupled with appropriate control tasks assessing the participant's ability to judge the synchronicity of two stimuli (Knapp et al., 1997)—provide better tests of cardiac interoceptive accuracy than the 2AFC-HDT but are not commonly used. This may be due to practical limitations: both measures are time-consuming, and performance is thus constrained by participants' ability to focus for an extended period. However, promising research has demonstrated that the MCS appears to produce valid results even when only five tones are presented, thus addressing the issue of the measure being time-consuming (Brener et al., 1994). Nonetheless, both the 6AFC and MCS lack granularity, given the limited choices in delays available to participants, meaning that the delay at which a

participant perceives as synchronous with their heartbeat cannot be estimated precisely. Moreover, there is a possibility that results may be impacted by the granularity of participants' perceptions. If the difference between several delay options is too small for participants to distinguish between, it allows for the possibility of false negative results, even when participants can detect their heartbeats. For instance, tones at delays of both 200ms and 300ms may feel coincident with the participants' heartbeats if they actually detect their heartbeat at around 250ms. This could result in participants switch between selecting 200ms and 300ms as synchronous across trials. Although the most typically used analysis (chi-square comparison to a uniform distribution), would theoretically allow for participants to be deemed as interoceptive if they are consistently selecting some delays over others, the threshold for this would be much higher if responses were split between two options rather than just one. This could result in some false negatives for interoceptive participants with a lower discriminatory resolution. At present, we do not know enough about discriminatory resolution for cardiac signals to estimate the scale of this issue.

Another method of measuring cardiac interoceptive accuracy includes infusing participants with differing amounts of isoproterenol, a drug that results in rapid and transient increases in heart rate, and asking them to rate the intensity of their heart rate or breathing sensations (Khalsa et al., 2009). Participants' actual changes in heart rate can then be correlated with their subjective ratings. However, the task instructions ask participants to focus both on breathing and heart rate, whilst only the objective heart rate is measured. This may be a limitation if participants entirely focus on breathing, which is then compared to their heart rate. Whilst there is a relationship between one's breathing and heartbeat, they are separable. Moreover, whilst objective heart rate measures concern only the frequency of beats, participants are instructed to attend to the intensity, which could also encompass how powerfully they feel it beating. In addition to this, the task requires specialist equipment and

is invasive in nature. It may also not be appropriate for participants who are disposed to panic, as the isoproterenol may induce a panic attack.

The phase adjustment task (PAT; Plans et al., 2021) is a new test of cardiac interoception that both overcomes the use of pre-defined delays to test perception of heartbeats and is quicker than the 6AFC and MCS, therefore representing an advance on the HDT variants described above. The task presents participants with auditory tones that are at the same frequency as their heartbeats, but out of phase with those heartbeats. Participants use a dial to change the phase relationship between the tone and their heartbeat until they perceive them as synchronous. Interoceptive accuracy is determined by the consistency of their responses. For further details of the PAT, please see the following section of this chapter. The PAT is amenable for smartphone administration, increasing its acceptability for use in clinical populations (especially those who may not wish to travel to novel laboratory environments). Another advantage the PAT has is that it explicitly instructs participants to focus on one location where they can feel their heartbeat across all trials. Due to differences in the pulse arrival time to these different areas, we could expect to see differences in the perceived timing depending on which location participants are focusing on. Thus, consistency scores may be impacted if participants switch between locations across trials. Instructions are rarely reported in enough detail for the 6AFC and MCS to ascertain whether participants know to consistently focus on one location across trials, whereas they are explicitly told to do this in the PAT. The PAT also spans the entire inter-beat interval (the time differences between each heartbeat), thus reducing the likelihood of individual differences in the inter-beat interval (IBI; the time between heartbeats) influencing performance (Knapp-Kline & Kline, 2005). The design of the task also prohibits false positives.

The PAT still requires further validation. Moreover, the task itself is quite difficult. One difficulty is that participants must keep all of their fingers of one hand on the phone

screen to prevent them from cheating and taking their pulse, however, this can also prove difficult, especially among those with larger fingers and smaller phone screens and potentially distract participants. Another issue is the task relies upon multimodal sensory integration that may interfere with heartbeat perception and result in false negatives. This could be overcome using an appropriate control task, but at present, none exist that are entirely comparable to the heartbeat detection tasks. Although in its current instantiation, a screening task is used, in which participant's heartbeats are sonified, and participants are required to synchronize the two auditory signals as in the PAT, this screening task is not entirely matched to the PAT, in that it involves the integration of two signals from within the same domain (auditory), rather than the integration of interoceptive and exteroceptive signals, which is more difficult. Therefore, the task is not a complete control and cannot completely reduce the risk of false negatives. However, Ring and Brener argue that this need not be of concern, as the nervous system is continually engaged in parallel processing (Molholm et al., 2002), meaning our acuity is extremely precise for the integration of interoceptive and exteroceptive information, and thus performance on these tasks should not be too constrained by this requirement (Ring & Brener, 2018).

The heart rate discrimination task (Legrand et al., 2022) is a new measure in the cardiac domain that does not attempt to measure the accuracy of perception of cardiac signals, or the degree of attention paid to them. Instead, the task aims to measure the accuracy and precision of beliefs about heart rate. Whether individual differences in these parameters are of clinical relevance is yet to be established, but in its current instantiation, the heart rate discrimination task requires a well-matched control task to address questions relating to clinical relevance.

Desmedt and colleagues provide a comprehensive review of all novel measures of cardiac interoceptive accuracy (Desmedt, Luminet, Walentynowicz, et al., 2023). Such a

review is beyond the scope of this chapter, but in short, many of these novel measures can still be informed by guessing, lack large-scale validation, do not separate interoceptive accuracy from response biases, or may be contaminated by factors such as heart rate, memory capacity, and multimodality processing. Novel computational methods may prove beneficial in splitting out response bias from interoceptive accuracy, whilst the development of appropriate control tasks could be instrumental in resolving the latter two concerns.

Respiratory. Respiratory interoception has been measured by both detection and discrimination tasks (Harrison et al., 2021; Webster & Colrain, 2000; Zhao et al., 2002), in which participants are required to detect whether respiration is restricted using filters, or to judge which of two filters provides more respiratory resistance, respectively. In both tasks, findings may be confounded by the fact that trials require long respiratory restrictions, which have an intrinsic negative value that may affect performance. In addition to this, these tasks lack granularity, meaning that assessing thresholds requires many trials and perceptual ability cannot easily be quantified.

More recently, a task has been developed that overcomes these issues, by measuring individuals' ability to detect small differences in lengths of short respiratory occlusions (Van Den Houte et al., 2021), also reducing the time taken to run the experiment. An adaptive staircase paradigm is then used to assess what the just noticeable difference is for participants. Initial psychometrics for this measure have been promising. However, pragmatically, the administration of any respiratory interoceptive accuracy task still requires the use of specialist equipment, such as clinical breathing filters, which makes it difficult to administer remotely or at scale. These measures may also be inappropriate for use within the clinical populations that may most benefit from studies of interoception, such as those who suffer from panic attacks, due to the risk of the task triggering a panic attack. Moreover,

many of these tasks rely on memory, as participants must remember and compare different conditions. This could result in false negatives.

Another method has been proposed as a way one might measure the "propensity to utilize interoceptive cues to gauge respiratory output" (Murphy, Catmur, et al., 2018). In this task, participants are asked to estimate what percentage of their initial maximum exhalation they reach in subsequent exhalations with various target values. This is compared to the actual value of their subsequent exhalations compared to their maximum exhalation. This is done first in a condition where participants have access to interoceptive and exteroceptive cues, and then in a condition where they are played white noise through headphones so they cannot use auditory cues about their breathing. The difference in these scores is assumed to indicate the reliance on exteroceptive over interoceptive cues when both are available. The task requires only a peak flow meter, which is more accessible than the specialist equipment discussed above. However, it is feasible that participants may have comparable ability to utilise either interoceptive or exteroceptive cues, resulting in a similar score across conditions, even though they may rely more on exteroceptive cues when they are available. This task could be advanced upon with the inclusion of an exteroceptive-only condition- for example, if participants have the sounds of their exhalations from previous trials played back to them and they make judgements from that. The scores in this condition could be compared to the scores where both interoceptive and exteroceptive information were available to indicate the degree to which the availability of interoceptive information improved performance.

Gastric. Gastric signals comprise those indicating the degree of satiety, fullness and hunger. Unfortunately, the most valid methods to assess their perception are invasive in nature, involving forced gastric distension, chemical stimulations in the oesophagus and stomach, thermal stimulations in the upper gut, and electrical stimulations in the rectum

(Desmedt, Luminet, Walentynowicz, et al., 2023), and are cumbersome and unpleasant to administer (Whitehead & Drescher, 1980). As discussed within the context of respiratory interoceptive accuracy tasks, the intrinsic negative value of this task may add noise to results.

As a slightly less invasive alternative, water-load tests are frequently performed in healthy populations (van Dyck et al., 2016). Here, the amount of water required for participants to report satiety is compared to the amount required for them to report a full stomach. Water-load tests do not exactly provide a measure of the accuracy of perception of gastric signals, however, and may better be considered a measure of beliefs about interoceptive signals and their relevance (in this case, the degree of water required to feel satiated), rather than the accuracy of interoceptive perception.

Another measure of gastric interoception requires participants to drink glucose solutions of varying concentrations and report fullness whilst their blood glucose is measured (Young et al., 2019). Smith and colleagues recently proposed another alternative, which involves participants ingesting a capsule that the experimenter can make vibrate at specific times (Smith, Mayeli, et al., 2021). Participants are then asked to tap every time they feel a vibration in their stomach. However, this task is limited in that performance appears to depend on the position of the pill within the stomach, which cannot be tightly controlled (Mayeli et al., 2023).

A limitation of all these tasks of gastric interoceptive accuracy is that they each only consider one of several cues that are considered in tandem to inform one about hunger and satiety, such as blood glucose levels, gastric distension, and gastric contractions. As such, the tasks lack real-world generalisability. To give an example, even if participants detect their blood glucose is high in Young and colleague's task, they may not feel "full" if they have only been drinking and if their stomach still feels empty. Additionally, most of these tasks

may be inappropriate for use within certain clinical populations, such as those with eating disorders, for whom sensations of fullness or the ingestion of glucose may be especially triggering.

Situational approach

A further avenue for research that has largely been neglected so far is to explore the situations in which participants can demonstrate interoceptive accuracy within specific domains. In the sense that one can either feel a certain sensation at a given instance or not at a given time, moment-by-moment interoceptive accuracy is binary. Over multiple trials, one can ascertain a more continuous measure of how often they feel the heartbeats that occur. However, if all contextual factors remain the same, then someone who can detect one heartbeat at time one should also be able to detect a subsequent beat at time two. This is why many measures define participants as either interoceptive or not as a binary outcome. What may be interesting is determining the range of contexts within which participants may be interoceptive. This allows for the differentiation of, for example, participants who can never feel certain interoceptive sensations, versus those who can feel it under conditions of physical arousal which heighten interoceptive sensations (such as after exercise for heart rate, or after a period of fasting for gastric sensations). There are still issues with this that would need to be addressed though- for example, exercise or fasting may affect different participants differently, in a manner which is difficult to control for.

Accuracy and bias

Interoceptive accuracy may be affected by response bias. Response bias can be defined as when participants are systematically more likely to respond one way than another. For example, in the event of an ambiguous cardiac signal, participants could be biased to over-report heartbeats. This is slightly different to accuracy. When one is inaccurate, this

could either be because of bias, but it could also be because of random error. Tasks performance being influenced by bias is not *always* strictly a limitation of measures of interoceptive accuracy as it does not affect the validity of what they intend to measure. However, there are occasions when having a bias enhances performance, rather than causing it to decrease. It is in these scenarios where response bias becomes an issue. One example of this has been previously discussed in the chapter: the fact that those who have a bias for overestimating their heartbeats are more likely to perform well in the HCT. This enhanced performance does not arise from them being better able to accurately perceive cardiac signals but from their response biases.

Some measures are unlikely to be affected at all by response bias, such as the PAT. Other measures that employ a signal detection approach allow for the extrication of the effects of both biases and accuracy. These include certain tasks of respiratory interoceptive accuracy (Harrison et al., 2021; Petersen et al., 2014; Slotta et al., 2023) and the HDT. This is especially useful if one considers that interoceptive biases may be of clinical relevance. Indeed, there is evidence to suggest that interoceptive biases may be related to negative affect and illness anxiety (Petersen et al., 2015; Slotta et al., 2023), although other studies have found no association between interoceptive bias and anxiety (Harrison et al., 2021). Accordingly, measures that can provide indices of both accuracy and bias may provide an advancement on those that only measure accuracy.

Confounders- do they matter?

Some researchers have postulated that certain factors may affect interoceptive accuracy, and thus act as confounders that affect the validity of tasks. For example, the fact that BMI, gender, blood pressure, body fat, fitness and heart rate variability may affect cardiac interoception (Domschke et al., 2010), that breathing behaviour or morphology of the

throat may affect respiratory interoception, or that body fat around the stomach may affect gastrointestinal interoception (Desmedt, Luminet, Walentynowicz, et al., 2023). However, just because these factors may affect performance, it does not necessarily make them confounders in this context. Desmedt and colleagues argue that interoceptive accuracy measures are intended to reflect an individual's ability to detect an internal signal, which implies the intensity of the signal should remain constant between individuals. Using the example of BMI, if a higher BMI reduces the intensity of cardiac signals (Rouse et al., 1988), then by their logic it would be a confounder. However, one could also argue that what is of clinical relevance is whether an individual can detect and interpret internal signals, and therefore if one is unable to accurately perceive interoceptive signals because of their BMI, then they lack interoceptive accuracy within their own body. In this case, BMI is a factor that might influence interoceptive accuracy, rather than a confounder.

Conclusion

To summarise, many of the most used tasks of interoceptive accuracy and self-report questionnaires about interoception lack validity. Although novel measures are emerging, at present these require further amendment and validation.

Empirical evaluation of the 2AFC-HDT

Introduction

As discussed in the preceding section, the HDT task is a commonly used measure of cardiac interoceptive accuracy. In its simplest and most-used form (Adams et al., 2022; Hickman et al., 2020), this task requires participants to determine whether a series of auditory or visual stimuli are in or out of sync with their heartbeats (2AFC design). Other variants present tones at multiple delays after cardiac contraction and assess the *consistency* of synchronicity judgements (e.g., 6AFC and MCS; Brener et al., 1993; Clemens, 1984; Yates et al., 1985).

These three variants show only "adequate" convergent validity (Abma et al., 2016; Brener et al., 1993; Ring & Brener, 2018). One study reported correlation coefficients of .72 between the 6AFC and MCS, -.59 between the 2AFC and MCS, and -.55 between the 2AFC and 6AFC (the latter two negative correlations indicate that participants who are judged to be able to detect their heartbeats on the 2AFC are more consistent using the other two variants; Brener et al., 1993). These correlations are not as high as one might expect given the similar task demands of matching the timing of an external stimulus to one's heartbeat. The same study also demonstrated a discrepancy in the percentage of participants categorized as interoceptive across tasks, with 54% in the MCS, 50% in the 6AFC, and only 33% in the 2AFC.

A plausible explanation for the lower-than-expected correspondences and the reduced proportion of interoceptive participants in the 2AFC-HDT relates to the use of pre-defined delays in the 2AFC-HDT. In the 2AFC-HDT, researchers present tones that they expect to be perceived as synchronous with heartbeats approximately 200ms after the r-wave (the signature of myocardial electrical depolarization triggering ventricular contraction) and

asynchronous stimuli approximately 500ms after the r-wave. In early instantiations of the task, delays of approximately 128ms were chosen based on the logic that this was how long it would take for the pulse to reach the neck, where participants were assumed to detect their heartbeats (Körmendi & Ferentzi, 2024; Whitehead et al., 1976). This delay was also assumed to coincide with the contraction of the heart, which was when several researchers assumed that individuals perceived their heartbeat (Katkin, 1985; Schandry & Specht, 1981; Whitehead et al., 1977). The delay choices were then supported by research using the MCS and 6AFC (Betka et al., 2020; Brener et al., 1993; Brener & Kluvitse, 1988; Yates et al., 1985), which show that *on average*, participants perceive tones presented at delays between 100 and 300ms from the r-wave as synchronous with their heartbeats.

In variants of the 2AFC-HDT procedure that detect heartbeats using the arrival of the pulse wave at the finger instead of the r-wave, delays are adjusted to account for the pulse arrival time, which is the time from the r-wave to when the pulse is detected at the finger. These tasks typically use delays of 0ms (from the detection of the pulse wave at the finger) as synchronous and 300ms as asynchronous (Betka et al., 2018; Ewing, Manassei, Gould Van Praag, et al., 2017), although in practice accounting for the pulse arrival time is difficult given the substantial individual variability in the pulse arrival time according to factors such as age and arm length (Budiman et al., 2022; see Chan et al., 2007; Wibmer et al., 2014, for variation in reported finger pulse arrival times).

Other variants of the 2AFC-HDT define delays as proportions of the IBI (Knoll & Hodapp, 1992; Lyyra & Parviainen, 2018; Mul et al., 2018; Salomon et al., 2016). These methods attempt to account for the fact that it might be the point in the cardiac cycle (e.g., systole vs. diastole) rather than the absolute delay in milliseconds, that determines when participants perceive their heartbeats. At present, not enough is known about how individuals perceive their heartbeats to judge which method is better. It is logical to assume, however,

that methods which present tones at a set proportion of the IBI are more susceptible to variance in participants' heart rates, as those with a faster heart rate will have to discriminate between tones series which have a reduced absolute difference in synchrony, potentially making the task more difficult.

As detailed above, studies using MCS and 6AFC designs show *on average* individuals perceive stimuli presented at delays of approximately 200ms as synchronous with their heartbeats. However, it is notable that there is a significant range in the delays producing synchronous judgements across individuals from 0 to 400ms (Brener et al., 1993; Clemens, 1984; Yates et al., 1985). This illustrates a problematic assumption of the 2AFC-HDT: that the point within the cardiac cycle when participants perceive an external stimulus as synchronous with their heartbeat is the same for all individuals. Consider the following scenario: one interoceptive individual perceives external stimuli as synchronous with their heartbeat when presented 100 to 200ms after the r-wave, whereas another interoceptive individual perceives external stimuli as synchronous when presented 400 to 500ms after the r-wave. Although the delays at which they perceive synchrony differ, they can both perceive their heartbeats and so should both be determined to be interoceptive (Ring & Brener, 2018). However, if participants are deemed interoceptive only if they consistently select the delay deemed as synchronous by the researcher (as is the case for most studies employing the 2AFC-HDT task; e.g., Garfinkel, Tiley et al., 2016; Herman et al., 2019), the 2AFC-HDT would only categorise the former individual as interoceptive. As such, the 2AFC-HDT is liable to false negatives (i.e., individuals who are truly interoceptive are determined not to be interoceptive). This may explain the lower-than-expected correspondence between the 2AFC and the MCS and 6AFC forms of the HDT and the reduced rate of individuals deemed interoceptive when assessed using the 2AFC-HDT (Brener et al., 1993), as well as introduce noise into studies attempting to determine the relationship between cardiac interoceptive

accuracy and any other variable of interest when cardiac interoceptive accuracy is measured using the 2AFC-HDT.

If synchronicity judgements were binary, such that participants either experienced tones presented at the predefined synchronous or the predefined asynchronous delays as synchronous, then the potential for false negatives could be removed through the adoption of a consistency-based approach to scoring. Under this approach, a consistent tendency to experience tones presented at either the predefined synchronous or asynchronous delays would be judged as evidence of cardiac interoceptive accuracy (e.g., if both positive and negative values are used to indicate consistent selections using *d*' or *A*'; see Brener et al., 1993). However, in practice, some participants could experience tones as synchronous with their heartbeats when presented at delays equidistant between the predefined synchronous and asynchronous delays. It would still be possible for this subset of participants to be judged as not interoceptive when they truly are ("false negatives"), as they have no basis to consistently choose tones at either the predefined synchronous or asynchronous delays as synchronous with their heartbeats.

There is little existing data to determine the extent to which the issue of false negatives on the 2AFC-HDT is a practical problem, rather than merely a potential problem with the assumptions of the test that may not occur in practice. However, the data that does exist suggests that false negatives are a practical issue for the 2AFC-HDT. For example, Brener and colleagues (1993) found that six out of 24 participants who completed the 2AFC-HDT, and one out of eight interoceptive participants, judged tones presented at the "asynchronous" r + 384ms interval to be more synchronous with their heartbeats than the "synchronous" r + 128ms interval. Wittkamp and colleagues (2018) used the 2AFC-HDT and reported *d*' values ranging from 4.48 (indicating a consistent preference for synchronous conditions) to -4.48 (indicating a consistent preference for asynchronous conditions).

Collectively, these findings suggest that at least some participants perceive 2AFC-HDT stimuli pre-defined as *asynchronous* to be *synchronous* with their heartbeats. By way of an illustration of this problem, Ring and Brener (2018) used data from the MCS task to plot the percentage of simultaneous judgements for each of the delays at which tones are presented in the task (0 to 500ms at 100ms intervals) for 13 interoceptive participants. This plot showed that although there was an overall tendency for tones presented at a delay of 200ms to be most frequently judged as synchronous, there was a spread of delays selected by participants, from 100 to 400ms. Despite the utility of these data in showing that false negatives in the 2AFC-HDT are a practical problem, the sample was very small, the MCS delays do not span the entire cardiac cycle, and analyses are restricted to 100ms bins. Therefore, despite evidence that demonstrates that false negatives exist, existing data is insufficient to provide a reasonable estimate of their frequency.

Estimating the frequency of false negatives is made even harder due to the inconsistency in test administration. Of particular relevance is the substantial variability in the delays used as synchronous in the literature. For example, synchronous presentations range from 0ms (Hina & Aspell, 2019; Lyyra & Parviainen, 2018; Mul et al., 2018) to 300ms following r-wave detection (Gardner et al., 1990; Schirmer-Mokwa et al., 2015). Similarly, asynchronous delays vary from 380ms (Lombardo & Epstein, 1986) to 600ms (Schirmer-Mokwa et al., 2015). This heterogeneity may result in differences in the proportion of false negatives across studies, difficulty combining data in meta-analyses, failures to replicate, and potential bias arising from systematic differences in preferences across groups, resulting in erroneous conclusions regarding group differences (or lack thereof) in cardiac interoceptive accuracy.

The potential issue of false negatives on the 2AFC HDT has yet to be examined in a large sample. Existing research is limited by the 100ms intervals used in the MCS and 6AFC

HDT, which provide limited granularity and do not always span the entire IBI. Accordingly, the present study aimed to conduct a secondary analysis of a large corpus of recent work that has employed the PAT (Plans et al., 2021). Our first analysis aimed to estimate the number of participants deemed interoceptive by the PAT that would have likely been deemed non-interoceptive by the 2AFC-HDT, because the delay at which they perceive tones to be synchronous is closer to the 'asynchronous' delay in the 2AFC-HDT, or at a point that could have been perceived as equidistant between the two 2AFC-HDT pre-defined delays. The second analysis aimed to examine whether false negatives could be present even when predefining delays as a proportion of the IBI, by plotting the delays selected as synchronous by interoceptive participants as a proportion of their IBI on polar histograms. Based on prior results, we hypothesised that there would be a subset of interoceptive individuals who would likely be falsely classified as non-interoceptive on the 2AFC-HDT. However, we made no predictions about the proportion of individuals affected by this issue, given the relatively small sample sizes previously employed.

Methods

Participants

Data was pooled from previous studies (Spooner, Bird, Clemente, et al., 2024; Spooner, Bird, Irigoras Izagirre, et al., 2024; Todd et al., 2024) and unpublished work. Although there were slight variations in the inclusion and exclusion criteria across studies, broadly, participants were over the age of 18 and spoke English fluently. After initial data cleaning to remove inattentive or unengaged respondents (those with four or fewer heart rate values, or those containing 0 responses) and those with too few trials (<17 as per Spooner, Bird, Irigoras Izagirre et al., 2024; Plans et al., 2021), we applied an additional exclusion criterion for any participants who completed more than 24 trials (n = 4) indicating an issue in

terminating the task. If participants had completed the PAT on more than one occasion, due to taking part in several studies or longitudinal studies, the first valid completion was retained. This resulted in the removal of 81 task datasets. The sample size before only the interoceptive participants were selected was 357. Please see the Appendix (Appendix 1) for THE demographic information of the participants included in analyses from each study (before only the interoceptive participants were selected), and Table 2.1 for an overview of the inclusion and exclusion criteria within each study, details of ethical approval and recruitment methods.

Table 2.1

Inclusion and exclusion criteria

Study	Inclusion criteria	Exclusion criteria	Filtering criteria	Ethics approval	Recruitment methods
Study 1 (Todd et al., 2024)	29 to 65 years old Owning a compatible smartphone* Willingness to download the free application	Significant hearing loss that prevents response to auditory stimuli Presence of a diagnosed cardiac condition Pregnancy Previous diagnosis of a pain condition Presence of chronic pain symptoms in the last three months	Removal of participants with duplicate IP addresses And who completed the study in a time that was implausible (i.e., <10 minutes) or did not provide sufficient data (i.e., 80% complete).	Anglia Ruskin University (PSY-S19-023)	Prolific Academic

Study 2	Enforced in Testable: approval rating <95%, aged 18-65 years old, English as a first language Enforced in Gorilla: 18 years old or older	Uncorrected visual or auditory issues Cardiac problems Mental health conditions	Royal Holloway University 3090	Testable and posters around the Royal Holloway University Campus
Study 3	Enforced in Testable: approval rating <95%, aged 18 to 65 years old, English as a first language Enforced in Gorilla: 18 years old or older, has a compatible smartphone	Uncorrected visual or auditory issues Cardiac problems Mental health conditions	Royal Holloway University 2186	Testable

Study 4	Aged 18–35 years; at least one night of heavy drinking in the past 30 days (4 or 5 alcoholic drinks in less than 2 h for females and males, respectively); consume more than 5 drinks per week; one hangover in the past 30 days; one blackout episode in the past 6 months. Access to an Apple iPhone	No self-reported history of cardiovascular or mental health disorder.	Rutgers University Institutional Review Board Pro2021002431	Word of mouth, online fora (i.e., Reddit), student lists, and through flyers pasted in bars and breweries in Continental USA.

	meeting eligibility requirements			
Other studies	Enforced in testable: aged 18 to 60 years old, English as a first language, passed screener General: owns a compatible smartphone or iPad Cisgender	Mental health conditions Cardiac conditions	Royal Holloway University 3090 and 3056	Testable

Note. * Compatible smartphones include any iPhone with one camera lens running iOS 13 or higher. For Study 4, we only included data from the initial baseline session where participants were told to refrain from drinking alcohol for 24 hours.

The PAT

For a comprehensive outline of the PAT, please see Plans et al. (2021). Briefly, a smartphone application was used to administer the task, allowing for remote administration. This application used a heartbeat detection algorithm developed in swiftUI by BioBeats Ltd (Cropley et al., 2017; Morelli et al., 2018). As participants place their index finger over the camera lens and flash, photoplethysmography is used to detect the change in colour at the finger as the participant's pulse wave arrives.

Participants are instructed to sit somewhere quiet, with the volume of their phone maximised, and with "do not disturb" mode enabled. Subsequently, participants are required to watch a short video depicting the task instructions (see Appendix 2 for the instructions). Briefly, participants are instructed to hold their phone with one hand, with their index finger over both the camera lens and the flash. If participants move their finger, an on-screen prompt is displayed asking them to reposition their finger before recording continues.

During each trial, participants listen to a series of tones (with a duration of 156ms) at the same frequency as their heart rate, triggered by the detection of participants' heartbeats by the application. The starting delay (with respect to heartbeat detection by the application) for tones is random across trials and participants are asked to rotate a virtual dial to advance or retard the tones in time until they perceive the tones to be synchronous with their heartbeats. Each turn of the dial advances or delays the tones by a set proportion of the time between heartbeats. This is to ensure that those with slower heart rates do not have to work harder to make the tones synchronous with their perceived heartbeats. Participants complete two practice trials, followed by 20 main task trials.

Scoring

As heartbeats are periodic, simply using the standard deviations of delays cannot establish consistency (Plans et al., 2021). To circumvent this problem, we expressed each delay as a periodic function of the period. This was achieved by expressing the delay as an arc of length $2\pi d/p$ on a unitary circumference. The use of a standardized circumference for each participant allows individual variabilities in heart rate to be accounted for. In the PAT, each movement of the dial corresponds to $\alpha = d + 2\pi d/p$. This allows for the similarity of angles to be used to define cross-trial consistency.

By expressing angles as complex numbers of *modulus = 1* and *argument* $= 2\pi d/p$, consistency can be computed as the modulus of the sum of the angles divided by the number of considered angles, $consistency(d, p) = \frac{1}{n} mod(\sum_{j=1}^{n} e^{i2\pi \frac{d_j}{p_j}})$. If angles are close to one another, the modulus of their average will be close to 1; if they are randomly positioned the modulus will approach 0 (Plans et al., 2021).

Interoceptive accuracy was determined by the consistency of delays chosen across trials. To classify participants as interoceptive or not, a Gaussian mixture model with two mixtures was applied to the consistency values. This assumes that the data encompasses two sub-populations: those who are interoceptive, and those who are not. Using an expectation-maximization algorithm, the model produced a distribution for each subpopulation. To calculate the probability of each participant belonging to either the interoceptive or non-interoceptive group, z-scores were calculated for each participant for each sub-population ($z = (x - \mu)/\sigma$). This was calculated as $P_{NI}(z) = 1 - F_{NI}(|z|) + F_{NI}(-|z|)$ and $P_I(z) = 1 - F_I(|z|) + F_I(-|z|)$, with F_{NI} and F_I being the cumulative distribution functions with a mean of 0 and unitary variance. $P_{NI}(z)$ and $P_I(z)$ provide the probability of obtaining a z-score at least as

extreme as the value z that was observed for each participant, $F_{NI}(z)$ and $F_{I}(z)$ provide the probability lower than z (Plans et al., 2021).

From these probabilities, Bayes Factors were calculated, as a ratio of the probability of belonging to the interoceptive sub-population versus the non-interoceptive sub-population. In this study, participants were deemed interoceptive at Bayes Factor 30, indicating very strong evidence that they belong to the interoceptive rather than the non-interoceptive sub-population. This conservative cut-off was employed to reduce the number of false positives so that participants were more likely to be genuinely interoceptive.

Scoring was performed on the first 17 trials of the PAT task (excluding the practice trials), as this was the minimum number of trials participants had to have completed to be included, and ensured the same number of trials were included for each participant. The threshold of 17 valid trials was the pre-registered threshold for one of the studies (Spooner, Bird, Irigoras Izagirre, et al., 2024) and was thus applied for all studies (see pre-registration: https://osf.io/j4dtr). Overall, there were 86 interoceptive participants, which resulted in data from 1462 trials (see Table 2.2 for demographic information).

Table 2.2

Demographic information for the interoceptive participants across all studies

Sex (%)	Age in years M(SD)	Presence of mental health condition	Presence of physical health condition	Missing all demographic data (n)
Male: 18.8% Female: 81.2% *	28.4 (10.6)	82.5% no 10.5% yes 7.0% not recorded	40.7% no 9.3% yes 50.0% not recorded	1

Note. This table reflects the demographic information of the participants across studies who were deemed to be interoceptive at BF30. * 19 (22%) of these participants reported on gender rather than sex.

Analysis

IBI. We recorded two different measures of IBI that would reflect the IBI participants could have used to inform their delay choices: the final IBI of the trial, just before the participant clicked "continue" to confirm their delay choice, and the median IBI across the last angle array they selected (beginning from when they moved within 5° of their final selected angle). For the latter, the number of heartbeats spent within this angle array ranged from 1 to 90, with an average of 7 (*SD* = 7). As illustrated in Table 2.3, there is very little difference between these values.

Table 2.3

Descriptive Statistics of the Different IBIs

	Mean	SD	Median	Min	Max
Final IBI	0.789	0.170	0.748	0.465	1.684
Median IBI final angle array	0.788	0.170	0.747	0.450	1.682

Note. The values here reflect the statistics for comparing the values across all trials for all participants. All values are in seconds. The average difference between the final angle array average IBI and the final IBI was -1ms (*SD* = 0.2ms), max difference of 187ms. Twenty-two (1.5%) trials had a difference larger than 50ms. There was a correlation of .99 between the two IBIs across all trials.

Analysis approach 1: Comparing delays at which tones are judged to be synchronous to delays used in previous 2AFC-HDT studies. Our first analysis approach aimed to estimate how many interoceptive participants identified by the PAT would have likely been deemed to be non-interoceptive in commonly used instantiations of the 2AFC-HDT because the participants perceive synchronicity at delays that are closer to the asynchronous than synchronous delays typically used to present tones. As the PAT uses the arrival of the pulse wave at the finger to trigger tones when implemented as a smartphone application, rather than the r-wave from an electrocardiogram (ECG), we considered delays used in 2AFC-HDT studies using the arrival of the pulse at the finger to present tones at different delays.

We found seven studies that recorded the arrival of the pulse at the finger and used absolute delays in milliseconds in the asynchronous and synchronous conditions (Betka et al.,

2018; Ewing, Manassei, Gould Van Praag, et al., 2017; Hart et al., 2013; Herman et al., 2019; Kandasamy et al., 2016; Leganes-Fonteneau et al., 2019; Palser, Fotopoulou, et al., 2018). In all studies, the synchronous condition was defined as 0ms after the arrival of the pulse at the finger, and the asynchronous condition was defined as 300ms after this. All but one of these studies calculated an interoception score based on either the ratio of "correct" versus "incorrect" trials, or percentages of "correct" trials. Critically, this assumption of a correct answer means that only participants who consistently select tones after a 0ms delay are judged as interoceptive. One of the studies calculated *d*' (Ewing, Manassei, Gould Van Praag, et al., 2017). Although *d*' values can indicate consistent selection of either 0 or 300ms delays, with positive values suggesting a preference for 0ms and negative values suggesting a preference for 300ms, it is not clear from the paper whether significant negative *d*' values were considered to imply interoceptive accuracy. Based on the correlation analyses reported in the paper, it seems likely that the authors only considered consistent selection of tones at 0ms delay as indicating cardiac interoceptive accuracy, a choice that would compound false negatives. Therefore, in the majority, if not all, of these studies, participants could only be judged to be interoceptive if they perceived their heartbeats closer to tones presented at 0ms delay following the arrival of the pulse at the finger than at 300ms delay.

Only one of these studies reported instructing participants that the tones should *follow* their heartbeats, by specifying that the tones would either be synchronous or delayed with respect to their heartbeat (Kandasamy et al., 2016). Even this instruction is quite vague, and therefore participants in all or nearly all these studies could have made synchronicity judgements based on the proximity to the heartbeat triggering the tone *or* the following heartbeat. To provide the most conservative analysis of the 2AFC-HDT (i.e., the analysis that would err on the side of underreporting, rather than overreporting, the proportion of false negatives on the 2AFC-HDT), our analysis allowed for the possibility that participants judge

synchronicity to either heartbeat. For each interoceptive participant, on a trial-by-trial basis, we calculated both positive and negative values that indicated the angular distance from the preceding and following heartbeats, respectively, from the PAT data. The median value across trials of whichever distance was smaller on each trial (making the analysis conservative) was converted into a delay at which tones were perceived as synchronous using the participant's IBI (e.g., if 0/360° represents a 0ms delay, an angle of 180° corresponds to a delay of 500ms if the IBI (period) is 1000ms). When calculating this delay, the final IBI on the specific trial on which the participant selected the median delay was used. If there were multiple trials with the same delay choice, the IBI from the first of these was used. We also repeated analyses using: 1) the participant's median final IBI across trials; 2) the median IBI from the final angle array on the specific trial used in the first analysis, and 3) the median final angle array across trials; as a robustness check.

Estimated participant selection of 0ms versus 300ms delay on the 2AFC-HDT was established using the following logic (with MD indicating median selected delay):

*If 0<MD<150 or 0>MD>150-(0.5*IBI), then 0ms is selected. If not, 300ms is selected.*

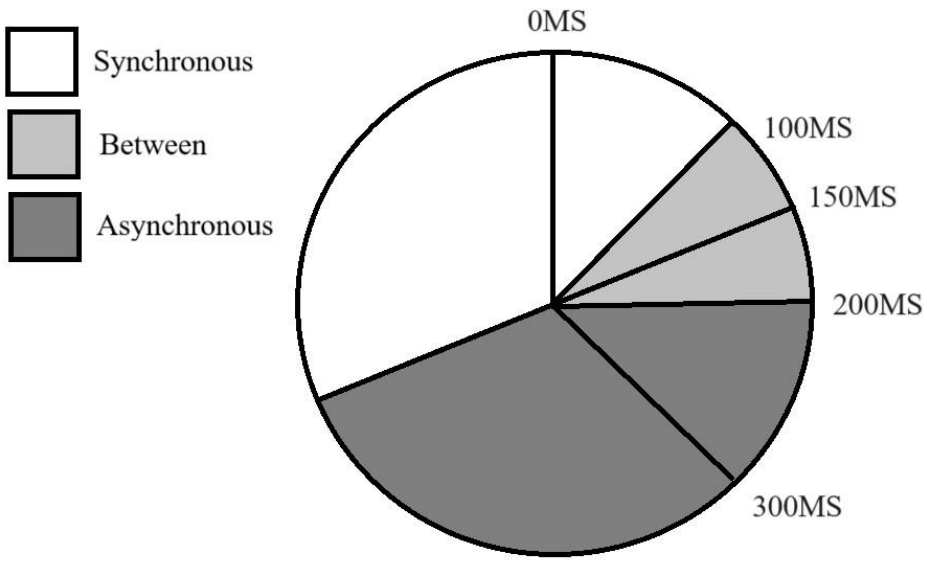
Heartbeats are temporally extended stimuli, and human perceivers do not have infinitely precise temporal perception. If heartbeats are perceived exactly between the predefined synchronous and asynchronous delays in the 2AFC-HDT, then participants may respond at chance when selecting which tones are synchronous with their heartbeats. However, due to a lack of temporal precision when perceiving heartbeats, the same may be true if heartbeats are perceived close to the midpoint between the predefined synchronous and asynchronous delays. There is not much information on the temporal precision with which individuals can perceive heartbeats, however, there is research suggesting the specificity of discrimination (*SD* of the interval perceived to be synchronous) ranges from 43 to 167ms

(Brener & Kluvitse, 1988). This suggests that within an approximately 100ms window, participants may be unable to differentiate delays. Based on this, we set a region of 100ms around the midpoint between synchronous and asynchronous delays (from 100 to 200ms, around the midpoint of 150ms) as an "ambiguous" window. If the delay at which a participant perceives synchronicity is within this range, they may randomly select between responses on the 2AFC-HDT.

Overall, our logic can be illustrated by Figure 2.1.

Figure 2.1

Categorising delay choices as synchronous, asynchronous, or between



Note. This illustrates the logic used to establish whether a participant's median delay at which they perceive synchronicity with their heartbeats corresponds to the synchronous or asynchronous delays in the 2AFC-HDT, or whether it could be perceived as falling between the two. The figure is an approximation, and not to scale.

Analysis approach 2. Illustrating delay preferences. The second analysis approach aimed to illustrate the spread of interoceptive participants' delay preferences within a cardiac cycle using polar histograms. The full 360° of the histograms reflect the entire IBI, with the detected heartbeat occurring at 0°, and the angles representing the delays the participant selected, reflecting the periodic nature of heartbeats.

This methodology allows us to address a potential limitation of our first analysis. Specifically, the majority of 2AFC-HDT tasks use ECG recordings of the heartbeat, meaning that the delay is calculated from the r-wave, rather than the arrival of the pulse at the finger (as is the case with the PAT when implemented as a smartphone application). Given the lack of large-scale studies reporting the pulse arrival time to the finger and the variability in the data that does exist, it would be difficult to estimate the delay between the r-wave and each participant's selected delay for our data. By plotting all delays on a standard cardiac cycle, we can understand whether perceived synchronicities are distributed across the cardiac cycle. If this were the case, the critique of adopting predefined delays at which participants *should* perceive tones as synchronous with the heartbeat would apply regardless of the point in the cycle at which the heartbeat is measured (i.e., regardless of whether delays are measured from the r-wave using ECG or the arrival of the pulse wave at the finger).

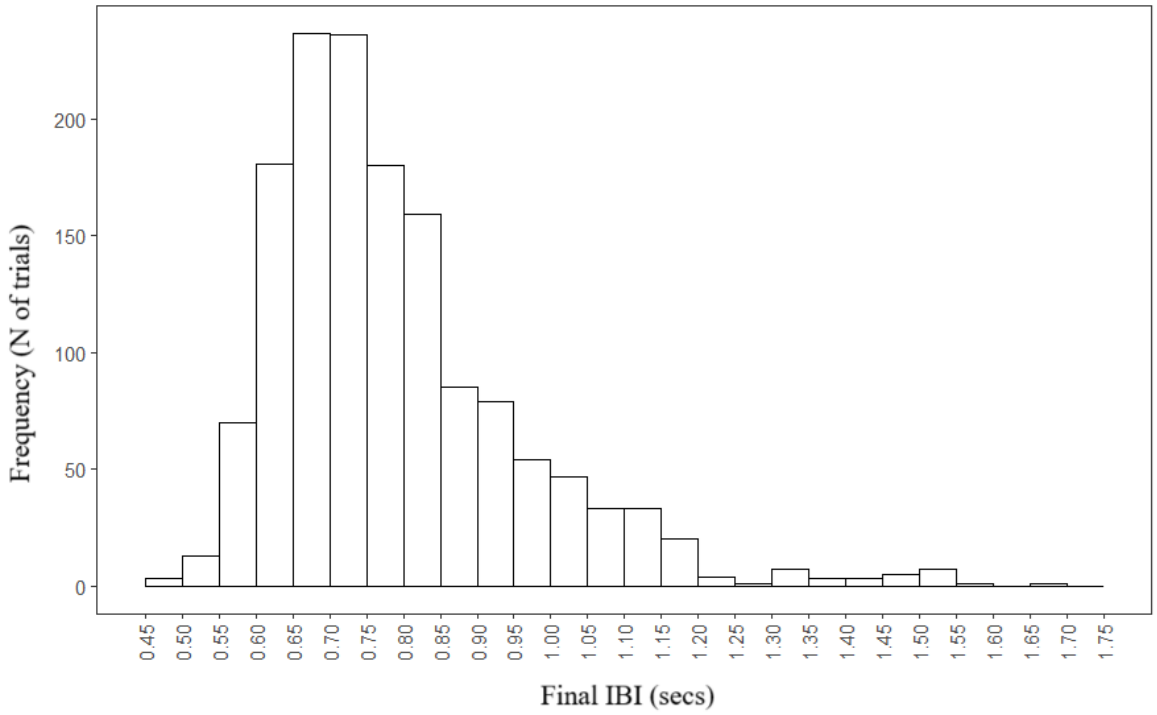
As the histogram reflects the entire IBI, the delay in milliseconds indicated by each angle varies according to the duration of the IBI. Data is plotted as a proportion of the final IBI for the trial, but there was very little difference between these values and the median of the IBI from the final angle array (see Table 2.3).

The range of delays at which participants perceive tones as synchronous with their heartbeats are plotted first regardless of their IBI (i.e., as a point in the cardiac cycle). This is relevant to variants of the 2AFC-HDT in which the predefined synchronous and

asynchronous tones are defined not in terms of a number of milliseconds from either the r-wave or detection of the pulse wave at the finger, but instead as a proportion of IBI. If significant variation is observed in this plot, it supports the claim that predefining periods when participants should perceive their heartbeat as a proportion of their IBI will result in false negatives. The second plot is similar to the first, but the plot only includes trials where the IBI was between 650 to 750ms (the modal interval, see Figure 2.2). The second plot allows variation in the delay with which participants perceive tones to be synchronous with their heartbeats to be interpreted in terms of milliseconds, as variance in the plot translates to a variable delay in milliseconds at which participants perceive tones to be synchronous with their heartbeats.

Figure 2.2

Distribution of final IBIs across trials



Results

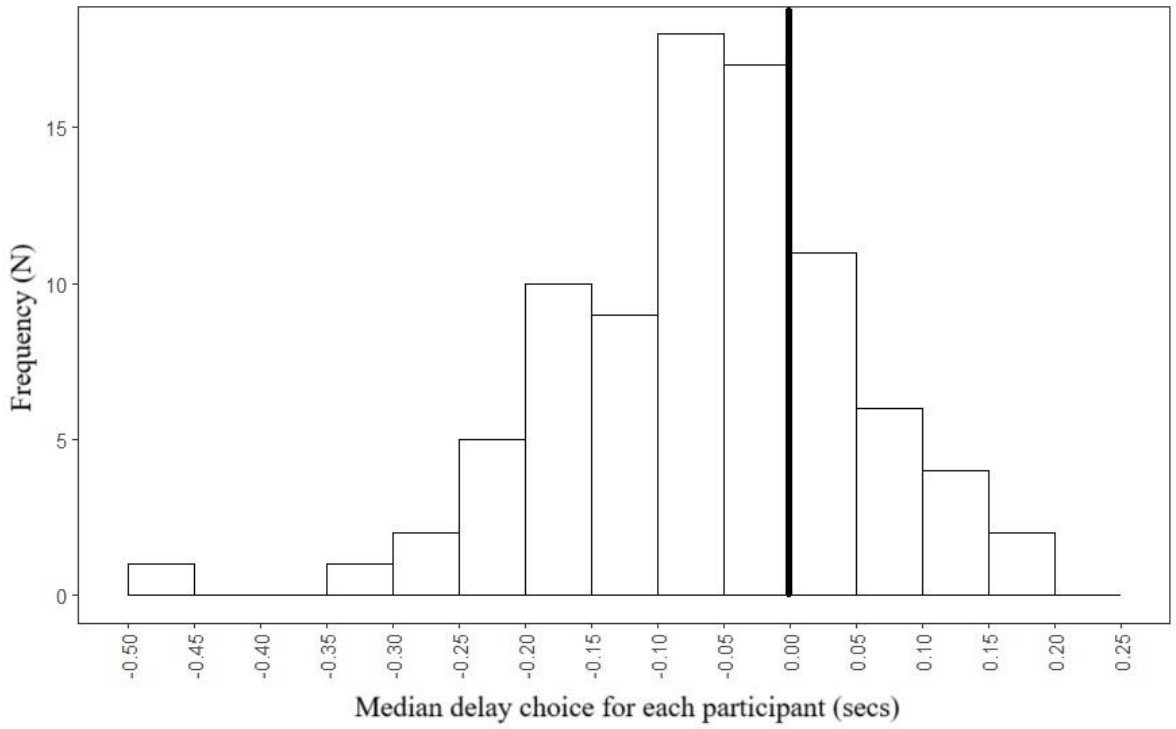
Analysis approach 1

A histogram of participants' median delay choices is illustrated in Figure 2.3. Out of the 86 interoceptive participants, one (1.2%) would have likely selected delays that were closer to 300ms than 0ms in the 2AFC-HDT and thus would have consistently selected tones presented at the predefined asynchronous delay. When this analysis was repeated using the median last IBI across all trials instead of the IBI from the specific trial the median delay was extracted from, this increased to two (2.3%) participants. When the analysis was repeated using the median IBI from the final angle array for the specific trial the median delay was selected from rather than the final IBI, no participants showed this pattern, while when using the median of all the median IBIs from the final angle array, this rose to two (2.3%) participants.

In addition to the participants identified in the analyses above, a further six participants (7.0%) showed preferences that were within 50ms on either side of the midpoint between the two conditions and therefore would likely randomly select between tones at the two predetermined delays in the 2AFC-HDT. In total, 7% to 9.3% of the interoceptive participants are likely to be incorrectly classed as non-interoceptive in the 2AFC-HDT due to the use of predefined delays.

Figure 2.3

Histogram of participants' median delay choices



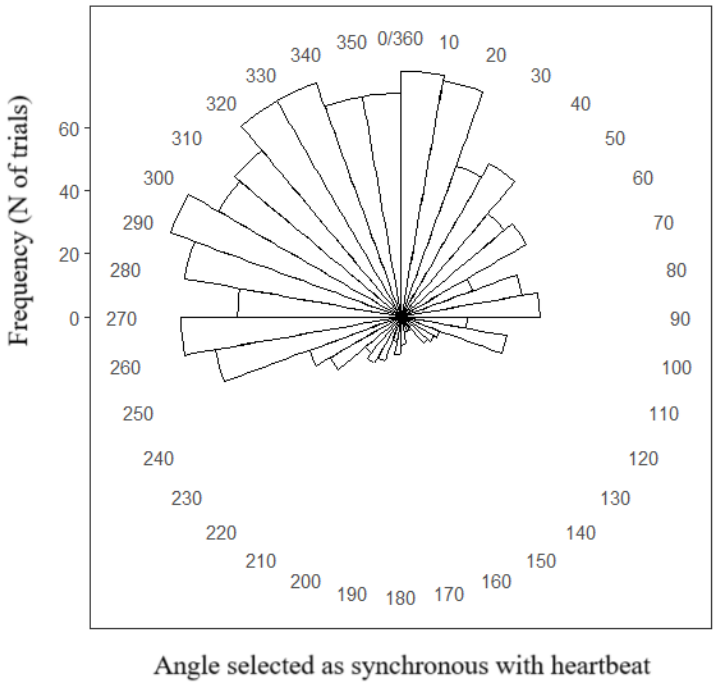
Note. The vertical line indicates the detection of the heartbeat at the finger. The median delay selected indicates the point at which participants perceived tones to be synchronous with their heartbeats.

Analysis approach 2

Figure 2.4 demonstrates significant variability in the point in the cardiac cycle at which participants perceive their heartbeat. All periods in the cardiac cycle are represented, although the bulk of the data occupies approximately 180° (half of the cardiac cycle). The same is true of Figure 2.5, which represents the same data for which the IBI is restricted to between 650-750ms. Across both plots, there is considerable variation in either the point in the cardiac cycle or the delay in milliseconds at which heartbeats are perceived.

Figure 2.4

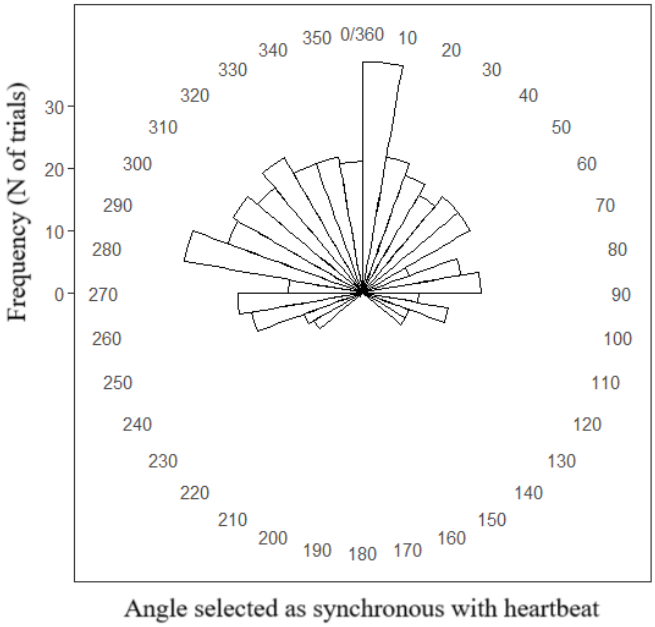
Delay choices within the cardiac cycle across all trials



Note. The angles selected by all interoceptive participants across all trials (n = 1462), with the full circle representing one complete cardiac cycle (i.e., one inter-beat interval), and 0/360° reflecting the detection of the heartbeat at the finger.

Figure 2.5

Delay choices within the cardiac cycle on trials where the IBI was between 650 to 750ms



Note. The angle selected by all participants across all trials where the final IBI was between 650 to 750ms (n = 476), with 0/360° reflecting the detection of the heartbeat at the finger.

Discussion

The aim of this study was to estimate the extent to which false negatives are possible in the 2AFC-HDT, specifically due to participants perceiving their heartbeat at a different point in the cardiac cycle to the pre-determined synchronous condition typically used in the 2AFC-HDT. First, we provided an estimate for the proportion of our interoceptive participants who consistently perceived their heartbeats at delays closer to the "asynchronous", rather than the "synchronous" delays of the 2AFC-HDT. These participants, although interoceptive, would likely be falsely deemed non-interoceptive when assessed using the 2AFC-HDT (i.e., a false negative result). We observed that approximately 2% of interoceptive participants consistently selected delays that would be classified as

asynchronous on the 2AFC-HDT task. An additional 7% of participants selected delays likely to be perceived as equidistant between the 2AFC-HDT conditions and therefore are also likely to be incorrectly classified as non-interoceptive using the 2AFC-HDT. Therefore, we estimate that approximately one in ten of our interoceptive participants would have been misclassified by the 2AFC-HDT, although estimates vary from 7% to 9.3% depending on the exact IBIs used for calculation. Note that these estimates are deliberately conservative (i.e., this estimate is likely to underestimate the true rate of false negatives due to the use of predetermined delays).

As the PAT measures the arrival of the pulse at the finger, we used the delays from studies which instantiate the 2AFC-HDT using detection of the pulse at the finger. This decision was made given the huge variability of (and therefore the difficulty of estimating) the pulse arrival time to the finger. As most 2AFC-HDT studies measure heartbeats using the r-wave, it could be argued that the issue illustrated here may not apply to all versions of the 2AFC-HDT. In addition, it could be argued that these results do not apply to studies which present tones at set proportions of the participant's IBI. However, our subsequent analyses address these issues. Delays selected as synchronous by interoceptive participants were distributed around the cardiac cycle, both when considering the selected proportion of the IBI and when considering a delay measured from *any* point in the cardiac cycle (i.e., from the pulse arrival at the finger or the r-wave). These results suggest that false negatives are likely to occur when using any instantiation of the 2AFC-HDT.

The varying frequencies of participants' preferred delays around the cardiac cycle also illustrate the issue with the range of delays that are defined as synchronous and asynchronous in previous studies using the 2AFC-HDT. The interaction between the choice of delays used in particular studies, and the specific delays preferred by each sample of participants in each study, would influence the proportion of false negatives and the specific participants who are

deemed interoceptive versus non-interoceptive. As such, the range of delays used, and of preferred delays across participants, impedes comparison across studies using the 2AFC-HDT.

This is the first time the extent of the 2AFC-HDT's false negative problem has been explored in a large sample; however, it is broadly consistent with previous findings employing smaller samples. Brener and colleagues (1993) found that of 8 interoceptive participants who completed the 2AFC-HDT, one appeared to prefer the "asynchronous" (384ms from r-wave) interval over the "synchronous" (128ms from r-wave). Our results also replicate the spread of delays selected by interoceptive participants shown by Ring and Brener (2018), but in a larger sample, and in a more fine-grained manner that spans the entire cardiac cycle. Therefore, despite differences in analytic approaches and methods of measuring heartbeats, the results of the two studies are consistent. The presence of false negatives may explain the lower-than-expected correlations of the 2AFC with the MCS and 6AFC forms of the HDT in Ring and Brener (2018), as well as the fact that a higher proportion of participants were deemed interoceptive when using the MCS and 6AFC compared to the 2AFC.

If *d*' is used to score participants' preferences in the 2AFC-HDT, and negative values beyond a critical value are taken as evidence of interoception, participants who demonstrate a consistent preference for the asynchronous condition could correctly be deemed to be interoceptive. However, this would not resolve the issue for those participants whose preferred delay falls at a time that is perceived as equidistant between the two delay options (7% of the present sample).

This would not be as much of an issue if participants were told that the tones could only follow (and not precede) their heartbeats. In the original Whitehead study (1977),

participants were told "to indicate whether he believed immediate or delayed feedback had occurred." Although this insinuates that the stimuli must follow the heartbeat, it is arguably not explicit, and it is unclear how frequently this instruction is used in contemporary research. Indeed, the majority of studies using the HDT do not report instructing participants that the stimuli must follow their heartbeat (for an exception to this, see Wiens & Palmer, 2001). Moreover, limitations in participants' perceptual acuity may limit their ability to distinguish whether a tone precedes or follows their perceived heart rate. At present, there is limited data on participants' sensitivity to disambiguate between external stimuli and heartbeats.

It is noteworthy that these data do not suggest that *only* one in ten participants would be falsely classified as non-interoceptive on the HDT – they suggest that approximately this number (which is a conservative estimate and may vary across studies) would be *misclassified based on their delay preference*. A participant may be able to perceive their heartbeat and fail the 2AFC-HDT because they have generally poor timing ability affecting their ability to make synchronicity judgments, have difficulty concentrating throughout the test, have poor motivation, are distracted by pain, are made anxious by the test procedure, or any one of several other reasons. The way to identify these participants would be to employ a control task which has the same task demands as the 2AFC-HDT but does not require the ability to perceive one's heartbeats for successful task performance. Any participant who fails both the 2AFC-HDT and the control task would be unclassifiable, one cannot determine which of the task requirements they cannot meet and so they cannot be classified as interoceptive or non-interoceptive. Unfortunately, a significant proportion of studies using the HDT do not include such a control task, and therefore the percentage of participants incorrectly classified as non-interoceptive despite being truly interoceptive is likely higher than illustrated by our study. Although we cannot say with confidence how severe the increase in those incorrectly classified will be, in a laboratory study using a non-interoceptive

control task where participants completed a version of the PAT where the heartbeats were replaced by a second tone, 17% of participants were judged to have failed the task (Spooner, Bird, Clemente et al., 2024; Spooner, Bird, Irigoras Izagirre et al., 2024).

The problem with using predetermined delays in the 2AFC-HDT does not only impact the measurement of cardiac interoceptive accuracy but also cardiac interoceptive “insight” (also called interoceptive metacognition or awareness; Suksasilp & Garfinkel, 2022), a measure which has been claimed to have clinical relevance (e.g., Nord & Garfinkel, 2022). Interoceptive insight refers to the degree of correspondence between an individual's assessment of their interoceptive accuracy and their actual interoceptive accuracy. When assessed with the 2AFC-HDT, it is typically measured as the correspondence between the accuracy of judgements and confidence in those judgements, across trials (e.g., Livermore et al., 2022). The use of predetermined delays affects the calculation of both accuracy on the 2AFC-HDT and confidence judgements on each trial. For example, "the participant was directed to respond to whether the tones were in or out of time with their heartbeats, and how confident they were in that answer using a visual analogue scale ranging from ‘total guess’ to ‘complete confidence’ on a scale of 1 to 10” (from Livermore et al., 2022; p2292). A participant who can accurately perceive their heartbeat may perceive neither of the delays as being synchronous, but they can select which of the two is more synchronous with their heartbeat, resulting in a consistent selection of, for example, the delay predetermined to be synchronous. This participant would be fully justified in responding with little confidence that the tones were synchronous, even though they performed well on the task. The participant would be judged as lacking interoceptive insight, even though their interoceptive insight is perfect.

A limitation of this study should be acknowledged, in that we assumed a 100ms window within which participants could not judge differences in synchronicity. This figure

was based on data from Brener and Kluvitse (1988), but this was the only data we found that we could use to estimate this window. In our first analysis, we could only compare participants' delay choices to other 2AFC-HDT tasks which measure the heartbeat at the finger, limiting generalizability to the majority of 2AFC tasks which measure from the r-wave. Moreover, in our second analysis, we could only plot trial-by-trial data, rather than participants' preferences. Although one must be consistent to be deemed interoceptive, analysis allows for some leeway such that participants may have a couple of trials where they are distracted or make a mistake and therefore select a delay that is inconsistent with the other trials. It may be that these trials cause the spread of delays deemed as synchronous. However, we could not use the median here, as if a participant had values that were close together, but were on either side of 0°, the use of only positive angles would mean median values that were not reflective of the actual spread of delay choices. Despite these limitations, it is important to note that this study is not intended to provide definitive numbers of the number of interoceptive participants which would have been incorrectly classified as non-interoceptive because of the point in the cardiac cycle they detect their heartbeat at, but rather to illustrate with data that it is possible and to estimate how many people it may affect given certain assumptions. Future research could explore this further by conducting both the 2AFC and PAT on a sample of participants and comparing results.

Conclusion

To conclude, this study demonstrates that the 2AFC-HDT may result in false negatives, whereby interoceptive participants are incorrectly misclassified as non-interoceptive due to individual differences in the point in the cardiac cycle at which participants perceive heartbeat-tone synchrony. Additionally, the use of the 2AFC-HDT will also likely result in incorrect calculation of interceptive insight or awareness. This is

important as it affects the validity of conclusions drawn when comparing interoceptive and non-interoceptive groups, as determined by the 2AFC-HDT.

Chapter Three: The Phase Selection Task Pilot

Kiera Adams, David Plans, Geoffrey Bird and Jennifer Murphy designed this study. Geoffrey Bird and Jennifer Murphy designed the PST task and Kiera Adams, Jennifer Murphy and Geoffrey Bird adapted it. Kiera Adams developed the scoring system for the PST. Kiera Adams recruited participants and conducted experiments. Kiera Adams developed the analysis script with feedback from Jonathan Bird. Kiera Adams interpreted the results and wrote up the pilot study. Feedback was given by Jennifer Murphy, Geoffrey Bird and Caroline Catmur.

Abstract

Cardiac interoceptive accuracy is the most researched aspect of interoception, however, measures that are commonly used to assess it lack validity. This preliminary pilot study of 32 participants assessed the feasibility of a novel measure of cardiac interoceptive accuracy: the Phase Selection Task (PST). This task, which is delivered by smartphone, requires participants to rank three tone series from most to least in sync with their heartbeats. Within each series, each tone is delayed with respect to each heartbeat by a consistent amount, with the delays differing in each series. Across trials, the delays for each of the three series are the same. The consistency of each participant's ranking across trials is assessed against chance, with above-chance consistency considered to be evidence that participants can perceive their heartbeats. This pilot showed that the task was feasible for remote deployment. However, a small number of participants found the anti-cheat mechanism to be overly sensitive, and some minor technical issues emerged that require further refinement. Most participants could complete the control task, suggesting the non-interoceptive specific task demands were not too challenging. Over half of the participants who successfully completed the control task were deemed interoceptive. Overall, this suggests that with some slight technical refinements, the PST could provide a valid and time-efficient way to measure cardiac interoceptive accuracy remotely.

Introduction

Limitations of existing measures of cardiac interoceptive accuracy have been outlined in the previous chapter. In response to these limitations, a novel measure of cardiac interoceptive accuracy was recently developed (Plans et al., 2021). The Phase Adjustment Task can be delivered by smartphone, allowing access to sample groups that may be less likely to visit laboratories. The PAT is less liable to false positives than the HCT, and less liable to false negatives than the HDT (see Chapter Two).

Despite the initial promise of the PAT, there are limitations to the task. It can still be time intensive, as participants have unlimited time to try and perceive their heartbeats, and it lacks a matched multisensory control task. To enhance confidence in the construct validity of the PAT, it would also be useful to have another task to validate it against. If there is a high concordance between the two tasks, it would suggest they both measure the same underlying construct despite differences in the specific tasks and differing levels of non-interoceptive task demands.

To address these concerns, the Phase Selection Task (PST) was developed. Like the PAT, the PST is delivered by smartphone. The PST asks participants to rank three tone series from most to least in sync with their heartbeats, where each series contains individual tones that are delayed from each heartbeat by a consistent amount which is different across each series. As such, it can be seen as a hybrid of the 2AFC and MCS, with fewer delays than the MCS (resulting in a shorter administration time) and avoiding the issue of having predefined "correct" answers in the 2AFC (thereby reducing the rate of false negatives, see Chapter Two). Provided that results from the PAT and PST show good convergence, the PST could provide a useful alternative to the PAT in time-constrained experiments.

This study aimed to conduct a small-scale pilot to establish the feasibility of the PST task. Specifically, we aimed to establish whether the app was usable for participants, whether participants felt they understood the instructions for the task, whether a sufficient number were able to pass the anti-cheat tests and the control task, and to establish the average time taken to complete both the PST and the control task.

Methods

Participants

Inclusion criteria were that all participants were between 18 and 60 years of age, and fluent English speakers. There were no exclusion criteria. Participants were either recruited from the Research Participation Scheme recruitment system at the University of Oxford and received credits for participation or responded to social media adverts and participated for an honorarium of £10. Participants were given the option of completing the study remotely, if they had an appropriate iPhone with one camera lens running iOS 13 or higher, or of coming into the laboratory and completing the study on a test phone. Ten participants elected to complete the study in the laboratory, and the remaining participants completed it remotely.

Ethical approval was obtained from the local ethics committee (R62582/RE001) and informed consent was obtained from participants.

The App

A smartphone application was used to administer the PST task. This app used a heartbeat detection algorithm developed in swiftUI by BioBeats Ltd (Cropley et al., 2017; Morelli et al., 2018). In short, as participants place their index finger over the camera lens and flash, the phone's camera detects the change in colour at the finger as the participant's pulse wave arrives. Therefore, it indicates the timing of heartbeats, delayed by the pulse arrival time. This app has been demonstrated to be of comparable accuracy to laboratory-based measures, such as ECG and pulse oximeter, and does not produce a tactile heartbeat sensation

in the finger due to pressure, meaning participants cannot perform the task using somatosensory information (Murphy, Brewer, et al., 2019).

PST. Before each trial participants had to place the index finger of one hand over the flash and camera lens and put five digits of the other hand on a blue box (5.8 by 4.0cm as measured on an iPhone 8) on the screen. Once their digits had been recognised as having been placed on the screen, they were able to slide them over the screen to a comfortable position but could not lift them off the screen. This was to prevent participants from cheating by measuring their pulse with their hand, thus circumventing the major risk associated with remote testing of cardioception.

There were two practice trials, followed by 18 experimental trials. Before each trial, there was a 500ms delay, plus a randomly selected delay between 270 and 1350ms (so the trials did not always start at the same point in the cardiac cycle). Participants' heart rates were recorded across the entire trial. In each trial, participants were played three sets of tones, each containing three individual tones that were triggered by their heartbeats. Each tone lasted for 80ms. One of the series of tones had a 0ms delay from when the pulse was detected at their finger by the app, one had a 150ms delay, and the final set of tones had a 300ms delay. Each series of tones was presented alongside a distinct coloured shape. The specific combinations of shapes and colours and the orders in which they appeared varied randomly across participants. After each tone series was presented, there was a delay of 1000ms + half the average inter-beat interval (IBI) before the next series was presented. The following heartbeat triggered the next series of tones. Each potential order of these three sets appeared three times, presented in a random, counterbalanced manner.

Participants were asked to order the shapes in terms of how close they perceived each series of tones to be to the time of their heartbeats (“When asked, please rank each of the shapes in terms of how synchronous they were with your heartbeat by tapping on them. The

shape you tap on first should be the one you perceive as most synchronous with your heartbeat, and the one you tap on last should be the least synchronous."; for full instructions please see Appendix 3). If they lifted their fingers off the screen during the trial, the anti-cheat mechanism would be activated, and that trial would be voided. That order of tone series would be entered back into the pool of potential upcoming trials and would be repeated. If the anti-cheat mechanism was triggered too many times (over six times in a task, a third of all experimental trials), then the task was ended. Participants were able to take breaks if needed between the trials and did not need to hold their fingers on the screen during this time.

Control. The control task was matched in format to the PST, with both requiring participants to integrate and compare the timing of sensory stimuli across two modalities and store this information in working memory to compare to other stimulus pairs. However, rather than comparing the timing of heartbeats to the timing of sounds, the control task required participants to compare the timing of visual information (a flashing shape with a duration of 40ms) to the timing of sounds. Therefore, it relied on the comparison of two exteroceptive stimuli, rather than one interoceptive and one exteroceptive. If participants failed both the control task and the PST this suggests that non-specific task demands, rather than a lack of cardiac interoceptive accuracy, may be affecting PST performance. For example, if participants struggled to hold their fingers on the screen whilst completing the task, or to hold their delays in working memory over the length of a trial, this would be seen in both the control and PST tasks.

Procedure

Participants completed the tasks either remotely while supervised by the experimenter via videocall, or in the lab. Thus, they could ask questions if they needed to, and any issues with the app could be identified by the experimenter.

Following consent, demographic and screening information were collected on Qualtrics. This included: gender, age, height, weight, frequency of physical exercise in a week, smoking habits, whether they have a heart condition, whether they take medication for a heart condition, whether they have mental health issues, whether they take medication for mental health issues, whether they are pregnant, and whether they own an iPhone with one camera running iOS 13 or higher. We did not exclude participants based on the screening questions at this stage.

The app portion of the study lasted 25 to 30 minutes. A baseline measure of participants' heart rate was taken for 20 seconds before the task began. If their heart rate was above 90 beats per minute, participants were told that they were not able to complete the task at this time. The reason for this was that their IBI would be too short for proper task administration, as they may misinterpret the tone with the longest delay as synchronous with their next heartbeat. If their heart rate was flagged as too fast, they were asked to relax and try again in a few minutes. If this issue continued, participants were able to leave with full credits or compensation, and a note was made. Participants were told to sit down somewhere where they wouldn't be distracted, and taken through the instructions on the app. The experimenter was either in the next room in the laboratory or was online on the call with their video turned off. This allowed participants to ask questions if required.

After completing the counterbalanced tasks, they were redirected back to Qualtrics, where they completed the debrief questions. For both the PST and the control task, participants were given an open question about what they thought they were asked to do. Seven-point Likert scales (ranging from strongly disagree to strongly agree) were then provided to assess the degree to which they agreed with the statements outlined in Table 3.1.

Table 3.1

Debriefing questions

Task	Question
PST	
	I found I could feel my heartbeat during the heartbeat and sound task.
	I found that I could detect differences in the delays between my heartbeat and the sounds.
	I found it possible to order the shapes and sounds in order of time after my actual heartbeat.
Control	
	I found it possible to order the shapes and sounds in order of time after the flashing shape.
	I found that I could detect differences in the delays between the flashing shape and the sounds.
General	The app was easy to use.
	It was easy for me to learn how to use the app.
	I understood how to move between screens.
	I found that I could feel my heartbeat whilst using the app.
	I can usually feel my heartbeat, even if not using the app.
	I encountered no glitches or technical issues in the task.

Note. The potential answers were as follows: Strongly disagree; Disagree, Somewhat disagree; Neither agree nor disagree; Somewhat agree, Agree; Strongly agree.

If the participant reported experiencing glitches or technical issues (meaning if their responses ranged from strongly disagree to neither agree nor disagree) then they were given a text box to report details of these issues.

Analysis

There is no objectively "correct" answer, due to individual differences in the point of the cardiac cycle at which participants may detect their heartbeats (see Chapter Two). Instead, consistency was measured.

A participant was deemed to be interoceptive if they put the same delay in the same position more than 10 times. The number 10 was derived from a binomial probability calculator (https://stattrek.com/online-calculator/binomial) given 18 trials, with a one in three chance of a participant being successful for that given position, with a multiple comparison correction for the three positions. The same criteria were also used in the control task.
The reason that the participants only had to be consistent across one position, rather than across all three, was that it is feasible that participants perceive their heart to beat exactly at the mid-point between two of the delays. For example, a participant may detect their heartbeat at exactly 75ms, which is halfway between the delay of 0ms and the delay of 150ms. This means, whilst being interoceptive, they may switch between selecting 0ms and 150ms as the closest to the heartbeat. Although we give the instructions to participants that they should select the tones that follow rather than perceive the heart rate, this assumes that participants can follow these instructions and that their acuity is enough that they can determine which follows or precedes with an accuracy of 75ms. To ensure the PST remains valid even when these assumptions are violated, we can just look at one position. In this case, although an interoceptive participant may alternative between 0ms and 150ms in the first and second positions, they should consistently put 300ms in the third position. This overcomes some of the issues of the 2AFC discussed in the previous chapter.

A participant was deemed to be not interoceptive if they were not interoceptive in the PST, but they passed the control.

To assess whether participants who were deemed interoceptive in the task also self-reported feeling their heartbeats and thus considered themselves to be interoceptive, agreement statements were collapsed into did not agree (strongly disagree, disagree, somewhat disagree, neither agree nor disagree), and agree (strongly agree, agree, somewhat agree). The decision to combine neither agree nor disagree and the disagree responses was due to the limited number of participants in our analyses, and to allow for comparisons between those who explicitly agreed with the statement and those who did not. Chi-square analyses were conducted using SPSS (version 29.0.2.0 (20)) to assess whether interoceptive or non-interoceptive participants were more likely to agree with the statements: "I found it possible to order the shapes and sounds in order of time after my actual heartbeat," I found that I could feel my heartbeat whilst using the app" and "I can usually feel my heartbeat, even if not using the app." It is important to note that a lack of concordance would not necessarily imply that the task is not valid, rather it could indicate that participants simply have low meta-cognitive awareness of their cardiac interoceptive accuracy. Nonetheless, if those who felt they were able to feel their heartbeats and complete the task accurately actually performed better, it would enhance our confidence that the task is indeed measuring what it is intended to measure.

Results

Of the 32 participants, four participants activated the anti-cheat mechanism too many times (over six times) and were therefore unable to complete the required 18 trials for both the PST and control task. One participant did not enter the same participant ID into Qualtrics as on the PST. As their demographic information could not be matched, they were not included in analysis. None of the participants were pregnant, and none reported having a

heart condition, although one participant reported taking medication for a heart condition. Seven reported having diagnosed mental health conditions, and of these, three participants were taking medication for these mental health conditions. See Table 3.2 for the demographic information of the 27 participants who completed enough trials for us to assess their performance.

Table 3.2

Demographic information

	Total (N)	*Interoceptive (n)*	*Non-interoceptive (n)*
Gender			
Female	19	13	6
Male	8	3	5
Other	0	0	0
Age			
18 to 21	16	12	4
22 to 25	7	2	5
Over 25	3	1	2
Did not answer	1	1	0
Body mass index			
Underweight (<18.5)	3	1	2
Normal (18.5 – 24.9)	19	12	7
Overweight (25 – 29.9)	3	1	2
Obesity Class I (30 – 34.9)	0	0	0
Did not answer	2	2	0

Note. The interoceptive participants are those whose PST performance indicated that they were able to accurately detect their heartbeat (see requirements in Analysis section).

Of those who completed both the control task and the PST, 25 passed the control task, and 16 passed the PST. Only one participant passed the PST but failed the control task. The average time taken to complete the study (including demographic and screening questionnaires and both tasks) was around 40 minutes.

We were underpowered to be able to compare interoceptive and non-interoceptive participants on demographic features, however, we report on the demographic information of both groups in Table 3.2 for completeness. Further work is required to confirm measurement invariance of the PST across different groups.

In terms of the feasibility of use, most participants either agreed or strongly agreed that the app was easy to use, easy to learn how to use, and they understood how to move between screens (~60%, 70% and 75%, respectively). For a more detailed breakdown of these results, see Table 3.3.

Table 3.3

Participant responses to the debriefing questions

Question	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
I found I could feel my heartbeat during the heartbeat and sound task.	4	5	4	2	13	3	1
I found that I could detect differences in the delays between my heartbeat and the sounds.	0	9	7	2	8	5	1
I found it possible to order the shapes and sounds in order of time after my actual heartbeat.	0	9	7	2	8	3	1

I found it possible to order the shapes and sounds in order of time after the flashing shape.	1	0	4	6	10	10	1
I found that I could detect differences in the delays between the flashing shape and the sounds.	1	0	3	4	13	10	1
The app was easy to use.	0	0	6	3	4	14	5
It was easy for me to learn how to use the app.	0	1	2	1	5	16	7
I understood how to move between screens.	0	1	0	1	5	13	12

I found that I could feel my heartbeat whilst using the app.	3	5	7	3	8	6	0
I can usually feel my heartbeat, even if not using the app.	2	8	6	4	8	2	2
I encountered no glitches or technical issues in the task.	1	2	1	2	0	12	14

Note. This table shows the number of participants who responded with each of the options for the debriefing questions, including participants who could not be matched to the app data.

Although responses to the question "I found it possible to order the shapes and sounds in order of time after my actual heartbeat" did not significantly differ according to whether or not participants were deemed interoceptive based on their task performance (χ^2 (1, N = 27) = 1.03, p = .309, φ = .20) this may have been due to the limited power of the analysis. Indeed, when inspecting the response frequencies, the expected pattern was observed (see Table 3.4), whereby more of those who agreed with the statement were interoceptive, and most of those who disagreed with the statement were not interoceptive.

Table 3.4

I found it possible to order the shapes and sounds in order of time after my actual heartbeat.

	Agree	Disagree
Interoceptive	9	7
Not interoceptive	4	7

Responses to the question "I can usually feel my heartbeat, even if not using the app" were also assessed (see Table 3.5). The reason for this was to provide an indication of the real-world applicability of findings of the PST and the degree to which performance on the task reflects how much participants can feel their heartbeat in day-to-day life. Again, it is important to note that this is not necessary for the validity of the task per se, as the task measures if participants can detect their heartbeat when instructed to focus on it, rather than if individuals are both able to and have a high proclivity for detecting their heartbeat in everyday life. There was no significant association between whether participants were interoceptive and the degree to which they agreed with this statement (χ^2 (1, N = 27) = 0.00, p = .952, φ = .00).

Table 3.5

I can usually feel my heartbeat, even if not using the app.

	Agree	Disagree
Interoceptive	6	10
Not interoceptive	4	7

We conducted the same analysis on the question “I found that I could feel my heartbeat whilst using the app” (Table 3.6). Deviation between this and task performance could indicate either that participants were not adequately paying attention to the post-task questionnaire, or that non-specific task demands constrain the performance of participants who can actually feel their heartbeats. Once again, this was not significant (χ^2 (1, N = 27) = 1.03, p = .309, φ = .20), although the frequencies did show the expected pattern.

Table 3.6

I found that I could feel my heartbeat whilst using the app.

	Agree	Disagree
Interoceptive	9	7
Not interoceptive	4	7

Open answers to the questions “In the task where we asked you to focus on your heartbeats and sounds, what do you think you were asked to do?” and “In the task where you were asked to focus on a sound and a flashing shape, what do you think you were asked to do?" demonstrated that participants had all been adequately focused on the instructions on the tasks and that they had an adequate understanding of what they were required to do. An example of an answer for the experimental task was "Try and rank the tones that closest aligned with my heartbeat, from closest to furthest” and an answer (from a different participant) for the control task was “Rank the shapes based on the delay between the flash and tone, shortest delay to longest.”

Four participants reported technical issues with the app. One participant reported that the sound did not work on the “test beep” screen, one participant reported the app crashing and activating the anti-cheat, despite their fingers being on the screen, one participant reported the screen freezing and two participants reported irregularities in the tones reflecting their heartbeats, however, this may not have been an issue with the app and may have reflected genuine heartbeat variability.

Three participants made other comments. One highlighted that the hand positioning required was awkward, which is in keeping with the number of participants excluded for activating the anti-cheat. One suggested further instruction could have been added for the PST task, and one advised that an instruction to turn on "do not disturb" mode on their phones would have helped.

Discussion

The results of this pilot demonstrated that the PST appears to be a feasible way to measure cardiac interoceptive accuracy remotely. Most participants agreed that the app was easy to use and that it was easy for them to learn how to use it. Data was collected seamlessly, and enough participants completed the control task to allow us to establish that the task demands are not too high, and there are no ceiling or floor effects.

Four participants did, however, report technical issues with the app. Two participants noted that there was some irregularity in the tones reflecting their heart rate. This is not necessarily an app issue, as it is normal for individuals to experience slight arrhythmia in their heart rate. The other technical complaints referred to over-sensitivity of the anti-cheat mechanism and the app crashing or freezing.

It also appeared from our results as though the anti-cheat mechanism may be too difficult for participants, and when deployed on a large scale, its use in its current instantiation may result in a large amount of valid data being discarded. The difficulty of the

anti-cheat mechanism is reflected by the fact that four participants had to be excluded for activating the anti-cheat mechanism over six times, and another participant left a comment about this in the post-task questionnaire. This issue may be especially pertinent if participants are using a phone with a smaller screen, or if they have longer nails. Amending this anti-cheat mechanism could address both these concerns, as well as the technical issue that one participant experienced. One way to do this would be to enlarge the square that participants must keep their fingers in.

Validity was tentatively explored by assessing whether participants who self-reported feeling as though they were able to feel their heartbeats were deemed interoceptive based on their PST performance. These results should be interpreted with caution. Firstly, due to the small sample size, these results are very underpowered. Moreover, rather than purely indicating the validity of the PST, the agreement between self-reports of the ability to detect interoceptive signals and objective accuracy reflects meta-cognition of interoceptive accuracy. It may be that some participants have poor metacognition, which causes a discrepancy between self-report and objective performance, rather than suggesting that the task is not actually measuring interoceptive accuracy. Nonetheless, cohesion between these measures would provide a degree of support for the validity of the task. Although none of these results were significant, inspection of the frequency of responses indicated that all results were in the predicted direction, in that those who self-reported feeling their heartbeats did seem better represented in the group of interoceptive participants compared with the non-interoceptive participants. A larger-scale replication of this would be beneficial.

The proportion of interoceptive participants in this pilot is slightly higher than one would expect, given results obtained from the HDT in the past. In this study, 60% of participants who passed the control task were deemed interoceptive by the PST. However, in their review, Brener and Ring (2016) found that only 25% of participants in the 2AFC HDT

task were deemed interoceptive, which was substantiated by later empirical findings that only 27% of participants were deemed interoceptive using the MCS task (Ring & Brener, 2018). To some degree, this is expected. The 2AFC allows for false negatives (see the previous chapter) and the MCS is very time-consuming, which may have affected participants' focus by the end of the task. Both of these factors may have resulted in an underestimation of the number of interoceptive participants. Indeed, Brener and Ring (2016) highlight that the low estimates of interoceptive participants are more likely to suggest a limitation of the tasks than the genuine proportion of those who are interoceptive.

Considering our results, we decided to make some minor changes to the PST. Firstly, we have decided to increase the size of the coloured box participants must place their fingers on to reduce the likelihood of cheating. This is to make hand placement easier and reduce the number of participants who are excluded due to accidentally activating the anti-cheat. We have also decided to increase the time taken to measure resting heart rate to two minutes and have increased the heart rate limit to 100 beats per minute to reduce the number of participants who are unable to proceed with the tasks due to their heart rates being too high. Given we only require consistency in one ranking for one delay, it will not affect results if participants mistake the tone as triggered by the second rather than the first heartbeat. Finally, to make the control task slightly easier, we have added 200ms onto each delay, consistent with the average pulse arrival time that would delay each heartbeat from the tones in the PST itself. This also increases comparability between the PST and control task.

Overall, this is extremely promising. As with the PAT, the use of the PST would allow more remote populations or clinical samples that would normally be unable to travel to laboratories to be reached, thus enhancing the generalisability of interoception research. The existence of another, appropriate measure of cardiac interoceptive accuracy also allows for

the cross-validation of the PAT. The shorter timeframe of this task also makes it less demanding for participants than the PAT.

In conclusion, the PST shows promise as a novel measure of cardiac interoceptive accuracy. Following some minor technical refinements, concordance with the PAT can be assessed.

Chapter Four: Validation of the Negative Interoceptive Evaluation Scale

Content adapted from: **Adams, K.A.,** Catmur, C. & Bird, G. (Submitted to Royal Society Open Science). Validation of the Negative Interoceptive Evaluation Scale

Kiera Adams designed the Negative Interoceptive Evaluation Scale, with feedback from Geoffrey Bird. Kiera Adams designed the study, which Geoffrey Bird approved. Kiera Adams recruited participants, created the analysis scripts, and conducted and interpreted analyses. Kiera Adams wrote the chapter, with feedback from Geoffrey Bird and Caroline Catmur.

Abstract

At present, there are no validated measures that aim solely to assess the evaluation of interoceptive signals. As negative interoceptive evaluation has been theorised to be linked to numerous mental health conditions, including anxiety, eating disorders and somatoform disorders, it is important to be able to operationalise this construct to allow for empirical testing of theories. This chapter presents the validation of a novel measure of interoceptive evaluation: the Negative Interoceptive Evaluation Scale (NIES). Convergent and divergent validity, test-retest reliability and internal consistency were tested, and the factor structure was explored with both exploratory and confirmatory factor analyses. Overall, the NIES appears to be a promising measure of interoceptive evaluation, although further work is required to determine the factor structure.

Introduction

Interoceptive evaluation, also known as interoceptive attribution (Suksasilp & Garfinkel, 2022), is a broad construct that has largely been overlooked within the interoception (but not clinical) literature. It refers to the way in which one interprets interoceptive signals, including their valence, the degree of arousal they induce, and the subsequent emotions and cognitions associated with them. Both the subjective experience of noticing a certain interoceptive signal and the judgements one makes about what that signal reflects are part of this evaluation. Interoceptive evaluation can be either adaptive or non-adaptive. Adaptive evaluation is where certain interoceptive signals are interpreted in a suitable (usually aversive) way as to induce behavioural changes that facilitate homeostasis. For example, the state of hunger should be suitably aversive so that it motivates one to eat. However, if the state of hunger is so aversive that it causes one to have a panic attack, that is a non-adaptive response. Although overly positive *or* negative evaluations could theoretically both be non-adaptive, the limited extant literature tends to focus on overly negative evaluations.

Negative interoceptive evaluation is closely related to "catastrophising"- a key concept within the literature on panic disorder (Clark, 1986). When one catastrophises, one worries excessively about the worst possible outcome indicated by a certain interoceptive sensation. This causes anxiety, which causes physiological symptoms that further exacerbate anxiety in a positive feedback loop. An individual may detect that their heart is beating faster than normal, which they interpret as a sign that there is something wrong- for example, that they are having a heart attack (Ehlers, 1993). This causes an increase in anxiety, which causes one's heart to beat even faster, thus resulting in even greater anxiety.

There is limited empirical evidence, and some theoretical work, to suggest that negative evaluation of interoceptive signals may be important in a range of mental health

conditions. In line with the catastrophising pathway discussed above, the association between negative interoceptive evaluation and panic and other anxiety disorders is the best-supported association (Clark et al., 1997; Ohst & Tuschen-Caffier, 2018; Paulus, 2013; Taylor et al., 1991). Recent qualitative research (see Chapter Six) highlights the importance of interoceptive evaluation in the experiences of anxiety of autistic adolescents specifically. Preliminary evidence also suggests eating disorders may be associated with interoceptive evaluation (Herbert, 2020; although see Brown et al., 2017) and indeed there is a logical basis to expect that a relationship may exist between an overly negative interpretation of gastric-related signals and unhealthy eating habits. There is also some evidence that those with depression may also worry more about interoceptive signals (Fissler et al., 2016), although other studies have shown no relationship between anxiety about physical sensations and symptoms of depression (Olthuis et al., 2014; Smári et al., 2003). Although there is currently a lack of empirical evidence, an overly negative preoccupation with somatic symptoms is part of the definition of somatic symptom disorders and illness anxiety disorders (Wolters et al., 2022). Indeed, self-reported worrying about interoceptive sensations appears to relate to the degree of symptom reporting in those with somatoform disorders (Flasinski et al., 2020).

Existing research operationalises interoceptive evaluation in vastly different ways, impeding comparisons across studies. One measure, used by Clark and colleagues (Austin et al., 2006; Clark et al., 1997) is the Body Sensations Interpretation Questionnaire, which asks participants why they might be experiencing a certain sensation (e.g.: “You feel short of breath. Why?”). The original version allowed participants to give an open-ended response, whereas the modified version also asks participants to rank three options in terms of which they would be most likely to assume. This measure was employed by the seven studies assessed in Ohst and Tuschen-Caffier’s meta-analysis (2018), which found an association between interoceptive evaluation and anxiety. The issue with this measure is that there are

reasons one may be more inclined to interpret bodily signals negatively that do not pertain directly to interoceptive evaluation. For example, one might have a health condition that makes negative interpretations more likely to reflect reality. It also only considers one's interpretation of what the signals could mean, rather than one's subjective experience of them when they occur. For example, when prompted, a participant may answer that their heart beating fast could indicate a panic attack, but this might not reflect how they feel when they actually notice their heart beating fast, nor whether that causes them anxiety. Moreover, there is limited data available on the psychometric properties of this measure.

Herbert and colleagues (2012) assessed interoceptive evaluation by asking participants to rate their evaluations of cardiac interoceptive signals whilst completing a heartbeat counting task. The disadvantage of this is that it only focuses on one interoceptive sensation and may be confounded by either participants' anxiety because of the unfamiliar experimental setting and evaluation apprehension, or their sense of reassurance that the experimenter would intervene if there was anything seriously wrong. Other studies have used self-report questionnaire measures designed for other purposes, such as the Anxiety Sensitivity Index (ASI; Reiss et al., 1986). The issue with using the ASI as a proxy for interoceptive evaluation is that it was designed as a measure of propensity to panic. As such, it only focuses on the interoceptive domains that are linked to the physiological symptoms of panic attacks and therefore may over-estimate the relationship between anxiety and interoception.

The Not Worrying subscale of the Multidimensional Assessment of Interoceptive Awareness (MAIA; Mehling et al., 2012, 2018) has also been used to assess interoceptive evaluation, although this only considers one type of evaluation (worrying) and has been demonstrated to have questionable internal consistency (Brown et al., 2017; Mehling et al., 2018). At present there is not a validated measure that has explicitly been produced to assess

purely interoceptive evaluation across a range of interoceptive domains.

The objective of this study was to assess the properties of a novel questionnaire measure of negative interoceptive evaluation (i.e., the degree to which individuals interpret signals across the main interoceptive domains as negative), the Negative Interoceptive Evaluation Scale (NIES). Initial analyses included item-total correlations, test-retest reliability, correlations to assess convergent and divergent validity, and comparisons across groups expected to differ in their extent of negative interoceptive evaluation. With respect to convergent validity, positive correlations with the ASI (both total and physical subscale scores) were expected, as this measure has been used previously to assess negative interoceptive evaluation, and negative correlations with the Not Worrying sub-scale of the MAIA, which has been used as a measure of adaptive interoceptive evaluation. With respect to divergent validity, a lack of convergence between measures of interoceptive accuracy and attention (the Interoceptive Accuracy Scale and the Interoceptive Attention Scale, respectively) was expected. Subsequently, an exploratory factor analysis (EFA) on the NIES was conducted, to establish the factor structure of the questionnaire. Considering both the results of the EFA and feedback from participants, adjustments were made to the questions. A confirmatory factor analysis (CFA) was then performed using the number of factors garnered from the exploratory factor analysis as a basis to assess the fit of the identified factor structure. Finally, the internal consistency of the measure and each of its factors was assessed.

Methods

Development of the Negative Interoceptive Evaluation Scale (NIES)

A comprehensive literature review was conducted to ascertain which signals are most unanimously deemed to be interoceptive. As outlined in Chapter One, these were judged to be: cardiac, respiratory, gastric, fluid regulation, and internal pain signals. Gastric signals

were split into hunger and satiety for this questionnaire. The literature review also informed the three types of negative evaluation included in the NIES: not liking a signal, the degree to which the signal is anxiogenic, and the degree to which the signal suggests that something is wrong. Qualitative interviews were simultaneously conducted with thirteen autistic adolescents who are affected by significant anxiety about their experiences of interoception (see Chapter Six), and findings were used to inform the questionnaire design. Two experts in the field of interoception also provided feedback on the proposed questions, which were amended accordingly.

Participants

All participants were over 18 years old and fluent English speakers. There were no exclusion criteria, to increase the representativeness of the sample. Based on guidelines suggesting that at least 15 participants should be recruited for each item in factor analyses (Field, 2005), we aimed to recruit a minimum of 315 participants for the EFA, and another 315 participants for the CFA. The study was conducted in line with the Declaration of Helsinki ('World Medical Association Declaration of Helsinki', 2013), and ethics approval was granted by the Oxford University Medical Science Division's Ethics Committee (reference: R86431/RE003).

Participants were recruited via social media, the Oxford Psychology Research Participation portal and using existing research databases. Participants were entered into a prize draw to win one of sixteen £25 vouchers. If completing the questionnaire at two time points, participants received two entries into the prize draw. The experiment was also hosted on the Research Participation Scheme, the University of Oxford's system whereby undergraduate students can participate in studies in exchange for course credits.

From the initial EFA sample (Sample One), 18 participants who had consented for their contact details to be retained for future studies were randomly selected to provide

feedback. They received £3 in Amazon vouchers for their time.

Measures

- The Interoceptive Accuracy Scale (IAS): Participants rate the degree to which they agree with 21 statements pertaining to the accuracy of their perception of somatic signals on a five-point scale. Internal consistency and test-retest reliability are good for this measure (α = .88, *r*(115)= .754; Murphy et al., 2020).
- The Interoceptive Attention Scale (IATS): This measure includes 21 items about attention to interoceptive signals. Participants are required to rate their agreement on a five-point Likert scale. Internal consistency is excellent for this measure (α = .905) and test-restest reliability is good (*r*(144)= .709; Gabriele et al., 2022).
- The Negative Interoceptive Evaluation Scale (NIES): This measure includes 21 items assessing the negative ways in which one might interpret eight different interoceptive signals. Participants are required to rate their agreement on a five-point Likert scale.
- The Anxiety Sensitivity Index 3: The Anxiety Sensitivity Index-3 (ASI-3; Taylor et al., 2007) is a 16-item scale specifying different concerns someone could have regarding their anxiety. The participant rates each item by selecting one of five phrases: “very little” (zero points), “a little” (one point), “some” (two points), “much” (three points), and “very much” (four points). This scale includes a subset of six items pertaining to anxiety surrounding physical sensations which was also used. This has been shown to have good psychometric properties in both clinical and non-clinical samples (Taylor et al., 2007).
- The Multidimensional Assessment of Interoceptive Awareness Version 2 (MAIA-2) “Not Worrying” subscale: this includes five questions assessing the tendency not to worry or experience emotional distress with sensations of pain or discomfort. Items are rated on a six-point scale from never to always (Mehling et al., 2018). Internal

consistency is questionable ($\alpha = .67$) and authors do not report on the test-retest reliability.

Procedure

The study was hosted on Qualtrics. Participants were directed to an information sheet via a link sent by email. After confirming they had read this, they could move on to a short screening questionnaire that confirmed they were over 18 and spoke fluent English. Provided these criteria were met, they were directed to the consent forms, followed by the main questionnaire.

Participants were asked for their gender, age, ethnicity, and any mental or physical health conditions they had. This was followed by the questionnaire measures listed in the preceding section. The duration of participation was between ten and twenty minutes.

28 days after completing the survey, a subset of 222 participants who completed the initial survey after the first round of recruitment were sent another link and a unique participant code. They were asked to enter this code at the beginning of the questionnaire and to complete the NIES a second time. Participants were told that they had seven days in which they could complete the questionnaire before the link would expire.

After data collection for the first time point had finished for Sample One, 18 participants who were randomly selected to provide feedback were sent a link to another Qualtrics questionnaire with three open-text questions:

1. What do you think this questionnaire was trying to measure?
2. Was there anything about this questionnaire that you did not understand or that was not clear?

3. This questionnaire was meant to assess how negatively you interpret bodily signals. Based on this, do you think there are any improvements we could make to our questionnaire?

A second sample was recruited for the CFA (Sample Two). Sample Two completed the same questions as Sample One. The NIES questionnaire was the same, although one question had been removed (see Results) and the order of questions had been randomly shuffled, to assess whether the associations found in the EFA replicated and were therefore independent of order effects.

Analysis

All analysis was conducted on IBM SPSS Statistics (version 29.0.2.0 (20)) unless otherwise specified, where r (Version 4.3.2) and RStudio (2023.12.1 + 402) were used. All data is publicly available (https://osf.io/tcre5/?view_only=c397bae6ff6746229d90ada1f9f36746).

To reduce the likelihood of bots, all participants who completed the survey in under 180 seconds, who entered nonsensical answers into the open text questions, or whom Qualtrics flagged as potential bots in their ReCAPTCHA check were removed. Those who had not completed the questionnaire were also removed. This resulted in the removal of 733 responses from Sample One, and 385 responses from Sample Two. Participants who had more than four missing responses were also removed, as there is a chance they may be poorer responders who were paying less attention, and this may have impacted the multiple imputation used for missing data. In Sample One, this resulted in four participants being removed, and in Sample Two, this resulted in three participants being removed.

0.14% of data was missing from responses in Sample One (82 out of 64,855 items), and 0.22% in Sample Two (143 out of 64,855 items). To check data was missing at random,

the Multiple Imputation Function on SPSS was used. As data was randomly missing across items and participants, imputation methods were selected. Missing data was imputed using the missForest package on r (Stekhoven, 2022), which is a nonparametric imputation method based on a random forest (Breiman, 2001). Data was imputed for every item for the IAT, IAS, NIES, ASI-3 and MAIA-2. Age was not imputed- as only two values were missing in both samples these participants were removed from the correlation analysis between age and the NIES. Initially, the imputation was conducted with responses as integers, but as this gave a higher normalised root mean square error (NRMSE) value indicating a poor model fit (over 0.5 for both datasets) this was repeated with the Likert scale responses coded as factors, using the proportion of false classification (PFC) as a measure for model fit. For the first and second samples, PFC values of 0.307 and 0.318, respectively, indicated the fit of the imputed values was acceptable. A comparison of histograms between imputed and non-imputed datasets showed no differences. Subsequent analyses were conducted on the imputed dataset.

Item breakdown. For all items of the NIES, mean, standard deviation, normality and floor and ceiling effects were investigated.

Item total correlations. Correlations between individual items and the total NIES scores were calculated for both datasets.

Convergent/divergent validity. All continuous variables were checked for normality, both visually, using Q-Q plots and histograms, and using statistics for skewness, kurtosis, Kolmogorov Smirnov and Shapiro-Wilk statistics. To ensure homoscedasticity, residuals were plotted against the fitted values in r using the ggplot2 package, and the Breusch-Pagan test was run using the bptest function in the lmtest package (Hothorn et al., 2022). Where data met assumptions of normality and homoscedasticity, Pearson's correlations were conducted. Where they did not, Spearman's were conducted.

Convergent validity was assessed by correlating scores from the NIES with relevant questionnaire responses assessing closely related constructs- namely the total score and Physical Subscale of the ASI-3 and the "Not Worrying" Subscale of the MAIA-2. To establish divergent validity, scores were correlated with scores for the IATS and IAS. This was conducted in both Sample One and Sample Two. The correlation between age and the NIES was also assessed to measure measurement invariance.

Test-retest reliability. Variables were checked for normality and homoscedasticity in the manner described above. For those in Sample One who completed responses at two time points, responses for the NIES at Time 1 were correlated with their responses at Time 2.

Group analyses. T-tests were used to assess whether interoceptive evaluation differed between those with and without self-reported mental health conditions (firstly, only including those with a diagnosis, then secondly including those that reported self-diagnosing) and between those with and without physical health conditions. A one-way analysis of variance (ANOVA) was used to assess whether scores differed across genders, and where significant, post-hoc Tukey's tests were used to conduct follow-up analysis.

EFA. To ascertain the number of factors to input into the model, multiple methods were used, and the results were compared (Fabriger & Wegener, 2012). We conducted a principal component analysis (PCA), scree plot, parallel analysis and map test then performed a factor analysis using the number of factors identified by each of these methods to see which resulted in the most interpretable model (meaning the model which was closest to every item loading onto one factor only). Both the PCA and scree plot were conducted using SPSS's pre-set options. For both the parallel analysis and map test, an analysis script was used (https://drive.google.com/file/d/1Zof__tq89Yys9o7MV1AxvkV4-fPP7-NP/view). For the parallel analysis, 1000 datasets were computed, and the desired percentile was 95%.

For the method of extraction for the factor analysis, principal axis factoring was used. The extraction was based on a set number of factors, which had been determined using each of the four methods discussed above. For the method of rotation, we wanted to allow factors to correlate, so we used an oblique method. We began with the promax, as this is one of the most used oblique methods of rotation. We then repeated results with a direct oblimin rotation but as results were not notably different, we report just the promax results here. The cut-off value for extracting factors out of the pattern matrix was 0.30. If an item loaded with a higher value than this, it was said to load onto that factor.

CFA. CFA was conducted using the lavaan package in r (Rosseel, 2012). The code used can be found at: https://github.com/KieraAdamsDPhil/Validation-of-NIES. We used weighted least square mean and variance adjusted (WLSMV) estimators, as has been recommended for Likert scale items. All items were scaled prior to analysis. We ran CFA on various models identified through the EFA, to compare which fit this second sample best. As AIC and BIC are not provided with WLSMV estimated models, we compared our models using the χ^2 value, the CFI, TLI, RMSEA and SRMR values. We report the robust values to allow for deviations from normality.

Internal consistency. Cronbach's alpha was used to test the internal consistency of the questionnaire, for both the total score and within the factors identified through the CFA.

Multiple comparison adjustments. For all *p* values for analyses aside from the assumption checks, the significance level was adjusted using the Benjamini-Hochberg false discovery rate correction (FDR; Benjamini & Hochberg, 1995). This is less conservative than the Bonferroni approach and would result in fewer false negatives given the large number of comparisons conducted in this study. The FDR was applied separately to the two different data sets.

Results

Sample One

763 participants were recruited for the first study (see Table 4.1).

For gender, two responses were missing. 562 participants (73.7%) were female, 154 participants (20.1%) were male, and 45 (5.9%) were classified as other (this included: agender, genderfluid, non-binary, genderqueer and queer).

431 participants had no mental health condition (56.5%) and 332 had a mental health condition (43.5%). When including those who disclosed that they suspected or self-diagnosed a mental health condition, then 358 (46.9%) had a mental health condition, meaning 405 (53.1%) did not. 214 participants had a physical health condition (28.0%) and 549 did not (72.0%). See the Appendix (Appendices 4 and 5) for a breakdown of the mental and physical health conditions reported by participants.

582 (76.3%) of participants were white, 113 (14.8%) of participants were Asian or Asian British, 11 (1.4%) were Black or Black British, 20 (2.6%) were classified as other and 37 (4.8%) either chose “prefer not to say” or left the question blank.

Table 4.1

Questionnaire results for Sample One

	Possible range	Mean (*SD*)	Median	Range	Skewness	Kurtosis	Kolmogorov Smirnoff	Shapiro Wilk	*N*
Age *	18 upwards	25.94 (8.87)	24	18 - 74	2.202	6.122	.185 (<.001)	.777 (<.001)	761
IATS Total	21 - 105	53.43 (12.19)	54	21 - 94	-.022	.048	.037 (.016)	.996 (.031)	763
IAS Total	21 - 105	74.09 (10.60)	74	28 - 105	-.150	.728	.032 (.068)	.993 (.002)	763
NIES Total Time 1	21 - 105	68.80 (12.42)	70	27 - 105	-.186	.104	.060 (<.001)	.994 (.006)	763
NIES Time 1 only for those who completed at Time 2	21 - 105	65.53 (13.87)	67	27 - 99	-.090	-.163	.062 (.200)	.994 (.829)	149
NIES Time 2	21- 105	65.07 (13.50)	66	28 - 98	-.170	-.265	.053 (.200)	.994 (.825)	149
ASI-3 TOTAL *	0 - 72	33.56 (15.84)	33	0-72	.098	-.760	.057 (<.001)	.985 (<.001)	763

ASI-3 Physical *	0 - 24	9.98 (6.29)	9	0-24	.329	-.807	.084 (<.001)	.964 (<.001)	763
MAIA-2 Not Worrying	0 - 25	12.28 (5.25)	13	0 - 25	-.059	-.617	.063 (<.001)	.988 (<.001)	763

Note. Due to the large sample size, the Kolmogorov Smirnoff and Shapiro Wilk tests were likely to be oversensitive to violations of normality. Thus, histograms and Q-Q plots were also visually inspected. Those with a non-normal distribution based on these visual checks are indicated with an asterisk.

Sample Two

413 participants were recruited for the second study (see Table 4.2).

For gender, 4 responses were missing. 303 participants, (73.4%) were female, 97 participants (23.5%) were male, and 9 (2.2%) were classified as other (this included: demi-girl and non-binary).

258 participants had no mental health condition (62.5%) and 155 had a mental health condition (37.3%). When including those who disclosed that they suspected or self-diagnosed a mental health condition, then 173 (41.6%) of the sample, had a mental health condition, meaning 240, (58.1%) did not. 121 participants had a physical health condition (29.3%) and 292 did not (70.7%). Please see the Appendix (Appendices 6 and 7) for a breakdown of the mental and physical health conditions reported by participants.

271 (65.6%) of participants were white, 86 (20.8%) of participants were Asian or Asian British, 3 (0.7%) were Black or Black British, 33 (8%) were classified as other and 20 (4.8%) either chose “prefer not to say” or left the question blank.

Table 4.2

Questionnaire results for Sample Two

	Possible range	Mean (*SD*)	Median	Range	Skewness	Kurtosis	Kolmogorov Smirnoff	Shapiro Wilk	*N*
Age *	18 upwards	28.18 (9.39)	26	18 - 59	.970	.313	.139 (<.001)	.898 (<.001)	411
IATS Total	21 - 105	54.84 (12.47)	55	21 - 86	.021	-.390	.036 (.200)	.996 (.292)	413
IAS Total	21 - 105	77.41 (10.91)	78	42 - 104	-.108	.112	.042 (.082)	.994 (.134)	413
NIES Total	20 - 100	63.93 (12.19)	64	26 - 96	-.185	-.103	.035 (.200)	.996 (.314)	413
ASI-3 Total *	0 - 72	32.70 (16.03)	32	2 - 72	.274	-.637	.057 (.003)	.981 (<.001)	413
ASI-3 Physical *	0 - 24	9.62 (5.90)	9	0 - 24	.490	-.521	.092 (<.001)	.962 (<.001)	413
MAIA-2 Not Worrying *	0 - 25	12.65 (2.07)	13	5 - 19	.052	.417	.098 (<.001)	.961 (<.001)	413

Note. * indicates distributions that were not normal when histograms and Q-Q plots were inspected.

Item breakdown

Table 4.3

Item breakdown Sample One

Number	Question	Mean (*SD*)	Item-total correlation	*p*
1	I don't like it when I can feel my heart beat fast.	4.0 (1.0)	.46	<.001
2	I don't like it when I feel out of breath.	4.3 (1.0)	.38	<.001
3	I don't like it when I feel hungry.	3.6 (1.1)	.31	<.001
4	I don't like it when I feel full.	3.6 (1.2)	.35	<.001
5	I don't like it when I feel thirsty.	4.0 (1.0)	.27	<.001
6	I don't like it when I feel like I need to use the bathroom.	3.7 (1.1)	.31	<.001
7	I don't like it when I can feel a pain inside of my body (like a stomach-ache).	4.5 (0.7) *	.19	<.001
8	If I focus on my heart beating, it makes me feel anxious.	3.4 (1.4)	.56	<.001
9	If I focus on my breathing, it makes me feel anxious.	3.1 (1.5)	.58	<.001
10	If I focus on how hungry I am, it makes me feel anxious.	2.6 (1.3)	.64	<.001
11	If I focus on how full I am, it makes me feel anxious.	2.7 (1.4)	.63	<.001

12	If I focus on how thirsty I am, it makes me feel anxious.	2.6 (1.3)	.56	<.001
13	If I focus on whether I need the bathroom, it makes me feel anxious.	3.0 (1.3)	.52	<.001
14	If I focus on pain inside my body, it makes me feel anxious.	3.4 (1.3)	.56	<.001
15	When I feel my heart beating fast, it makes me feel like something is wrong.	3.8 (1.1)	.56	<.001
16	When I feel out of breath, it makes me feel like something is wrong.	3.4 (1.3)	.60	<.001
17	When I feel hungry, it makes me feel like something is wrong.	2.2 (1.1)	.54	<.001
18	When I feel full, it makes me feel like something is wrong.	2.4 (1.3)	.54	<.001
19	When I feel thirsty, it makes me feel like something is wrong.	2.4 (1.2)	.50	<.001
20	When I feel like I need to use the bathroom, it makes me feel like something is wrong.	2.2 (1.1)	.53	<.001
21	When I feel a pain inside of my body, it makes me feel like something is wrong.	4.0 (1.0)	.41	<.001

Note. This shows the means and standard deviations for each of the questions. * indicates a potential ceiling effect. All correlations are Spearman's to account for non-normal distributions.

The distribution of responses for each question was visually inspected and means and standard deviations are presented in Table 4.3. Both Questions 2 and 7 could be interpreted as being subject to ceiling effects, with Question 7 showing a more marked effect. Correlations between the items and total NIES scores are also shown in Table 4.3. All items are above the cut-off of 0.2 which would indicate an acceptable level (Kline, 1986), aside from Question 7. This indicates that this question may not address the same underlying construct as the rest of the questionnaire and should perhaps be removed. It is interesting to note that correlations are much higher for the questions that ask about feeling anxious or feeling like something is wrong when compared to the questions that ask about not liking a signal. This could be because the first two address thoughts pertaining more directly to anxious cognitions, and as there are double the number of these questions compared to those about not liking interoceptive signals, the total score is more likely to reflect this.

We wanted to be conservative when removing questions, as every question removed would result in a greater difference between the EFA and CFA items. As such, we decided to only remove Question 7 at this stage.

The same item-by-item breakdown for the second sample is shown in Table 4.4.

Table 4.4

Item breakdown Sample Two

Number	Question	Mean (*SD*)	Item-total correlation	*p*
1	I don't like it when I can feel my heart beat fast.	3.9 (1.0)	.45	<.001
2	I don't like it when I feel out of breath.	4.3 (0.9)	.35	<.001
3	I don't like it when I feel hungry.	3.8 (1.1)	.33	<.001
4	I don't like it when I feel full.	3.5 (1.2)	.42	<.001
5	I don't like it when I feel thirsty.	4.0 (1.0)	.27	<.001
6	I don't like it when I feel like I need to use the bathroom.	3.7 (1.1)	.34	<.001
7	If I focus on my heart beating, it makes me feel anxious.	3.3 (1.4)	.50	<.001
8	If I focus on my breathing, it makes me feel anxious.	2.9 (1.4)	.52	<.001
9	If I focus on how hungry I am, it makes me feel anxious.	2.6 (1.3)	.58	<.001
10	If I focus on how full I am, it makes me feel anxious.	2.7 (1.4)	.56	<.001
11	If I focus on how thirsty I am, it makes me feel anxious.	2.5 (1.2)	.60	<.001

12	If I focus on whether I need the bathroom, it makes me feel anxious.	3.0 (1.4)	.56	<.001
13	If I focus on pain inside my body, it makes me feel anxious.	3.4 (1.2)	.52	<.001
14	When I feel my heart beating fast, it makes me feel like something is wrong.	3.8 (1.2)	.56	<.001
15	When I feel out of breath, it makes me feel like something is wrong.	3.4 (1.3)	.58	<.001
16	When I feel hungry, it makes me feel like something is wrong.	2.2 (1.2)	.59	<.001
17	When I feel full, it makes me feel like something is wrong.	2.3 (1.3)	.60	<.001
18	When I feel thirsty, it makes me feel like something is wrong.	2.4 (1.2)	.61	<.001
19	When I feel like I need to use the bathroom, it makes me feel like something is wrong.	2.2 (1.2)	.58	<.001
20	When I feel a pain inside of my body, it makes me feel like something is wrong.	4.0 (1.0)	.38	<.001

Note. This shows the means and standard deviations for each of the questions. All correlations are Spearman's to account for non-normal distributions.

All items in the adapted version of the NIES used for the CFA sample show correlations that are all above 0.2, and all are significant. This suggests that the NIES does measure an underlying construct to which all of the items relate. Once again, the correlations are higher for the questions which ask about one feeling anxious, or thinking something is wrong.

Convergent and divergent validity

Table 4.5 shows all of the effect sizes for the convergent and divergent validity assessments.

Table 4.5

Convergent and divergent validity across both samples

	Sample One					Sample Two				
	r	*p*	*N*	Size	Type	*r*	*p*	*N*	Size	Type
Convergent										
MAIA-2 Not Worrying	-.48	<.001	763	Moderate	Pearson	-.22	<.001	413	Small	Spearman
ASI-3 Total	.63	<.001	763	Large	Spearman	.64	<.001	413	Large	Spearman
ASI-3 Physical	.57	<.001	763	Large	Spearman	.56	<.001	413	Large	Spearman
Divergent										
IATS	.38	<.001	763	Moderate	Pearson	.39	<.001	413	Moderate	Pearson
IAS	-.22	<.001	763	Small	Pearson	-.20	<.001	413	Small	Pearson
Age	-.13	<.001	761	Small	Spearman	-.05	.355	413	N/A	Spearman

Convergent validity. ***Sample One.*** We found a significant negative correlation between the Not Worrying scale of the MAIA-2 and negative interoceptive evaluation, as measured by the NIES ($r(761) = -.48$, $p < .001$). According to Cohen (1988), this indicates a moderate (approaching large) correlation. The Bruesch Pagan test and graphs indicated no violation of the assumption of homoscedasticity ($\chi^2 = 0.53$, $p = .465$).

There was a large, significant positive correlation between general anxiety sensitivity, as measured by the ASI-3, and negative interoceptive evaluation, as measured by the NIES ($r(761) = .63$, $p < .001$). Similarly, there was a large, significant positive correlation between anxiety sensitivity for physical signals specifically and negative interoceptive evaluation ($r(761) = .57$, $p < .001$).

Sample Two. There was a significant, small, negative correlation between the Not Worrying subscale of the MAIA-2 and the NIES ($r = -.22$, $p < .001$) and large, significant, positive correlations between the total ASI-3 scores and the NIES ($r = .64$, $p < .001$) and the scores for the Physical subscale of the ASI-3 and the NIES ($r = .56$, $p < .001$).

Divergent validity. ***Sample one.*** We found a significant, moderate positive correlation between interoceptive attention, as measured by the IATS, and negative interoceptive evaluation, as measured by the NIES ($r(761) = .38$, $p < .001$). The Bruesch Pagan test and graphs indicated no violation of the assumption of homoscedasticity ($\chi^2 = 0.45$, $p = .503$).

A small but significant negative correlation was found between interoceptive accuracy, as measured by the IAS, and negative interoceptive evaluation, as measured by the NIES ($r(761) = -.22$, $p < .001$). The Bruesch Pagan test and graphs indicated no violation of the assumption of homoscedasticity ($\chi^2 = 1.83$, $p = .176$).

There was a small, significant negative correlation between age and negative interoceptive evaluation ($r(759) = -.13$, $p < .001$).

Although all correlations were significant, it is important to note that because of our large sample size, even very small correlations would be deemed significant. Despite their statistical significance, the correlations between age and interoceptive accuracy and negative interoceptive evaluation are both small.

Sample Two. We found a significant, moderate, positive correlation between interoceptive attention, as measured by the IATS, and negative interoceptive evaluation, as measured by the NIES ($r(413) = .39$, $p < .001$). Bruesch Pagan tests indicated no violation of the assumption of homoscedasticity ($\chi^2 = 1.65$, $p = .198$).

A small but significant negative correlation was observed between interoceptive accuracy, as measured by the IAS, and negative interoceptive evaluation, as measured by the NIES ($r(413) = -.20$, $p < .001$). Again, the Bruesch Pagan was non-significant ($\chi^2 = 1.03$, $p = .310$). There was no significant correlation between age and the NIES ($r = -.05$, $p = .355$).

Test-retest reliability

222 participants were asked to complete the NIES again 28 to 35 days after initially completing it. Of those who were asked, 67.1% of participants who completed the measure at Time 1 also completed the measure at Time 2 ($N = 149$). There was no significant difference in any demographic variables between those who did and did not respond to requests to complete the follow-up. There was a significant correlation between Time 1 and Time 2 scores ($r(147)=.67$, $p < .001$), which can be seen as reflecting good test-retest reliability (Cicchetti, 1994). The Breusch-Pagan test was non-significant ($\chi^2 = 2.08$, $p = .149$).

Group comparisons

Sample One. An independent samples t-test was conducted to assess the difference in NIES scores between participants who self-reported a physical health condition and those who did not. When adjusting for multiple comparisons, there was no significant difference between those who did not report a physical health condition ($N = 549$, $M = 68.2$, $SD = 12.3$)

and those who did (N = 214, M = 70.2, SD = 12.6) (t(761) = 2.00, p = .023). Levene's test demonstrated no violation of the assumption of homoscedasticity (F(1,761) = .00, p = .967).

An independent samples t-test did find a significant difference in the NIES scores of those who self-reported no diagnosed mental health conditions (N = 431, M = 65.9, SD = 12.1) and those who did report a diagnosed mental health condition (N = 332, M = 72.6, SD = 11.9), (t(761)= 7.64, p < .001). Those with self-reported, diagnosed mental health conditions scored significantly higher on the NIES. The Cohen's d value was .56, indicating a moderate effect size. Again, Levene's test indicated no violation of the assumption of homoscedasticity (F(1,761) = 1.12, p =.290).

The same was found when including those who self-reported self-diagnosed mental health conditions. Those with self-reported mental health conditions (N = 358, M = 72.6, SD = 11.9) scored significantly higher than those with no mental health conditions (N = 405, M = 65.4, SD = 11.9) (t(761)= 8.30, p <.001). The Cohen's d value was .60, again indicating a moderate effect size. Once more, Levene's test suggested no violation of the assumption of homoscedasticity (F(1,761), .77, p = .382).

A one-way ANOVA was conducted to assess whether NIES scores differed across gender groups. It indicated a significant difference across the groups of male (N = 154, M = 64.9, SD = 12.3), female (N = 562, M =69.5, SD = 12.3) and nonbinary (N = 45, M = 72.7, SD = 11.5) (F(2,758) = 11.27, p < .001, ηp^2 = .029). Post-hoc Tukey's tests indicated that non-binary participants scored significantly higher than male participants (p < .001), but not female participants (p = .201) and that female participants scored significantly higher than males (p < .001).

Further analysis was conducted to assess the independence of the effects of mental health and gender. This was achieved by conducting an ANCOVA between gender and NIES

total scores whilst controlling for the presence of mental health conditions. The association was attenuated but remained significant ($F(2,758) = 6.81$, $p = .001$, $\eta p^2 = .017$).

Sample Two. An independent samples t-test was conducted to assess the difference in NIES scores between participants who self-reported a physical health condition and those with did not. When adjusting for multiple comparisons, there was no significant difference between those who did report a physical health condition ($N = 121$, $M = 65.91$, $SD = 12.37$) and those who did not ($N = 262$, $M = 63.11$, $SD = 12.04$), ($t(411) = 2.14$, $p = .017$). A non-significant Levene's test indicated no violation of the assumption of homoscedasticity ($F(1,414) = .020$, $p = .888$).

An independent samples t-test indicated a significant difference in the NIES scores of those who self-reported no diagnosed mental health conditions ($N = 258$, $M = 60.41$, $SD = 11.76$) and those who did report a diagnosed mental health condition ($N = 155$, $M = 69.79$, $SD = 10.56$), with those with a mental health condition scoring significantly higher ($t(411) = 8.151$, $p < .001$). The Cohen's d value was .83, indicating a large effect size. Levene's test indicated no violation of the assumption of homoscedasticity ($F(1,411) = 2.03$, $p = .155$).

The same was found when including those who self-reported self-diagnosed mental health conditions. Those with self-reported mental health conditions ($N = 173$, $M = 69.14$, $SD = 11.04$) scored significantly higher than those with no mental health conditions ($N = 240$, $M = 60.17$, $SD = 11.60$) ($t(411) = 7.92$, $p < .001$). The Cohen's d value was .79, indicating a moderate, approaching large, effect size. Once more, Levene's test suggested no violation of the assumption of homoscedasticity ($F(1,411) = .436$, $p = .509$).

A one-way ANOVA was conducted to assess whether NIES scores differed across gender groups. It indicated no significant difference across the groups of male ($N = 97$, $M = 61.84$, $SD = 12.90$), female ($N = 303$, $M = 64.44$, $SD = 11.96$) and non-binary ($N = 9$, $M =$

70.22, *SD* = 11.96) ($F(2,406) = 2.92$, $p = .055$). No post-hoc tests were conducted due to non-significance. Levene's test suggested no violation of the assumption of homoscedasticity ($F(2,406) = 2.601$, $p = .075$).

Questionnaire feedback

See Table 4.6 for the demographic information of the 18 participants randomly selected from Sample One to provide feedback. This sample included four autistic participants, seven with anxiety disorders, one with a somatoform disorder, and two with eating disorders, meaning we were able to get feedback from those who are members of certain groups that negative interoceptive evaluation may be especially important for.

Table 4.6

Demographic information for the feedback sample

Age	Gender	Ethnicity	Presence of a mental health condition	Presence of a physical health condition
26.83 (10.34)	66.7% Female 33.3% Male	77.8% White 22.2% Asian or Asian British	66.7%	27.8%

Overall, the feedback was positive. All the participants appeared to understand the purpose of the questionnaire, with answers including "how people experience bodily sensations and the amount of distress that they cause," "how people feel their internal body actions and how that makes us feel," and "interoception and how this relates to anxiety around bodily sensations." Two of the participants thought the questionnaire was just about anxiety.

Most participants found the questionnaire easy to understand, and no participants reported having difficulty understanding. One participant said, “trying to separate my feelings towards the event or action causing the sensation from my feeling just towards the sensation required a bit of thought/introspection.” This does not indicate a problem with the questionnaire- rather it is exactly as we might expect from a questionnaire assessing interoceptive evaluation. One participant worried they had misunderstood the question about not liking pain (Question 7).

Few participants chose to provide any feedback. Of those that did, some examples of the suggestions were as follows:

- “*Maybe include something about the degree to which we are able to understand our bodily sensations/recognise them as what they are and whether this contributes to the negative evaluation*” This suggestion relates more to interoceptive accuracy, and the use of the NIES alongside the IAS would address this.
- “*Maybe the option for people to explain why certain bodily sensations make them feel nervous? e.g. someone might know they they are susceptible to panic attacks or be aware of a family history of heart attacks etc*” Authors discussed this and thought it was a good suggestion. In the EFA and CFA questionnaires, participants had already indicated if they had physical or mental health conditions. We would suggest when using the NIES in instances where this is not already done, having an open text box for participants to input mental or physical health conditions that might affect their interpretation of bodily signals.

Overall, participant feedback supported the appropriateness of our questionnaire. See the Appendix (Appendix 8) for full feedback.

EFA

Initial checks were performed to establish the appropriateness of a factor analysis for this data. This included examination of the correlation matrix and checking that there were sufficient correlations that were .30 or above, whilst also ensuring there was not too high a level of collinearity or singularity (Watkins, 2018). As the highest correlation was .67, with many correlations being higher than .30, this requirement was met. The KMO measure of sampling adequacy was .823, which according to Kaiser's criteria is meritorious (Kaiser & Rice, 1974), and the Bartlett-test of sphericity indicated that our matrix appeared to be factorable as it differs significantly from an identity matrix ($\chi^2(210) = 5392.023$, $p < .001$). The anti-image correlations were all .64 to .88, meaning that they all met the criteria of being greater than .5. The determinant of the correlation matrix was .001, which is higher than the criterion value of .00001. The data was deemed highly suitable for factor analysis.

There was a high degree of variability in the number of factors suggested by the different models (see Table 4.7). Out of all of the models (two, three, four and five factor), the only one without cross-loading was the three-factor model suggested by the scree plot (see Figure 4.1). Overall, the three-factor model explained 37.2% of variance. See Table 4.8 for a breakdown of the variance that each factor explains, Table 4.9 for the communalities and factor loadings of all items, and Table 4.10 for the factor correlation matrices associated with the Scree plot-based model.

Table 4.7

Number of factors suggested by each method

Method	Number of factors
PCA	5
Scree	3
Parallel Analysis	4
Map Test	2

Figure 4.1

Scree Plot for initial EFA

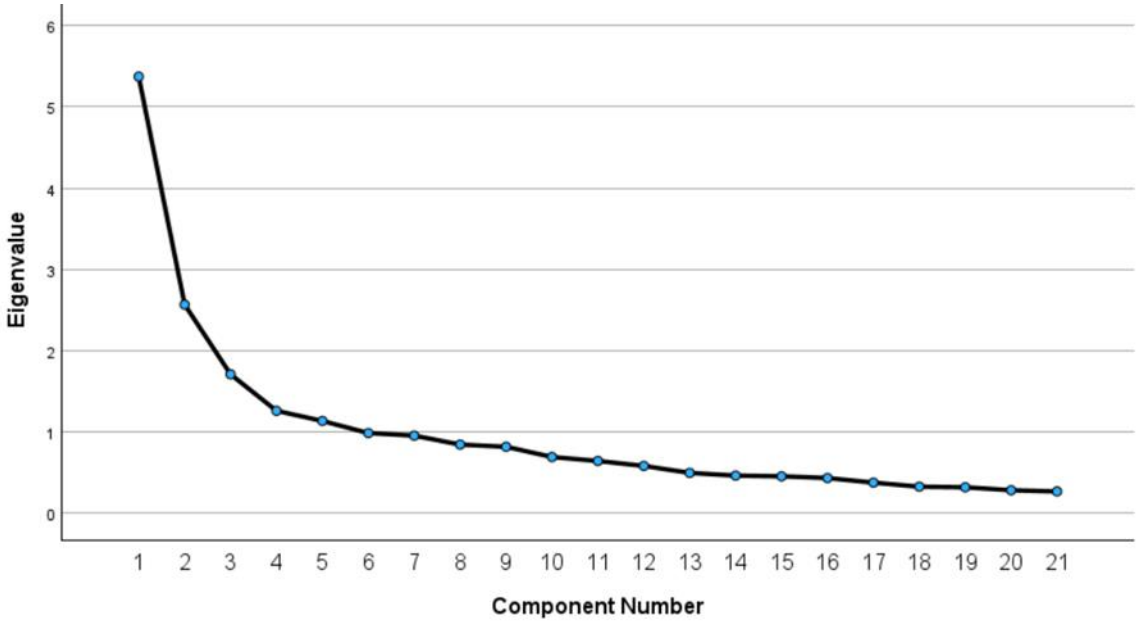


Table 4.8

Variance explained

	Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings
Factor	Total	% of Variance	Cumulative %	Total
1	4.801	22.863	22.863	3.951
2	2.002	9.535	32.398	3.775
3	1.004	4.783	37.180	1.768

Table 4.9

Communalities and factor loadings of all items

	Communalities		Pattern Matrix		
	Initial	Extraction	1	2	3
NIES1	0.424	0.433	-0.109	**0.683**	0.032
NIES2	0.354	0.297	-0.107	**0.541**	0.112
NIES3	0.215	0.166	-0.013	0.091	**0.380**
NIES4	0.287	0.127	**0.354**	0.050	-0.100
NIES5	0.249	0.228	0.084	-0.122	**0.468**
NIES6	0.243	0.219	0.156	-0.086	**0.420**
NIES7	0.187	0.229	-0.243	0.171	**0.444**
NIES8	0.538	0.533	0.083	**0.721**	-0.151
NIES9	0.509	0.491	0.116	**0.675**	-0.138

NIES10	0.514	0.449	**0.604**	0.124	0.033
NIES11	0.572	0.560	**0.756**	0.065	-0.172
NIES12	0.489	0.427	**0.600**	-0.057	0.179
NIES13	0.417	0.330	**0.411**	0.039	0.278
NIES14	0.404	0.372	0.111	**0.456**	0.224
NIES15	0.529	0.557	0.000	**0.751**	-0.022
NIES16	0.487	0.468	0.119	**0.628**	0.012
NIES17	0.467	0.344	**0.539**	0.079	0.040
NIES18	0.538	0.568	**0.806**	-0.067	-0.207
NIES19	0.428	0.389	**0.547**	-0.105	0.238
NIES20	0.407	0.363	**0.552**	-0.030	0.151
NIES21	0.298	0.259	-0.058	**0.378**	0.291

Note. Extraction Method: Principal Axis Factoring

Rotation Method: Promax with Kaiser Normalisation

a. Rotation converged in 5 iterations

The numbers highlighted in bold are above the cut-off value of .30.

Table 4.10

Factor correlation matrix

	Factor		
	1	2	3
1	1.000	0.364	0.289
2	0.364	1.000	0.241
3	0.289	0.241	1.000

The model can be broken down as illustrated in Table 4.11:

Table 4.11

Three factor model suggested by EFA

Factor One	Factor Two	Factor Three
Neutral internal signals making one anxious	Evaluation of anxiety-related internal signals	Not liking neutral internal signals
I don't like it when I feel full.	I don't like it when I can feel my heart beat fast.	I don't like it when I feel hungry.
If I focus on how hungry I am, it makes me feel anxious.	I don't like it when I feel out of breath.	I don't like it when I feel thirsty.
If I focus on how full I am, it makes me feel anxious.	If I focus on my heart beating, it makes me feel anxious.	I don't like it when I feel like I need to use the bathroom.
If I focus on how thirsty I am, it makes me feel anxious.	If I focus on my breathing, it makes me feel anxious.	*I don't like it when I can feel a pain inside of my body (like a stomach-ache).*
If I focus on whether I need the bathroom, it makes me feel anxious.	If I focus on pain inside my body, it makes me feel anxious.	
When I feel hungry, it makes me feel like something is wrong	When I feel my heart beating fast, it makes me feel like something is wrong	
When I feel full, it makes me feel like something is wrong.	When I feel out of breath, it makes me feel like something is wrong.	
When I feel thirsty, it makes me feel like something is wrong.	When I feel a pain inside of my body, it makes me feel like something is wrong.	

When I feel like I need to use the bathroom, it makes me feel like something is wrong.	

Note. The italics reflect items that do not fit the pattern.

There were some instances where the loadings did not perfectly fit the factor labels identified, as illustrated by the questions in italics. One was the question identified as demonstrating ceiling effects (Question 7- "I don't like it when I feel a pain inside of my body"). We would expect this to appear in Factor 2, however, it appeared as part of Factor 3. The other exception to the labels identified above was that Question 3 ("I don't like it when I feel full") loaded onto Factor One rather than Factor Three as expected.

After having observed strong ceiling effects for Question 7 ("I don't like it when I feel a pain inside my body), observing that it did not fit into the expected factor structure, and noted that some participants expressed confusion about the item in their feedback, we considered removing it. We ran the EFA analysis again, removing Question 7. It is important to note that as the context of the questions asked in a questionnaire matters, one would ideally repeat the questionnaire with the intended question removed on a new sample to truly assess the factor structure of the questionnaire without the question. However, this was done as an exploratory measure, before removing the question for the CFA sample.

The data was again deemed suitable for factor analysis. The same number of factors was indicated for each of the four analysis methods (see Table 4.12). Once again, the scree plot indicated a structure of three factors was the most interpretable (see Figure 4.2). Overall, the three-factor model explained 38.2% of variance. See Tables 13 to 15 for a breakdown of the model.

Table 4.12

Number of factors suggested for each method

Method	Number of factors
PCA	5
Scree	3
Parallel Analysis	4
Map Test	2

Figure 4.2

Scree plot for EFA with Question 7 removed

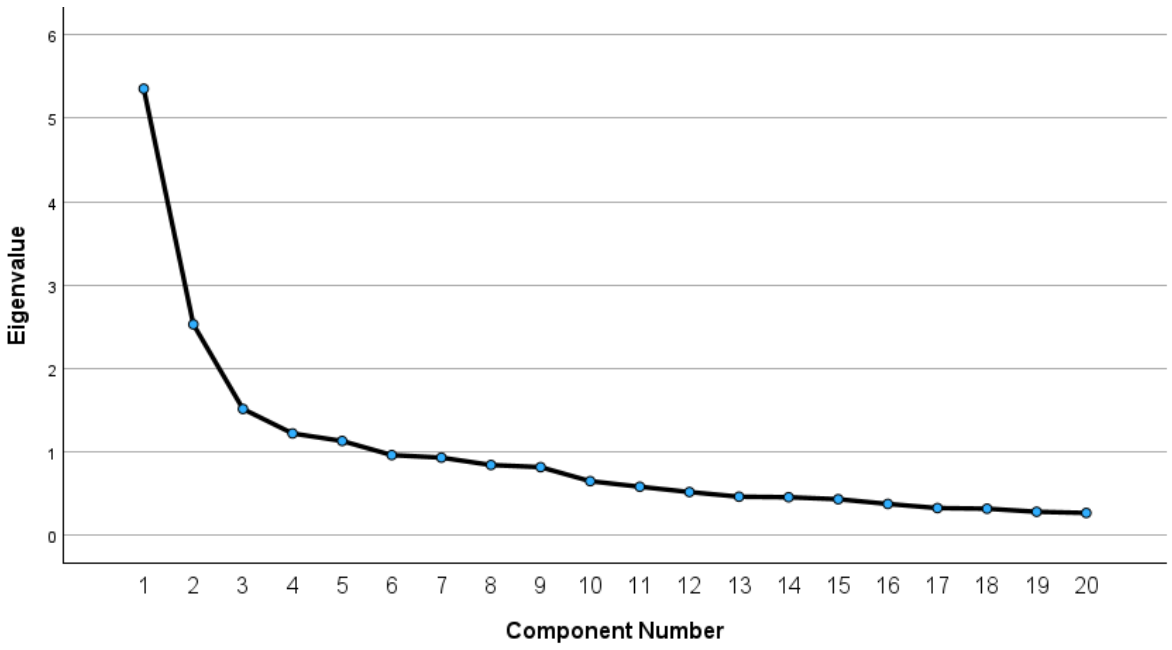


Table 4.13

Variance explained

	Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings
Factor	Total	% of Variance	Cumulative %	Total
1	4.797	23.986	23.986	3.828
2	1.981	9.904	33.890	3.607
3	0.858	4.290	38.181	3.255

Table 4.14

Communalities and factor loadings of all items

	Communalities		Pattern Matrix		
	Initial	Extraction	1	2	3
NIES1	.424	.435	**.699**	-.097	-.036
NIES2	.349	.290	**.562**	-.130	.046
NIES3	.205	.126	.135	-.215	**.381**
NIES4	.287	.193	.018	**.511**	-.169
NIES5	.240	.144	-.067	-.147	**.461**
NIES6	.218	.163	-.041	-.084	**.459**
NIES8	.538	.532	**.710**	.173	-.187
NIES9	.509	.492	**.665**	.194	-.163
NIES10	.514	.441	.121	**.422**	.251

NIES11	.571	.645	.016	**.843**	-.082
NIES12	.489	.431	-.041	.297	**.452**
NIES13	.415	.363	.074	.052	**.538**
NIES14	.394	.359	**.482**	-.042	.257
NIES15	.529	.561	**.764**	-.018	-.025
NIES16	.487	.471	**.641**	.055	.055
NIES17	.467	.337	.081	**.311**	.298
NIES18	.538	.639	-.126	**.878**	-.073
NIES19	.427	.409	-.083	.195	**.532**
NIES20	.404	.379	-.031	.224	**.461**
NIES21	.280	.227	**.409**	-.192	.260

Note. Extraction Method: Principal Axis Factoring

Rotation Method: Promax with Kaiser Normalisation

b. Rotation converged in 5 iterations

The numbers highlighted in bold are above the cutting value of .30.

Table 4.15

Factor correlation matrix

	Factor		
	1	2	3
1	1.000	.380	.349
2	.380	1.000	.594
3	.349	.594	1.000

The removal of Question 7 changed the factor structure slightly (see Table 4.16). Whilst the factor “Evaluation of anxiety-related interoceptive signals” largely remained the

same, the remaining two factors were split into “evaluation of gastrointestinal related signals” and “evaluation of thirst and need to use the bathroom.” The only question that did not fit with this structure was Question 3- “I don’t like it when I feel hungry,” which appeared to load into the third factor. Other than this, these results suggest a fairly clear split into different domains of interoception based on the physiological systems or aspects of homeostasis associated with them. We decided to also test this factor structure with a CFA, to assess whether it is a better fit than the factor structure suggested by our initial EFA.

Table 4.16

Three factor model suggested by EFA (with Question 7 removed)

Factor One	Factor Two	Factor Three
Evaluation of anxiety-related internal signals	Evaluation of gastric related signals	Evaluation of other alimentary signals
I don't like it when I can feel my heart beat fast.	I don't like it when I feel full.	*I don't like it when I feel hungry.*
I don't like it when I feel out of breath.	If I focus on how hungry I am, it makes me feel anxious.	I don't like it when I feel thirsty.
If I focus on my heart beating, it makes me feel anxious.	If I focus on how full I am, it makes me feel anxious.	I don't like it when I feel like I need to use the bathroom.
If I focus on my breathing, it makes me feel anxious.	When I feel hungry, it makes me feel like something is wrong.	If I focus on how thirsty I am, it makes me feel anxious.
If I focus on pain inside my body, it makes me feel anxious.	When I feel full, it makes me feel like something is wrong.	If I focus on whether I need the bathroom, it makes me feel anxious.

When I feel my heart beating fast, it makes me feel like something is wrong	When I feel thirsty, it makes me feel like something is wrong.
When I feel out of breath, it makes me feel like something is wrong.	When I feel like I need to use the bathroom, it makes me feel like something is wrong.
When I feel a pain inside of my body, it makes me feel like something is wrong.	

Note. The italics reflect items that do not fit the factor pattern.

CFA

A CFA was used to test whether the factor structure identified by the EFA was a good fit for the NIES in a new dataset and to test which of the different models fit the CFA data the best.

As Question 7 (“I don’t like it when I feel a pain inside of my body”) was removed from the questionnaire for the sample used for the CFA, all participants answered only a 20-item version. We tested the models suggested by the initial EFA, using the 21-item questionnaire, and the secondary EFA we ran, removing Question 7, to establish which would be a better fit. We also tested models that had been edited slightly to ensure all items were in factors where they “made sense” theoretically.

- Model 1: Using the factor structure suggested by the initial EFA.
- Model 2: Using the factor structure as suggested by the initial EFA, but with the question “I don't like it when I feel full*”* moved from the Neutral Interoceptive

Signals Making One Anxious factor to the Not Liking Interoceptive Signals factor, as one might expect.

- Model 3: As suggested by the second EFA that we ran with the question "I don't like it when I feel a pain inside my body" removed.
- Model 4: As suggested by the second EFA that we ran with the question "I don't like it when I feel a pain inside my body" removed but with the question "I don't like it when I feel hungry" moved from the Evaluation of Thirst and Needing to Use the Bathroom factor to the Gastric factor, as one might expect.

The statistics for each model are presented in Table 4.17.

Table 4.17

CFA model fit statistics

	χ^2	*df*	χ^2/df	CFI	TLI	RMSEA	SRMR
Criteria			<3	>.90 for acceptable, >.95 for good		<.06	<.07
Model 1	1029.45	167	6.15	0.61	0.55	0.16	0.10
Model 2	1071.27	167	6.41	0.60	0.55	0.16	0.10
Model 3	965.85	167	5.78	0.63	0.57	0.16	0.10
Model 4*	963.60	167	5.77	0.63	0.58	0.16	0.10

Note. * indicates our best performing model. The CFI, TLI and RMSEA SRMR are all robust values.

Given that none of our models met the criteria for a good model fit, we chose to modify Model 4 (our best performing model) with the inclusion of correlated residuals between the questions: "if I focus on my heart beating it makes me feel anxious" and "if I

focus on my breathing, it makes me feel anxious." We chose to do this as the modindices function indicated this would make the largest difference of all the item-to-item residual correlations (*mi* = 47.753), and it is also reasonable, as the two items are in the same factor (anxiety-related signals) and both have the same question wording. For the statistics of this updated model (Model 5), see Table 4.18.

Table 4.18

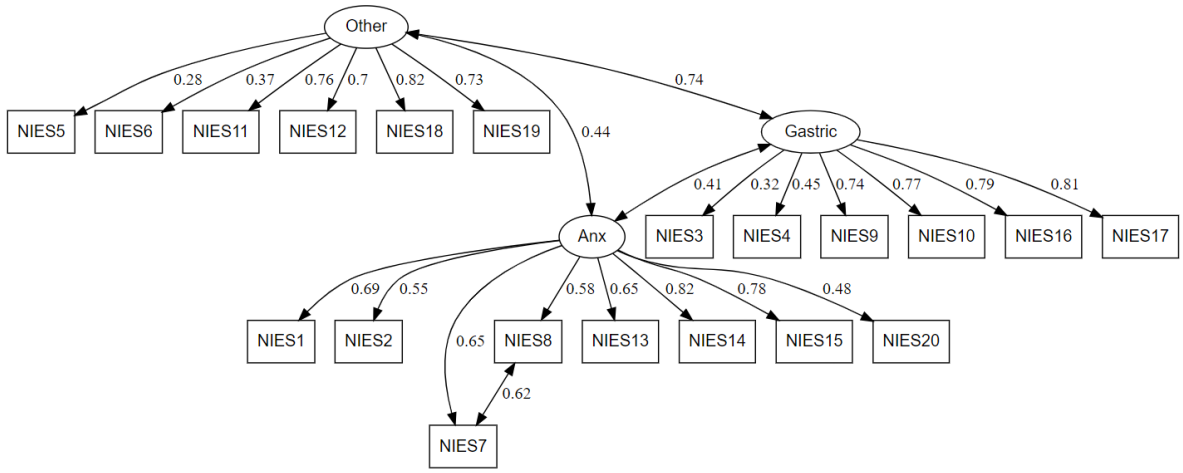
Model fit statistics when including correlated residuals

	χ^2	*df*	χ^2/df	CFI	TLI	RMSEA	SRMR
Criteria			<3	>.90 for acceptable, >.95 for good		<.06	<.07
Model 5	860.79	166	5.15	0.66	0.61	0.15	0.09

The model was slightly improved but still was a poor model fit. A path diagram of Model 5 is presented in Figure 4.3.

Figure 4.3

Path diagram of Model 5



Note. Path diagram of the best fitting model (Model 5) and coefficients. Anx = anxiety-related internal signals. Gastric = gastric related signals. Other = other alimentary signals.

Factor correlations

As an exploratory validity check, we assessed whether the expected relationships were demonstrated between the different factors and the ASI (total and physical subscale). As expected, correlations between the ASI scores and anxiety-related signals scale were higher than for the other two scales (Table 4.19).

Table 4.19

Correlations between factors and the ASI

	Anx	Gastric	Other
ASI Total	.689	.371	.382
ASI Physical	.700	.264	.264

Note. All correlations were Spearman's, due to non-normally distributed data. All correlations were significant at $p < .001$. Anx = anxiety-related internal signals. Gastric = gastric related signals. Other = other alimentary signals.

We wanted to further explore why our CFA did not have a good fit. To begin with, we conducted another EFA on the CFA data. EFA was conducted in the same manner as in the main analysis. All checks showed that the data was factorable and thus appropriate for EFA. For brevity, the full EFA is not reported here. In short, the scree plot indicated a three-factor solution that was similar to that identified in the initial EFA. All of the questions about heartbeat and breathing loaded onto the first factor. All questions about fullness and hunger loaded onto the second factor, except the question "I don't like it when I feel hungry," which loaded onto the third factor. All the questions about thirst and needing to use the bathroom loaded onto the third factor, however, the question "When I feel thirsty, it makes me feel like something is wrong" exhibited co-loading and also loaded onto the second factor. Both items relating to pain (if I focus on pain inside my body, it makes me feel anxious and when I feel a pain inside of my body, it makes me feel like something is wrong) did not load onto any factors. We speculated that these items may be problematic and affect model fit. However, removing the two questions on pain from Model 5 during CFA did not improve the fit of the model. We also ran some further CFA tests on alternative factor structures. This included a one-factor model, with all items included, and a two-factor model, split into anxiety-related

signals and non-anxiety-related signals. Both of these models showed a poorer fit than the three-factor models illustrated above.

The next step was to explore whether differences between the EFA and CFA sample demographics may have led to different response patterns. This would be a possibility if certain groups (e.g.: men versus women) interpreted questions differently and would illustrate a lack of measurement invariance. We used t-tests to compare the two samples on the age of participants and scores on the NIES and used chi-square to test whether gender distributions, the presence of mental and physical health conditions, and ethnicity differed across groups. Due to our large sample size, nearly all of the comparisons (except for the presence of physical health conditions across groups) were significant. Therefore, we considered the effect sizes (Cohen's *d* for the t-tests, and *W* for the chi-square; Cohen, 1988). The differences in gender, mental health and physical health were all below the threshold for a small effect. There was a small difference in age across samples ($d = 0.25$), a small difference in ethnicity ($W = 0.14$), and a small difference in NIES scores ($d = 0.40$). Therefore, it is a possibility that the factor structure of responses may differ according to age, ethnicity, or total NIES scores.

Internal consistency

We first conducted the Cronbach's alpha test on our EFA sample, to test the internal consistency of the questionnaire including all items. Internal consistency was good (.85; Konting et al., 2009). A similar value was obtained from the CFA data (.85).

Further analysis was conducted on each of the three factors of our best-performing models (Models 4 and 5; see Table 4.20).

Table 4.20

Internal consistency of the three factors

Factor	Cronbach's Alpha	Interpretation
1 (anxiety-related signals)	.81	Good
2 (gastric related signals)	.74	Acceptable
3 (other alimentary related signals)	.74	Acceptable

Discussion

Prior to the development of the Negative Interoceptive Evaluation Scale (NIES), there were no questionnaires designed specifically to measure negative interoceptive evaluation, despite theories linking this concept to various mental health conditions. We developed the NIES as a 21-item, self-report measure of negative interoceptive evaluation, later reducing the number of items to 20. This was administered to two large diverse samples of participants.

Our measure appears to have good convergent and divergent validity. The NIES has the expected correlations with other measures that have been used as proxies for interoceptive evaluation, namely the ASI-3 (total score and Physical subscale score) and the MAIA-2 Not Worrying Scale. Given that these measures are limited in the domains of interoception that are assessed, focusing specifically on anxiety-related signals, one would not expect perfect correlations. However, the correlations we did observe were all in the expected direction, and all moderate to large (with the exception of the correlation between the MAIA-2 and NIES in Sample Two, which was small). The correlation with the IAS was small, suggesting the NIES measures a construct that is sufficiently distinct from interoceptive accuracy. The correlation

with the IATS was moderate. Despite the presence of a moderate correlation, it is not necessarily an issue for divergent validity as the collinearity is not concerningly high, especially as there is a logical basis to expect a correlation between a greater degree of interoceptive attention and more negative evaluation. If one interprets interoceptive sensations as a sign of danger, it makes sense that one would attend to them to a greater degree, as humans are hypervigilant to signals that indicate a threat (Berdica et al., 2018).

Only a small correlation was observed with age, illustrating a reassuring degree of measurement invariance. However, our sample was only comprised of adults. Some have speculated that puberty may be an important time for the development of interoception (Murphy et al., 2017), so further research could perhaps explore the relationship between age and this scale in a sample of adolescents.

The correlations from the first sample were replicated in an independent second sample. With the exception of the MAIA-2 Not Worrying Subscale, which had a moderate negative correlation in the first sample and a small negative correlation in the second, and age, which was not significant in the second sample, all the correlation sizes were the same, and the r values were extremely similar, illustrating the replicability of the findings.

Of all the group comparisons we conducted, we had the most reason to assume that interoceptive evaluation would differ between those with and without a mental health condition. This predicted group difference was found. Meanwhile, interoceptive evaluation did not differ between those with and without physical health conditions. This may be surprising, as one might expect that those with physical health conditions may be more likely to relate a somatic sensation to their illness and therefore perceive it as a sign that something is wrong. However, a possible explanation is that those with physical health conditions only negatively interpret sensations that directly relate to their condition, not generalising to other sensations. It could also be that they are very familiar with the sensations that indicate their

condition, so do not misattribute other sensations. Moreover, it might be that a relationship between negative interoceptive evaluation and physical health conditions does exist with specific health conditions (perhaps those that affect a broader range of bodily systems), but this was obscured in our analysis given the wide range of conditions that were grouped together. Further research would benefit from a more in-depth exploration of this.

Although there have been gender differences observed in interoceptive accuracy (with women showing poorer interoceptive accuracy than men, Prentice & Murphy, 2022), and interoceptive "awareness" (defined by authors as noticing interoceptive sensations; Grabauskaitė et al., 2017), there is not enough work to drive hypotheses regarding gender differences in interoceptive evaluation. In our first sample, we found that gender was related to negative interoceptive evaluation, with more negative evaluation in non-binary and female participants compared to male participants. This existed even when controlling for mental health. The group difference was not observed in the CFA sample, which could in part be due to the smaller sample size.

Test-retest reliability was good, but not excellent. However, this does not necessarily reflect an issue with the questionnaire. Although one might assume interoceptive evaluation is a trait, there is preliminary evidence to suggest state-like elements to it. For example, initial qualitative results (see Chapter Six) suggest that within autistic adolescents specifically, there is a bidirectional relationship between anxiety and interoceptive evaluation, meaning that when one is anxious, one's evaluation becomes more negative. Therefore, this lower degree of test-retest reliability may reflect genuine state-like fluctuations in interoceptive evaluation.

The EFA suggested a three-factor structure for the questionnaire, with the identified factors defined as: "neutral internal signals making one anxious," "evaluation of anxiety-related internal signals," and "not liking neutral internal sensation". The distinction between these factors was mostly clean (with only two questions not fitting into the pattern), and the

existence of these factors makes intuitive sense. However, there was a lack of consistency across EFA methods as to exactly how many factors could be identified in the data, with the four methods employed suggesting between two and five factors. We chose the three-factor model as it had no co-loading and therefore was deemed the most interpretable. We decided to remove the question about not liking pain coming from inside of the body, due to ceiling effects that may have affected the model, and participant feedback that indicated confusion about this question. We repeated an EFA without this question included. The same variability in the suggested number of factors was observed across methods, with the three-factor model again being the most interpretable. However, with the removal of the question, the suggested three-factor structure changed slightly. This time, factors represented questions asking about the evaluation of anxiety-related signals, gastric signals, and other alimentary signals (thirst and needing the bathroom). This factor structure equally makes logical sense, and the loadings were clean (with only one question not fitting the expected pattern).

During the CFA, we compared both factor structures described above, as well as cleaned versions of both, where one question for each was moved so it better fit the factor labels. There was not much difference in the fit between models and each showed a poor fit, however, the model that had the best fit was the model where the factors were: evaluation of anxiety-related internal signals (heartbeat, breathing and pain), evaluation of gastric signals (hunger and fullness), and evaluation of other alimentary signals (thirst and needing to use the bathroom). None of our models met the criteria for a good model fit. This was the case even when we included residual covariance between two of the items in our model. This does raise some concern. If the questions are consistently interpreted by participants, one would expect to see a replicable factor structure between the EFA and the CFA.

The poor fit of our model could be because our samples exhibited small differences in NIES scores, age and ethnicity. These might have influenced how participants answered the

items. It is important to explore this further to ascertain measurement invariance. Another likely reason for this is that one might logically expect different levels of factors to emerge from the NIES. For example, questions could be split according to whether they pertain to not liking interoceptive signals, interoceptive signals making one feel anxious, or interoceptive signals making one feel as though something is wrong. Items could also be split according to the dimension of interoception that they address. The fact that there are various, logical factor structures may explain the range of factors suggested by each of the different EFA models, and why the factor structures specified in the CFA do not fit the data well. Based on this, our factor analysis results do not necessarily invalidate our questionnaire. Given that the total internal consistency and item-total correlations of the questionnaire are good, it does still support the notion that the questionnaire is measuring a clear underlying construct or latent variable, however, the underlying factor structure of the NIES remains unclear. Another point to consider is that the cut-offs that would indicate a good model fit are arbitrary, and recent work has illustrated that they are often not appropriate (Groskurth et al., 2024). Consequently, the lack of a good model fit was not seen as a reason to discard the NIES.

Despite a lack of clarity in the factor structure, there was one consistent pattern across all our models: the clear distinction between anxiety-related and non-anxiety-related signals. By "anxiety-related" we mean signals that one would expect to trigger anxious responses (increased heart rate, faster breathing and pain responses). Other work has suggested that this distinction may be relevant to interoception (Vlemincx et al., 2023). The replicability of this factor suggests that an alternative to using the entire NIES may be limiting questions to only those about anxiety-related signals (cardiac, respiratory and internal pain). The issue with this, however, is that the measure becomes very similar to the ASI and is not reflective of interoception as a whole. As the NIES was intended to advance on the ASI as a measure of negative interoceptive evaluation, this seems somewhat pointless.

Within each factor, there was acceptable to good internal consistency, and the factors varied in their correlations with expected constructs. Specifically, the factor pertaining to anxiety-related signals correlated more with the ASI-3 (total and physical subscale scores) than the other two factors.

In summary, the NIES has good internal consistency, strong convergent and divergent validity and good test-retest reliability. However, the factor structure is not clear cut and further research on this is warranted. This could arise from the fact that there are multiple different ways the NIES can be factorised.

Limitations and future directions

An important limitation to address pertains to attempts to validate questionnaire measures in general. With questionnaire measures, especially those addressing constructs which have no current widely accepted measures, it is impossible to empirically "validate" the measure by comparing it to a "gold-standard" measure. As we currently do not have a validated measure of pure interoceptive evaluation upon which to compare the NIES, the closest we can get is by comparing results to other partially related measures and across groups where we might expect certain relationships to exist. However, a lack of the predicted effects would not necessarily undermine the validity of the novel measure. It may be that the existing measures are not valid, or that the theorised relationships between groups do not exist. The same can be said for both internal consistency and test-retest reliability, which are only relevant if one expects a construct to be homogenous and stable, respectively. As we do not yet know much about interoceptive evaluation, it is difficult to make predictions about what would be expected. Similarly, although a factor analysis may be useful in confirming whether certain expected factors exist within our data, and one can explore whether these factors "make sense", factor analysis should not be seen as validation. Bearing this in mind, it is important to take a holistic approach to assessing the validity of these questionnaires.

Relatedly, further work is required to increase our confidence in the appropriateness of the NIES. One way to do this would be to correlate NIES scores with changes in participants' state anxiety before and after tasks that require them to focus on interoceptive sensations. One would expect higher changes in state anxiety when one is required to focus on interoceptive sensations in those with higher NIES scores.

It would also be useful to compare results across specific clinical groups, to establish whether results differ. This study was limited to physical and mental health conditions being defined very broadly, which precluded conclusions being reached about exactly which conditions may relate to more negative interoceptive evaluation. It would be especially useful to look at group comparisons between those with and without specific conditions that have been linked to interoceptive evaluation in the literature, such as anxiety disorders, eating disorders, depression and somatic symptom disorder. Related to this issue, due to the self-reporting of physical and mental health conditions, it may be that some participants self-reported a "condition" that others did not self-report, as they did not consider it as such. An example here is autism. This may lead to some inconsistency in reporting that may have affected results. Another limitation is that we have not currently assessed the appropriateness of the NIES in groups with varying reading comprehension levels.

Conclusion

In conclusion, the NIES appears to be a promising measure of interoceptive evaluation, and one which appears to relate to mental health. This addresses a pressing need in the field. Further validation of this measure is required, especially within specific relevant patient groups (for example, those with anxiety or eating disorders). Moreover, further validation of the factor structure is required, as the structure identified here is not a good fit.

Chapter Five: The association between anxiety and interoception

Content adapted from: **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 and biobehavioral reviews, 140,* 104754. https://doi.org/10.1016/j.neubiorev.2022.104754

Kiera Adams took over this project, which initially began before the commencement of her DPhil. Jennifer Murphy designed the project and pre-registered it, and Kiera Adams adapted the design with the inclusion of new moderators and completed the PRISMA checklist. Charlotte Peart and Alexandra Edwards conducted the abstract and title review, light full-text screening, and initial content extraction. Kiera Adams checked and corrected this. Kiera Adams completed the final data extraction, which was checked by Jennifer Murphy and Geoffrey Bird. Ines Mendes and Lynn Ellett rated the properties of the anxiety measures used across studies. Kiera Adams designed analysis plans, developed the analysis script, conducted the analysis and interpreted results. Kiera Adams wrote the chapter and received feedback from all co-authors.

Abstract

Anxiety is often conceptualised as the prototypical disorder of interoception. Whilst theoretical models predict an association between interoceptive accuracy and anxiety, empirical work has produced mixed results. This chapter presents a pre-registered systematic review (https://osf.io/2h5xz) and meta-analysis of 55 studies, obtained via a Pubmed search on 9th November 2020, examining the relationship between state and trait anxiety and objectively measured cardiac interoceptive accuracy as assessed by heartbeat counting and discrimination tasks. Potential moderators of this relationship - the age, gender and clinical diagnoses of participants, the anxiety measures used, and the study design - were also explored. Overall, we found no evidence for an association between cardiac interoceptive accuracy and anxiety, with none of the factors examined moderating this finding. We discuss the implications these findings have for future research, with a particular focus on the need for further investigation of the relationship between anxiety and other facets of interoception.

Introduction

As discussed in Chapter One, there is significant theoretical work linking anxiety to interoception. Ehlers' account (1993) first suggested that in panic disorder (PD) specifically, increased sensitivity to (e.g., a higher accuracy of perceiving) cardiac signals may lead to an increased potential for panic attacks (Ehlers, 1993; Ehlers & Breuer, 1992). More recently, Garfinkel and colleagues suggested that anxiety may be characterised by poor interoceptive accuracy, coupled with increased attention to interoceptive signals (Garfinkel, Tiley, et al., 2016). Although these theories differ in the exact nature of the predicted association, both suggest a relationship between interoceptive accuracy and anxiety.

Despite good theoretical reasons to expect that anxiety should relate to atypical interoceptive accuracy, empirical evidence is equivocal (see Chapter One). Given such mixed results and the considerable disease burden of anxiety disorders (Baxter et al., 2013; Mojtabai et al., 2015; Scott et al., 2016), it is important to establish whether a relationship exists between anxiety and cardiac interoceptive accuracy, the direction of this association, and which, if any, specific moderators may explain the mixed results in the literature. Such understanding is necessary to establish the clinical relevance of interoceptive accuracy for anxiety disorders, which in turn may provide new opportunities for intervention.

Previous studies have identified several factors that appear to influence the direction of the association between anxiety and cardiac interoceptive accuracy. One of these factors includes the specificities of the tasks used to measure cardiac interoceptive accuracy. The most used task to measure cardiac interoceptive accuracy is the heartbeat counting task (HCT), where participants are required to count their heartbeats over a series of intervals which are then compared to an objective estimate to determine accuracy (Schandry, 1981). The delivery of this task can vary according to whether participants are explicitly told not to

guess the number of beats. When participants are explicitly told not to guess, individuals with panic or anxiety disorders appear to perform similarly to controls, with superior performance in panic or anxiety patients only observed when more liberal instructions are utilized (Ehlers et al., 1995; Smith, Feinstein, et al., 2021). This may suggest that reports of superior performance in those with anxiety are driven by factors other than interoceptive accuracy, such as a greater knowledge of their typical heart rate (Ring & Brener, 1996) or a bias towards over-reporting. This illustrates the main limitation of the HCT: the potential for false positives (see Chapter Two). It is important to note that even if participants are given instructions not to guess, there is no way of ensuring that they will not.

Concerns regarding the validity of the HCT have prompted an increased focus on the heartbeat detection task (HDT; Brener et al., 1993; Clemens, 1984; Whitehead et al., 1977; Yates et al., 1985; see Chapter Two). Although both the HCT and HDT are measures of cardiac interoceptive accuracy, a previous meta-analysis suggests only a small relationship between these two tasks (4.4% of the shared variance; Hickman et al., 2020). Unsurprisingly given this small association, the measures also often show different relationships with outcome variables; for example, Garfinkel and colleagues found that whilst trait anxiety was negatively correlated with performance on the HDT, this relationship was not observed with the HCT (Garfinkel, Tiley, et al., 2016). Conversely, Ewing and colleagues found that those with a clinical diagnosis of anxiety were significantly poorer than controls on the HCT, though no difference was found with the HDT (Ewing, Manassei, van Praag, et al., 2017). Given such discrepancies, coupled with evidence for lack of correspondence between the HCT and HDT, it remains an outstanding question whether the relationship between anxiety and interoception varies as a function of the cardiac interoceptive accuracy task employed, as previous meta-analyses have only focused on the HCT (Desmedt, Van Den Houte, et al., 2022).

The inclusion of clinical groups and the measurement of anxiety may also moderate relationships between anxiety and cardiac interoceptive accuracy. Indeed, some have argued that a relationship between anxiety and cardiac interoceptive accuracy is only observed in clinical samples that demonstrate the most extreme manifestations of anxiety symptoms (Domschke et al., 2010; but see Desmedt et al., 2022 for conflicting evidence using the HCT). Others have highlighted the possible influence of depression on the relationship between anxiety and interoceptive accuracy (Desmedt, Van Den Houte, et al., 2022). Despite strong associations between anxiety and depression (Dobson, 1985; Lamers et al., 2011), depression and anxiety are often reported to have different relationships with cardiac interoceptive accuracy- depression is thought to be associated with worse interoceptive accuracy and anxiety is usually thought to be associated with better interoceptive accuracy (Dunn et al., 2010; Pollatos et al., 2009; but see Garfinkel, Tiley et al., 2016). The influence of depression on the relationship between anxiety and cardiac interoceptive accuracy may be particularly important to consider where anxiety is assessed by self-report measures as evidence suggests such measures may be non-specific and pick up on negative affect and symptoms of depression (Clark & Watson, 1991; Lovibond & Lovibond, 1995). However, Desmedt and colleagues did not find that the anxiety measure used moderated the relationship between anxiety and performance in the HCT (Desmedt, Van Den Houte, et al., 2022).

The demographics of participants may also alter the strength of the association between cardiac interoceptive accuracy and anxiety. Both interoceptive accuracy and anxiety change across development, with some evidence that interoceptive ability decreases in adolescence (Murphy et al., 2017) around the time that anxiety disorders typically have their onset (Kessler et al., 2007). As such, one might expect to see a stronger relationship in adults compared to children, after the point where anxiety disorders typically emerge. Additionally,

considering the evidence for sex differences in interoception, with males showing superior accuracy on cardiac tasks (Prentice & Murphy, 2022), the gender ratio of participants could also be a moderator. As recent evidence suggests that interoceptive accuracy relates to certain emotion regulation strategies in men only (Lischke et al., 2020), it may be that the association between cardiac interoceptive accuracy and anxiety is greater in males. Finally, there is a suggestion that any association between interoception and psychopathology may manifest more prominently during states of cardiac perturbation, meaning that a greater relationship may be observed if participants' heart rates are manipulated to be either higher or lower than baseline (Paulus et al., 2019).

Given mixed results in the literature regarding the relationship between anxiety and cardiac interoceptive accuracy, the aim of this pre-registered systematic review and meta-analysis was to confirm the presence or absence of a relationship between anxiety and cardiac interoceptive accuracy and assess the extent to which previously identified factors moderate such relationships. In terms of moderators, we focused on assessing the impact of the interoceptive accuracy task employed, the instructions used (for the HCT), the measurement of anxiety, whether clinical groups were included, depression (where anxiety was assessed using questionnaires), the age and gender of participants, and whether heart rate (HR) was manipulated on the pooled effect size.

Method

Search strategy

The systematic literature search was conducted in accordance with the 2020 Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) guidelines (Page et al., 2021). The PRISMA checklists and extracted data can be found at https://osf.io/gvk97/.

The search strategy was pre-registered prior to the search and can be found at https://osf.io/2h5xz. Cardiac interoceptive accuracy was focused on given its primacy in the literature and because discrepant results concerning interoceptive accuracy are commonly reported in this domain. We searched PubMed for studies that were available online before 9th November 2020 using the following search terms:

((Interoception OR visceroception OR interoceptive OR ((awareness OR perception OR discrimination OR Detection OR tracking OR counting) AND Heartbeat OR Heart beat OR Cardiac)) AND Anxiety).

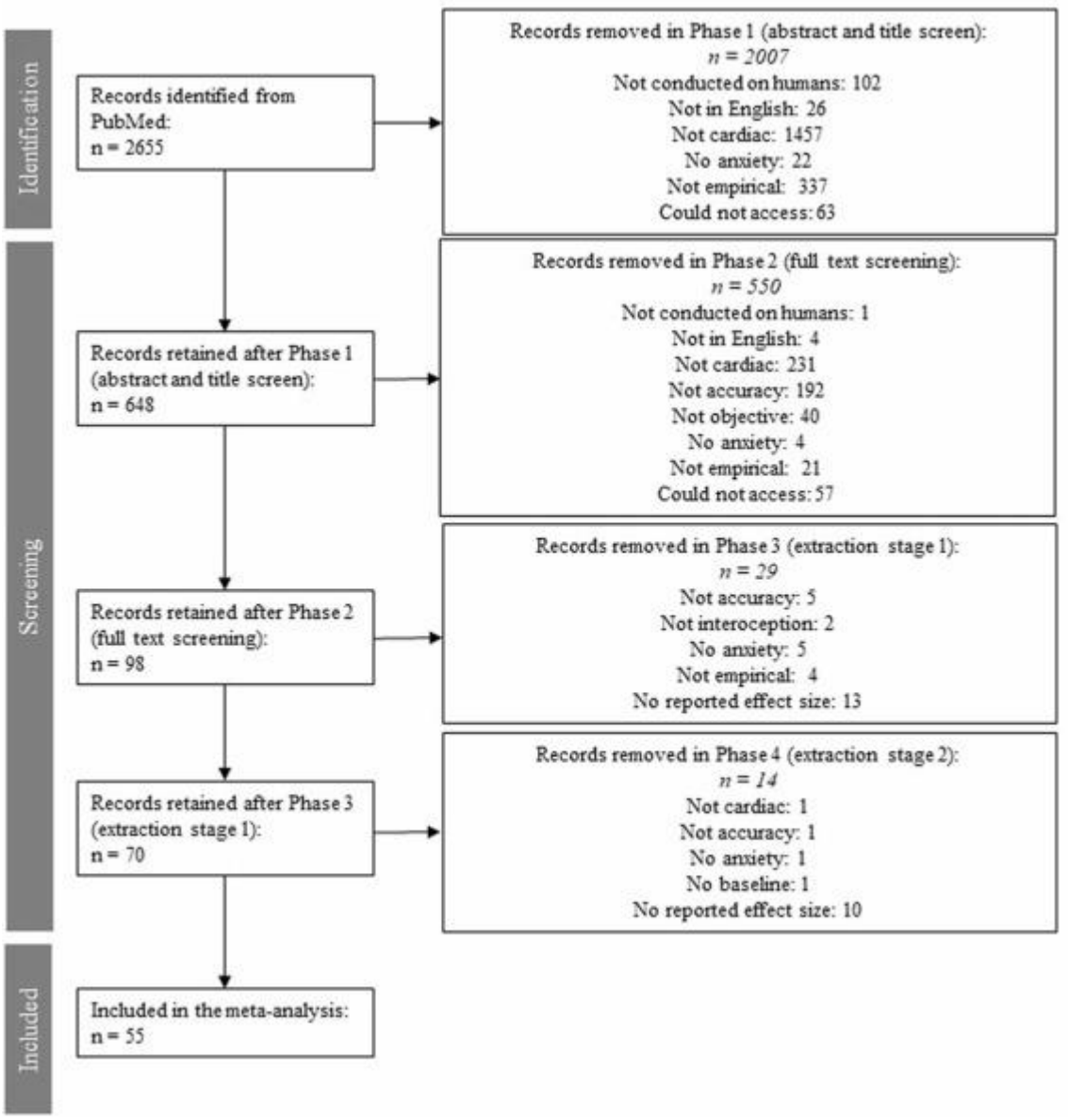
This search returned 2655 studies.

Study selection and extraction

Identified articles were screened in four phases by four reviewers (CP, AE, KA and JM). First, two reviewers (CP and AE) assessed the titles and abstracts for relevance for the meta-analysis. Interrater reliability was acceptable (80.3% agreement) and all studies where disagreement was observed were carried forward to full text screening. This initial screening resulted in the removal of 2007 articles. Studies were removed if they were not conducted on humans, were not in English, did not assess cardiac interoception or anxiety, did not present empirical data about the relationship between anxiety and cardiac interoception or the full text could not be retrieved (Figure 5.1). The remaining 648 articles were submitted for full-text screening.

Figure 5.1

Flow diagram of studies included in the meta-analysis



Note. This figure depicts the number of papers excluded at each phase and the reasons why. For papers where multiple reasons were given for exclusion, just one reason is presented here for illustrative purposes.

Second, light full-text screening was completed by CP and AE. Again, interrater reliability was acceptable (70.68%). Disagreement was resolved by a third reviewer (JM), resulting in the removal of 550 articles. Reasons for removal were consistent with the abstract

screening phase. Additionally, studies were removed if they did not utilise objective measures of interoception or if the measures did not assess cardiac interoceptive accuracy. Third, detailed data extraction was completed by three reviewers (CP, AE and KA) and checked for accuracy by KA.

Extracted data included details of the type and version (where applicable) of the cardiac interoceptive accuracy task employed, the type of anxiety measure used, whether strict 'no guessing' instructions were used for the HCT, whether participants' heart rates were manipulated, whether studies included participants with a clinical diagnosis of anxiety, depression, or any clinical group, the age of participants and the proportion of male and female participants. In addition to the number of participants, where correlations were reported the *r* value was taken as the effect size of interest. For group designs (or studies where proportions of good vs. poor perceivers were reported), *F*, *T* or chi-square test statistics were extracted or manually calculated from extracted means and standard deviations or cross-tabulations. A further 43 studies were removed during the extraction stages for reasons consistent with the full-text screening phase, in addition to articles that either did not report effect sizes for the simple relationship between anxiety and cardiac interoceptive accuracy or did not report sufficient data for this to be calculated, did not include assessment of anxiety directly (e.g., focused solely on disorders, such as Obsessive Compulsive Disorder, which are no longer included as an anxiety disorder in the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (American Psychiatric Association, 2013) or cortisol measurements), or did not include a baseline control condition (e.g., only scores following cardiac interoceptive training were reported). Following a discussion with GB, a further paper was removed because it involved intensity estimations of cardiac and breathing sensations simultaneously, and thus did not focus solely on cardiac interoceptive accuracy (Khalsa et al., 2015). This resulted in 54 articles that were included in the meta-analysis. We opted not to

contact authors for unreported data given previous experiences of poor response rates for previous meta-analyses (Prentice & Murphy, 2022). Where correlations were calculated across different participant groups (e.g., case and controls), or where both the HCT and HDT were employed, multiple effect sizes were extracted from papers.

Of the studies included, 49 utilised the heartbeat counting task (or a similar variant) and 6 utilised the heartbeat detection task (or a similar variant). 33 reported correlations between questionnaire measures of anxiety and cardiac interoceptive accuracy, and 22 utilised a case-control design and compared means across groups. In these 22 studies, nine were grouped by clinical diagnosis and 13 by other methods (e.g., questionnaire cut-offs, or the occurrence of a panic attack after a caffeine challenge). 26 studies included at least one clinical group, whilst 20 studies' samples were comprised solely of typically developing individuals, and sufficient detail was not reported in eight. Most studies focused on adult populations, with only four studies reporting data collected in children.

Measures used: interoceptive accuracy

HCT and variants. In the HCT (Schandry, 1981), participants are instructed to count the number of heartbeats they believe to have occurred over a series of intervals while their heartbeat is objectively recorded. The durations used vary across trials, and participants are not told how long each interval will last. Most commonly, either an accuracy or error score is calculated from the ratio of correct responses. Four studies that used the heartbeat tapping task (where participants tap or press a button as they feel each heartbeat; De Pascalis et al., 1984) were also included in the HCT category, as both tasks involve similar processes (Brener & Ring, 2016) and three of the four studies using the heartbeat tapping task calculated accuracy using the same equation as the majority of studies using the HCT (e.g., Schandry, 1981). Two further tasks were also classed as variants of the HCT, as they both

require a determination of the number of heartbeats in a given period. Steptoe and Vögele (1992) asked participants to rate the degree to which they felt their heart racing on a 10 cm scale, following various tasks designed to increase heart rate. Similarly, Näring and Van Der Staak (1995) instructed participants to rate the degree to which they felt their "pounding heart" and "fast pulse" on a five-point scale following tasks designed to increase heart rate. Accuracy was determined by correlating these ratings with actual heart rate.

HDT and variants. In the 2AFC HDT (used by all studies included in this meta-analysis; Whitehead, 1977), participants are required to judge whether an external stimulus, such as a series of tones or flashes of light, is presented synchronously or asynchronously with their heartbeat. The delays chosen as synchronous and asynchronous vary across studies, with synchronous defined as between 0 to 300ms after the r wave, and asynchronous ranging from 384 to 600ms. As with the HCT, the scoring of these tasks varied, and included error scores, accuracy scores, and d'.

Measures used: anxiety questionnaires

17 different anxiety questionnaires were used to assess anxiety across 48 studies (Table 5.1, Table 5.2, Table 5.3).

Assessing the impact of depression on anxiety scores. As pre-registered, we sought to examine the extent to which depression may influence the relationship between anxiety and interoception. Specifically, we sought to examine whether the extent to which non-specific items included in anxiety questionnaires, which may tap symptoms of depression (Lovibond & Lovibond, 1995), may modulate relationships. Scoring of questionnaires was conducted by two trained clinical psychologists (IM and LE) who rated each item from each questionnaire according to the extent to which items were 1) specifically tapping anxiety symptoms 2) may pick up on symptoms of depression and 3) may pick up on symptoms of

other conditions (e.g., were unlikely to be anxiety or depression). Interrater agreement was acceptable overall (74.78%), with discrepancies resolved through discussion. Scores were then summed for each questionnaire (omitting the few items where an 'other' response was given) and a score was provided reflecting the extent to which the majority of items assessed anxiety specifically or may be influenced by depression. Only the STAI-T was identified as having more than 50% of items that may pick up on symptoms of depression.

Table 5.1

Anxiety questionnaires

Measure	What does it assess	Number of items	Test-retest reliability (*r*)	Internal consistency (*r*)	Influenced by depression?
State-Trait Anxiety Inventory- State (STAI-S) (Spielberger et al., 1983)	State anxiety	20	0.70 (Barnes et al., 2002)	0.91 (Barnes et al., 2002)	No
State-Trait Anxiety Inventory- Trait (STAI-T) (Spielberger et al., 1983)	Trait anxiety	20	0.88 (Barnes et al., 2002)	0.89 (Barnes et al., 2002)	Yes
KSP-ANX			Not found	Not found	Not found
Self-consciousness scale- social anxiety subscale (Fenigstein et al., 1975)	Social anxiety (trait)	22	0.73 (Fenigstein et al., 1975)	Not found	No
Anxiety Sensitivity Index (ASI) (Reiss et al., 1986)	Anxiety sensitivity (trait)	16	0.72 (Rodriguez et al., 2004)	0.88 (Carter et al., 2009)	No

Measure	What does it assess	Number of items	Test-retest reliability (*r*)	Internal consistency (*r*)	Influenced by depression?
PROMIS Anxiety scale short form 4a (Pilkonis et al., 2011)	Anxiety (trait)	4	Not found	0.89 (Kroenke et al., 2014)	No
Anxiety Sensitivity Index 3 (ASI3) (S. Taylor et al., 2007)	Anxiety sensitivity (trait)	18	0.76 (Ghisi et al., 2016)	0.93 (Wheaton et al., 2012)	No
Mood and Anxiety Symptom Questionnaire Short Form-Anxious Arousal (D. Watson & Clark, 1991)	Anxiety (trait)	17	Not found	0.77 (Lackner & Fresco, 2016)	No
Social Anxiety Scale for Children-Revised (SASC-R) (La Greca & Stone, 1993)	Social anxiety (trait)	24	0.42 (Fear of Negative Evaluation subscale); 0.36 (Social Avoidance and Distress (New) subscale); 0.36 (Social Avoidance and Distress (General) subscale) (Storch et al., 2003)	0.86 (Fear of Negative Evaluation subscale); 0.78 (Social Avoidance and Distress (New) subscale); 0.69 (Social Avoidance and Distress (General) subscale) (La Greca & Stone, 1993)	No
Revised Children's Anxiety and Depression Scale (RCADS) – Total Anxiety Subscale (Chorpita et al., 2005)	Anxiety (trait)	37	0.75 (Separation Anxiety Disorder subscale); 0.80 (Social Phobia subscale); 0.65 (Obsessive Compulsive Disorder subscale); 0.76 (Panic Disorder subscale); 0.79 (Generalised Anxiety Disorder subscale) (Chorpita et al., 2000)	0.78 (Separation Anxiety Disorder subscale); 0.87 (Social Phobia subscale); 0.82 (Obsessive Compulsive Disorder subscale); 0.88 (Panic Disorder subscale); 0.84 (Generalised Anxiety Disorder subscale) (Chorpita et al., 2005)	No

Measure	What does it assess	Number of items	Test-retest reliability (*r*)	Internal consistency (*r*)	Influenced by depression?
Screen for Child Anxiety Related Disorders (SCARED)- panic/ somatic subscale (Birmaher et al., 1997)	Panic (trait)	13	.61 (Behrens et al., 2019)	.84 (Hale et al., 2011)	Low agreement (scored for SCARED overall)
SCARED- general subscale (Birmaher et al., 1997)	General anxiety (trait)	9	.62 (Behrens et al., 2019)	.81 (Hale et al., 2011)	Low agreement (scored for SCARED overall)
SCARED- separation subscale (Birmaher et al., 1997)	Separation anxiety (trait)	8	.59 (Behrens et al., 2019)	.72 (Hale et al., 2011)	Low agreement (scored for SCARED overall)
SCARED- social phobia (Birmaher et al., 1997)	Social anxiety (trait)	7	.60 (Behrens et al., 2019)	.78 (Hale et al., 2011)	Low agreement (scored for SCARED overall)
SCARED- school phobia (Birmaher et al., 1997)	School avoidance (trait)	4	.60 (Behrens et al., 2019)	.62 (Hale et al., 2011)	Low agreement (scored for SCARED overall)
Child Anxiety Sensitivity Index (CASI) (Silverman et al., 1991)	Anxiety sensitivity (trait)	18	0.76 (Silverman et al., 1991)	0.87 (Silverman et al., 1991)	No
Fear of Negative Evaluation Scale (FNE) – German (Vormbrock & Neuser, 1983; Watson & Friend, 1969)	Social anxiety (trait)	20	0.90 (Reichenberger et al., 2016)	0.94 (Reichenberger et al., 2016)	Not found
Beck Anxiety Inventory (BAI)	Anxiety (trait)	21	0.67 (Fydrich et al., 1992)	0.94 (Fydrich et al., 1992)	No

Measure	What does it assess	Number of items	Test-retest reliability (*r*)	Internal consistency (*r*)	Influenced by depression?
(A. T. Beck et al., 1988)					
The Short Health Anxiety Inventory (SHAI) (Warwick & Salkovskis, 1989)	Health anxiety (trait)	18	0.76 (Salkovskis et al., 2002)	0.95 (Salkovskis et al., 2002)	No
Sheehan Patient-Rated Anxiety Scale (SPRAS) Not Found	Anxiety (trait)	35	0.91 (Mimura et al., 2011)	0.96 (Mimura et al., 2011)	No
IPAT- anxiety (Cattell, 1969)	Anxiety (trait)	40	0.94 (Levitt & Persky, 1962)	0.78 (men); 0.84 (women) (Bendig, 1966)	Not found
Depression and Anxiety Stress Scale – anxiety (DASS-A) (Lovibond & Lovibond, 1995)	Anxiety (trait)	14	0.79 (Brown et al., 1997)	0.89 (Brown et al., 1997)	No
Whitely Index (WI) (Pilowsky, 1967)	Health anxiety (trait)	14	0.90 (Medical outpatients) (Speckens et al., 1996)	0.80 (medical outpatients); 0.78 (general practice); 0.76 (general population) (Speckens et al., 1995)	No
Whitely Index (WI) Two-factor version (Schwarz et al., 2007)	Health anxiety (trait)	10	Not found	0.63–0.72 (Schwarz et al., 2007)	No

Note. In some studies, the total, or an average of, the STAI-S and STAI-T was taken. We specify when this is the case. For the "influenced by depression" column, two clinicians independently rated the items from these scales according to whether or not they related to

anxiety specifically, may pick up on symptoms of depression, or appeared to tap other symptoms entirely. Due to the low number of items scored as “other” (one for the SASC-R, one for the SHAI, four for the RCADS, five for the WI, five for the WI 2 factor version, three for the SPRAS and three for the SCARED), these items were ignored when calculating the extent to which scales were influenced by depression.

Table 5.2

HCT studies

Author	*N*	Age	F: M ratio	Anxiety patients	Depression patients	Clinical group	Interoception task	No guessing instructions?	HR manipulation ?	Anxiety measure	ES type	Extracted ES	*r*
Abrams et al., 2018	61	Adults	0.57	NR	NR	NR	HCT	No	No	ASI-3 and STAI-T	*r*	0.08 (comp)	0.08
Ardizzi & Ferri, 2018	36	Adults	0.39 +	No	No	No	HCT	No	No	SCS - Social anxiety subscale	*r*	-0.25	-0.25
Borg et al., 2018	42	Adults	1.00	NR	NR	Fibromyalgia	HCT	No	No	STAI-S	*r*	0.07	0.07
Borg et al., 2018	42	Adults	1.00	NR	NR	Fibromyalgia	HCT	No	No	STAI-T	*r*	0.15	0.15
De Pascalis et al., 1984	16	Adults	0.50	NR	NR	NR	HCT - Tapping	No	No	IPAT-anxiety (Italian)	*r*	-0.51*	-0.51
Di Lernia et al., 2018	30	Adults	1.00	No	No	No	HCT	No	No	STAI-S (Italian)	*r*	-0.18	-0.18
Di Lernia et al., 2018	30	Adults	1.00	No	No	No	HCT	No	No	STAI-T (Italian)	*r*	-0.03	-0.03
Di Lernia et al., 2020	60	Adults	0.78	NR	NR	Chronic pain	HCT	Yes	No	STAI-S (Italian)	*r*	-0.28	-0.28
Dunn et al., 2010	111	Adults	0.73	Yes	Yes	Anxiety and depression	HCT	No	No	HC MASQ-S and STAI-T	*r*	0.18 (comp)	0.18
Ehlers & Breuer, 1992	111	Adults	0.74	Yes	NR	PD	HCT	No	No	CD - panic	*t*	2.42*	0.23
Ehlers & Breuer, 1992	96	Adults	0.73	Yes	NR	IP	HCT	No	No	CD - infrequent panic	*t*	0.39*	0.04
Ehlers & Breuer, 1992	73	Adults	0.77	Yes	NR	SP	HCT	No	No	CD - phobia	*t*	5.10*	-0.12
Ehlers & Breuer, 1992	65	Adults	0.74	Yes	NR	PD	HCT	No	No	STAI -T (German)	*r*	0.03	-0.03

Ehlers & Breuer, 1992	50	Adults	0.72	Yes	NR	IP	HCT	No	No	STAI -T (German)	*r*	-0.13	0.13
Ehlers & Breuer, 1992	27	Adults	0.81	Yes	NR	SP	HCT	No	No	STAI -T (German)	*r*	0.34	-0.34
Ehlers & Breuer, 1992	46	Adults	0.74	No	NR	No	HCT	No	No	STAI -T (German)	*r*	0.12	-0.12
Ehlers & Breuer, 1992	65	Adults	0.74	Yes	NR	PD	HCT	No	No	STAI-S (German)	*r*	-0.08	0.08
Ehlers & Breuer, 1992	50	Adults	0.72	Yes	NR	IP	HCT	No	No	STAI-S (German)	*r*	-0.31	0.31
Ehlers & Breuer, 1992	27	Adults	0.81	Yes	NR	SP	HCT	No	No	STAI-S (German)	*r*	0.25	-0.25
Ehlers & Breuer, 1992	46	Adults	0.74	No	NR	No	HCT	No	No	STAI-S (German)	*r*	0.11	-0.11
Ehlers, 1995	169	Adults	0.75	Yes	Yes	PD, IP, SP, depression	HCT	No	No	STAI-T (German)	*r*	0.09	0.09
Ehlers et al., 2000	180	Adults	NR	Yes	NR	Anxiety	HCT	No	No	CD - any anxiety disorder	χ^2	12.71	0.27
Eley et al., 2004	79	Children	0.57	Yes - Q	NR	No	HCT	No	No	CASI and SCARED	*t*	-1.13 (comp)	0.34
Eley et al., 2007	576	Children	0.57	Yes - Q	NR	NR	HCT	No	No	CASI and SCARED	*r*	0.06(comp)	0.06
Ewing et al., 2017	66	Adults	0.81	Yes	No	Anxiety	HCT	No	No	CD - any anxiety disorder	*t*	2.39*	-0.29
Furman et al., 2013	25	Adults	1.00	No	Yes	Depression	HCT	No	No	BAI	*r*	0.01	0.01
Furman et al., 2013	36	Adults	1.00	No	No	No	HCT	No	No	BAI	*r*	-0.22	-0.22
Gaebler et al., 2013	42	Adults	NR	Yes	Yes	SeAD, PD, OCD, depression	HCT	No	No	CD - social anxiety	*t*	3.06	-0.44
Garfinkel, Tiley et al., 2016	40	Adults	0.10	NR	NR	ASC	HCT	No	No	STAI-T	*r*	-0.17	-0.17

Garfinkel, Tiley et al., 2016	40	Adults	0.10	NR	NR	ASC	HCT	No	No	STAI-S	*r*	-0.22	-0.22
Krautwurst et al., 2014	100	Adults	0.79	No	No	No	HCT	No	No	WI - 2 factor version (German)	*r*	-0.22	-0.23
Krautwurst et al., 2016	105	Adults	NR	Yes	No	PHA	HCT	No	No	CD - PHA	*t*	0.92	0.09
Lackner & Fresco, 2016	101	Adults	0.74	Yes	Yes	Anxiety, depression, diabetes, HBP, ADHD	HCT	No	No	MASQ-S anxious arousal	*r*	-0.23	-0.23
Leonidou et al., 2020	60	Adults	0.80	Yes - Q	NR	Yes - unspecified	HCT	No	No	SHAI (Greek)	*F*	0.22	0.05
Li et al., 2020	36	Adults	0.67	Yes	No	GAD	HCT	No	No	CD - GAD	*t*	1.00	-0.18
Mallorquí-Bagué et al., 2014	36	Adults	0.53	No	No	Hypermobility	HCT	No	No	STAI-S	*r*	0.28	0.28
Masdrakis et al., 2008	21	Adults	0.61 +	Yes	No	PD	HCT	No	No	Panic attack after caffeine challenge	Odds ratio	1.78	0.16
Michal et al., 2014	24	Adults	0.46	No	No	No	HCT	No	No	STAI-average (German)	*r*	-0.18	-0.18
Michal et al., 2014	26	Adults	0.46	Yes	Yes	Anxiety, depression, DPD	HCT	No	No	STAI-average (German)	*r*	-0.06	-0.06
Näring & Van Der Staak, 1995	24	Adults	0.00	Yes - Q	NR	No	HR rating		Yes	STAI-S (Dutch)	*F*	10.28	0.55
Pfleiderer et al., 2014	48	Adults	1.00	Yes - Q	No	No	HCT - Time1	No	No	ASI	*t*	1.15*	0.16
Pile et al., 2018	29	Children	0.48	NR	NR	TD, and other comorbidities	HCT	No	No	RCADS	*r*	0.40	0.40
Pile et al., 2018	25	Children	0.48	No	No	No	HCT	No	No	RCADS	*r*	-0.13	-0.13
Pollatos, Herbert, et al., 2007	18	Adults	0.00	No	No	No	HCT	No	No	STAI-S	*r*	-0.15	-0.15
Pollatos, Traut-Mattausch, et al., 2007	18	Adults	0.00	No	No	No	HCT	No	No	STAI-T	*r*	0.47	0.47

Pollatos, Traut-Mattausch et al., 2007	102	Adults	0.66	No	No	No	HCT	No	No	STAI-T	*r*	0.28	0.28
Pollatos, Matthias, et al., 2007	36	Adults	0.72	NR	NR	NR	HCT	No	No	STAI-T	*r*	0.30	0.30
Pollatos et al., 2009	119	Adults	0.82	No	No	No	HCT	No	No	STAI-T	*r*	0.20	0.20
Pollatos & Georgiou, 2016	23	Adults	1.00	No	No	No	HCT - control for BMI	No	No	STAI-T (German)	*r*	0.50	0.50
Pollatos & Georgiou, 2016	23	Adults	1.00	Yes	Yes	Anxiety, depression, BN	HCT - control for BMI	No	No	STAI-T (German)	*r*	-0.16	-0.16
Rost et al., 2017	98	Adults	0.83	NR	NR	Fibromyalgia	HCT	No	No	DASS-A	*r*	0.02	0.02
Santangelo et al., 2018	43	Adults	NR	No	No	Parkinson's	HCT	No	No	BAI (Italian)	*r*	0.01	0.01
Schaefer et al., 2014	52	Adults	0.73	No	NR	MUS	HCT	No	No	WI (German)	*t*	2.33*	0.32
Schandry, 1981	34	Adults	0.41 +	NR	NR	NR	HCT	No	No	STAI-S (German)	*t*	2.88	0.44
Schmitz et al., 2012	40	Children	0.53	NR	NR	High social fear + control group	HCT	No	No	SASC-R	*t*	0.80	-0.13
Schultchen et al., 2019	52	Adults	0.46	Yes	Yes	OCD, depression, AN, SD, SP, PD	HCT	No	No	STAI-T	*r*	-0.21	-0.21
Solano López & Moore, 2019	76	Adults	0.71	Yes	Yes	PH, depression, anxiety, diabetes, DL, asthma, COPD, OA, obesity, cataract, glaucoma	HCT	No	No	PROMIS Anxiety scale short form 4a	*r*	0.17	0.17
Steptoe & Vögele, 1992	29	Adults	1.00	No	No	No	HR rating		Yes	STAI-T	Cohen's *d*	0.49*	-0.20

Stern et al., 2020	23	Adults	NR	NR	NR	NR	HCT- tapping	No	No	STAI-average	*r*	-0.36	-0.36
Stevens et al., 2011	48	Adults	0.52	No	No	No	HCT	No	Yes	FNE (German)	*t*	2.81	0.38
Stewart et al., 2001	30	Adults	0.70	Yes - Q	No	No	HCT	No	No	ASI	*t*	3.02	0.49
Sturges et al., 1998	24	Adults	1.00	Yes - Q	NR	NR	HCT	No	Yes	ASI	*F*	0.60	0.13
Suschinsky & Lalumière, 2014	31	Adults	1.00	Yes	NR	Anxiety	HCT	No	Yes	SPRAS and ASI	*t*	1.33	-0.24
Tabor et al., 2019	38	Adults	0.50	No	No	No	HCT	Yes	No	ASI-3	*r*	-0.07	-0.07
Tan et al., 2018	40	Adults	0.55	No	No	No	HCT	No	No	STAI-T (Chinese)	*r*	0.28	0.28
Tan et al., 2018	40	Adults	0.55	No	No	No	HCT	No	No	STAI-S (Chinese)	*r*	0.17	0.17
Watson et al., 2019	54	NR	0.85	No	No	No	HCT	No	No	STAI-T	*r*	0.01	0.01
Watson et al., 2019	54	NR	0.85	No	No	No	HCT	No	No	STAI-S	*r*	0.02	0.02
Yoris et al., 2015	34	Adults	0.44	Yes	NR	AD with panic in the last month	HCT- tapping	No	No	Panic attack in the previous month	*F*	1.76	-0.23
Yoris et al., 2017	40	Adults	0.58	Yes	NR	PD	HCT - tapping	No	No	CD - panic disorder	*t*	0.57*	-0.09
Zoellner & Craske, 1999	58	Adults	0.48	Yes	NR	IP	HCT	Yes	No	ASI and CD	*t*	2.16	0.28

Note. F:M = female to male; NR = not reported; Q = questionnaire; PD = panic disorder; IP = infrequent panic; SP = specific phobia; PHA = pathological health anxiety; HBP = high blood pressure; BN = bulimia nervosa; MUS = medically unexplained symptoms; TD = tic disorder; SeAD = separation anxiety disorder; OCD = obsessive-compulsive disorder; SoD = somatic disorder; AD = anxiety disorder; GAD = generalised

anxiety disorder; COPD = chronic obstructive pulmonary disease; OA = osteoarthritis; PH = primary hypertension; DL = dyslipidaemia; CD = clinical diagnosis; + = ratio calculated before exclusions; * = effect size calculated from data provided in the paper

Table 5.3

HDT studies

Author	*N*	Age	F: M ratio	Anxiety patients	Depression patients	Clinical group	Interoception task	HR manipulation ?	Anxiety measure	ES type	Extracted ES	r
De Pascalis et al., 1991	78	Adults	1.00	NR	NR	Not reported	HDT	No	STAI-S	*r*	0.08	0.08
De Pascalis et al., 1991	78	Adults	1.00	NR	NR	Not reported	HDT	No	STAI-T	*r*	-0.10	-0.10
Ewing et al., 2017	66	Adults	0.81	Yes	No	Anxiety	HDT	No	CD	*t*	-0.36	-0.16
Garfinkel, Tiley et al., 2016	40	Adults	0.10	NR	NR	ASC	HDT	No	STAI-T	*r*	-0.47	-0.47
Garfinkel, Tiley et al., 2016	40	Adults	0.10	NR	NR	ASC	HDT	No	STAI-S	*r*	-0.35	-0.35
Lyyra & Parviainen, 2018	50	Adults	0.54	No	No	No	HDT	No	KSP-ANX	*r*	0.34	0.34
Michal et al., 2014	24	Adults	0.46	No	No	No	HDT	No	STAI-T	*r*	-0.14	-0.14

Author	*N*	Age	F: M ratio	Anxiety patients	Depression patients	Clinical group	Interoception task	HR manipulation ?	Anxiety measure	ES type	Extracted ES	r
Michal et al., 2014	26	Adults	0.46	Yes	Yes	DPD, anxiety, depressio n	HDT	No	STAI-T	*r*	0.09	0.09
Schirmer-Mokwa et al., 2015	13	Adults	0.46	No	No	No	HDT	No	STAI-T	*r*	-0.01	-0.01
Schirmer-Mokwa et al., 2015	12	Adults	0.50	No	No	No	HDT	No	STAI-T	*r*	0.37	0.37
Schirmer-Mokwa et al., 2015	13	Adults	0.54	No	No	No	HDT	No	STAI-T	*r*	0.43	0.43

Note. Where it says "comp" in the effect size, it means the average was taken from all reported correlations within that sample group, to avoid issues with non-independence. F:M = female:male; NR = not reported; ASC = autism spectrum condition; AN = anorexia nervosa; DPD = depersonalisation derealisation disorder; GAD = generalised anxiety disorder; OCD = obsessive-compulsive disorder; E.S. = effect size; CD= clinical diagnosis

Heart rate manipulation

Whether or not participants' heart rate was manipulated was coded as a binary variable. In two studies, participants completed several tasks that may increase their heart rate, including mental arithmetic, the cold pressor test or squeezing a rubber ball (Näring et al., 1995; Steptoe & Vögele, 1992). One study induced hyperventilation in participants (Sturges et al., 1998), another played sexual and non-sexual films to participants (Suschinsky & Lalumière, 2014), and in a final study, participants were told to prepare to give a public speech (Stevens et al., 2011). In all of these studies, all or the majority of the manipulations resulted in an increased heart rate.

Clinical diagnoses

For conditions other than anxiety (e.g., depression), we only classified the study as including clinical groups if the participants had received a diagnosis from a clinician. For anxiety, we took this approach in our initial moderator analyses, with follow-up analyses extending inclusion to studies that classified individuals as anxious where questionnaire scores were above established clinical thresholds.

Analysis strategy

Separate meta-analyses were conducted for the HCT and HDT given evidence that the two measures are not highly correlated (Hickman et al., 2020). For the HCT, separate meta-analyses were conducted for state and trait measures of anxiety. Two studies utilised a composite score of the STAI including both state and trait anxiety (Michal et al., 2014, Stern et al., 2020). These studies were included in both state and trait models and excluded for comparisons. Separation of state and trait anxiety was not conducted for the HDT as only two studies examined the relationship between the HDT and state anxiety (Garfinkel, Tiley et al., 2016; De Pascalis et al., 1991). In these initial analyses, data from both clinical groups and

typically developing participants was accepted, as a greater distribution of scores would be expected for studies including clinical populations. In primary analyses, we excluded papers reporting data in child populations, which were examined later in separate meta-analyses. One study did not report the ages of participants (Watson et al., 2019) but was included in the adult group given recruitment methods focused on university students. Where pre-registered or a sufficient number of papers were identified, specific moderators were examined. As several studies reported the relationship between multiple measures of anxiety and cardiac interoceptive accuracy, for these studies composite scores were created using the method described by Borenstein and colleagues (2009).

All analyses were conducted on R version 4.1.2, using the devtools, dmetar and meta packages (Balduzzi et al., 2019; Harrer et al., 2019; Wickman et al., 2021). All test statistics were converted to *r* values, which were chosen as the majority of studies examined correlations between cardiac interoceptive accuracy and anxiety. R-to-z transformations were implemented by the software before analyses. As the direction of scoring varied across studies (with differences in whether error or accuracy scores were computed), the sign on all effect sizes was adjusted such that a positive score indicates better cardiac interoception with increasing anxiety. As substantial heterogeneity was expected, random-effects models were employed (Field, 2001; Hunter & Schmidt, 2000) and the conservative Sidik-Jonkman estimator was applied (Sidik & Jonkman, 2007). Heterogeneity was investigated for each meta-analysis, and Q and I^2 statistics are reported. A Q statistic indicates whether variation is greater across or within studies and is calculated by summing the weighted squared differences between the effect sizes of each study and the fixed effect estimate. I^2 indicates the percentage of variation in effect sizes across studies that is due to heterogeneity rather than chance. Values of 25%, 50% and 75% indicate low, moderate, and high heterogeneity respectively. Funnel plots which mapped the relationship between effect sizes and standard

error were created as one method to explore publication bias. Additionally, Egger's tests were used to assess the asymmetry of the funnel plot. This could only be done in instances where there were 10 or more studies included (Borenstein et al., 2009). We also conducted influence analyses using the leave-one-out method to establish the influence of individual studies on the overall pooled effect sizes.

Results

HCT and variants with trait anxiety

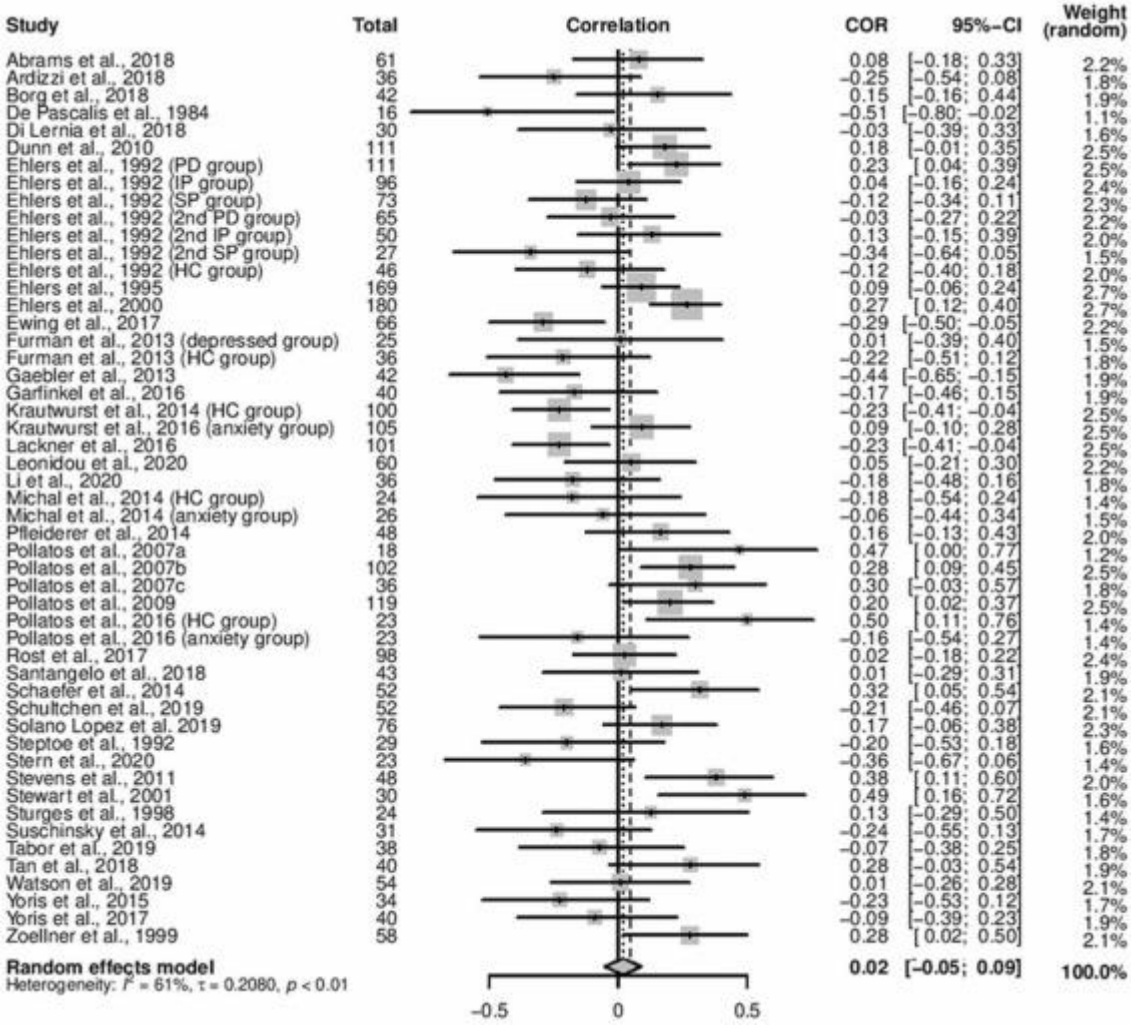
In this category, we included studies using any questionnaire measure of anxiety that was not state-dependent (i.e., participants reported their anxiety levels over an extended period), in addition to studies that adopted between-group designs comparing anxious and non-anxious individuals. All diagnoses and questionnaire measures (general, health anxiety and social phobia) were accepted. For these, the lowest level of heart rate perturbation (baseline) was taken if there were multiple conditions. This resulted in 51 effect sizes for this analysis.

No significant association was observed between trait anxiety and HCT variants, with a pooled effect size of .02 ($p = .609$, see Figure 5.2). There did appear to be evidence of slight publication bias, whereby negative correlations were favoured, as indicated by a significant Egger's test ($p = .043$), although the funnel plot was mainly symmetrical (Figure 5.3). In terms of heterogeneity, a significant Q statistic ($Q = 129.68$, $p < .001$) and an I^2 value of 61% both indicated moderate-high heterogeneity in the total sample. Ten studies were identified as outliers that significantly deviated from the 95% confidence interval of the pooled effect size (De Pascalis et al., 1984; Ehlers et al., 2000; Ewing et al., 2017; Krautwurst et al., 2014; Lackner & Fresco, 2016; Pollatos, Traut-Mattausch et al., 2007; Pollatos & Georgiou, 2016 (non-anxious group); Stevens et al., 2011; Stewart et al., 2001). However, the removal of

these papers did not meaningfully change the pattern of results observed ($r = .02$, $p = .508$). In the total sample, influence analysis indicated effect sizes ranged from 0.01 to 0.03, all of which were non-significant, suggesting little influence of individual studies on the pooled effect size.

Figure 5.2

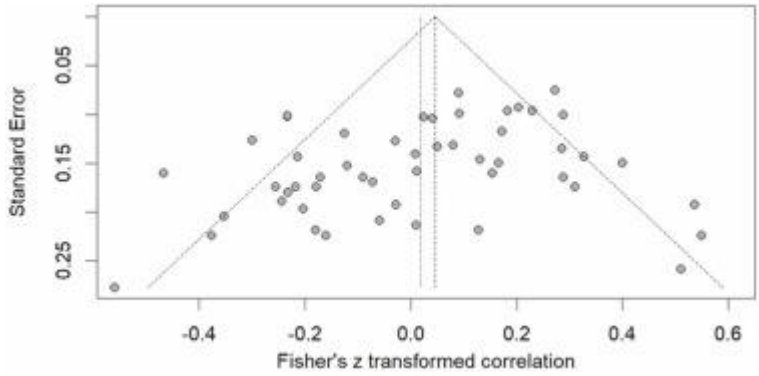
Forest plot: trait anxiety and HCT variants



Note. PD = panic disorder; IP = infrequent panic; SP = specific phobia; HC = healthy control. Pollatos et al., 2007a refers to: Pollatos, Herbert et al., 2007; Pollatos et al., 2007b refers to: Pollatos, Traut-Mattausch, et al., 2007; Pollatos et al., 2007c refers to: Pollatos, Matthias, et al., 2007 (citations edited for space).

Figure 5.3

Funnel plot: trait anxiety and HCT variants



Note. Funnel plot to demonstrate the risk of publication bias in studies using HCT and variants with trait anxiety measures. The degree of asymmetry indicates the risk of publication bias, with higher asymmetry indicating a greater risk. This plot indicates a low risk.

Further analyses suggested that the pooled effect size was similar when examining studies reporting correlations with questionnaire measures and those using group designs (Q =.02, p = .887), with no significant relationship observed for studies using group (k = 20, r = .02, p = .712) or correlational designs (k = 31, r = .01, p = .761). The effect size was also relatively unchanged when excluding questionnaire measures that did not assess general 'trait' anxiety, restricting inclusion to studies assessing anxiety using the ASI, ASI-3, BAI, STAI-T (excluding those that created composite scores with the STAI-S), SPRAS, PROMIS, MASQ-S, IPAT or DASS (k = 32, r = .04, p = .327). Similarly, restricting analyses to the most used measure – the STAI-T – revealed no significant association (k = 18, r = .07, p = .267).

HCT and variants with state anxiety

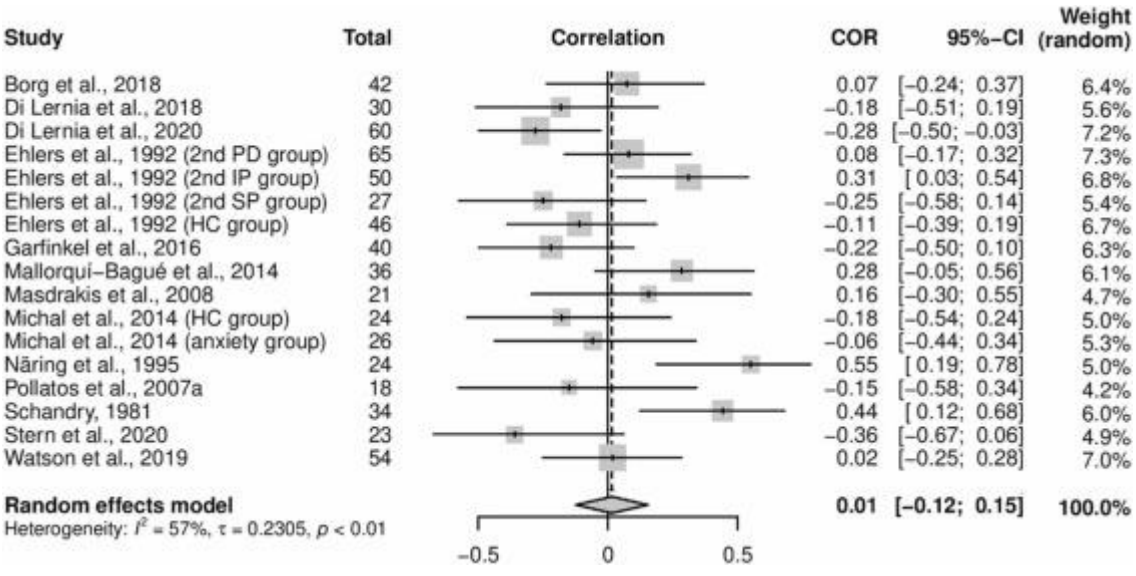
In this category, we included studies using any questionnaire measure of state anxiety (uniformly assessed by the STAI-S), in addition to one study that compared HCT

performance in individuals who suffered a panic attack following a caffeine challenge to those who did not (Masdrakis et al., 2008).

No significant association was observed between state anxiety and HCT variants, with a pooled effect size of .02 ($p = .835$, see Figure 5.4). There did not appear to be evidence of publication bias, as indicated by the symmetry of the funnel plot (Figure 5.5) and a non-significant Egger's test ($p = .971$). A significant Q statistic ($Q = 37.12$, $p < .005$) and an I^2 value of 56.9% both indicated moderate to high heterogeneity in the total sample. In terms of outliers, two studies significantly deviated from the 95% confidence interval of the pooled effect size (Näring et al., 1995; Schandry, 1981), but the removal of these papers did not alter the overall results obtained ($r = -.05$, $p = .454$). Effect sizes ranged from − 0.02 to 0.04 in the influence analysis, all of which were non-significant. Results remained non-significant when considering only studies that used the most common measure of state anxiety (the STAI-S; $k = 13$, $r = .05$, $p = .562$).

Figure 5.4

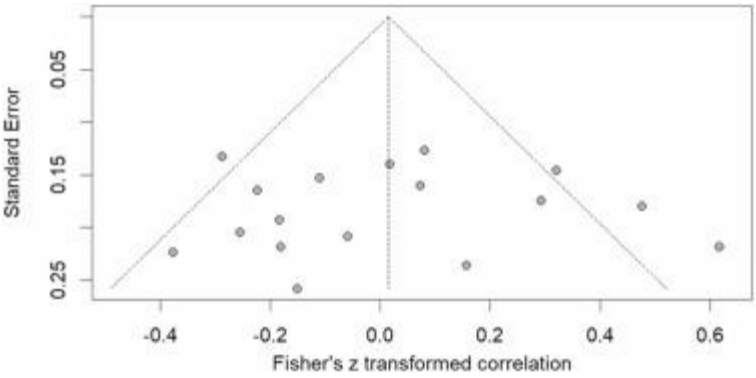
Forest plot: state anxiety and HCT variants



Note. PD = panic disorder; IP = infrequent panic; SP = specific phobia; HC = healthy control.

Figure 5.5

Funnel plot: state anxiety and HCT variants



Note. Funnel plot to demonstrate the risk of publication bias in studies using HCT and variants with state anxiety measures.

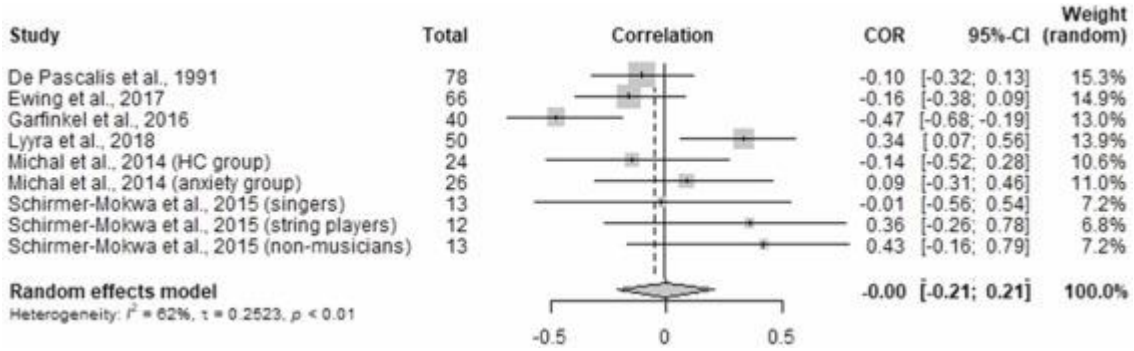
HDT and variants with trait anxiety

In this category, we included studies using any questionnaire measure of trait anxiety as assessed by questionnaires, in addition to studies that adopted between-group designs comparing anxious and non-anxious individuals.

Again, no significant association was observed between anxiety and HDT variants, with a pooled effect size of − .00 (p = .977, see Figure 5.6). There did not appear to be evidence of publication bias, as indicated by the symmetry of the funnel plot (Figure 5.7) and a non-significant Egger's test (p = .383). In terms of heterogeneity, a significant Q statistic (Q = 21.30, p < .01) and an I^2 value of 62.4% both indicated moderate heterogeneity in the total sample. One study was identified as an outlier that significantly deviated from the 95% confidence interval of the pooled effect size (Garfinkel, Tiley et al., 2016). When this paper was removed, there was not a notable change in the pooled effect size (r = .06, p = .528). Effect sizes from the influence analysis were again all non-significant, ranging from − 0.06 to 0.06, demonstrating that individual studies had little effect on the pooled effect size.

Figure 5.6

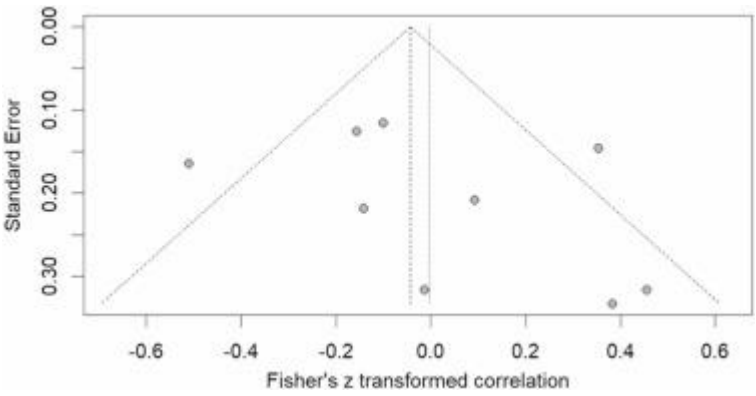
Forest plot: trait anxiety and HDT variants



Note. HC = healthy control.

Figure 5.7

Funnel plot: trait anxiety and HDT variants



Note. Funnel plot to demonstrate the risk of publication bias in studies using HDT and variants with trait anxiety measures.

Further analysis indicated that results were still non-significant when only including studies that used the most commonly used measure, the STAI-T ($k = 7$, $r = -0.04$, $p = .758$).

Moderation analysis

Given previous suggestions that several factors may moderate the relationship between cardiac interoception and anxiety, several further analyses were conducted (see Table

5.4). As outliers, and the specific measures of anxiety and cardiac interoceptive accuracy employed, had little influence on results for primary analyses these were retained in subgroup analyses. It is important to note that in a number of these comparisons, $k < 10$, meaning they were underpowered to detect significant effects.

Table 5.4

Moderation analysis

Analysis	Category	Group	*Q*	*p*	*k*	PES	*p*
		Pre-registered					
HCT instructions	HCT + trait		0.34	.560			
		No guessing instructions *			2	0.12	.484
		No no guessing instructions			48	0.02	.634
Influence of depression	HCT + trait		0.21	.646			
		Anxiety measure captured depression			18	0.07	.267
		Anxiety measure captured anxiety			18	0.03	.690
		Exploratory					
Inclusion of P's with AD	HCT + trait		2.55	.110			
		AD included			22	-0.03	.563
		AD not included			22	0.09	.109
	HCT + state		0.03	.866			
		AD included *			5	0.07	.516
		AD not included *			7	0.04	.715
	HDT + trait		2.11	.147			
		AD included *			2	-0.08	.574
		AD not included *			5	0.20	.140
Inclusion of P's with AD OR above threshold	HCT + trait		0.67	.412			
		High anxiety included			26	0.00	.978
		High anxiety not included			18	0.07	.300
	HCT + state		1.31	.252			
		High anxiety included *			6	0.14	.254
		High anxiety not included *			6	-0.03	.707
Inclusion of P's with PD	HCT + trait		0.01	.939			
		PD included			11	0.02	.758
		PD not included			34	0.03	.504
	HCT + state		1.41	.235			
		PD included *			3	0.18	.066
		PD not included *			8	0.01	.948

Inclusion of P's with depression diagnosis	HCT + trait		1.38	.240			
		Depression included *			9	-0.06	.452
		Depression not included			20	0.06	.532
Inclusion of P's with other clinical diagnosis	HCT + trait		1.17	.280			
		Other diagnosis included			12	-0.04	.543
		Other diagnosis not included			34	0.05	.298
	HCT + state		0.38	.537			
		Other diagnoses included *			5	-0.05	.639
		Other diagnosis not included			10	0.04	.688
	HDT + trait		1.79	.180			
		Other diagnoses included *			2	-0.22	.436
		Other diagnosis not included *			5	0.20	.140
Age of participants	HCT + trait		0.73	.392			
		Children *			5	0.12	.279
					51	0.02	.608
Female to male ratio	HCT + trait		0.98	.323			
		More females			37	0.05	.165
		More males *			9	-0.05	.594
	HCT + state		0.18	.674			
		More females			10	0.01	.883
		More males *			6	0.08	.590
	HDT + trait		1.56	.212			
		More females *			4	0.08	.578
		More males *			4	-0.18	.225
Heartrate manipulation	HCT + trait		1.05	.304			
		Manipulation *			3	-0.12	.378
		No manipulation			48	0.03	.504

Note. PES = pooled effect size, AD = anxiety disorder (clinical diagnosis), PD = panic disorder (clinical diagnosis), * = $k < 10$, likely underpowered.

Pre-registered moderators

Instructions for HCT. For studies which considered trait anxiety, two conducted the HCT with explicit no guessing instructions provided, whereas 48 studies either did not use no guessing instructions or did not report that they used no guessing instructions. Overall, there was no significant difference ($Q = 0.34$, $p = .560$) in the pooled effect sizes, and neither the

studies that gave no guessing instructions ($k = 2$, $r = .12$, $p = .484$) nor those that did not ($k = 48$, $r = .018$, $p = .634$) showed a significant pooled effect size.

We were unable to run this analysis on the studies looking at state anxiety, as only one study employed no guessing instructions.

Influence of depression. Clinician ratings of the questionnaires used to measure anxiety highlighted that only the STAI-T captured depression in over 50% of the items. Therefore, this analysis compared effect sizes from studies using the STAI-T to all other questionnaire measures of anxiety. For this analysis, we removed any effect sizes that were created using composite measures of the STAI-T and other measures of anxiety. This included those using an average of the STAI-S and STAI-T. Of the studies using the HCT, 18 used the STAI-T and 18 studies used other measures. Results indicated no significant difference in the effect sizes across studies employing the STAI-T and other measures ($Q = 0.21$. $p = .646$); neither studies using the STAI-T ($k = 18$, $r = .07$, $p = .267$) or other measures ($k = 18$, $r = .03$, $p = .690$) showed a significant pooled effect size.

We could not run this analysis on studies using the HDT, as there were no studies using anxiety measures that had been confirmed to not capture depression.

Exploratory moderators

Group versus correlational design. For the HCT and state measures of anxiety, there did appear to be a significant difference in the effect sizes between studies employing group versus correlational designs ($Q = 9.01$, $p = .003$), with correlational studies showing a non-significant, negative effect size ($k = 14$, $r = .06$, $p = .374$) and group studies showing a significant, positive effect size ($k = 3$, $r = .43$, $p = .004$). However, it is important to note that this analysis was extremely under-powered, with only three studies employing group studies.

We could not conduct this comparison for studies using the HDT and trait measures of anxiety, as there was only one study employing a group design.

Inclusion of anxiety diagnoses. Given suggestions that relationships may be more likely to be observed where individuals with anxiety are included in samples (Domschke et al., 2010), we also assessed whether such sample characteristics may moderate relationships. For these analyses, we excluded papers where such sample characteristics were not reported. In studies using the HCT and trait measures of anxiety, there was no significant difference between studies including participants who had clinical anxiety diagnoses and those who did not (Q = 2.55, p = .110; studies including participants with clinical anxiety: k = 22, r = -.03, p = .563; studies not including participants with clinical anxiety: k = 22, r = .09, p = .109). There was also no difference between participants who either had clinical anxiety or scores above established thresholds for high anxiety on questionnaire measures and participants who did not (Q = 0.67, p = .412; studies including anxious participants, as measured by diagnosis or questionnaire: k = 26, r = .00, p = .978; studies not including anxious participants: k = 18, r = .07, p = .300).

For studies using the HCT and state measures of anxiety, there was no difference in studies that included participants with clinical diagnoses of anxiety and those that did not (Q = 0.03, p = .866). Both were non-significant (studies including participants with clinical anxiety: k = 5, r = .07, p = .517; studies not including participants with clinical anxiety: k = 7, r = .04, p = .715). The same was found when including participants who scored over thresholds that indicated high anxiety (Q = 1.31, p = .252; including anxious participants: k = 6, r = .14, p = .254; not including anxious participants: k = 6, r = -.03, p = .707).

For studies using the HDT and trait measures of anxiety, there was also no significant difference between studies including participants with a clinical anxiety diagnosis and those

without (Q = 2.11, p = .147; studies including participants with a clinical anxiety diagnosis: k = 2, r = -.08, p = .574; studies not including participants with a clinical anxiety diagnosis: k = 5, r = -.20, p = .140). No additional studies included participants scoring over a threshold for high anxiety.

Inclusion of individuals with panic disorder diagnoses. Given suggestions that relationships may be specific to, or more pronounced in, individuals with PD (Ehlers & Breuer, 1992), we also assessed whether relationships would be more likely to be observed where samples contained individuals with PD. For this, we included data from individuals with infrequent panic attacks. Of the studies using the HCT and trait measures of anxiety, there was no significant difference in the effect sizes obtained from studies including participants with PD and studies that did not (Q = 0.01, p = .939). Neither studies including those with PD (k = 11, r = .02, p = .758) or those without PD (k = 34, r = .03, p = .504) were significant.

Of the studies using the HCT and state measures of anxiety, there was also no significant difference (Q = 1.41, p = .235), with both those including participants with PD (k = 3, r = .18, p = .066) and those not including participants with PD (k = 8, r = .01, p = .948) not being significant. We could not run an analysis on studies using the HDT, as no panic patients were included in any of the studies.

Inclusion of individuals with a depression diagnosis. As depression is often reported to show differing relationships with cardiac interoception when compared to anxiety (Dunn et al., 2010; Pollatos et al., 2009), we also examined depression diagnosis as a moderating factor. As before we removed studies that did not report whether participants with depression were included or excluded. For the studies employing the HCT and trait measures of anxiety, there was no significant difference between studies with clinically depressed

patients versus no clinically depressed patients (Q = 1.38, p = .240). Neither group was significant (k = 20, r = .06, p = .532 for studies not including participants with depression, k = 9, r = -.06, p = .452 for studies including participants with depression). The analysis could not be run on the HCT and state measures studies as no studies reported depressed participants. Similarly, only one study of the studies using the HDT included participants with depression.

Inclusion of any clinical groups. As included papers often employed samples with other clinical diagnoses, we also examined associations removing studies that included samples that had clinical diagnoses that were not anxiety or depression. Within studies using the HCT and trait measures of anxiety, there was no significant difference (Q = 1.17, p = .280) between studies that did include participants with clinical diagnoses other than depression and anxiety (k = 12, r = -.04, p = .543) and those that did not (k = 34, r = .05, p = .298). The same was found within studies using the HCT and state measures of anxiety (Q = 0.38, p = .537; studies including other clinical participants: k = 5, r = -.05, p = .639; studies not including other clinical participants: k = 10, r = .04, p = .688) and studies using the HDT and trait measures of anxiety (Q = 1.79, p = .180; studies including other clinical participants: k = 2, r = -.22, p = .436; studies not including other clinical participants: k = 5, r = .14, p = .529).

Age of participants. Given that interoceptive accuracy varies across the developmental timeline (Murphy et al., 2017), we also examined the relationship in child populations, to see if it differed from what was observed in adult populations. This was conducted only for the HCT and trait anxiety given the number of studies conducted in child populations. No significant relationship was observed between trait anxiety and HCT performance in children (k = 5, r = .12, p = .279).

Female: Male ratio. Given evidence of sex differences in interoception (Prentice & Murphy, 2021), and suggestions that the relationship between interoceptive accuracy and mental health may vary as a function of sex (Lischke et al., 2020), we also examined if relationships varied according to the proportion of female relative to male participants in samples. For this analysis, studies not reporting sample characteristics were excluded and studies with equal numbers of male and female participants were excluded. One study only reported demographics before exclusions (Ardizzi & Ferri, 2018), so this ratio was used. When considering studies that examined HCT performance and trait anxiety, there was no significant difference in the effect sizes between studies employing designs with a female versus male majority (Q = 0.98, p = .323). No significant pooled effect size was observed for studies with a female (k = 37, r = .05, p = .165) or male majority (k = 9, r = -.05, p = .594). The same was found for studies using the HCT and state measured of anxiety (Q = 0.18, p = .674; studies with a female majority: k = 10, r = .01, p = .883; studies with a male majority: k = 6, r = .08, p = .590) and the HDT and trait measures of anxiety (Q = 1.56, p = .212; studies with a female majority: k = 4, r = .08, p = .578; studies with a male majority: k = 4, r = -.18, p = .225).

Manipulation to increase HR. HR manipulation appeared to have no significant effect on the relationships observed in studies using the HCT and trait measures of anxiety (Q = 1.05, p = .304). Neither studies with (k = 3, r = -.12, p = .378) or without manipulation (k = 48, r = .03, p = .504) were significant. We could not assess the effect of HR manipulation in studies using the HCT and state measures of anxiety or the HDT and trait measures of anxiety.

None of the pre-registered moderators (whether no guessing instructions were used for the HCT and whether anxiety questionnaires captured depression) significantly impacted the pooled effect size. None of the exploratory moderators (whether any participants had a

diagnosis of any anxiety disorder, PD, depression or any other clinical condition, the ratio of female to male participants, or whether participants' HRs were manipulated to make them more detectable) significantly impacted the pooled effect size. Only one significant relationship emerged; for the HCT and state anxiety, the pooled effect size from studies using group designs was significantly larger than studies that employed correlational designs. However, these results should be interpreted with caution as only three studies were included in the group design category.

Discussion

The aim of this pre-registered systematic review and meta-analysis was to establish the presence or absence of a relationship between cardiac interoceptive accuracy and anxiety. Overall, we found no evidence for an association. Neither trait nor state anxiety related to performance in the HCT task, and trait anxiety did not relate to performance on the HDT task. Our analyses indicated moderate to high levels of heterogeneity across categories assessed, that did not seem to be accounted for by the examined moderators. Such findings are inconsistent with theoretical proposals that implicate interoception in the aetiology of anxiety (Domschke et al., 2010; Ehlers, 1993; Paulus & Stein, 2010; Garfinkel, Tiley et al., 2016).

Research conducted following the literature search for this meta-analysis mostly substantiates our findings. Recent studies using the HCT have found no association with trait (Noda et al., 2022), state (Wiśniewski et al., 2023), or combined state and trait anxiety (Vabba et al., 2023), although one study found that participants with comorbid depression and anxiety performed better on the HDT than those with only depression (Ironside et al., 2023).

One explanation for our null findings may be the inadequacy of both the HCT and HDT to provide a valid measure of cardiac interoceptive accuracy. The limitations of both

measures have been well discussed in both the wider literature (Brener et al., 1993; Desmedt et al., 2018; Murphy, Millgate, et al., 2018; Ring et al., 2015), and in Chapter Two of this thesis.

Another explanation is that there is simply no association between anxiety and cardiac interoceptive accuracy. Indeed, whilst methodological limitations of both the HCT and the HDT may underlie the results obtained, it is notable that studies using novel tasks or analytic approaches that overcome these limitations have also produced mixed results with some studies reporting relationships between anxiety and cardiac interoceptive accuracy (Smith, Feinstein, et al., 2021) and others not (Plans et al., 2021). Whilst task differences or differences in the anxiety measures employed may account for these discrepant results, such discrepancies are consistent with the mixed results noted throughout this meta-analysis. As such, the results of this meta-analysis are not consistent with the proposal that anxiety relates to cardiac interoceptive accuracy and recent work remains inconclusive.

Although several different factors have been proposed to moderate this association, including participants' clinical diagnoses and the instructions given for the HCT task (Domschke et al., 2010; Ehlers et al., 1995; Smith, Feinstein, et al., 2021), we found no evidence that these moderators significantly impacted the association between anxiety and interoception. Whilst some of these moderator analyses may have been insufficiently powered, for the six comparisons where sufficient power was present, we found no evidence of a moderating effect. Such results suggest that discrepant results for the HCT and anxiety are not explained by the use of correlational vs group designs, depression influencing anxiety scores on questionnaires, the inclusion of individuals with an anxiety diagnosis, above threshold symptom reporting, a panic diagnosis, or indeed any of the other diagnoses explored in this meta-analysis (see Table 5.2, Table 5.3).

The results from this meta-analysis are difficult to interpret considering recent evidence that cardiac interoceptive training that improves accuracy also reduces anxiety, as measured by the STAI-T, in autistic samples (Quadt et al., 2021). However, one possibility is that anxiety relates to other dimensions of interoception that may be influenced by the intervention, but not interoceptive accuracy. As discussed in Chapter One, there is evidence that anxiety may relate to other dimensions of interoception; for example, anxiety seems to relate to greater self-reported interoceptive attention (Vabba et al., 2023), more negative evaluations of interoceptive states (Paulus, 2013; Slotta et al., 2023; Taylor et al., 1991, Chapter Four), and greater physiological reactivity and perceived intensity of interoceptive states (Teed et al., 2022). It is plausible that it is these facets specifically, or the interaction between these facets and interoceptive accuracy, that are anxiogenic (Garfinkel, Tiley, et al., 2016; Palser, Fotopoulou, et al., 2018). Interoceptive accuracy training may result in downstream effects on these other facets of interoception, thus reducing anxiety through those mechanisms. Alternatively, as training protocols often include psychoeducation aimed at attention to, and interpretations of, interoceptive signals (e.g., Quadt et al., 2021), it may be that the benefits of interoceptive training on anxiety are via those mechanisms, and not via interoceptive accuracy at all. Further studies comparing interoceptive training protocols are required to isolate the mechanism by which anxiety relates to interoception.

It is also worth acknowledging that whilst our results suggest no relationship with cardiac interoceptive accuracy, a relationship between anxiety and interoceptive accuracy may be observable in other domains of interoception (e.g., respiratory, gastric) given evidence that interoceptive accuracy dissociates across interoceptive domains (Ferentzi et al., 2018; Garfinkel, Manassei, et al., 2016; Pollatos et al., 2016). For example, like cardiac interoceptive accuracy, there are good theoretical reasons to expect a relationship between anxiety and respiratory interoceptive accuracy (Paulus, 2013) and there is initial evidence of

such an association (Harrison et al., 2021; Teed et al., 2022; Tiller et al., 1987; although see Slotta et al., 2023). As evidence suggests respiratory and cardiac interoceptive accuracy dissociate (Garfinkel, Manassei, et al., 2016), it remains a possibility that associations between anxiety and interoception may be observed in the respiratory domain and indeed other domains of interoception.

Despite the importance of these results for understanding the relationship between anxiety and cardiac interoceptive accuracy, it is important to acknowledge the limitations of this study. First, many of our moderator analyses were underpowered ($k < 10$; Schwarzer et al., 2015). As such, we cannot conclusively say that these moderators do not impact the relationship between anxiety and cardiac interoceptive accuracy. Second, due to small numbers, it was not possible to explore the moderating effect of certain factors for the HDT and HCT with state anxiety. Third, due to the variety of clinical conditions included in the papers, we could not explore the moderating effects of specific conditions other than anxiety, PD and depression and only considered the effects of including participants with any clinical diagnosis. Our meta-analysis was also unable to examine the relationship between cardiac interoceptive accuracy and specific aspects of anxiety (e.g., anxiety sensitivity) due to the small number of studies identified. Future research employing fine-grained assessment of anxiety symptoms is required to examine this further. Finally, PubMed was the only database searched due to the large number of records identified. Whilst we checked for additional papers using citation searches, it is possible that a small number of studies may have been missed.

Despite these limitations, this pre-registered meta-analysis triangulated findings from numerous studies to provide a unifying overview of the relationship between anxiety and cardiac interoceptive accuracy. More work is required to investigate whether certain factors moderate this relationship and whether relationships between interoception and anxiety are

observed for other domains and dimensions of interoception. Such work is important for understanding discrepancies between theoretical and empirical work, and for understanding the mechanism by which interoceptive training improves anxiety.

Conclusion

Our findings suggest there is no relationship between trait anxiety and cardiac interoceptive accuracy, as measured by either the HCT or HDT, or between state anxiety and cardiac interoceptive accuracy, as measured by the HCT. There was no evidence that any of the variables we explored moderated this relationship.

Chapter Six: Interoception and anxiety in autistic adolescents

Content adapted from: **Adams, K.A.,** Smith, J., Brown, M., Bird, G. & Waite, P. (2025). Experiences of interoception and anxiety in autistic adolescents: A reflexive thematic analysis. *Autism.* https://doi.org/10.1177_13623613251314595

Kiera Adams designed this project in collaboration with Polly Waite, who acted as Kiera's supervisor for this project. Kiera Adams conducted the relevant PPI work, amending the design considering feedback from her PPI participants, which included co-authors Jonahs Smith and Mary Brown. Kiera Adams conducted all thirteen interviews and transcribed them. Kiera Adams conducted initial code and theme generation, which was amended following discussions with Polly Waite, Geoffrey Bird, Jonahs Smith and Mary Brown, and participant feedback. Kiera Adams wrote up the chapter, with input from Polly Waite and Geoffrey Bird.

Abstract

Most autistic adolescents experience anxiety. Interoception, defined as one's ability to detect and interpret bodily signals, might contribute to this. This exploratory, qualitative study aimed to facilitate a better understanding of interoceptive experiences in autistic adolescents and how this relates to anxiety. Semi-structured one-on-one interviews were conducted with thirteen UK-based, autistic adolescents, aged 14 to 17, who reported experiencing significant, day-to-day anxiety. Interviews were analysed using a reflexive thematic analysis approach. Interoception appeared relevant to experiences of anxiety in most participants, although this influence was exerted through various routes. Whilst hypoawareness of interoceptive signals led to a reduced ability to detect and regulate anxiety, hyperawareness became overwhelming and anxiety-inducing. Overly negative evaluations of interoceptive signals, including catastrophising about the signals indicating danger, also triggered anxiety. Relationships between interoception and anxiety appeared bidirectional, with anxiety also affecting interoceptive awareness and evaluation. Interoceptive experiences were variable across and within participants, with factors such as arousal and cognitive overload affecting experiences. Future research should focus on assessing the feasibility of interventions based on the mechanisms identified here.

Introduction

Compared to their neurotypical counterparts, autistic adolescents face a higher risk of developing anxiety disorders, with research estimating that autistic children and adolescents are twice as likely as neurotypicals to have an anxiety disorder (Costello et al., 2005), and an estimated 40% of autistic youths experiencing one or more anxiety disorder (van Steensel et al., 2011). Autistic adolescents also have higher levels of trait anxiety than non-autistic adolescents (White et al., 2009). In general, adolescence is a pivotal time for the development of anxiety disorders and symptoms, potentially due to the increasingly complex social systems and physiological changes occurring (Pfeifer & Allen, 2021). Qualitative research has highlighted that in autistic populations specifically, these conventional triggers for anxiety exist alongside, and interact with, autism-specific anxiety triggers such as sensory hypersensitivity (Acker et al., 2018) and masking (Chapman et al., 2022). Concerningly, autistic youth appear to benefit less from conventional interventions for anxiety disorders than neurotypicals (van Steensel & Bögels, 2015), highlighting a need for a better understanding of the phenomenology of anxiety in autistic young people.

There are numerous routes by which autism may result in increased levels of anxiety (Rodgers & Ofield, 2018; Wood & Gadow, 2010), including via autism-related stressors such as peer rejection, social confusion and aversive sensory experiences, as well as external environments being inadequately accessible. Another potential route is through interoception (Trevisan, Mehling, et al., 2021), which has been implicated in both autism and anxiety. Whilst it has been suggested that adolescence may be a critical time in the development of interoception, thus coinciding with the peak emergence of anxiety disorders in autistic individuals (White et al., 2010), there is currently a lack of research on this (Murphy et al., 2017).

Interoception has been theorised to be different in autistic compared to non-autistic individuals (Hatfield et al., 2019; Quattrocki & Friston, 2014). However, research findings have been equivocal, and some dimensions of interoception, such as evaluation, have not been researched at all within the context of autism. In their meta-analysis, Williams and colleagues (2023) concluded that in general, autistic participants appeared to show poorer objectively measured cardiac interoceptive accuracy but greater self-reported confidence in their cardiac interoceptive accuracy. However, they highlighted some pressing limitations of research so far, including modest effect sizes and a lack of replicability of findings when using better-validated measures.

In terms of studies specifically on adolescents, findings have been mixed, and there is a lack of consistency in the age groups defined, with many studies combining children and adolescents into one sample. In terms of cardiac interoceptive accuracy, one study reported no difference between autistic and non-autistic adolescents aged 11 to 17 years (Pickard et al., 2020). Contrastingly, Palser and colleagues (2018), Mash and colleagues (2017) and Nicholson and colleagues (2019) each found their blended sample of autistic children and adolescents (aged 6 to 18, 8 to 17, and 10 to 16 years, respectively) had worse cardiac interoceptive accuracy than non-autistics, whereas Schauder and colleagues found the opposite pattern (aged 8 to 17; Schauder et al., 2015). Moving our focus to interoceptive attention, Pickard and colleagues (2020) found that autistic adolescents (aged 11 to 17 years) paid more attention to interoceptive signals than non-autistic adolescents, whereas Butera and colleagues (2023) and Palser and colleagues (2018) found no difference between autistic and non-autistic under 18-year-old youths.

Some potential reasons for the mixed findings obtained so far may be the heterogeneity of autism, the broad and variable age ranges employed across studies, or the limitations of measures of interoception. In the case of the former, it may be that heightened

or diminished interoception is only observed in some sub-groups (such as those with co-occurring conditions such as alexithymia; Proff et al., 2022). Limitations of the most used measures of interoception have been comprehensively outlined in Chapter Two of this thesis.

Interoception (especially within the cardiac domain) has been suggested to be causal in anxiety. However, there is little consensus as to the exact mechanism by which interoception affects anxiety (see Chapter One). Mirroring this theoretical discrepancy, empirical findings have also been mixed (see Chapter Four).

Within autistic populations specifically, findings of an association between interoception and anxiety are more consistent, but they are limited in number. In autistic adults, the discrepancy between reduced interoceptive accuracy and higher interoceptive attention was predictive of state *and* trait anxiety (Garfinkel, Tiley, et al., 2016). Moreover, a randomized control trial has shown a reduction in anxiety experienced by autistic adults who underwent an intervention to improve cardiac interoceptive accuracy and reduce the discrepancy between interoceptive accuracy and attention (Quadt et al., 2021). In the single study involving adolescents, Palser and colleagues replicated Garfinkel's findings in their sample of autistic six- to 18-year-olds and found that interoceptive attention alone additionally predicted anxiety (Palser, Fotopoulou, et al., 2018).

Given the multi-dimensionality of interoception and the heterogeneity of autism, a qualitative approach may be especially beneficial in understanding the relationships between autism, anxiety and interoception. Rather than aiming to reduce experiences to overarching patterns, qualitative research instead explores individual experiences in depth. It allows for a rich, bottom-up approach to theory generation accounting for the fact that interoception may influence anxiety by multiple different, potentially bi-directional, routes.

Qualitative research focusing on experiences of anxiety in autistic individuals has not specifically focused on the role of interoception, although findings do suggest it may be

relevant. One study of eleven verbal, autistic young adults and ten caregivers and professionals found that autistic young adults highlighted the importance of somatic markers in their experiences of anxiety. Moreover, some caregivers reported that they had used heart rate monitors as an effective strategy for managing anxiety in autistic individuals, suggesting a potential role of cardiac interoception as a moderator of anxiety (Trembath et al., 2012). Similarly, in another study utilizing focus group interviews with autistic adults interpreted through thematic analysis, a "somatic" theme emerged, where numerous participants reported feeling bodily symptoms of anxiety which triggered anxious thoughts (Halim et al., 2018), which was echoed in Spain and colleague's interviews with six autistic men (2020). In autistic adolescents specifically, Jackson and colleagues (2022) concluded that somatic markers of anxiety (such as a rapid heartbeat and shortness of breath) were critical in experiences of anxiety and could trigger further panic in six autistic girls.

Of the few qualitative studies that specifically focus on interoception in autism, all participants were adults. One study used a content analysis approach on forum discussions of autistic adults describing interoceptive experiences (Trevisan, Parker, et al., 2021). They reported a general hyposensitivity to internal signals such as thirst and hunger, but incidences of hyper-attention to anxiety-related bodily sensations, like one's breathing rate or heartbeat. Another content analysis of autobiographical texts of ten autistic individuals found that whilst there were instances of both interoceptive hyper and hyposensitivity discussed, only *persistent* hyposensitivity was described (Elwin et al., 2012). The methodology of this study precluded the researchers from asking specific questions about the participants' experiences, and the reliance on autobiographies resulted in the sample having strong verbal proficiency, limiting representativeness. Moreover, the use of content analysis in both studies mean data is analysed in a more quantitative manner than with other qualitative methods, leaving scope for a more purely qualitative exploration of the topic.

To our knowledge, ours is the first study to explore the role of interoception in autistic adolescents' experiences of anxiety. The unique developmental characteristics of adolescence mean it is important to look at this age group specifically. We aimed to both characterise interoceptive experiences within our sample and explore how this related to their experiences of anxiety. This study addresses a pressing need for autism research that is constructed around the lived experiences of autistic individuals. Moreover, the use of qualitative methodology allows for a clearer understanding of the complexities of associations between autism, anxiety, and interoception that quantitative research has not been able to capture.

Methods

Study design

We employed one-on-one interviews given the sensitive topics to be discussed. We chose a reflexive thematic analysis approach, to allow for a view into the lived experiences of those interviewed, whilst acknowledging the role that the researcher plays in interpreting this. We pre-registered the study (https://osf.io/brmpu/?view_only=68b6a32636e94e208cdfca35cef485fa), which was approved by the Medical Sciences Division at the University of Oxford (reference: R85616/RE001) and carried out in accordance with the Declaration of Helsinki.

Positioning statement

At the time of interviewing, KA, the principal researcher, was 25 years old. She is a white, able-bodied, cis-gendered female with diagnosed generalised anxiety disorder. She experienced intense sensory hypersensitivity as a child and during therapy sessions the possibility of being autistic has been raised, although she has not received a formal diagnosis. At the time of this research, she was in the second year of her DPhil. KA began this study with no specific expectations, due to the mixed findings in the literature.

JS and MB are both autistic adolescents. Both describe themselves as mixed race. JS is male and MB is female. Both volunteered to be part of the PPI work conducted and then provided further feedback on the themes identified in the analysis. Neither had a previous research background.

GB is a professor of cognitive neuroscience with a wide range of research interests including autism and interoception.

PW is a female researcher and clinical psychologist by background, with expertise in qualitative methods. Her research focuses on developing an understanding of, and psychological therapies for, adolescents with anxiety disorders.

Participants and sampling

The eligibility criteria for this study were that all participants were between the ages of 14 to 17, spoke fluent English, had a diagnosis of autism, and self-reported experiencing significant anxiety that impacted their day-to-day lives.

Before conducting the interviews, we estimated our sample size would be between 15 to 20 participants. The primary source we used to guide this was the “information power” approach, proposed by Malterud and colleagues (2016). Considerations included the heterogeneity of autism and planned diversity of our sample, the fact that young, autistic participants may struggle with the communication required by the interview format, and the interviewer's (KA's) relative inexperience with conducting interviews. These factors could all necessitate a larger sample size. However, the interviewer’s lived experiences with anxiety and the steps taken to increase accessibility (e.g., the introductory call and individual adjustments made to the interview) could both enhance the amount of discussion in interviews, thus allowing for a slightly smaller sample size. We also considered sample sizes used in previous interview-based studies on autistic adolescents to inform our estimates (e.g.: Acker et al., 2018).

Plans for the sample size were updated as we collected, familiarized, and analysed our data. For example, after having coded the first four interviews, KA and PW discussed reducing the sample size. This was because during two of these interviews, participants elected to talk for over half an hour more than the assumed interview time of an hour, resulting in extremely detailed and in-depth data.

Rather than attempting to reach “data saturation,” which is incompatible with the underpinnings of reflexive thematic analysis (Braun & Clarke, 2021), we continued to conduct interviews until we reached “theoretical sufficiency” (defined as when the researcher has reached a sufficient or adequate depth of understanding to build a theory; Low, 2019). This resulted in a final sample of 13, at which point KA and PW both felt there was sufficient information power.

Participants’ background characteristics are reported in Table 6.1.

Table 6.1

Demographic information for interviewees

Participant ID	Age (years)	Ethnicity	Gender	Parent/carer 1 job NS-SEC analytical code	Parent/carer 2 job NS-SEC analytical code	School type	Region
Redduck	17	White	Female	1.1	5	Hybrid (mainstream private & home study)	South East
Bluecat	16	White	Female	6	4	Mainstream state	Scotland
Whiteseal	17	White	Female	2	8	Homeschooled	East
Greyfrog	16	White	Female	4	N/A	N/A*	South West
Pinkrat	17	White	Male	2	2	Homeschooled	South East
Greensnake	16	White	Female	3	5	Mainstream state	South East
Blackturtle	16	Black	Male	2	N/A	Mainstream state	London
Tealswan	17	White	Female	3	2	Further education college	London
Tanbat	16	Black	Male	5	8	Mainstream state	London
Orangedog	17	White	Male	2	1.2	Mainstream grammar (moving to specialist)	South West

Redtoad	17	White	Male	8	N/A	Didn’t disclose	London
Goldfox	15	White	Male	3	N/A	Mainstream	London
Yellowgoat	15	White	Female	Unspecified	5	Specialist	London

Note. * Employed. NS-SEC (National Statistics Socio-economic Classification) analytical codes: 1.1: Large employers and higher managerial and administrative occupations; 1.2: Higher professional occupations; 2: Lower managerial, administrative and professional occupations; 3: Intermediate occupations; 4: Small employers and own account workers; 5: Lower supervisory and technical occupations; 6: Semi-routine occupations; 7: Routine occupations; 8: Never worked and long-term unemployed.

Initially, any participant who indicated interest and fulfilled the eligibility criteria was interviewed. After eight interviews, we began to purposively sample using the demographic information from the screening questionnaires, selecting future participants on the grounds of ethnicity, gender, county and parental occupation, allowing for a diverse sample. Of particular importance was recruiting those from groups under-represented in autism research, such as non-male, non-white and participants from lower socio-economic backgrounds. Our participants were selected from an overall group of 19 who completed the screening questionnaire. Participants came from eight different counties within the UK.

Public and patient involvement (PPI)

Initial PPI work with two autistic adolescents and two parents of autistic adolescents (none of whom participated in the study) was conducted to inform the research design. See Tables 6.2 and 6.3 for PPI participant demographics.

Table 6.2

Demographic information for the adolescent PPI group

Participant ID	Age (years)	Ethnicity	Gender	Parent/carer 1 job	Parent/carer 2 job
cyanwolf	16	Mixed or multiple ethnicities	Female	1.2	8
mauvebird	17	Mixed or multiple ethnicities	Male	1.2	2

Note. NS-SEC analytical codes: 1.1: Large employers and higher managerial and administrative occupations; 1.2: Higher professional occupations; 2: Lower managerial, administrative and professional occupations; 3: Intermediate occupations; 4: Small employers and own account workers; 5: Lower supervisory and technical occupations; 6: Semi-routine occupations; 7: Routine occupations; 8: Never worked and long-term unemployed.

Table 6.3

Demographic information for the parent PPI group

Participant ID	Age of adolescent (years)	Ethnicity	Gender of adolescent	Job
pinkgoose	14	White	Female	3
redhorse	14	White	Male	8

Note. NS-SEC analytical codes: 1.1: Large employers and higher managerial and administrative occupations; 1.2: Higher professional occupations; 2: Lower managerial, administrative and professional occupations; 3: Intermediate occupations; 4: Small employers and own account workers; 5: Lower supervisory and technical occupations; 6: Semi-routine occupations; 7: Routine occupations; 8: Never worked and long-term unemployed.

Feedback was given on the initial research design, our topic guide, how to make the interview setting more accessible, and how to ensure a diverse sample was recruited. The following changes were made based on our PPI work:

- The wording of certain questions was changed to avoid appearing patronizing. "How would you describe anxiety to an alien who has never felt it before?" was changed to "How would you describe anxiety to someone who has never felt it before?"
- We used visual aids that participants could use to non-verbally communicate. These included a sheet with a pause button and one with the image of a toilet.
- Considering the comments of one parent of an autistic adolescent who had Irlen's syndrome, we chose to avoid the high contrast of black-on-white backgrounds (using a green colour instead).
- We decided to have a scheduled five-minute break in the middle of the interview.

The two autistic adolescents from the PPI work were included as co-authors and reviewed the analyses after the themes were constructed.

Topic guide

The topic guide (see Appendix 9) was developed by KA, who amended it after discussion with co-authors and the PPI group. Broadly, the questions covered participants' experiences of anxiety, their ability to detect emotions, how aware participants were of interoceptive signals, how signals were interpreted, and how this related to their experiences of anxiety.

Procedure

Participants were recruited through adverts on social media (TikTok, Facebook and Twitter), and via the email list of a specialist, UK-based tutoring agency. An information sheet was provided for parents, and an information PowerPoint for the adolescents, following which consent was gained. Participants then completed a short screening questionnaire (see Appendix 10), involving questions for both the adolescent and parent/guardian. This included demographic information to allow for purposive sampling and questions regarding the adolescent's autism diagnosis (month of diagnosis and the name of the organisation or professional that made this diagnosis) that were intended to provide validation of the diagnosis. Participants were informed that completing the questionnaire would not ensure they would be able to participate in the interviews, however, all potential participants who completed the screening questionnaire would be entered into a prize draw to win one of two £25 Amazon vouchers, regardless of whether they were selected for the interview.

The selected participants were informed by email and were able to choose whether they wanted to have their interview conducted remotely over Microsoft Teams, or in person in Oxford. If neither of these options were feasible, participants were given the option to interview over the phone. Four participants chose to do the interviews in person, and the rest chose to do it on Microsoft Teams. Participants were also able to request a short introductory Teams meeting with the interviewer ahead of the meeting, to allow for rapport building, minimize participant anxiety, and determine accessibility requirements (Robertson et al., 2018). Eight participants chose to have an introductory call.

In line with the adjustments used by Trembath and colleagues (2012), autistic adolescents were given the option to have a supportive individual with them during the interview. However, none chose to do this. Participants were also encouraged to express their views through writing or drawing (via Miro if online) if they felt more comfortable. This was

emphasised especially to one participant, who expressed during the introductory call that they sometimes experienced selective mutism. No participants used these tools to communicate in the interviews. Stim/sensory toys and tissues were placed on the table and participants were told they could use them at any time. Participants were told that they could take breaks at any time, come back to questions later, or avoid questions completely, and that they could do anything that made them comfortable, including avoiding eye contact. Interviews were completed between May to October 2023. Interviews took 45 minutes to an hour and a half, and participants who were interviewed were compensated £20 in Amazon vouchers. All interviews were voice recorded. See Figure 6.1 for a photograph of the interview set-up.

Figure 6.1

The interview set up



The topic guide was used as a flexible basis for the interview, but the question wording and order were contextual and participants were able to direct the discussion to topics they felt were especially important regarding their experiences of anxiety and interoception. At the end, participants were asked if there was anything else they would like to add. If something novel was brought up and suggested to be important in one interview, the interviewer adapted the interview schedule for the next interviews.

Transcription was conducted immediately following the interview and in line with the guidance given by Braun and Clarke (2013). Transcriptions were mostly verbatim, including all filler words, however, in instances of immediate repetitions of a word, this was replaced with ((stutters)). Identifying information was removed and indicated with [REMOVED]. Code names were given to any mentioned names in the interview. Transcription was performed manually, as part of data familiarization. Participants were then sent the transcripts as part of the validation process. They were given two weeks to make any minor changes. Two participants added further information that was included in the final transcripts and indicated by italicised font. The full, anonymized transcripts can be accessed at the Oxford University Research Archive (https://ora.ox.ac.uk/objects/uuid:9b95afae-c932-4a4b-8230-2e7a264dfc8e). Quotations included in this chapter use ellipses used to indicate some speech has been cut for brevity.

Analysis

A bottom-up, inductive approach was taken given that there is no singular theory guiding research.

Analysis was conducted using the following six steps as guidance (adapted from Braun & Clarke, 2013):

- Data familiarisation

- Line-by-line analysis and preliminary code generation
- Grouping of codes into initial themes
- Theme assessment and theme map construction
- Defining and naming the themes
- Writing and finalizing analysis

Analysis was performed flexibly and iteratively during data collection, moving back and forth between the different stages. All coding was done in NVivo 14.

Assuring validity

KA kept a research journal throughout the process, facilitating reflection on her role in the construction of themes and conclusions. This included field notes, analytical insights, and emotional aspects of the research. It also allowed for an iterative approach to methodology, where changes to the topic guide and interview style could be made considering what was successful or not in previous interviews.

Most of the analysis and development of the thematic structure was conducted by KA, in collaboration with PW. This was further refined through discussion with GB. Participants were active collaborators in the research process. Between stages three and four of the analysis, the codes and themes identified were sent to participants and co-authors/PPI contributors, and they could provide further feedback. Two participants gave suggested amendments. One highlighted the potential role of technology in helping with interoception and cited their own experiences with this, and the other provided a more detailed explanation as to why they felt hyperaware of some signals, and hypoaware of others. Both points were integrated into the themes. Both autistic co-authors felt the generated themes were appropriate and felt their own lived experiences were reflected in some of the identified themes.

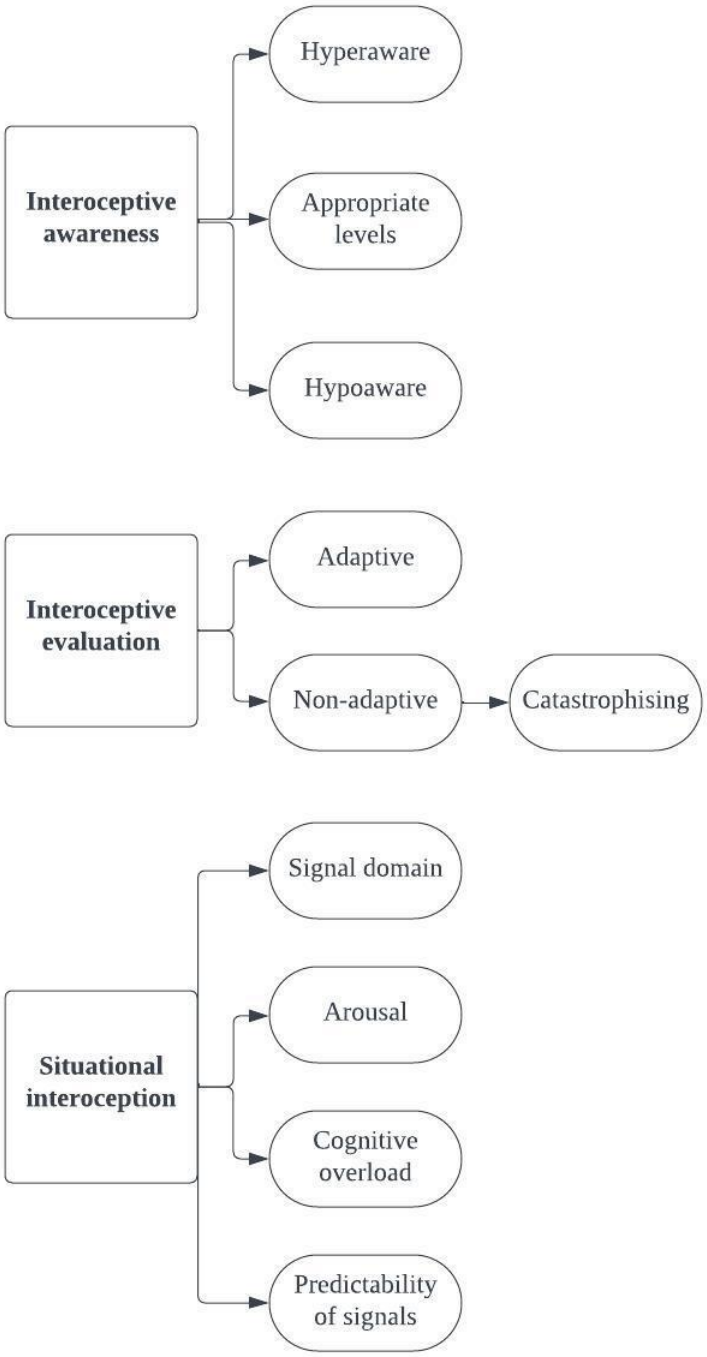
Results

Quotations are verbatim, with ellipses used to indicate some speech has been cut for brevity.

We constructed three main themes: interoceptive awareness, interoceptive evaluation, and situational interoception. There were a further ten sub-themes, as illustrated in Figure 6.2.

Figure 6.2

Themes and subthemes



Theme 1: Interoceptive awareness

Theme 1 reflects the degree to which participants reported being aware of interoceptive signals. The term "awareness" is used here to mirror the language used by participants. It can be defined as whether participants are able to detect signals that are informative about their bodily state in an adaptive manner. Adaptive awareness involves both an accurate detection of signals (interoceptive accuracy) and spending an appropriate proportion of the time being aware of these signals (interoceptive attention). Interoceptive awareness also encompasses the degree to which interoceptive signals take priority over other signals for attentional resources. Interestingly, participants appeared to refer more to how often, if at all, they noticed signals, rather than evaluating how accurate they felt they were when they could notice the signals.

Interoceptive awareness varied significantly across participants, spanning from constant hyperawareness to "typical" awareness to constant hypoawareness. It also varied within participants, with some participants experiencing both at different times.

Theme 1, sub-theme 1: Hyperaware.

"*I am almost constantly aware of my heart rate*" (Tealswan)

Participants who experienced hyperawareness of interoceptive signals described not being able to tune out from them. These participants were commonly hypersensitive to a range of sensory signals (e.g., bright lights and food smells), not just interoceptive signals. For them, awareness of interoceptive signals was frequent, largely involuntary, and problematic. Interoceptive signals were too frequently the object of consciousness, resulting in other information being drowned out.

"*I could be having a really nice quiet walk which to a lot of people would be super peaceful but because I'm hyperaware of how my heart is beating or how fast I'm breathing or you know every tiny bit of movement I get incredibly over-stimulated"* (Tealswan)

These participants described how interoceptive signals could become overwhelming and they felt unable to cope with them and anxious. This was especially challenging if there were other stressors present. One participant describes it as:

"*just another thing to keep track of when things are getting too much"* (Pinkrat)

Whereas another described:

"*it feels like you are going to pop you're going to explode*" (Orangedog, referring to the rare occasions he experienced interoceptive hyperawareness)

The overwhelming nature of some of these signals also resulted in both shutdowns and meltdowns in several participants.

"*Like normally the path will be I'll get anxious about something and then um (.) ... I'll go to break and it will be really noisy and that leads to me getting overwhelmed and then I'll go to another lesson and that's noisy and that escalates and it's just like too much and then I become more aware of everything like my heart my breathing and then I'll like shut down*." (Greensnake)

Shutdowns and meltdowns are behavioural responses characterised by a loss of control. Whilst a shutdown results in internal behaviours, like withdrawal and halted communication, a meltdown results in a more external response, including aggression and screaming. Interoception was identified as both a cause and a consequence of shutdowns and meltdowns, allowing for it to act as a maintenance factor in a vicious cycle. The consequences are discussed in more detail in Theme 3.

Theme 1, sub-theme 2: Appropriate levels. A smaller group of participants had levels of interoceptive awareness that were deemed appropriate, as their awareness allowed them to enact behaviours to maintain allostasis, whilst not inducing anxiety. These participants were able to tune in and consciously access information about internal signals if they focused on it, or if the signal was especially strong, but were not overly focused on these signals in a way that negatively impacted their day-to-day life.

> "*I wouldn't have noticed that if we haven't been talking about it obviously because now I'm focusing [I can]"* (Yellowgoat, referring to her heart rate)

Theme 1, sub-theme 3: Hypoawareness. When participants experienced hypoawareness of interoceptive signals, they were unable to detect internal signals, even if they tried, meaning they often struggled to meet their body's homeostatic needs. Frequently discussed examples were not knowing when to eat until they felt faint or not being aware of the need to use the bathroom until it was almost too late.

> "*I just cannot tell when I'm thirsty or not I have no idea*" (Goldfox)

"*Since I was little I've never naturally gone to the toilet (.) I've always had set times I'd go to the toilet and I would force it."* (Redduck)

"*I can go the whole day without eating anything at all and I'm feeling fine"* (Tanbat)

This difficulty recognizing one's needs led to a general sense of physical discomfort, as well as causing participants to worry that it affected their behaviour in a way that would lead to negative appraisal from others (for example, needing to run to use the bathroom at the last moment). In the case of interoceptive signals that indicate anxiety, such as heart rate, it led to participants not being able to recognise low-level anxiety and remove themselves from such situations, thus allowing anxiety to increase.

"*I sometimes only realise I'm getting overwhelmed when it gets quite extreme... that can lead to... a bigger meltdown than I would have had if I'd spotted it earlier*" (Greensnake)

Theme 2: Interoceptive Evaluation

The second theme of interoceptive evaluation, which refers to the interpretation of the interoceptive signals that participants were aware of, also varied across participants and appeared to exist along a spectrum.

Theme 2, sub-theme 1: Adaptive. Some participants could be described as having "adaptive" interoceptive evaluation, meaning that certain signals were suitably aversive to allow them to fulfil their role as motivational states (for example, feeling hungry would motivate them to eat), but the degree to which they found these signals aversive did not

negatively impact their day-to-day life or cause undue anxiety. These participants could recognise the diagnostic values of the signals.

"*Uh neutral or positive... they don't really uh fall on the negative side*" (Blackturtle)

One participant recognised that interoceptive signals can be evaluated as a positive rather than a negative in certain situations and that this can directly reduce anxiety. This suggests that sometimes these signals can sometimes be reassuring or a distraction.

"*sometimes these signals will in fact be the cause of anxiety or even help stimulate it. But there's other times where it will in fact be the thing that stops the anxiety or calms it."* (Redtoad)

Theme 2, sub-theme 2: Non-adaptive. Although theoretically, non-adaptive evaluations could encompass both those that were overly negative or positive, the latter was not identified from these interviews. Many participants described how they had overly negative evaluations of interoceptive signals, which impacted their everyday life and went beyond what could be seen as adaptive or proportional.

"*I absolutely detest feeling full up. Hate it hate it hate it hate it hate it"* (Orangedog)

"*I can't see a positive of feeling every time my heart beats to me all that's doing is making me more stressed out and over-stimulated*" (Tealswan)

Some participants directly linked this to their anxiety.

"*I'll find that the increase in heartbeat and the increase of the pulsating feeling definitely creates a feedback loop of anxiety*" (Redtoad)

Theme 2, sub-theme 3: Catastrophising. One subset of an overly negative, non-adaptive interoceptive evaluation was catastrophising, which several participants described. This is where one noticed a change in bodily signals that could indicate anxiety (e.g.: an increase in heart rate) and assumed this was a sign that something was wrong. This would further cause one to panic, resulting in a positive feedback loop which further increased one's heart rate and thus anxiety. For some, the interoceptive signals were interpreted as a sign of ill health, or something being wrong with their body.

"*It kind of feels like you're kind of dying*" (Whiteseal)

"*The physical response is exceptionally scary... It feels like you're dying*" (Orangedog)

"*You just find your heart like beating so fast you're (.) it makes me sometimes afraid like am I am I going to have like a heart attack*" (Tanbat)

One participant highlighted how a focus on these internal signals made her aware of how little control she had over her body. Lack of control was highlighted earlier in the interview as a key trigger for her anxiety.

"*When I get overwhelmed by the internal processes I get overwhelmed by the sense I have no control."* (Redduck)

A need for control is something often linked to the autistic experience (Kinnaird & Tchanturia, 2021). It is interesting to note here how it interacts with interoceptive experience in a complex way, resulting in an interpretation of interoceptive signals that may be unique to autism.

Theme 3: Situational Interoception

Theme 3 illustrates how both interoceptive awareness and evaluation were variable within as well as across participants.

"*It depends on the situation*" (Blackturtle, referring to how he feels on feeling his heart beating)

Theme 3, sub-theme 1: Signal domain. The signal domain refers to the type of interoceptive signal, for example, whether it is cardiac, respiratory or gastric. In the case of interoceptive awareness (Theme 1), some participants were hyperaware of certain signals, but hypoaware of others. In particular, the cardiac signal consistently appeared to be the most difficult to detect in day-to-day life, with few participants able to feel their heartbeat at rest.

Different interoceptive signals also seemed differentially aversive. Whilst signals typically associated with the symptoms of anxiety, such as elevated heart rate and faster breathing, were more commonly cited as anxiety-inducing, interoceptive signals relating to more general bodily needs (such as thirst or needing the bathroom) were more commonly seen as neutral. Interestingly, signals relating to appetite existed between these, with a

minority of participants finding these intensely aversive, but the majority finding them neutral.

Theme 3, sub-theme 2: Arousal. Almost unanimously, participants reported being more aware of their heartbeat when it was raised due to anxiety or exercise.

> "*After exercising is probably when it feels the most or when you know I'm stressed for whatever reason*" (Pinkrat)

> "*When I'm just sitting and relaxed it's um it's sometimes hard to like feel [your heartbeat] but when you're anxious or when you've done something like an activity or when you're feeling angry at someone [I can feel it]"* (Tanbat)

> "*When I do feel really anxious or have really high emotions I can feel it really really strongly*" (Greyfrog)

In some cases, this heightened awareness with anxiety was general to all interoceptive signals, not just heartbeat.

> "*In the potent situations I can feel absolutely everything in my body all at once*." (Orangedog)

This highlights a bidirectional link between anxiety and interoception, thus allowing for the positive feedback loops mentioned in the previous section.

Some expressed the idea that it was the change in state that triggered their awareness, rather than simply higher than normal levels.

"I can kind of feel when it changes" (Goldfox, referring to heart rate)

Even within participants, there were differences in how interoceptive awareness manifested when anxious. One participant (Tealswan) linked this to either shutdowns or meltdowns, both of which she experienced, although she was unsure about whether there were different triggers to these two scenarios.

"*I'm either super super present where like everything is a million degrees... or the like input that I'm being given is just nothing at all. Or like I'm completely dissociated I cannot really feel anything*" (Tealswan)

For some participants, shutdowns led to feeling completely detached from interoceptive signals, as though they could feel nothing at all.

"*I entirely shut down... zombie state... it's almost like being in someone else's body if that makes sense*" (Tealswan)

"*I feel completely... not in my body*" (Whiteseal)

By contrast, for those who experienced meltdowns, interoceptive signals (in much the same manner as other sensory signals) became increasingly more intense and unbearable.

"*I feel too much going on inside my body*" (Greensnake, when asked how meltdowns feel)

This further underscores the immense complexity of experiences of interoception in relation to anxiety.

Arousal was also likely to influence one's interoceptive evaluation (Theme 2). If participants were already feeling overwhelmed, they were more likely to interpret interoceptive signals negatively.

> "*I think it depends on what environment I'm in because there's not really anything else that's overwhelming me at the moment I don't mind but if I was in somewhere where there was someone talking to me someone else having a conversation next to me um some background noise and then I started feeling my heartbeat as well I think that's too many things for me to focus and process... that would cause me more anxiety*" (Greensnake)

One participant avoided exercise because any feeling of her heart increasing at all made her panic, even if she knew there was a non-threatening explanation for it. This highlights an indiscriminate negative evaluation of interoceptive signals, regardless of the situation.

> "*Like I don't really like running and stuff because like when I run... my heart beats fast and that like reminds me of being anxious*" (Bluecat)

Theme 3, sub-theme 3: Cognitive overload. Some participants found it harder to notice interoceptive sensations (Theme 1) when focusing on numerous other things already.

"*I think it's also like based on how overwhelmed I feel sometimes cause like if I feel really overwhelmed then I just won't feel like eating*" (Bluecat, when talking about her appetite)

Masking (defined as the automatic or voluntary suppression of natural autistic responses; Pearson & Rose, 2021) was another factor that affected some participants' ability to detect internal signals. Some participants felt more detached from themselves and less attuned to internal sensations when masking, as their attention was directed outwards rather than inwards. Relatedly, focusing on observing and imitating the behaviour of others was cognitively demanding and left little room to focus on interoceptive signals.

"*My ability to read myself lessens quite a bit because I'm so focused on other people and their actions and their experiences and I think that's why I ended up having so many um shutdowns in school... I wasn't reading the signs that I was getting overwhelmed."* (Greensnake)

One participant describes how they only began to experience hunger properly when they were home-schooled and could begin to focus on themselves rather than masking to fit into their external environment.

"*When I was masking at school like I wouldn't really feel hungry or thirsty and stuff and like I often wouldn't even go to... the toilet until I literally got home got through the front door and then like kind of unmasked more*" (Whiteseal)

"*I noticed since I started homeschooling... I'm doing my own thing that then I seem to kind of realise oh I need to go to the toilet I need to drink because I'm thirsty*" (Whiteseal)

For other participants who masked, masking resulted in a hyperawareness of certain, anxiety-related interoceptive signals due to the stress of vigilantly monitoring their behaviour.

"*heartbeat and breathing… might be heightened… in times where I'm like in social situations or masking because (.) I would be (.) like (.) using my body would kind of be using that as an indicator*" (Redtoad)

As well as masking, some participants experienced a kind of cognitive overload that came with extreme emotions that made it difficult to focus on bodily signals.

"*Trying to listen to your body sometimes does become very difficult if you are feeling very strong emotions"* (Greyfrog)

Theme 3, sub-theme 4: Predictability of signals. For one participant, the ability to detect signals (Theme 1) and how anxiety-inducing they were (Theme 2) depended on their predictability.

"*Because hunger and thirst or needing the toilet do not have a set order or schedule I often misread them as there is no pattern or association that can be made... But because heart rate and breathing are consistently there when I do tune into them it is*

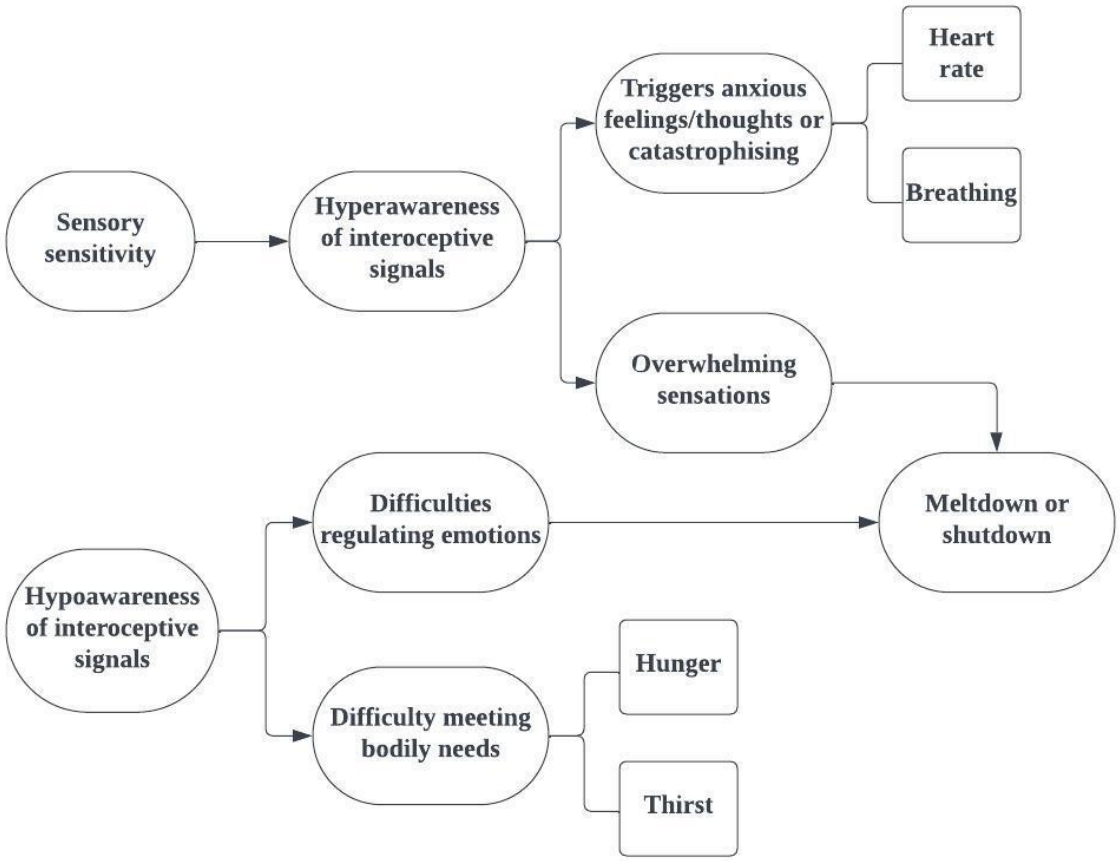
not confusing for me to interpret but can be overwhelming due to the sensory aspect of this" (Redduck)

This may be reflective of an intolerance of uncertainty that is known to be associated with autism, whereby predictability and certainty are preferred, whilst unpredictability is anxiety inducing (Jenkinson et al., 2020). Unpredictability prevents one from being able to plan for something, and allows for the potential of a shock, which can trigger further anxiety.

Overall, these themes have been combined into two potential mechanisms by which interoception may lead to anxiety in autistic adolescents. These are presented in Figure 6.3.

Figure 6.3

Pathways by which interoception may affect anxiety



Note. Two of the potential routes by which interoception appeared to influence participants' anxiety.

Discussion

We found that in our sample of thirteen autistic adolescents who experience significant anxiety, interoception played a role in the majority's experiences of anxiety, although the route by which interoception affected, or was affected by, anxiety varied both within and between participants. One route by which interoception resulted in anxiety was when participants felt hypoaware of certain signals, meaning they missed signs that they were becoming anxious and therefore missed opportunities to engage in positive emotion

regulation strategies at lower levels of distress. Hypoawareness would also result in participants not attending to their bodily needs sufficiently, which could result in general discomfort and fatigue that might further exacerbate anxiety. Another route was through interoceptive hyperawareness, which led to interoceptive signals becoming overwhelming. When participants were aware of signals, overly negative interpretations of these signals also could result in anxiety, especially when participants catastrophised about these signals and perceived them as a sign they were in danger. A key theme across these routes was the variability of interoceptive experiences, both within and between participants. Indeed, the context one was in (including factors such as cognitive overload, masking, and arousal) and the specific signals in question played a huge role in how one experienced interoceptive signals and thus the potential for it to cause anxiety.

In terms of how our findings relate to the extant literature, our overall finding of an association between anxiety and interoception echoes the suggestion made by Spain and colleagues when looking specifically at social anxiety in autistic adults (Spain et al., 2020). However, it is important to note that this association was not endorsed by all of our participants, with a minority expressing that interoception played no role in their experiences of anxiety.

There has not been much research on the role that hypoawareness might play in experiences of anxiety. Our finding that hyperawareness can cause anxiety is consistent with models proposed within the context of more general sensory sensitivity in autism (Green & Ben-Sasson, 2010; Hazen et al., 2014), as illustrated in Acker and colleagues' (2018) interpretative phenomenological analysis of interviews with thirteen male autistic adolescents. Our findings are consistent with the notion that sensory sensitivities for both internal and external information can occur similarly in autism, leading to similar relationships with anxiety.

Moreover, the importance of overly negative interoceptive evaluation in triggering anxiety identified here aligns with initial quantitative findings in neurotypical adults (Paulus, 2013; S. Taylor et al., 1991), as well as qualitative reports from autistic adults (Halim et al., 2018; Trembath et al., 2012) and autistic adolescent girls (Jackson et al., 2022). One participant in Halim and colleague's study describes how when they feel bodily signals associated with anxiety, their brain starts to look for something to be anxious about. This shows the role of higher-level cognitions and interpretations of interoceptive signals in making participants feel anxious, despite an initial lack of anxious emotions. Similarly, one participant in Trembath and colleagues' study described how they become anxious whenever they feel their heart beating fast, and in Jackson's study, one girl describes how she feels like she is suffocating. The latter is analogous to the catastrophising route discussed by participants in our study, whereby they take these interoceptive signals as a sign that they are in danger. Trevisan, Parker and colleagues (2021) directly explored interoception in their content analysis of forum posts by autistic adults. As in the current study, they found that some participants discussed being constantly worried due to an inability to distinguish benign bodily sensations from serious health issues.

The variability and situational nature of interoception that we observed in our study is consistent with Trevisan, Parker and colleagues' finding of both hypo- and hypersensitivity specifically to interoceptive signals in autistic adults (2021). Although they found that hyposensitivity was more common, hypersensitivity was mentioned by participants in the context of anxiety-related interoceptive signals such as heart rate and breathing. Whilst we also found that our sample of autistic adolescents responded slightly differently to anxiety and non-anxiety-related interoceptive signals, the way this linked to either hypo- or hypersensitivity was not as clear as in Trevisan's study. This could be due to differences in the age groups involved in the studies.

The finding that masking can affect interoception is consistent with previous qualitative work which suggests that sensory suppression is a key part of masking in autistic adults (Miller et al., 2021). This pathway from masking to either hypo or hyperawareness and therefore anxiety may be one of many potential reasons that masking is associated with poorer mental health (Cook et al., 2021).

Our findings could be understood with reference to monotropism (Murray et al., 2005). Monotropism describes a claimed autistic tendency towards intense and targeted attention ("hyperfocus"). Such hyperfocus may explain hyperawareness of interoceptive information in instances where interoceptive signals are the focus of attention, and hypoawareness when they are not.

The findings of this study have a variety of wider applications. Firstly, they highlight the importance of further research into interoceptive evaluation, which has been under-researched and typically measured using non-specific and unvalidated questionnaires, substantiating the usefulness of the Negative Interoceptive Evaluation Scale presented in Chapter Four. This measure should be further validated, to allow larger-scale quantitative investigations to be triangulated against qualitative findings. There is also potential for interventions for interoceptive evaluation, perhaps by using cognitive restructuring to encourage patients to focus on interpreting bodily signals neutrally and teaching about alternative, non-anxiety-related explanations for changes in certain signals such as heart rate and breathing. This is currently a key part of Cognitive Therapy, which is efficacious in neurotypical adolescents with social anxiety (Leigh & Clark, 2023) but has not yet been evaluated within autistic populations.

Our findings also suggest that existing technologies could be used by autistic adolescents with anxiety to help manage their symptoms. Smartwatches that detect bodily signals may be useful for those experiencing interoceptive hypoawareness who find

recognising anxiety difficult. In communications with one participant after the interview, they discussed how beneficial the interview had been in terms of encouraging them to learn to recognize certain bodily signals better. To do this, they had begun using a wearable fitness device to track things like their heart rate, which they said had positively impacted their mental health. However, given the variability in experiences of interoception noted here, it is vital to make clear that interventions that may work to reduce anxiety for some individuals may be anxiogenic for others. For example, whilst wearables may be beneficial for those who are hypoaware, they are likely to exacerbate issues for anyone who is hyperaware of signals already and liable to evaluate these signals in an overly negative way. For these individuals, psychoeducation about bodily signals and how changes do not always indicate a problem, as discussed above, may be more useful. A co-design approach should be a priority for researchers looking to develop interoception-based interventions, given the nuances of the relationship between anxiety and interoception.

Our study is the first study to focus on how interoceptive experiences in autistic adolescents relate to anxiety. Autistic young people were included in the design of the study and analysis, and our qualitative data allowed for a rich exploration of the lived experiences of our participants. Interpretation was also influenced by both KA's lived experiences of an anxiety disorder, as well as PW's experience as a clinical psychologist, both of which allowed for an understanding and rapport to be built in interviews that facilitated a more open discussion with participants. Meanwhile, the range of participant responses and variability in reported experiences prevented KA from anticipating certain responses, allowing for an open approach to interpretation. Despite these strengths, there are some limitations to address. Despite efforts to sample purposively, some groups (those of Asian heritage, non-binary participants, and those with severe intellectual disability) remained unrepresented due to time constraints. However, we did manage to include several often-under-represented groups.

Over half of our participants were autistic girls, we had two Black British participants and participants of a range of socioeconomic statuses (indicated by parental occupation and region of the UK). Additionally, participants attended a range of different types of schools, some of which indicated a higher level of support needs (several participants were either home-schooled or in specialist schooling). One participant experienced selective mutism.

The results of this qualitative study present a more complex picture than that typically explored in quantitative research of interoception, which may explain the inconsistency in quantitative findings of studies exploring the associations between interoception and anxiety and autism (Williams et al., 2023; Chapter Five). Given that autistic participants may be hyper-, hypo-, or typically aware of interoceptive signals, and that this can vary within individuals according to contextual factors, the likelihood of either null or conflicting findings due to different experimental settings and sampling variance is high. Despite the variation in the possible routes, we found overall that interoception related to experiences of anxiety in many of our participants, highlighting a potential route for intervention. It would be beneficial to triangulate these findings with future large-scale questionnaire studies, allowing us to better ascertain the generalisability of these results.

Chapter Seven: The associations between self-reported dimensions of interoception and anxiety and autism

Kiera Adams designed this project in collaboration with Geoffrey Bird, who acted as Kiera's supervisor for this project. Kiera Adams collected data, ran analyses and led interpretation and visualization of results. Kiera Adams wrote up the chapter, with feedback from Geoffrey Bird.

Abstract

Interoception has been linked theoretically to both autism and anxiety. However, we lack an understanding of which dimensions of interoception (if any) are associated with autism and anxiety, and the causal direction of any associations. This pre-registered study aims to further our understanding of whether self-reported interoceptive accuracy, attention and evaluation relate to autism and anxiety, at both the trait and diagnostic level. Questionnaire responses from 519 participants, including 232 autistic participants and 149 participants with an anxiety disorder, were analysed. Our findings show that higher trait anxiety is associated with greater self-reported interoceptive attention and more self-reported negative interoceptive evaluation, and unrelated to self-reported interoceptive accuracy. Anxiety disorders were also associated with greater self-reported interoceptive attention. Higher autistic traits were associated with higher self-reported interoceptive attention, lower self-reported interoceptive accuracy, and higher self-reported negative interoceptive evaluation. Autism diagnoses were associated with more self-reported negative interoceptive evaluation. Autism did not interact with any dimension of interoception to increase anxiety (at either the trait or diagnosis level).

Introduction

As outlined in previous chapters of this thesis, there is a large corpus of theoretical work linking both anxiety and autism to atypical interoception, however, the exact nature of these associations is still not well understood. This thesis has focused on three dimensions of interoception which have been linked to mental health and neurodevelopmental conditions: interoceptive accuracy, interoceptive attention, and interoceptive evaluation. At present, we know very little about the relative role of each of these dimensions in anxiety and autism. Part of this lack of understanding is due to inconsistent empirical findings and the use of invalid measures of interoception, or measures which conflate multiple dimensions (see Chapter Two). Moreover, it is difficult to compare, for example, the contribution of accuracy versus attention when the former is almost always objectively measured, and the latter subjectively measured. The aim of this study is to address this by exploring the role of self-reported interoceptive accuracy, attention, and evaluation in both diagnosed autism and anxiety disorders, as well as autistic and anxious traits.

Autism and interoception

In terms of prior work using self-report measures to understand interoception in autism, most studies have used questionnaires that are most akin to interoceptive attention, although they are more often described as capturing interoceptive "awareness". Results of one meta-analysis first discussed in Chapter One indicate no difference in self-reported interoceptive attention (as measured by questionnaires such as the Noticing Scale of the Multidimensional Assessment of Interoceptive Awareness (MAIA; Mehling et al., 2012) and the Awareness subscale of the Porge's Body Perception Questionnaire (BPQ, Porges, 1993)) between autistic and non-autistic participants. However, the analysis indicated high levels of heterogeneity across studies (Williams et al., 2023).

Focusing specifically on the MAIA, Yang and colleagues found that when self-consciousness, empathy and (critically) alexithymia were controlled for, there was a negative association between MAIA scores and autistic traits in their general population sample (Yang, Hu et al., 2022). However, this is difficult to interpret given that authors used the total MAIA scores as an indication of interoceptive "sensibility," even though the MAIA was designed to be used as a series of separate subscale scores. The use of a total score thus conflates several different constructs, including attention to interoceptive signals, as well as worrying about interoceptive signals (which would be considered as attention and evaluation within this study). Another study first identified three clusters within the MAIA using a multidimensional scaling analysis: awareness, active and reactive strategies (both of which were lower in autistic compared to non-autistic adults) and attention regulation (which was no different across autistic and non-autistic adults) (Mul et al., 2018). Of these, the awareness and attention regulation constructs most closely relate to interoceptive attention as defined within this thesis, whilst the active and reactive strategies factor fits less clearly within our framework.

With respect to studies using the BPQ, one demonstrated higher scores in a small sample of autistic adults compared to non-autistic adults (Garfinkel, Tiley, et al., 2016). Confusingly, however, the authors defined this measure both as capturing a "sensitivity to internal sensations" and an "increased belief in interoceptive aptitude," which are slightly different, with the first pertaining more to attention, and the second to accuracy. An assessment of the questions themselves would suggest the questionnaire appears to measure attention rather than the accuracy. However, additional problems arise because some of the signals discussed in this questionnaire may be considered exteroceptive (e.g.: sweat on the forehead), and most

focus on signals that are inherently negative, such as pain, limiting generalisability across all interoceptive signals.

Results of the qualitative study presented in Chapter Six of this thesis present a similarly complex picture of subjective interoceptive attention in autism. A key finding was that this was extremely heterogeneous across autistic adolescents, with some participants discussing how they almost always noticed interoceptive sensations, and others describing never being able to notice interoceptive sensations. This sheds some light on why previous findings that rely on group level differences may have been mixed. However, the design of the study presented in Chapter Six precluded any comparison between autistic and non-autistic participants and findings may not be applicable to autistic adults.

A theoretical paper by Trevisan and colleagues (Trevisan, Mehling, et al., 2021) aimed to address these mixed findings, by suggesting that autistic individuals may show lower levels of adaptive interoceptive attention to neutral signals, but higher levels of maladaptive interoceptive attention, or increased anxiety-driven somatization. They interpret the results of previous studies as supporting this idea, highlighting that a study using the BPQ, which captures the maladaptive form of interoceptive attention, found higher scores in autistic when compared to non-autistic participants, and that BPQ scores correlate with anxiety (Garfinkel, Tiley et al., 2016). On the other hand, the limited available studies that explore adaptive interoceptive attention (as measured by the MAIA and Interoception Sensory Questionnaire) and autism find no relationship (Fiene et al., 2018; Fiene & Brownlow, 2015; Mul et al., 2018). Trevisan and colleagues substantiate their claims with findings from their own qualitative study, where autistic participants reported experiences in line with the idea that autistic individuals show lower adaptive but higher maladaptive attention to interoceptive signals (Trevisan, Parker, et al., 2021). In terms of the framework employed throughout this thesis, Trevisan and colleague's theory conflates interoceptive

attention and evaluation. Their theory could either mean that autistic individuals would be expected to show lower interoceptive attention and higher negative interoceptive evaluation than non-autistic participants, or that there will be no overall difference between total interoceptive attention scores between autistic and non-autistic participants, but that when split into anxiety and non-anxiety-related signals, attention will be higher in the former and lower in the latter in autistic compared to non-autistic groups.

Few studies have considered self-reported interoceptive accuracy in autism. Some studies have used the Awareness Subscale of the Body Awareness Questionnaire (BAQ; Shields et al., 1989), which measures one's ability to identify inner sensations and detect non-emotional bodily processes. This can be seen as close to our construct of interoceptive accuracy, however, it also involves aspects of attention. Fiene and Brownlow (2015) found that autistic individuals scored significantly lower on this measure than non-autistic individuals with a large effect size. Contrastingly, other studies have found no difference in the awareness subscale of the BAQ between autistic and non-autistic adults (Jeon et al., 2024) and youth (Butera et al., 2023).

Interoceptive accuracy can also be assessed by taking confidence measures of task performance during interoceptive accuracy tasks. Meta-analytic results indicate that there is higher self-reported confidence in cardiac interoceptive accuracy task performance in autistic compared to non-autistic participants. However, results are reliant on only five effect sizes across four studies, two of which were based on adults, with the remaining three conducted on children (Williams et al., 2023). As results were restricted to the cardiac domain, it is unknown how cross-domain subjective interoceptive accuracy may be impacted by autism.

Chapter Six of this thesis suggested that interoception may show similar patterns to exteroceptive sensory processing in autism (Proff et al., 2022), in that there is the potential for both heightened and attenuated attention and accuracy, dependent on the individual and

context. This may explain some of the mixed results from studies exploring the relationship between self-reported interoceptive attention and autism. Based on this, one might expect to see a slightly different distribution in the interoceptive attention and accuracy scores in autistic when compared to non-autistic individuals. This distribution may be closer to bimodal in autistic individuals, reflecting a tendency for autistic individuals to show either hyper or hypo interoceptive accuracy and attention.

There is also a lack of empirical work considering self-reported interoceptive evaluation in autistic compared to non-autistic individuals. One study found no significant association between the Physical subscale of the Anxiety Sensitivity Index (ASI) and autistic traits, as measured by the autism quotient (AQ) (Baiano et al., 2022). In contrast, our qualitative study (Chapter Six) indicated that many of our autistic adolescent participants experienced overly negative interoceptive evaluation. Given differences in interoceptive attention between autistic and non-autistic participants vary according to whether negative or neutral signals were considered (Trevisan, Mehling, et al., 2021), interoceptive evaluation in autism may warrant more attention, as it may be that the potential for catastrophically negative interpretations of the negative signals underlies this difference.

Anxiety and interoception

A recent meta-analysis explored the relationship between trait anxiety and self-reported dimensions of interoception. The results showed that anxiety was associated with increased negative evaluations of, and a greater degree of attention to, bodily signals, as well as reporting of more frequent bodily sensations, which could reflect greater attention (Clemente et al., 2024). However, the authors expressed concern regarding the overlap between anxiety and interoception measures, which may potentially overinflate associations. For instance, several measures of anxiety (for example, the Anxiety Symptoms Questionnaire; Baker et al., 2019) also include questions that pertain to the experience of,

noticing of, or evaluation of interoceptive signals (e.g., “How frequently have you experienced Shortness of Breath, Chest Tightness or Pain, Pounding/Skipping/Racing Heartbeat in the last week?”). A discussion of specific studies that have found associations between self-reported interoceptive evaluation and anxiety can be found in Chapter Four of this thesis. As well as supporting the meta-analytic findings that trait anxiety is associated with negative interoceptive evaluation, these findings also suggest that anxiety disorders are associated with negative interoceptive evaluation. However, none of these studies have used measures specifically designed to assess negative interoceptive evaluation across multiple domains.

Although they did not specifically report on differences in subjective interoceptive accuracy, Garfinkel and colleagues reported that trait anxiety was associated with an increased propensity to overestimate ability in heartbeat discrimination tasks (Garfinkel, Tiley, et al., 2016). Contrastingly, Harrison and colleagues (2021) found that when participants were split into those with low and moderate anxiety, those with moderate anxiety reported lower perceived accuracy in a task related to respiratory interoception than those with low anxiety. The role of interoceptive accuracy in anxiety is discussed in more depth in Chapter Four.

Anxiety, autism and interoception

At the intersection of both these areas of research lies the possibility of more complex relationships between anxiety, autism and interoception. Some have suggested that interoception may be one contributing factor to the overlap between autism and anxiety (Garfinkel, Tiley, et al., 2016), suggesting that a certain profile of interoceptive traits that may trigger anxiety may be more common in autistic rather than non-autistic individuals (Trevisan, Mehling, et al., 2021). It could also be that interoceptive atypicality has more of an impact on anxiety in autistic rather than non-autistic individuals. Although purely speculative

at this stage, one might theorise that, for example, the dichotomous thinking commonly found in autism (Stark et al., 2021) may result in negative interoceptive evaluation having more of a detrimental effect on anxiety than in non-autistic individuals when certain interoceptive signals are inflexibly interpreted as a sign that something is wrong. Chapter Six of this thesis illustrated that for several autistic adolescents who experience significant anxiety in their day-to-day life, interoception directly played a role in their experiences of anxiety, especially regarding negative interoceptive evaluation. This study endeavours to extend these findings further, assessing relations between autism, anxiety and interoception in a large sample of adults, allowing for comparisons across different diagnostic groups.

Covariates

There are numerous potential confounding factors which must be considered when assessing associations between interoception and autism and anxiety. When comparing groups, age should be considered, given speculation that interoception varies across different developmental stages (Murphy et al., 2017), although there is currently a lack of research on this. Another potential confounder is gender. Objectively measured interoceptive accuracy appears to be influenced by gender (Prentice & Murphy, 2022), and it appears there is also a difference in self-reported interoceptive attention (Grabauskaitė et al., 2017) and negative interoceptive evaluation (Stewart et al., 1997) according to gender.

Both matrix and verbal reasoning, which can be seen as proxies for intelligence, may affect one's understanding of questionnaires and therefore affect questionnaire responses. Moreover, it is common practice when comparing autistic and non-autistic groups to control for the effects of IQ, as IQ is typically lower in autistic groups and could therefore act as a confounding variable (Wolff et al., 2022).

Alexithymia is known to be associated with autism (Kinnaird et al., 2019), anxiety (Karukivi et al., 2010) and interoception (Brewer et al., 2016) and may act as a confounder if

not controlled for (see Chapter One). Similarly, other mental health conditions are more likely to occur in both autistic participants and those with anxiety disorders, and may influence interoception (Brewer et al., 2021).

The current study

This study addresses gaps in the current literature by employing three self-report measures which measure distinct interoceptive dimensions (accuracy, attention and evaluation), allowing for the effects of each of these to be explored individually. This study also aimed to address limitations of prior research, by using a measure of autistic traits that is better able to capture presentation in women than the commonly used AQ (Baron-Cohen et al., 2001), thus increasing validity across genders. Moreover, analyses in this study controlled for relevant covariates, including alexithymia.

Hypotheses

Our hypotheses were as follows:

1. Trait level anxiety across participants will be positively associated with negative interoceptive evaluation and interoceptive attention, but not interoceptive accuracy, after accounting for relevant covariates.
2. Autistic traits will be positively associated with negative interoceptive evaluation, but not interoceptive accuracy or attention, after accounting for relevant covariates.
3. Participants with diagnosed anxiety disorders will show a) more negative interoceptive evaluation, b) a greater degree of interoceptive attention, and c) comparable self-reported interoceptive accuracy when compared to those without diagnosed anxiety disorders, after accounting for relevant covariates.
4. Autistic participants will show a) more negative interoceptive evaluation, b) comparable interoceptive attention, and c) comparable self-reported interoceptive

accuracy when compared to non-autistic participants, after accounting for relevant covariates.

5. Autistic participants will have more extreme high and low scores on both interoceptive accuracy and attention than non-autistic participants.
6. Negative interoceptive evaluation will moderate the association between autistic traits and trait anxiety, and autism and anxiety disorder diagnoses. More negative interoceptive evaluation will lead to a stronger association. There are no specific predictions as to whether interoceptive accuracy or attention will moderate the association between autistic traits and trait anxiety, or autism and anxiety disorder diagnoses.

Methods

Participants

This study was part of a larger project which aimed to explore the co-occurrence between autism and mental health conditions including depression, eating disorders and anxiety. The inclusion criteria for this study were that all participants were 18 years of age or older and spoke English fluently. There were no exclusion criteria. Participants were recruited through social media, existing laboratory databases, adverts on relevant websites (Autistica, BEAT, and Bodywhys), and on the Oxford Psychology Research portal. The study was also included on the Oxford University Research Participation Scheme, whereby psychology students can receive course credits in exchange for participation. All participants were entered into a prize draw, where ten participants were randomly selected to win £25 Amazon vouchers. Ethics approval was granted by the Oxford University Medical Science Division's Ethics Committee (reference: R86431/RE003).

Our expected sample size was calculated using G-Power, which suggested that a total sample size of 128 (64 in each group) would be sufficient to detect a medium effect size (0.25) using an analysis of covariance (ANCOVA) with 7 covariates. This sample size was more than sufficient for the planned linear multiple regressions including interaction terms, for which a total sample size of 77 was recommended to detect an effect size of 0.15 when including 6 covariates, as well as 3 predictors. We planned to recruit more participants than this, to account for the fact that some of the participants would not meet the criteria for group membership (for example, participants may self-report an autism diagnosis but not pass validation checks). 519 participants were recruited for this study overall, including 232 with validated autism diagnoses and 149 with a validated diagnosis of an anxiety disorder (see the Procedure section for how diagnoses were validated). 235 participants met the criteria to be included in the "no anxiety disorder" group, and 143 participants met the criteria to be included in the "not autistic" group.

Materials

This questionnaire was part of a wider study. For brevity, only the measures used within this chapter are described, all of which are self-report measures.

Toronto Alexithymia Scale (TAS-20): 20 item scale in which participants report how much they agree with a statement pertaining to alexithymic symptoms on a five-point Likert scale (Bagby et al., 1994). This questionnaire is reported to have acceptable internal consistency (Cronbach's alpha = .81) and good test re-test reliability at three weeks (r = .77) (Bagby et al., 1994).

Depression, Anxiety, Stress Scale (DASS-21) Anxiety subscale: Measures symptoms of anxiety (Lovibond & Lovibond, 1995). Participants rate the extent to which they have experienced seven negative emotional symptoms over the past week, on a four-point

severity/frequency scale. This measure was chosen due to its divergence from depression (see Chapter Five). Psychometric properties are also reported in Chapter Five.

The Comprehensive Autistic Trait Inventory (CATI): 42 items pertaining to the traits of autism, with seven items associated with each of the subscales: Social Interactions, Sensory Sensitivity, Repetitive Behaviours, Communication Difficulty, Cognitive Rigidity, and Social Camouflaging. Each item is rated on a five-point Likert scale. The CATI is suggested to be better for capturing autism in females than the AQ (English et al., 2021). It has a high internal consistency (Cronbach's alpha = .95) and a better predictive ability for classifying autism than the AQ (Youden's Index = .62 vs .56 to .59) (English et al., 2021).

Autism Quotient (AQ-Short): This includes 28 statements pertaining to symptoms of autism, and participants rate agreement on a 4-point Likert scale. Internal consistency is acceptable (Cronbach's alpha = .78) (Hoekstra et al., 2011). This measure was only used to ensure participants with high autistic traits were not included in the non-autistic control group.

The Interoceptive Accuracy Scale (IAS), Interoceptive Attention Scale (IATS), and Negative Interoceptive Evaluation Scale (NIES): scored as described in Chapter Four.

The International Cognitive Ability Resource (ICAR) matrix reasoning: 11 matrix reasoning items similar to those used in Raven's Progressive Matrices. The stimuli are three-by-three arrays of geometric shapes with one of the nine shapes missing. Participants are instructed to identify which of six geometric shapes presented as response choices will best complete the stimuli (Condon & Revelle, 2014).

ICAR verbal reasoning: 16 Verbal Reasoning items include a variety of logic, vocabulary and general knowledge questions (Condon & Revelle, 2014).There is limited psychometric information about both ICAR resources, however, one study finds that both the

ICAR matrix and verbal reasoning questionnaires are comparable to the current gold standard of measuring cognitive abilities (the Wechsler Adult Intelligence Scale—Fourth Edition; Wechsler, 2012) (Young, 2020).

Procedure

Recruitment materials included either a link or QR code that directed participants to the questionnaire, which was hosted on Qualtrics and could be completed on either smartphones or laptops. First, participants were shown an information sheet. They then completed a short screening questionnaire to ascertain their eligibility. Provided they were eligible, participants were then directed to the consent form, followed by the main questionnaire. Completion took around one hour. Participants were able to skip any questions they wished.

Participants were first asked their age and what gender they identified as (participants were able to type their gender into an open text box to allow for maximum inclusivity). They were then given the questionnaires outlined in the materials section. They were then asked a series of questions about their physical and mental health.

They were also asked if they had ever received a diagnosis of an anxiety disorder, autism, or an eating disorder. If they indicated “yes”, participants were asked to specify the clinic in which they were diagnosed, the clinician who diagnosed them, the year they were diagnosed and in the case of anxiety disorders and eating disorders, which disorder and whether the diagnosis is still active. A diagnosis was considered verified if the participant provided answers for at least two of the three questions for autistic participants, or four of these five questions for participants with anxiety or eating disorders. In this study, participants both with and without a diagnosis can participate for an equal honorarium, and as such, there is little incentive for participants to lie about having a diagnosis. They were also

asked to self-report any other current mental health or neurodevelopmental conditions in an open text box.

Data cleaning

1,177 responses were excluded as they failed Qualtrics' reCAPTCHA checks, gibberish had been entered into any of the open text options, response times were under 10 minutes in total, or they had not completed the questionnaire.

Participants with more than 10% missing responses in the same questionnaire (excluding the cognitive tests) or with more than 20 missing responses in total (excluding the cognitive tests) were removed, to avoid negatively impacting the validity of the missing value imputation. This resulted in the removal of eight participants.

Gender was dummy coded as two variables: male (1) and not male (0), and female (1) and not female (0), which allows for three categories of gender: male, female and other. Under the category of other, participants reported the following: non-binary, androgynous, agender, NB, demigirl, no gender, and fluid.

Cognitive ability was calculated as the total of both the ICAR matrix and verbal reasoning scores combined, providing a proxy measure that can be seen as comparable to intelligence quotient scores (IQ).

Presence of another current mental health or neurodevelopmental condition was coded in a binary manner. If participants listed any mental health or neurodevelopmental condition (as included in The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition; American Psychiatric Association, 2013) that was not an anxiety disorder or autism, it was coded as 1, if not, it was coded as 0. As we had collected data on whether participants had a validated and current eating disorder diagnosis, this was included within this variable.

Interoceptive accuracy was operationalised as the total IAS score, interoceptive attention as the total IATS score, negative interoceptive evaluation as the total NIES score, alexithymia as the total TAS-20 score, autistic traits as the total CATI score, and traits of anxiety as the total DASS-Anxiety score.

Participants were grouped into those with and without a current, validated anxiety disorder diagnosis (hypothesis three) and with and without a validated autism diagnosis (hypothesis four). Participants were defined as having an anxiety disorder if they self-reported an anxiety diagnosis **and** provided four out of five validation question responses. Participants are defined as not having an anxiety diagnosis if they do not self-report an anxiety disorder (current or otherwise) and score below 15 on the DASS-Anxiety scale (which indicates severe trait anxiety). Those who did not fit into either the validated anxiety diagnosis or no anxiety disorder group were removed for any analyses that included anxiety diagnosis as a variable. Participants are defined as autistic if the participant self-reported an autism diagnosis **and** provided two out of three validation question responses, and were compared to participants without an autism diagnosis, who scored below 26 on the AQ and 134 on the CATI. Those who did not fit into either the validated autism diagnosis or no autism diagnosis group were removed for any analyses that included autism diagnosis as a variable.

Analysis

Our analysis plan was pre-registered after the collection of data, but prior to looking at the data (https://osf.io/cuh98). Any analysis that was not pre-registered is specified.

Overall, there was a an extremely low level of missing data. 0.1% of responses to items of the TAS-20 were missing (14 items out of a possible 10,250); 0.1% of responses to items of the DASS (Depression and Anxiety sub-scales) (eight items out of a possible 7,364);

0.2% of responses to items on the AQ (23 items out of a possible 14,728); 0.2% of responses to items on the IATS (24 items out of a possible 11,046); 0.1% of responses to items on the IAS (16 items out of a possible 11,046); and 0.1% of responses to items on the NIES (25 items out of a possible 10,520). There were no missing responses for the CATI. Two participants had data missing for age, and five for gender. These participants were excluded from analysis, as the decision was made to not impute single item variables.

Missing items when calculating total questionnaire scores introduce a form of bias whereby participants who miss a question will have lower scores. Similarly, removing all participants with any missing items could also introduce biases and lead to a very small sample. To avoid the total scores being affected, data missing at the item level was multiply imputed using the forest function on r, and the total scores were calculated using these imputed values. Items were initially coded as integers, however, the normalized root mean squared error (NRMSE) value of 0.43 indicated a less-than-optimal goodness of fit. The imputation was run again whilst coding the items as factors, resulting in a proportion of false classification (PFC) value of 0.21. This was deemed acceptable. Distributions of both the imputed and non-imputed datasets were compared, and as they were almost identical this imputation was retained.

For the first two hypotheses, multiple linear regressions were conducted. The predictor variables were anxious traits (hypothesis one) and autistic traits (hypothesis two). Three models were tested for each hypothesis, with the dependent variables for each being interoceptive accuracy, attention, and negative interoceptive evaluation. Covariates were age, gender, IQ, presence of other mental health or neurodevelopmental conditions, and autistic traits for hypothesis one and anxious traits for hypothesis two.

Assumptions of the multiple linear regressions were tested as follows:

- The linearity of relationships between variables was tested with scatterplots.
- The normal distribution of residuals was tested by visually inspecting P-P plots.
- Multicollinearity was tested by checking variance inflation factor (VIF) and correlation matrices.
- Homoscedasticity was tested using a scatterplot of residuals versus predicted values.
- The presence of outliers that impacted results was tested by examining Cook's distance.
- Durban-Watson statistics were used to check the independence of observations.

For the third and fourth hypotheses, ANCOVA tests were conducted. For the third hypothesis, the independent variable was whether participants had a validated anxiety disorder. For the fourth, it was whether participants had a validated autism diagnosis. Three ANCOVAs were conducted for each hypothesis, with the dependent variables being interoceptive accuracy, interoceptive attention, and negative interoceptive evaluation. Covariates in all ANCOVAs included: age, gender, cognitive ability, alexithymia and the presence of other mental health or neurodevelopmental conditions. Autistic traits were included as covariates in hypothesis three, and anxious traits were included as covariates in hypothesis four.

Assumptions of the ANCOVA were tested as follows:

- Normality of the dependent variable was tested by visually inspecting histograms for each of the groups, and checking skewness, kurtosis, Shapiro Wilk and Kolmogorov Smirnoff values. If the variable only showed a minor deviation from normality, the ANCOVA was conducted, given the robustness of this test to slight deviations from normality. If the deviation from normality was more severe, the Quade test was used as a non-parametric alternative.

- Homogeneity of variance was tested with Levene's test.
- The linear relationship between covariates and the dependent variable at each level of the independent variable was visually checked using scatter plots.
- Homoscedasticity was tested with the Breusch Pagan test.
- The homogeneity of regression slopes was tested by re-running each model including the interactions between the dependent variables and covariates.
- Multicollinearity was checked using variation inflation factors (VIF) and correlation matrices.

For hypothesis five, distributions of scores of interoceptive accuracy, attention, and negative interoceptive evaluation were compared between the autistic and non-autistic group using histograms and two sample Kolmogorov Smirnov tests.

For hypothesis six, linear regression models and logistic regression models were conducted. For the linear regressions, autistic traits, the relevant measure of interoception (accuracy in one model, attention in another, and evaluation in a third), and the interaction between autistic traits and the relevant dimension of interoception were each included as predictors of anxious traits. This was then repeated with the same covariates as outlined in the previous models. The same assumptions as listed above for each were tested for the linear regression. For the logistic regression the same analysis was carried out, but instead of autistic traits, autistic diagnosis was used as a predictor, and instead of anxious traits, a validated anxiety disorder diagnosis was considered for the outcome.

Assumptions of the logistic regression were tested as follows:

- Multicollinearity was tested by checking VIF and correlation matrices.
- Linearity of the independent variables and log odds of the dependent variables was tested by plotting scatter plots.

- Outliers were checks with the Cook's distance.

The Benjamini-Hochburg adjustment was used to correct for multiple comparisons (Benjamini & Hochberg, 1995). All p values for the main analyses, pre-registered and exploratory analyses were included in this. None of the p values for the demographic information assumption checks or for control variables were included.

Aside from the multiple imputation, which used r, all analysis were conducted using SPSS (Version 29.0.2.0 (20)).

Results

Trait level analyses

Demographics. 519 participants were recruited for the trait level analyses (hypotheses one and two). Of these participants, 104 (20.0%) were male, 366 (70.5%) were females, and 49 (9.4%) were another gender. 149 (28.7%) had a current, validated anxiety disorder diagnosis. 232 (44.7%) had a validated autism diagnosis. 100 (19.3%) had both an anxiety disorder and autism, whilst 96 (18.5%) had neither an anxiety disorder nor autism. 246 (47.4%) had a mental health or neurodevelopmental condition other than an anxiety disorder or autism (Table 7.1 shows a breakdown of the reported mental health and neurodevelopmental conditions). See Table 7.2 for descriptive statistics of the continuous variables.

Table 7.1

Other mental health and neurodevelopmental conditions reported by the whole sample

Diagnosis	*n*
Depression	112
Attention deficit hyperactivity disorder	57
Post-traumatic stress disorder (including complex PTSD)	40
Anorexia	35
Obsessive compulsive disorder	22
Avoidant restrictive food intake disorder	18
Borderline personality disorder	10
Bulimia	10
Bipolar	9
Other specified feeding or eating disorder	9
Binge eating disorder	8
Emotional unstable personality disorder	7
Dyspraxia	4
Dyslexia	3
Orthorexia	3
Pre-menstrual dysphoric disorder	3
Body dysmorphic disorder	2
Cyclothymia	2
Schizoaffective disorder	2
Adjustment disorder	1
Alcoholism	1
Chronic pain syndrome	1
Dissociative disorder	1
Dysthymia	1
Mild neurocognitive disorder	1
Mixed mood disorder	1
Obsessive compulsive personality disorder	1
Oppositional defiant disorder	1
Pica	1
Postnatal depression	1
Psychotic disorder not otherwise specified	1
Schizophrenia	1
Substance use disorder	1
Tourette’s syndrome	1
Trauma	1
Trichotillomania	1

Table 7.2

Demographic information for the whole sample

	Mean	*SD*	Range	Skewness	Kurtosis	*KS*	*SW*
Age *	36.7	15.2	18-82	.727	-.402	.126 (<.001)	.923 (<.001)
IQ *	18.0	5.0	3-30	-.372	-.488	.090 (.013)	.992 (.006)
TAS-20 *	59.7	13.8	20-89	-.456	-.294	.080 (<.001)	.979 (<.001)
CATI *	154.7	30.2	56-209	-.746	.123	.085 (<.001)	.954 (<.001)
DASS-A *	6.9	4.6	0-21	.657	-.084	.117 (<.001)	.954 (<.001)
IATS	50.7	14.4	21-100	.282	-.012	.034 (.200)	.994 (.037)
IAS	71.8	14.0	21-105	-.177	.216	.042 (.027)	.992 (.005)
NIES	67.6	14.9	20-100	-.281	.006	.045 (.013)	.992 (.006)

Note. N = 519. *SD* = standard deviation. *KS* = Kolmogorov Smirnov. *SW* = Shapiro Wilk. IQ = combined scores from the ICAR Verbal and Matrix Reasoning tests. * = when histograms were visually inspected, the distribution of these variables was non-normally distributed.

Assumption checks. All continuous variables (control variables and the independent variables) were found to have linear relationships with the dependent variables based on

scatter plots. VIF values were all under 5, with the highest being for the dummy-coded male or not male variable (2.67 in all regressions). The only correlation coefficient over 0.7 was the correlation between the dummy coded male and female responses ($r = 0.77$ in all regressions). This was to be expected. Therefore, multicollinearity was not seen to be an issue. The P-P plots indicated a normal distribution of residuals. The scatterplot of residuals versus predicted values indicated that the assumption of homoscedasticity was met. The Durban-Watson statistics indicated independence of observations for each of the regressions (1.94 for the NIES regression, 1.98 for the IAS regression, and 1.82 for the IATS regression). Cook's distance was below one in all of the regressions, suggesting that no data points exerted a significant influence on the regression line. Therefore, all the assumptions for the multiple linear regressions were met.

NIES regression. A multiple regression was run to predict NIES scores from gender, age, presence of mental health or neurodevelopmental conditions other than autism or anxiety, IQ, TAS-20, CATI and DASS-A scores. This model statistically significantly predicted NIES ($F(8, 510) = 40.841$, $p < .001$, $R^2 = .391$). Of all the variables, those that were significant predictors of higher NIES scores were: not being male ($\beta = -.202$, $p < .001$), the presence of another mental health condition ($\beta = .086$, $p = .023$), higher CATI scores ($\beta = .287$, $p < .001$), and higher DASS-A scores ($\beta = .321$, $p < .001$).

IAS regression. A multiple regression was run to predict IAS scores from gender, age, presence of mental health or neurodevelopmental conditions other than autism or anxiety, IQ, TAS-20, CATI and DASS-A scores. This model statistically significantly predicted IAS ($F(8, 510) = 23.732$, $p < .001$, $R^2 = .271$). Of all the variables, those that were significant predictors of higher IAS scores were: higher age ($\beta = .130$, $p < .001$), being male ($\beta = .133$, $p = .031$), lower CATI scores ($\beta = -.139$, $p = .010$), and lower TAS-20 scores ($\beta = -.395$, $p < .001$). DASS-A scores were not a significant predictor ($\beta = .018$, $p = .687$).

IATS regression. A multiple regression was run to predict IATS scores from gender, age, presence of mental health or neurodevelopmental conditions other than autism or anxiety, IQ, TAS-20, CATI and DASS-A scores. This model statistically significantly predicted IATS ($F(8, 510) = 27.259$, $p < .001$, $R^2 = .299$). Of all the variables, those that were significant predictors of higher IATS scores were: lower age ($\beta = -.114$, $p = .004$), higher CATI scores ($\beta = .222$, $p < .001$), and higher DASS-A scores ($\beta = .395$, $p < .001$).

Summary. Higher autistic traits predicted higher negative interoceptive evaluation, lower interoceptive accuracy, and higher interoceptive attention. Higher anxious traits predicted higher negative interoceptive evaluation and interoceptive attention.

Autistic versus non-autistic

Of the original sample of 519 participants, those with a validated autism diagnosis and those with no reported autism diagnosis and CATI and AQ scores that were below established thresholds that would indicate a high degree of autistic traits were selected and formed the two groups for this analysis.

The autistic group had a total of 232 participants. Of these, 59 were male (25.4%), 149 were female (64.2%) and 24 (10.3%) were another gender. 100 had a validated anxiety disorder diagnosis (43.1%) and 155 had other mental health or neurodevelopmental conditions (66.8%; see Table 7.3). See Table 7.4 for the descriptive statistics of all continuous variables within the autistic group.

Table 7.3

Other mental health and neurodevelopmental conditions reported by the autistic group

Diagnosis	*N*
Depression	92
Attention deficit hyperactivity disorder	34
Post-traumatic stress disorder (including complex PTSD)	23
Anorexia	15
Obsessive compulsive disorder	11
Avoidant restrictive food intake disorder	8
Bulimia	6
Binge eating disorder	5
Borderline personality disorder	5
Bipolar	4
Emotional unstable personality disorder	3
Other specified feeding or eating disorder	3
Body dysmorphic disorder	2
Dyslexia	2
Dyspraxia	2
Pre-menstrual dysphoric disorder	2
Alcoholism	1
Dissociative disorder	1
Mixed mood disorder	1
Obsessive compulsive personality disorder	1
Orthorexia	1
Pica	1
Postnatal depression	1
Psychotic disorder not otherwise specified	1
Schizoaffective disorder	1
Schizophrenia	1
Trauma	1
Trichotillomania	1
Mild Neurocognitive Disorder	1

Table 7.4

Descriptive statistics for the autistic group

	Mean	*SD*	Range	Skewness	Kurtosis	*KS*	*SW*
Age *	39.5	14.2	18-76	.461	-.729	.103 (<.001)	.954 (<.001)
IQ *	18.3	5.1	3-30	-.366	-.324	.093 (<.001)	.975 (<.001)
TAS-20 *	64.5	11.9	35-89	-.471	-.374	.086 (<.001)	.970 (<.001)
CATI *	169.4	20.6	64-205	-1.110	2.626	.096 (<.001)	.941 (<.001)
DASS-A *	7.3	4.6	0-21	.572	-.292	.137 (<.001)	.957 (<.001)
IATS	51.0	14.0	21-93	.228	.042	.046 (.200)	.991 (.138)
IAS *	69.4	13.7	24-105	.057	.265	.054 (.099)	.990 (.117)
NIES *	70.8	15.1	20-100	-.440	.434	.059 (.046)	.983 (.007)

Note. $N = 232$. *SD* = standard deviation. *KS* = Kolmogorov Smirnov. *SW* = Shapiro Wilk. IQ = combined scores from the ICAR Verbal and Matrix Reasoning tests. * = when histograms were visually inspected, the distribution of these variables was non-normally distributed.

The non-autistic group had a total of 143 participants. Of these, 26 (18.1%) were male, 109 (76.4%) were female, and 8 (5.5%) were another gender. 20 (13.8%) had a

validated anxiety disorder diagnosis, and 29 (20.1%) had another mental health or neurodevelopmental condition (see Table 7.5). See Table 7.6 for the descriptive statistics of all continuous variables within the non-autistic group.

Table 7.5

Other mental health and neurodevelopmental conditions reported by the non-autistic group

Diagnosis	*N*
Depression	26
Attention deficit hyperactivity disorder	12
Anorexia	9
Post-traumatic stress disorder (including complex PTSD)	7
Bulimia	4
Other specified feeding or eating disorder	4
Obsessive compulsive disorder	2
Avoidant restrictive food intake disorder	1
Binge eating disorder	1
Bipolar	1
Body dysmorphic disorder	1
Borderline personality disorder	1
Cyclothymia	1
Dyspraxia	1
Dysthymia	1
Emotional unstable personality disorder	1
Substance use disorder	1

Table 7.6

Descriptive information for the non-autistic group

	Mean	*SD*	Range	Skewness	Kurtosis	*KS*	*SW*
Age *	33.0	15.9	18-77	1.252	.574	.180 (<.001)	.829 (<.001)
IQ *	18.4	4.5	6-26	-.556	-.195	.108 (<.001)	.965 (.001)
TAS-20 *	48.3	12.2	20-78	-.185	-.504	.053 (.200)	.988 (.235)
CATI *	117.7	20.8	56-147	-.833	.124	.148 (<.001)	.930 (<.001)
DASS-A *	5.0	4.0	0-19	.950	1.139	.130 (<.001)	.924 (<.001)
IATS *	46.2	13.7	21-79	.248	-.548	.067 (.200)	.981 (.049)
IAS	78.1	12.1	47-105	-.221	-.100	.062 (.200)	.988 (.273)
NIES	60.1	13.2	29-91	-.099	-.407	.059 (.200)	.991 (463)

Note. N = 143. *SD* = standard deviation. *KS* = Kolmogorov Smirnov. *SW* = Shapiro Wilk. IQ = combined scores from the ICAR Verbal and Matrix Reasoning tests. * = when histograms were visually inspected, the distribution of these variables was non-normally distributed.

Those in the autistic group were significantly older than those in the non-autistic group ($t(373) = 4.12$, $p < .001$, $g = 0.44$). There was no significant difference in IQ ($t(373) =$

0.19, p = .847, g = -0.02). Autistic participants had significantly higher TAS-20 scores (t(373) = 12.70, p < .001, g = 1.35), CATI scores (t(373) = 23.57, p < .001, g = 2.50), and DASS-A scores (t(373) = 4.95, p < .001, g = 0.52) than non-autistic participants. Of particular interest (but without accounting for covariates), autistic participants had significantly higher IATS scores than non-autistic participants (t(373) = 3.26, p < .001, g = 0.35), significantly lower IAS scores than non-autistic participants (t(373) = 6.25, p < .001, g = -0.66), and significantly higher NIES scores than non-autistic participants (t(373) = 7.00, p < .001, g = 0.74).

There was a higher proportion of autistic participants than non-autistic participants with anxiety disorders (χ^2 (1, N = 375) = 34.90, p < .001, φ = 0.30), and other mental health and neurodevelopmental conditions (χ^2 (1, N = 375) = 77.45, p < .001, φ = 0.45). Gender ratios were significantly different between autistic and non-autistic participants (χ^2 (2, N = 375) = 6.44, p = .011, φ = 0.13).

Assumption checks: Assumptions were checked for each of the three ANCOVAs conducted to compare scores in the IAS, IATS and NIES between the autistic and non-autistic participants. Histograms indicated that although not all the dependent variables were normally distributed, the deviation from normality was not such that ANCOVAs were deemed inappropriate. Scatter plots showed that within each group (autistic and non-autistic), there was a linear relationship between the continuous covariates and each of the dependent variables (the IAS, the IATS, and the NIES). There was no issue with multicollinearity, as again, all VIF values were under 5. As with the linear regressions, the only correlation above .70 was the correlation between the male and female dummy coded variables (r = .77). Breusch Pagan tests indicated no heteroscedasticity (IAS: χ^2(1) = .044, p= .833; IATS: χ^2(1) = .265, p = .607; NIES: χ^2(1) = .034, df = 1, p=.853). Levene's test indicated there was homogeneity of variance (IAS: F(1, 373) = .604, p = .437; IATS: F(1, 373) = .067, p = .796; NIES: F(1, 373) = 3.664, p = .056). The homogeneity of regression slopes was tested by re-

running each model including the interactions between the dependent variables and covariates, and as none of the interaction terms were significant, the use of ANCOVA was deemed appropriate.

NIES. A one-way ANCOVA was conducted to examine the effect of the presence of an autism diagnosis on NIES scores, while controlling for age, gender, anxious traits, alexithymia, presence of another mental health or neurodevelopmental condition and IQ. There was a significant difference in the NIES scores between participants with and without a validated autism diagnosis when controlling for covariates ($F(1, 366) = 8.50$, $p = .004$, $\eta^2_p = .02$). Those with a validated autism diagnosis ($M = 70.8$, $SD = 15.1$) had higher NIES scores than those without ($M = 60.1$, $SD = 13.2$). The effect size is small.

IAS. A one-way ANCOVA was conducted to examine the effect of the presence of an autism diagnosis on IAS scores, while controlling for covariates (see above). There was no significant difference in the IAS scores between participants with and without validated autism diagnoses when controlling for covariates ($F(1, 366) = 1.80$, $p = .180$, $\eta^2_p = .01$).

IATS. A one-way ANCOVA was conducted to examine the effect of the presence of an autism diagnosis on IATS scores, while controlling for covariates (see above). There was no significant difference in the IATS scores between participants with and without validated autism diagnoses when controlling for covariates ($F(1, 366) = 1.07$, $p = .303$, $\eta^2_p = .00$).

Summary. Those with an autism diagnosis had higher negative interoceptive evaluation than non-autistic individuals, but there was no difference in interoceptive accuracy or attention.

Anxiety disorder versus no anxiety disorder

The anxiety disorder group (those with a current, validated anxiety disorder diagnosis) were comprised of a total of 149 participants. Of these, 27 (18.1%) were male, 103 (69.1%)

were female, and 19 (12.8%) were another gender. 100 (67.1%) were autistic, and 107 (71.8%) had another mental health or neurodevelopmental condition (see Table 7.7). See Table 7.8 for the descriptive statistics of all continuous variables within the anxiety disorder group.

Table 7.7

Other mental health and neurodevelopmental conditions reported by the anxiety disorder group

Diagnosis	*N*
Depression	80
Attention deficit hyperactivity disorder	31
Post-traumatic stress disorder (including complex PTSD)	26
Anorexia	16
Obsessive compulsive disorder	13
Binge eating disorder	7
Avoidant restrictive food intake disorder	6
Borderline personality disorder	6
Bulimia	6
Emotional unstable personality disorder	6
Other specified feeding or eating disorder	6
Orthorexia	3
Body dysmorphic disorder	2
Dyslexia	2
Dyspraxia	2
Alcoholism	1
Bipolar	1
Cyclothymia	1
Mixed mood disorder	1
Obsessive compulsive personality disorder	1
Pica	1
Postnatal depression	1
Psychotic disorder not otherwise specified	1
Schizophrenia	1
Trichotillomania	1

Table 7.8

Descriptive information for the anxiety disorder group

	Mean	*SD*	Range	Skewness	Kurtosis	*KS*	*SW*
Age *	36.8	14.6	18-77	.669	-.459	.140 (<.001)	.930 (<.001)
IQ *	17.5	4.9	3-27	-.096	-.740	.091 (.004)	.970 (.002)
TAS-20	63.9	12.4	24-89	-.549	-.037	.078 (.026)	.973 (.005)
CATI *	164.8	25.9	87-204	-1.139	.902	.129 (<.001)	.904 (<.001)
DASS-A *	8.5	5.0	0-21	.429	-.366	.083 (.014)	.970 (.003)
IATS	54.9	14.2	21-100	.442	.411	.057 (.200)	.985 (.100)
IAS	68.8	13.5	38-103	.210	-.243	.044 (.200)	.992 (.535)
NIES	72.1	12.5	27-95	-.503	.360	.059 (.200)	.980 (.028)

Note. N = 149. *SD* = standard deviation. *KS* = Kolmogorov Smirnov. *SW* = Shapiro Wilk. IQ = combined scores from the ICAR Verbal and Matrix Reasoning tests. * = when histograms were visually inspected, the distribution of these variables was non-normally distributed.

Non anxious. The non-anxiety disorder group (those who didn't report having an anxiety disorder and scored below the established threshold for severe anxiety on the DASS-

A) comprised of a total of 235 participants. Of these, 62 (26.4%) were male, 154 (65.5%) were female, and 19 (8.1%) were another gender. 77 (32.8%) were autistic, and 67 (28.5%) had other mental health or neurodevelopmental conditions (see Table 7.9). See Table 7.10 for the descriptive statistics of all continuous variables within the non-anxiety disorder group.

Table 7.9

Other mental health and neurodevelopmental conditions reported by the non-anxiety disorder group

Diagnosis	*N*
Depression	31
Attention deficit hyperactivity disorder	21
Anorexia	13
Avoidant restrictive food intake disorder	7
Bipolar	6
Post-traumatic stress disorder (including complex PTSD)	6
Bulimia	3
Obsessive compulsive disorder	3
Dyslexia	2
Other specified feeding or eating disorder	2
Schizoaffective disorder	2
Binge eating disorder	1
Body dysmorphic disorder	1
Borderline personality disorder	1
Chronic pain syndrome	1
Dysthymia	1
Emotional unstable personality disorder	1
Pre-menstrual dysphoric disorder	1
Substance use disorder	1
Tourette syndrome	1

Table 7.10

Descriptive information for the non-anxiety disorder group

	Mean	*SD*	Range	Skewness	Kurtosis	*KS*	*SW*
Age *	37.4	17.0	18-82	.655	-.751	.143 (<.001)	.904 (<.001)
IQ *	18.6	4.8	4-30	-.459	-.297	.088 (<.001)	.972 (<.001)
TAS-20 *	55.5	13.7	20-84	-.299	-.469	.062 (.027)	.984 (.010)
CATI *	144.5	31.2	56-208	-.437	-.172	.052 (.200)	.981 (.003)
DASS-A *	5.2	3.8	0-14	.646	-.297	.128 (<.001)	.935 (<.001)
IATS *	46.9	14.5	21-92	.416	-.054	.057 (.063)	.981 (.003)
IAS	75.2	13.0	24-105	-.400	.837	.060 (.040)	.984 (.009)
NIES	62.9	15.8	20-100	-.068	-.147	.062 (.027)	.992 (.175)

Note. $N = 235$. *SD* = standard deviation. *KS* = Kolmogorov Smirnov. *SW* = Shapiro Wilk. IQ = combined scores from the ICAR Verbal and Matrix Reasoning tests. * = when histograms were visually inspected, the distribution of these variables was non-normally distributed.

There was no difference in the age of participants with and without anxiety disorders ($t(382) = 0.36$, $p = .722$, $g = .04$). Participants with anxiety disorders had lower IQ than those

without anxiety disorders ($t(382) = 2.17$, $p = .031$, $g = .23$). Participants with anxiety disorders had higher TAS-20 scores than those without ($t(382) = 6.07$, $p < .001$, $g = .64$). Participants with anxiety disorders had significantly higher CATI scores than those without ($t(382) = 6.62$, $p < .001$, $g = .69$). Participants with anxiety disorders had significantly higher DASS-A scores than those without ($t(382) = 7.32$, $p < .001$, $g = 1.23$). Participants with anxiety disorders had significantly higher IATS scores than those without anxiety disorders ($t(382) = 5.31$, $p < .001$, $g = .57$). Participants with anxiety disorders had significantly lower IAS scores than those without anxiety disorders ($t(382) = 4.63$, $p < .001$, $g = .48$). Participants with anxiety disorders scored significantly higher on NIES than those without anxiety disorders ($t(382) = 6.01$, $p < .001$, $g = .63$).

There was a higher proportion of participants with anxiety disorders than participants without anxiety disorders who had a diagnosis of autism (χ^2 (1, $N = 384$) = 43.30, $p = < .001$, $\varphi = .34$), and other mental health and neurodevelopmental conditions (χ^2 (1, $N = 384$) = 69.00, $p = < .001$, $\varphi = .42$). There was no significant difference in the gender ratios of participants with and without anxiety disorders (χ^2 (2, $N = 384$) = 4.87, $p = .088$, $\varphi = .11$).

Assumption checks. Assumptions were checked for each of the three ANCOVAs conducted to compare scores on the IAS, IATS and NIES between the anxiety disorder and non-anxiety disorder groups. Histograms indicated that although not all the dependent variables were normally distributed, the deviation from normality was not such that ANCOVAs were deemed inappropriate. Scatter plots showed that within each group (anxious and non-anxious), there was a linear relationship between the continuous covariates and each of the dependent variables (the IAS, the IATS, and the NIES). There was no issue with multicollinearity, as again, all VIF values were under 5, with the highest being for the dummy-coded male or not male variable (2.80). As with the linear regressions, the only correlation above .70 was the correlation between the male and female dummy coded

variables (r = .78). Breusch Pagan tests indicated no heteroscedasticity (IAS: $\chi^2(1) = .024$, p = .876; IATS: $\chi^2(1) = 2.708$, p = .100; NIES: $\chi^2(1) = .944$, p = .331). Levene's test indicated there was homogeneity of variance (IAS: $F(1,382) = .052$, p = .820; IATS: $F(1,382) = .269$, p = .604; NIES: $F(1,382) = 3.47$, p = .063).

NIES. A one-way ANCOVA was conducted to examine the effect of the presence of an anxiety disorder on NIES scores, while controlling for age, gender, autistic traits, alexithymia, presence of another mental health or neurodevelopmental condition and IQ. There was no significant difference in the NIES scores between participants with and without validated anxiety disorders when controlling for covariates ($F(1,375) = 3.76$, p = .053, η^2_p = .01).

IAS. A one-way ANCOVA was conducted to examine the effect of the presence of an anxiety disorder on IAS scores, while controlling for covariates (see above). There was no significant difference in the IAS scores between participants with and without validated anxiety disorders when controlling for covariates ($F(1,375) = 1.17$, p = .281, η^2_p = .00).

IATS. A one-way ANCOVA was conducted to examine the effect of the presence of an anxiety disorder on IATS scores, while controlling for covariates (see above). There was a significant difference in the IATS scores between participants with and without validated anxiety disorders when controlling for covariates ($F(1,375) = 5.10$, p = .024, η^2_p = .01). Those with an anxiety disorder (M = 54.9, SD = 14.2) had higher IATS scores than those without anxiety disorders (M = 46.9, SD = 14.5). Partial eta squared indicated a small effect size (.013).

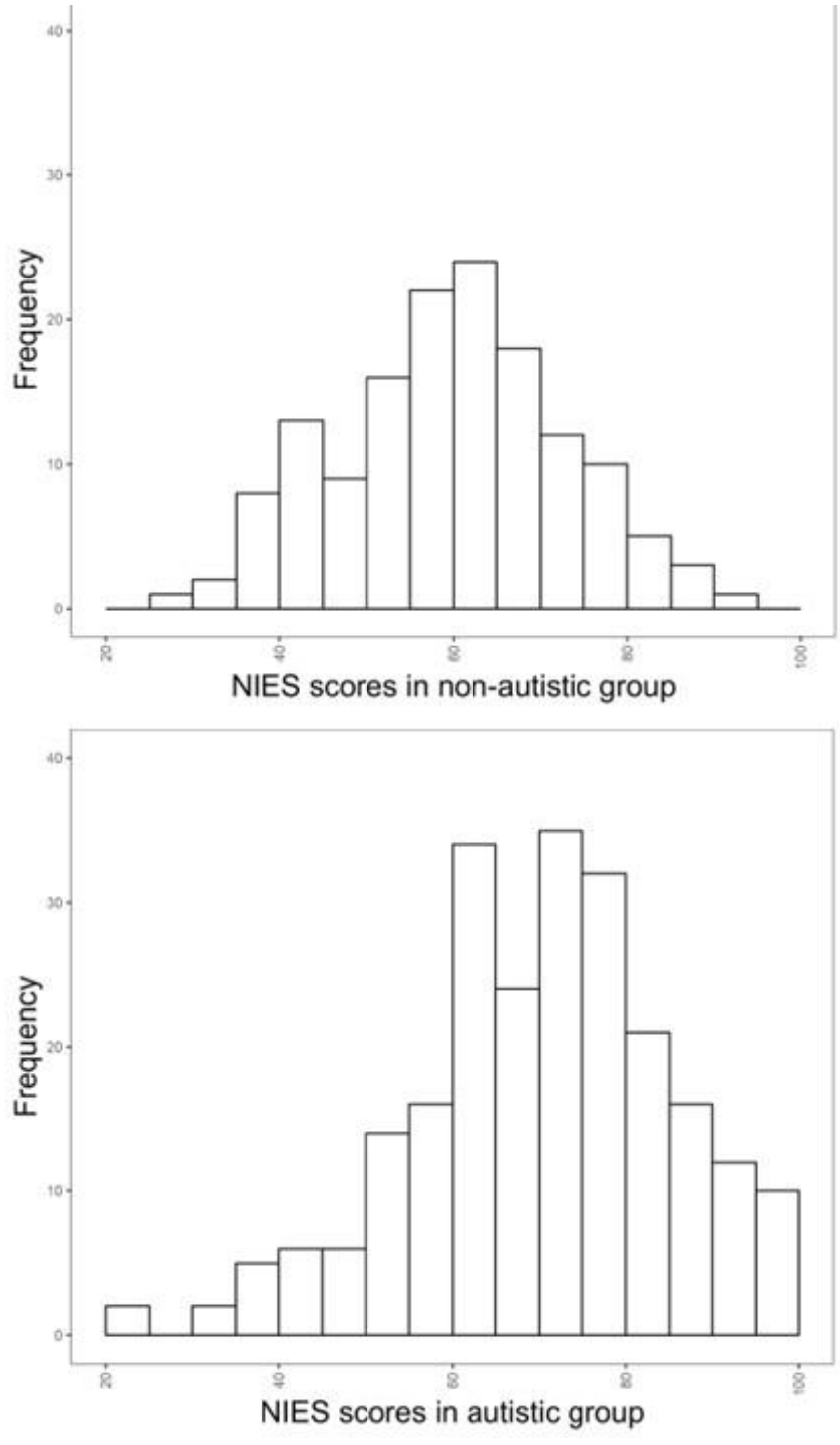
Summary. Those with a diagnosed anxiety disorder had higher interoceptive accuracy than non-anxious participants, but there was no difference in interoceptive accuracy or negative interoceptive evaluation.

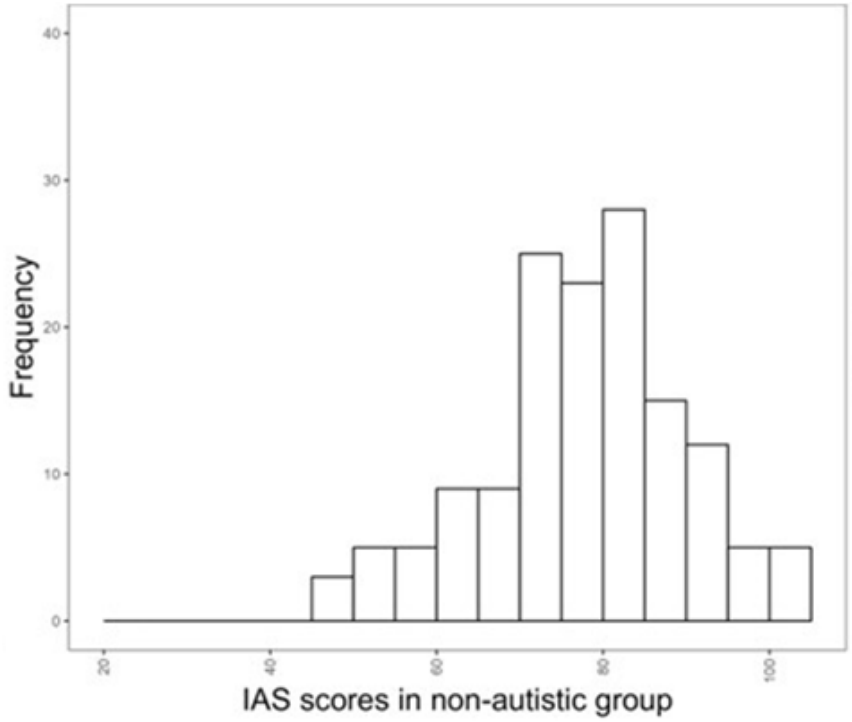
Distributions across groups

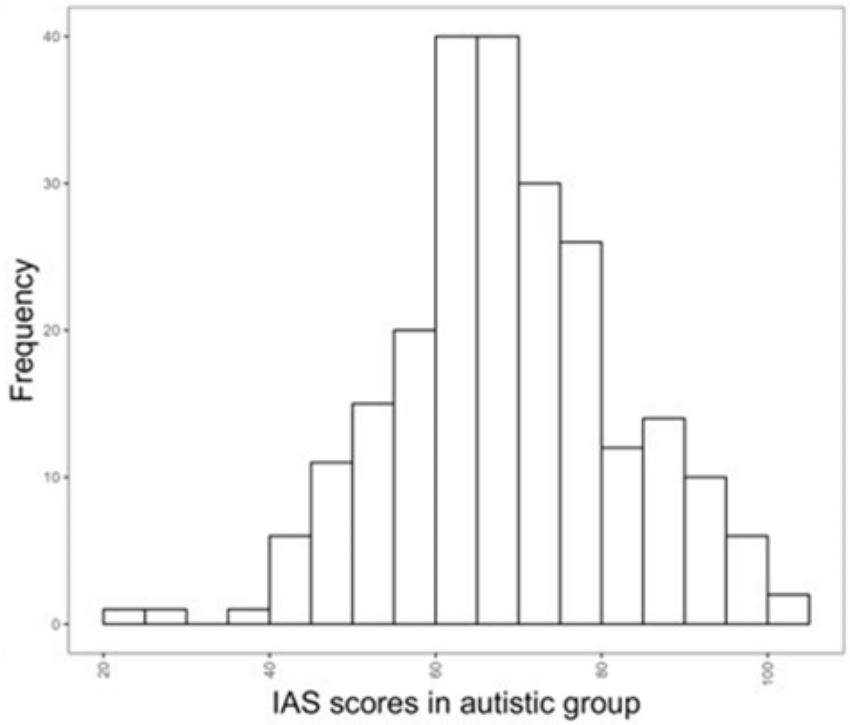
Two sample Kolmogorov Smirnov tests indicated differences between autistic and non-autistic participants in the distribution of scores across each of the three self-report measures of interoception (NIES: $Z = 3.17$, $p < .001$; IAS: $Z = 3.31$, $p < .001$; IATS: $Z = 1.76$, $p = .004$). However, given the large sample size, it is possible that even minor differences in distributions result in significant results. Therefore, the next step was to plot the distributions and compare them visually (Figure 7.1). All histograms indicate a unimodal distribution. The NIES scores are slightly more negatively skewed in autistic compared to non-autistic participants, reflecting the fact that a greater proportion of autistic participants report a higher degree of negative interoceptive evaluation. IAS scores are slightly more negatively skewed in the non-autistic versus autistic participants, which indicates that data points are more clustered on the side of the graph indicating higher interoceptive accuracy in non-autistic participants. The shape of the distribution of IATS scores is similar between autistic and non-autistic participants. However, none of the differences in distributions are marked based on the histograms.

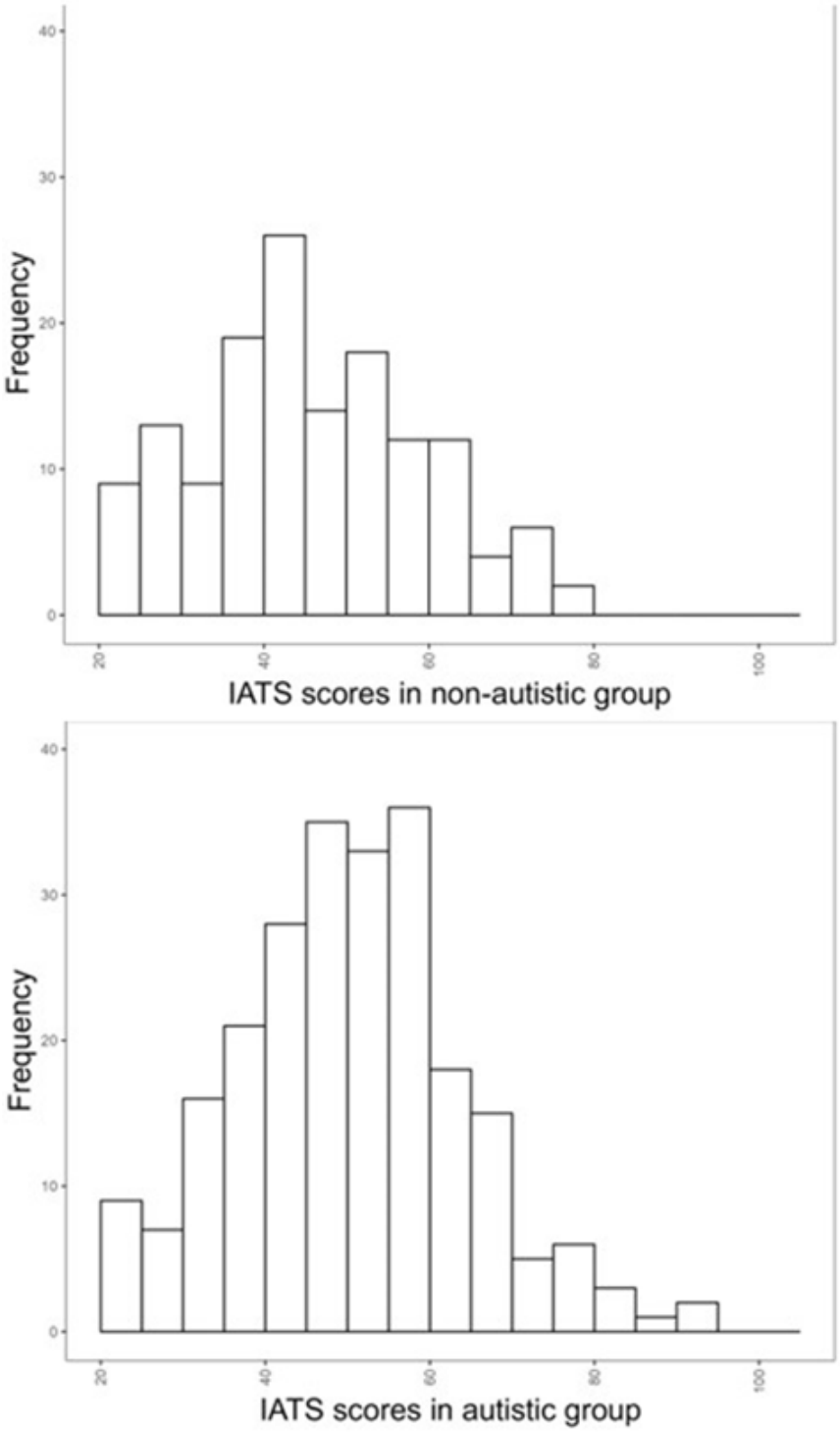
Figure 7.1

Interoception score distributions in autistic and non-autistic participants




Frequency
0
10
20
30
40
20
40
60
80
100
IAS scores in non-autistic group


Frequency
0
10
20
30
40
20
40
60
80
100
IAS scores in autistic group

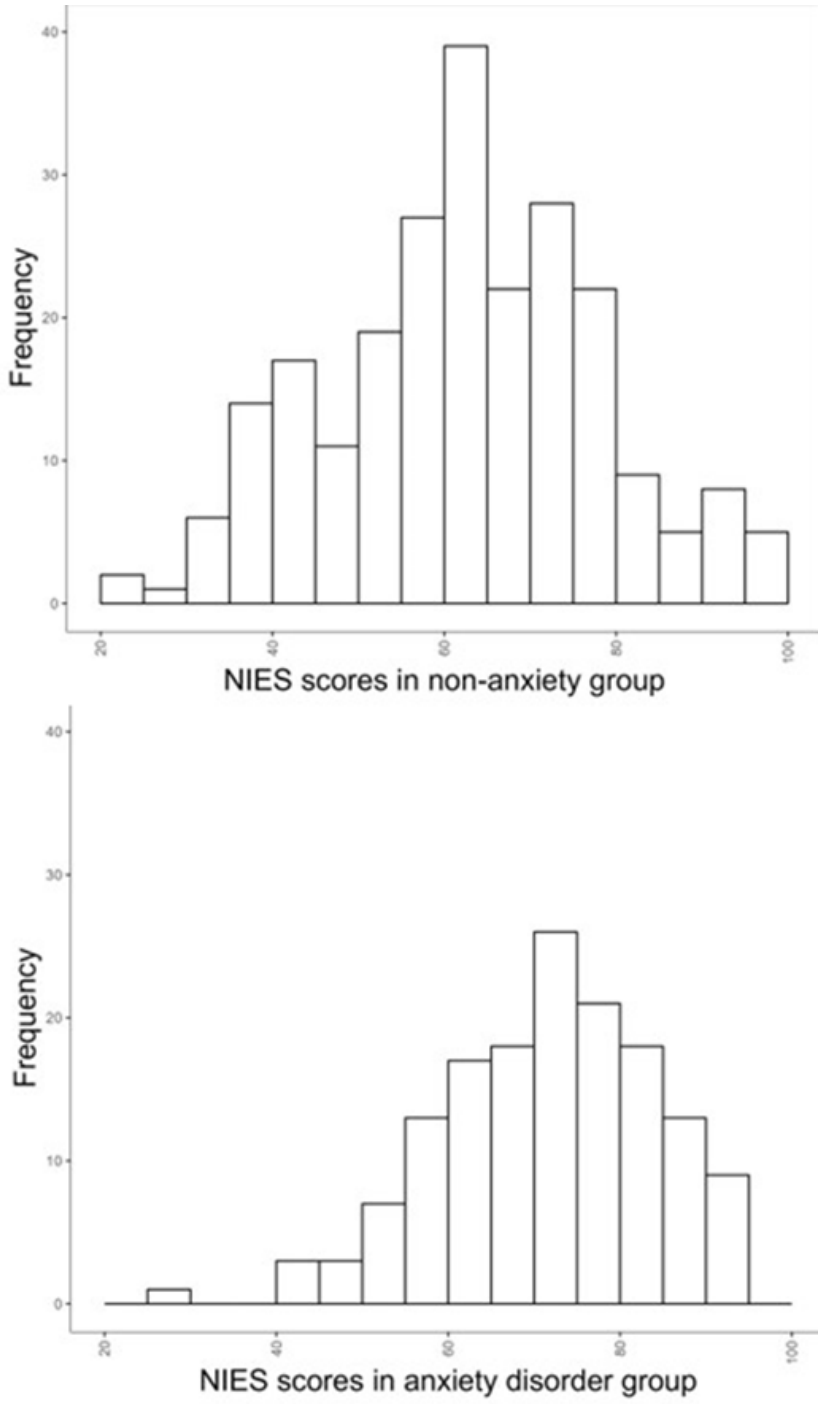


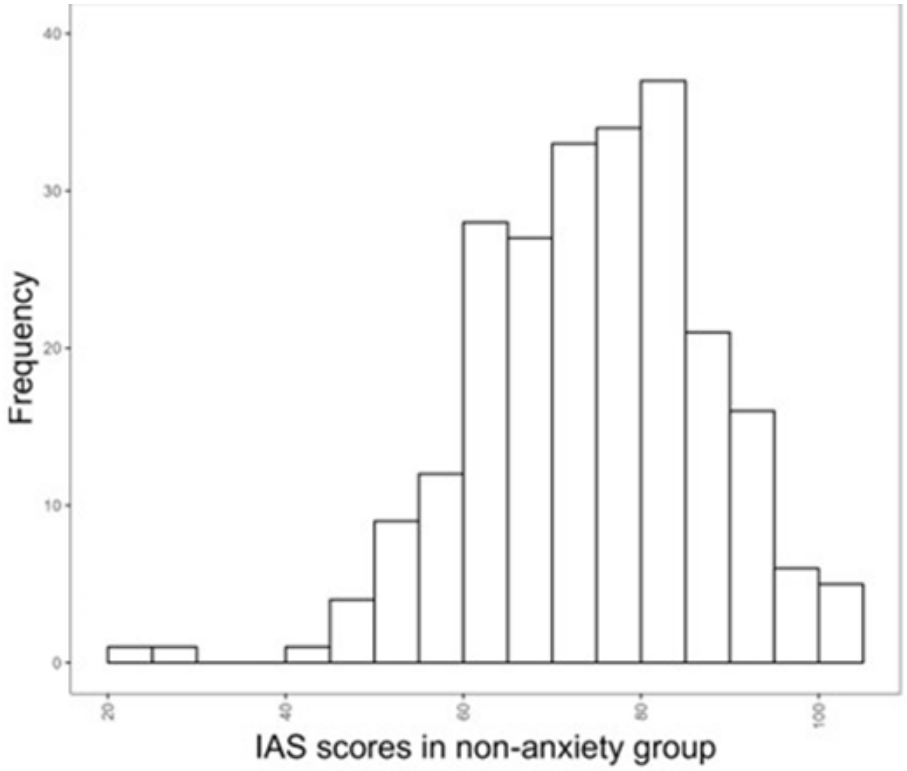
Note. This figure illustrates the differences in distributions of self-reported interoception dimensions between autistic (bottom) and non-autistic (top) participants. Scales are standardised across groups, therefore the overall frequencies are lower in the non-autistic group to reflect the reduced sample size.

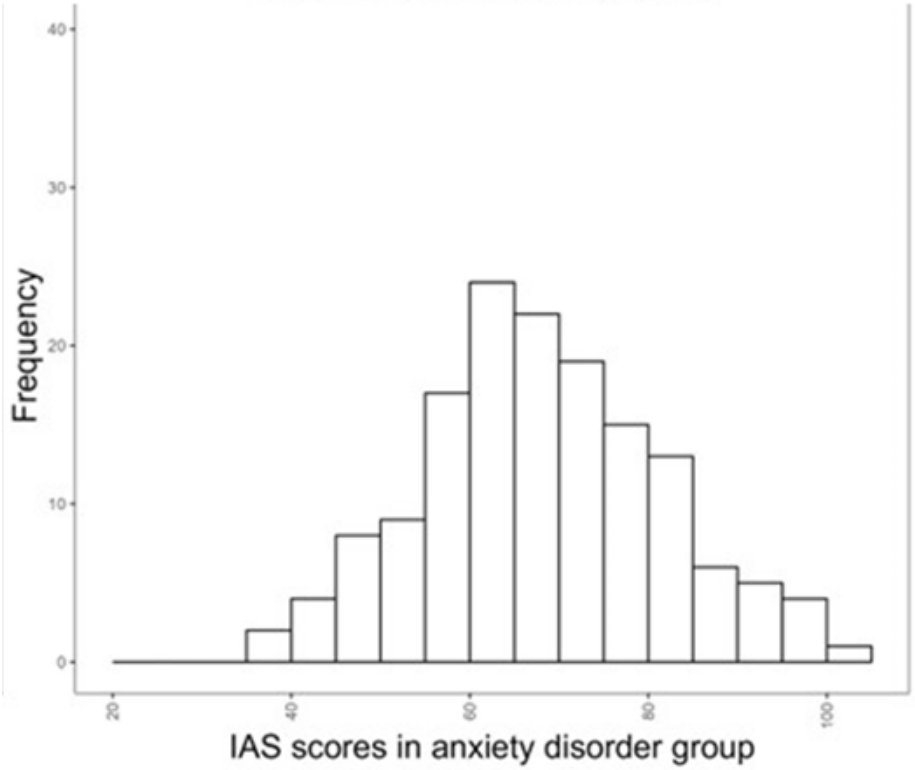
Although we did not pre-register a comparison of the distributions between anxious and non-anxious participants, or have any hypotheses about them, they are reported here for completeness. As with the comparisons between autistic and non-autistic samples, two sample Kolmogorov Smirnov tests were used to establish differences in the distribution of scores across each of the three self-report measures of interoception between anxious and non-anxious participants. All comparisons were significant (NIES: $Z = 2.95$, $p <. 001$; IAS: $Z = 2.15$, $p < .001$; IATS: $Z = 2.32$, $p < .001$), however, this could be due to the large sample sizes. Histograms were plotted to visually compare the distributions (see Figure 7.2). Again, all distributions are unimodal. The NIES scores were more negatively skewed in anxious participants, indicating that compared to the non-anxious participants, their scores were more clustered around the side indicating higher negative interoceptive evaluation. Although the distributions for the IAS scores are very similar, the distribution for the non-anxious participants is slightly more negatively skewed, meaning there is a slight tendency for the scores of non-anxious participants to cluster around the higher interoceptive accuracy scores. In contrast, for the IATS, the distribution for the non-anxious sample is slightly more positively skewed than for the anxious sample, meaning that the scores for the non-anxious sample are more clustered around the lower interoceptive attention scores.

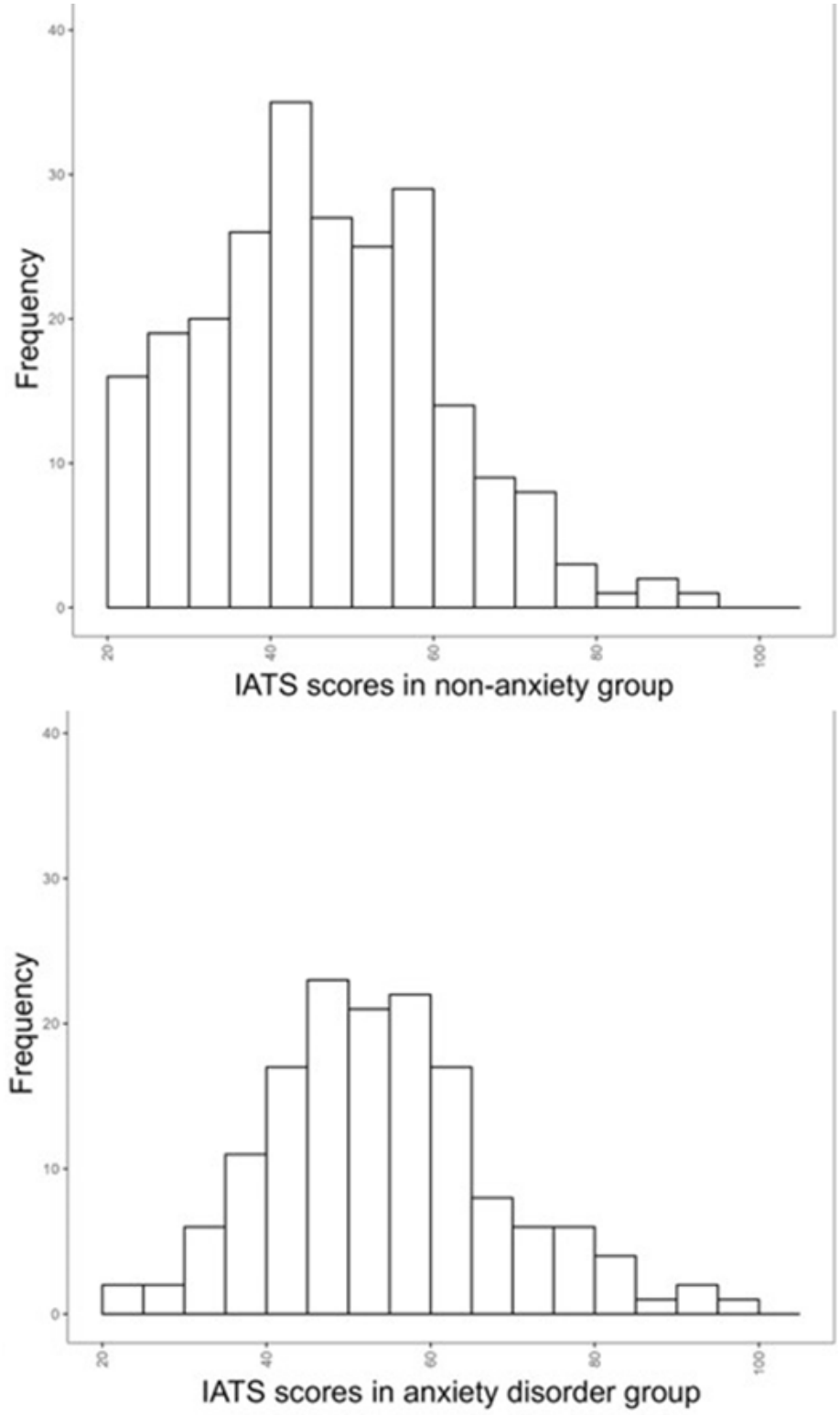
Figure 7.2

Interoception score distributions in participants with and without anxiety disorders




Frequency
0
10
20
30
40
20
40
60
80
100
IAS scores in non-anxiety group


Frequency
0
10
20
30
40
20
40
60
80
100
IAS scores in anxiety disorder group



Note. This figure illustrates the differences in distributions of self-reported interoception dimensions between participants with anxiety disorders (bottom) and those without (top) participants. Scales are standardised across groups, therefore the overall frequencies are lower in the non-autistic group to reflect the reduced sample size

Does autism moderate associations between interoception and anxiety?

All continuous variables (control variables and the independent variables) were found to have linear relationships with the dependent variables based on scatter plots, meaning that the assumption of linearity was met. VIF values were all under 5, with the highest being for the dummy-coded male or not male variable (2.73 for the NIES regression, 2.69 for the IAS regression, 2.68 for the IATS regression). The only correlation coefficient over 0.7 was the correlation between the dummy coded male and female responses (r = 0.76 for the NIES regression, 0.77 for the IAS and IATs regressions). Therefore, multicollinearity was not seen to be an issue. The P-P plots indicated a normal distribution of residuals. The scatterplot of residuals versus predicted values indicated slight heteroscedasticity, in that all three were slightly cone shaped, however this was not so marked that the regression was deemed inappropriate. The Durban-Watson statistics indicated independence of observations for each of the regressions (2.02 for the NIES regression, 2.07 for the IAS regression, and 1.96 for the IATS regression). Cook's distance was below 1 in each of the regressions, suggesting that no data points exerted a significant influence on the regression line.

NIES. A multiple regression was conducted to predict DASS-A scores from CATI and NIES scores, and the interaction between mean-centred CATI and NIES scores. The interaction term between mean-centred CATI and NIES scores was not a significant predictor of DASS-A scores (β = .06, p = .116). The same was true when gender, age, presence of mental health or neurodevelopmental conditions other than autism or anxiety and TAS-20 scores were entered into the regression as covariates (β = .05, p = .172).

IAS. A multiple regression was run to predict DASS-A scores from CATI and IAS scores, and the interaction between mean-centred CATI and IAS scores. The interaction term between mean-centred CATI and IAS scores was not a significant predictor of DASS-A

scores (β = -.07, *p* = .110). The same was true when covariates (see above) were entered into the regression (β = -.06, *p* = .153).

IATS. A multiple regression was run to predict DASS-A scores from CATI and IATS scores, and the interaction between mean-centred CATI and IATS scores. The interaction term between mean-centred CATI and IATS scores was not a significant predictor of DASS-A scores (β = .06, *p* = .113). The same was true when covariates (see above) were entered into the regression (β = .06, *p* = .132).

Summary. Interaction effects between autistic traits and interoceptive accuracy, attention and evaluation were not predictive of anxious traits.

Ahead of the logistic regression used to explore interaction effects, the sample was further reduced to include only participants with **both** either a validated autism diagnosis or those who were validated as not being autistic, and participants who had either a validated anxiety diagnosis or who were validated as not having an anxiety disorder. This resulted in a sample of 294. The outcome variable was having a validated anxiety diagnosis. The same covariates as outlined previously were included.

The anxiety disorder group within this sub-sample (those with a current, validated anxiety disorder diagnosis) was comprised of a total of 121 participants. Of these, 26 (21.5%) were male, 81 (66.9%) were female, and 14 (11.6%) were another gender. 99 (81.8%) were autistic, and 91 (75.2%) had another mental health condition. See Table 7.11 for the descriptive statistics of all continuous variables within the anxiety disorder group.

Table 7.11

Demographic information for the anxiety disorder group for the logistic regression

	Mean	*SD*	Range
Age	37.7	14.1	18-76
IQ	17.5	4.9	3-27
TAS-20	62.9	13.1	24-89
CATI	163.2	27.8	87-204
DASS-A	8.7	4.6	0-21
IATS	54.2	13.3	21-93
IAS	69.0	13.5	38-103
NIES	72.0	11.9	42-95

Note. $N = 121$. *SD* = standard deviation. IQ = combined scores from the ICAR Verbal and Matrix Reasoning tests.

Non anxious. The non-anxiety disorder group (those who didn't report having an anxiety disorder and scored below the established threshold for severe anxiety on the DASS-A) comprised of a total of 173 participants. Of these, 48 (27.7%) were male, 114 (65.9%) were female, and 11 (6.4%) were another gender. 77 (44.5%) were autistic, and 46 (26.6%) had other mental health or neurodevelopmental conditions. See Table 7.12 for the descriptive statistics of all continuous variables within the non-anxiety disorder group.

Table 7.12

Demographic information for the non-anxiety disorder group for the logistic regression

	Mean	*SD*	Range
Age	37.8	16.9	18-77
IQ	18.9	4.7	4-30
TAS-20	53.9	14.1	20-84
CATI	137.5	32.5	56-205
DASS-A	4.5	3.5	0-14
IATS	44.2	13.2	21-79
IAS	75.8	13.5	24-105
NIES	61.0	16.1	20-100

Note. $N = 173$. *SD* = standard deviation. IQ = combined scores from the ICAR Verbal and Matrix Reasoning tests.

Assumptions. VIF values were all under 5. The only correlation coefficient over 0.7 was the correlation between the dummy coded male and female responses ($r = 0.78$ in all regressions). Therefore, multicollinearity was not seen to be an issue. Cook's distance was below 1 for all participants, indicating no singular observation had a high influence over the model. None of the interaction terms between the natural log of the continuous variables and the variables themselves were significant in any of the models, indicating the assumption of linearity to the logit was not violated.

NIES. A logistic regression analysis revealed that the model significantly predicted the likelihood of participants having an anxiety disorder diagnosis ($\chi^2(9) = 95.623$, $p < .001$). The model explained 37% of the variance in the outcome (Nagelkerke $R^2 = .37$). However,

the interaction between NIES scores and an autism diagnosis was not a significant predictor (B = -.04, p = .141).

IAS. A logistic regression analysis revealed that the model significantly predicted the likelihood of participants having an anxiety disorder diagnosis ($\chi^2(9) = 88.390$, $p < .001$). The model explained 35% of the variance in the outcome (Nagelkerke $R^2 = .35$). However, the interaction between IAS scores and an autism diagnosis was not a significant predictor (B = -.01, p = .742).

IATS. A logistic regression analysis revealed that the model significantly predicted the likelihood of participants having an anxiety disorder diagnosis ($\chi^2(9) = 105.74$, $p < .001$). The model explained 41% of the variance in the outcome (Nagelkerke $R^2 = .41$). However, the interaction between IATS scores and an autism diagnosis was not a significant predictor (B = .05, p = .064).

Summary. The interaction between autism diagnoses and interoceptive accuracy, attention and evaluation did not predict anxiety diagnoses.

Exploratory analysis

The following analyses were not pre-registered but were included as exploratory analyses.

Is the interaction between interoceptive attention and evaluation associated with anxiety? It is possible that interoceptive evaluation may only exert an influence on one's anxiety if one pays attention to interoceptive signals. Therefore, the interaction between interoceptive attention and evaluation may be predictive of anxiety. We explored both whether this interaction term was predictive of trait anxiety (with a linear regression on the sample of 519 participants, see Table 7.2) or an anxiety disorder diagnosis (with a logistic regression, on the 149 participants with an anxiety disorder and the 235 participants in the no

anxiety disorder group, see Tables 7.8 and 7.10). Both analyses included all the covariates listed previously. Predictors were IATS and NIES scores and their interaction.

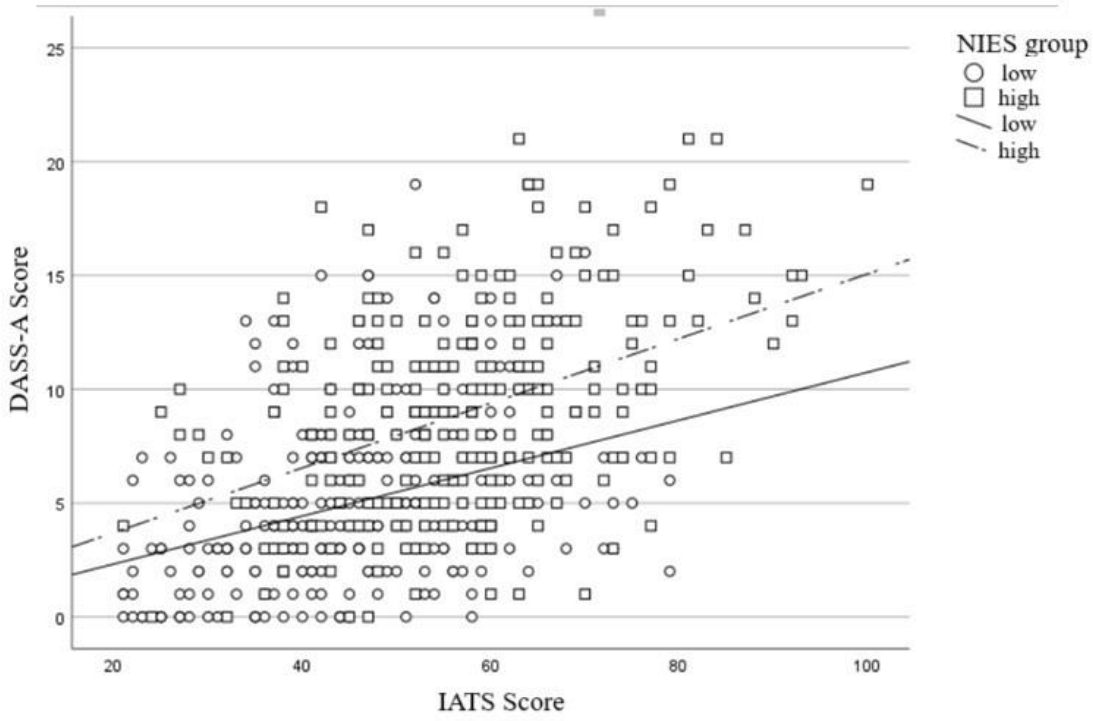
Anxious traits and the interaction between interoceptive attention and evaluation.

When conducting assumption checks, all continuous variables (control variables and the independent variables) were found to have linear relationships with the dependent variables based on scatter plots. VIF values were all under 5 and multicollinearity was not seen to be an issue. The P-P plots indicated a normal distribution of residuals. The scatterplot of residuals versus predicted values indicated that the assumption of homoscedasticity was met. The Durban-Watson statistics indicated independence of observations for each of the regressions (1.95). Cook's distance was below 1, suggesting that no data points exerted a significant influence on the regression line. Therefore, all the assumptions for the multiple linear regressions were met. A multiple regression was conducted to predict DASS-A scores from IATS and NIES scores, and the interaction between IATS and NIES scores, when controlling for the previously listed covariates. The interaction term between mean-centred IATS and NIES scores was a significant predictor of DASS-A scores ($\beta = .11$, $p = .002$). NIES and IATS scores were also both predictors independently ($\beta = .25$, $p < .001$ and $\beta = .25$, $p < .001$, respectively).

A graph was plotted to better understand what this interaction looks like. NIES scores were median split, creating a "low" and "high" group. Scatter plots were plotted to demonstrate the relationship between IATS scores and DASS-A scores in both the low and high NIES groups (see Figure 7.3). Overall, this confirms that in those with higher negative interoceptive evaluation, there is a stronger relationship between interoceptive attention and trait anxiety.

Figure 7.3

A graph to show the effect of the interaction between interoceptive attention and evaluation on anxious traits



Note. The low NIES group where those with a score lower than the median value. The high NIES group were those with a score higher than the median value.

Anxiety disorders and the interaction between interoceptive attention and evaluation. All assumptions of the logistic regression were met (see Analysis in Methods for a breakdown of these assumptions). The interaction term between the mean-centred IATS and NIES scores was not a significant predictor of the presence of anxiety disorders in participants (B = .00, p = .993). In this model, neither were IATS or NIES scores independently (B = .02, p = .116 and B = .01, p = .279, respectively).

Anxious traits with items related to bodily sensations removed. The measure of trait anxiety used in this study (the DASS-A) includes questions pertaining to interoceptive

sensations, which could artificially inflate associations between interoception and anxiety. To explore whether this may have affected our results, we repeated the analyses which included anxious traits as an outcome variable using the total DASS-A scores, minus the items which referred to internal bodily signals ("I experienced breathing difficulty" and "I was aware of the action of my heart in the absence of physical exertion"). There were two further questions that referred to other bodily sensations ("I was aware of dryness of my mouth" and "I experienced trembling"), but as these could be detected via exteroceptive pathways, they were not removed. It is important to note that this version of the DASS-A subscale has not been validated, which is why this is only included as an exploratory analysis. The linear regressions were run exactly as specified in the "Trait Level Analyses" section on the sample of 519, but with the updated DASS-A scores.

Assumption checks. All continuous variables (control variables and the independent variables) were found to have linear relationships with the dependent variables based on scatter plots. VIF values were all under 5 and the only correlation coefficient over 0.7 was the correlation between the dummy coded male and female responses (r= 0.77 for all models). Therefore, multicollinearity was not seen to be an issue. The P-P plots indicated a normal distribution of residuals. The scatterplot of residuals versus predicted values indicated that the assumption of homoscedasticity was met. The Durban-Watson statistics indicated independence of observations for each of the regressions (1.96 for the NIES regression, 1.98 for the IAS regression, and 1.83 for the IATS regression). Cook's distance was below 1 in all the regressions, suggesting that no data points exerted a significant influence on the regression line. Therefore, all the assumptions for the multiple linear regressions were met.

NIES regression. This model statistically significantly predicted NIES ($F(8, 510) = 38.781$, $p < .001$, $R^2 = .378$). Of all the variables, those that were significant predictors of higher NIES scores were: not being male ($\beta = -.195$, $p < .001$), the presence of another mental

health condition (β = .084, *p*=.028), higher CATI scores (β = .291, $p < .001$), and higher DASS-A scores (β = .295, $p < .001$).

IAS regression. This model statistically significantly predicted IAS ($F(8, 510) = 23.711$, $p < .001$, $R^2 = .271$). Of all the variables, those that were significant predictors of higher IAS scores were: higher age (β = .126, *p* =.002), being male (β = .132, *p* = .033), lower CATI scores (β = -.130, *p* = .017), and lower TAS-20 scores (β = -.392, $p < .001$).

IATS regression. This model statistically significantly predicted IATS ($F(8, 510) = 24.400$, $p < .001$, $R^2 = .277$). Of all the variables, those that were significant predictors of higher IATS scores were: lower age (β = -.113, *p* = .003), higher CATI scores (β = .229, $p < .001$), and higher DASS-A scores (β = .355, $p < .001$).

Overall, the significant relationships identified were no different to those identified when the total DASS-A scores were used.

Attention to anxiety-related and non-anxiety-related signals and autism. We wanted to explore whether there was empirical support for Trevisan and colleagues' claim that attention to anxiety-related signals might be heightened in autistic compared to non-autistic individuals, whilst attention to non-anxiety-related signals would be reduced (Trevisan, Mehling, et al., 2021). We split the IATS into questions pertaining to anxiety-related signals (heart beating fast, breathing fast, feeling nauseated, feeling pain from an injury, and feeling pain not from an injury) and all other signals, and created total scores for each. It is important to note that the factorisation of the IATS to form these sub-scales has not been validated, however, as this analysis is only exploratory it was deemed acceptable. For this analysis, we compared the 232 autistic and 143 non-autistic participants (Table 7.4 and Table 7.6).

Histograms indicated that although not all the dependent variables were normally distributed (the attention to anxiety-related signals was positively skewed), the deviation from normality was not such that ANCOVAs were deemed inappropriate. Scatter plots showed that within each group (autistic and non-autistic), there was a linear relationship between the continuous covariates and each of the dependent variables (attention to anxiety-related signals and attention to non-anxiety-related signals). There was no issue with multicollinearity, as again, all VIF values were under 5. Breusch Pagan tests indicated no heteroscedasticity for the non-anxiety-related ANCOVA ($\chi^2(1) = .080$, $p = .777$), although there was significant heteroscedasticity for the anxiety-related ANCOVA ($\chi^2(1) = 9.742$, $p = .002$). Levene's test indicated there was homogeneity of variance (anxiety-related: $F(1, 373) = .018$, $p = .893$; non-anxiety-related: $F(1, 373) = .493$, $p = .483$). The homogeneity of regression slopes was tested by re-running each model including the interactions between the dependent variables and covariates, and as none of the interaction terms were significant, the use of ANCOVA was deemed appropriate.

There was no significant difference in the IATS anxiety-related signals scores between participants with and without validated autism diagnoses when controlling for covariates ($F(1, 366) = 0.26$, $p=.611$). There was also no significant difference in the IATS non-anxiety-related signals scores between participants with and without validated autism diagnoses when controlling for covariates ($F(1, 366) = 5.57$, $p=.019$).

Negative interoceptive evaluation of anxiety-related signals and anxiety disorders. Finally, we wanted to explore whether negative interoceptive evaluation of anxiety-related signals differed significantly between participants with and without anxiety disorders. For this, we used the score of the "evaluation of anxiety-related" signals factor of the NIES that was identified in Chapter Four. We compared the 149 participants with an anxiety disorder and the 235 participants in the no anxiety disorder group (Tables 7.8 and

7.10). Assumption checks demonstrated that some assumptions had been violated. Namely, Breusch Pagan tests indicated heteroscedasticity ($\chi^2(1) = 5.90$, p=.015) and Levene's test indicated there no homogeneity of variance ($F(1,382) = 5.49$, p=.020). However, as this was an exploratory analysis, the decision was made to continue with the planned analysis despite these violations.

A one-way ANCOVA was conducted to examine the effect of the presence of an anxiety disorder on NIES anxiety-related signals scores, while controlling for covariates. There was a significant difference in the NIES anxiety-related signals scores between participants with and without validated anxiety disorders when controlling for covariates ($F(1,375) = 5.57$, p=.019). Partial eta squared indicated the effect size was small (0.15).

Discussion

The aim of this study was to gain a better understanding of how different self-reported dimensions of interoception relate to both autism and anxiety. Consistent with our first hypothesis, results indicated that higher anxious traits were predictive of higher self-reported negative interoceptive evaluation and interoceptive attention scores, but were unrelated to interoceptive accuracy scores, once relevant covariates were controlled for. The former two findings are consistent with the findings of a recent meta-analysis (Clemente et al., 2024), demonstrating a fairly robust effect regardless of the use of different measures of interoceptive accuracy and evaluation. Given that the role of cross-domain, subjective interoceptive accuracy in anxiety had not previously been explored, there are no prior findings to compare our results to. However, results are consistent with our findings of a lack of a relationship between objective measures of cardiac interoceptive accuracy and anxiety (see Chapter Four), and the mixed findings when considering confidence ratings of task performance in objective measures of interoception (Harrison et al., 2021; Garfinkel, Tiley et

al., 2016). Interestingly, we also found that the interaction between interoceptive attention and negative interoceptive evaluation was significant when predicting trait level anxiety, suggesting that those with higher self-reported interoceptive attention demonstrate a stronger relationship between negative interoceptive evaluation and trait level anxiety. This relationship was confirmed with a plotted graph.

As predicted in our second hypothesis, with relevant covariates controlled for, higher autistic traits were predictive of higher negative interoceptive evaluation. This is consistent with the frequent reports of negative interoceptive evaluation identified in our qualitative study of autistic adolescents (see Chapter Six), suggesting this may be applicable across both adolescence and adulthood in autism. However, in contrast to what was predicted in our hypothesis, higher autistic traits were also predictive of lower interoceptive accuracy and higher interoceptive attention. The finding of an association between higher autistic traits and higher interoceptive attention appears to contrast Yang and colleagues findings of a negative correlation between autistic traits and interoceptive awareness in the general population (Yang et al., 2023). However, as discussed in the introduction, this study used a total score of all the MAIA subscales to calculate interoceptive awareness, which conflates multiple independent dimensions of interoception.

Our third hypothesis was that participants with diagnosed anxiety disorders would show more negative interoceptive evaluation, a greater degree of interoceptive attention, and comparable interoceptive accuracy after controlling for relevant covariates. Our results provided partial support for this hypothesis. Self-reported interoceptive accuracy was not significantly different between participants with and without anxiety disorders when other relevant covariates were controlled for, and self-reported interoceptive attention was higher in those with diagnosed anxiety disorders when compared to those with no anxiety disorders or severely anxious traits. However, we found no evidence that negative interoceptive

evaluation was significantly different between these two groups. This is surprising given the role ascribed to the negative interpretation of anxiety-related bodily signals in cognitive behavioural accounts of anxiety disorders (Clark, 1986) and given our findings when looking at the relationship to anxious traits.

There are some potential explanations for the lack of an observed association between interoceptive evaluation and diagnosed anxiety disorders. First, it may be that the inclusion of several covariates masked any potential effects of anxiety disorders on interoceptive evaluation. Consistent with this is the fact that our descriptive t-tests showed that the anxiety disorder group had significantly higher NIES scores than the non-anxiety disorder group. Secondly, it might be that only evaluation of *anxiety-related* signals relates to anxiety disorders, which may explain why measures such as the ASI (Physical Sub-scale) have shown significant differences between anxious and non-anxious participants (Deacon & Abramowitz, 2006; Wheaton et al., 2012), whilst our measure which captures both anxiety-related and non-anxiety-related interoceptive signals showed a non-significant effect. Consistent with this is the fact that results were significant when considering only the evaluation of anxiety-related signals, although caution should be taken when interpreting these results, given violations of the ANCOVA assumptions. Another explanation is that we observed a relationship between anxious traits and interoceptive evaluation, but not anxiety disorders, because of a degree of conceptual overlap between measures of anxiety and interoceptive evaluation. However, given that the same association was found when the interoception related questions were removed from the total DASS-A scores, this seems unlikely. This will be discussed further in the limitations section. Interestingly, we also found no evidence that negative interoceptive evaluation may only cause anxiety in individuals who pay a high degree of attention to interoceptive signals, as the interaction term between interoceptive attention and evaluation was not significant in predicting the presence of an

anxiety disorder. Therefore, our results do not support the explanation that a lack of an association between negative interoceptive evaluation and anxiety disorders is due to the relationship only existing in a portion of the sample with higher interoceptive attention. A final explanation is that certain anxiety disorders may be more related to negative interoceptive evaluation than others. We were unable to explore this in this study, as the majority of our anxiety disorder sample had generalised anxiety disorder (N = 122).

Our fourth hypothesis was that autistic participants would show more negative interoceptive evaluation, comparable interoceptive attention, and comparable interoceptive accuracy when compared to non-autistic participants, after accounting for relevant covariates. Our results supported this hypothesis in its entirety. Our results suggest an alternative way of interpreting the suggestions made by Trevisan and colleagues (Trevisan, Mehling, et al., 2021). They suggest that "maladaptive" interoceptive attention to signals that cause anxiety may be higher in autism, whilst "adaptive" attention to neutral signals may be lower. In our exploratory analyses, we did not find support for this, as attention to only anxiety-related signals was no different between autistic and non-autistic participants when relevant covariates were controlled for. However, what Trevisan and colleagues label as "maladaptive attention" could actually be the more negative interoceptive evaluation we found in autistic individuals in our study. As discussed in Chapter One, a more negative evaluation of certain signals may cause an individual to focus their attention to this signal more, as they are vigilant to it. It may also be that there is no heightened maladaptive attention in autism, but that the questionnaires reviewed by Trevisan and colleagues captured negative interoceptive evaluation as well as attention.

There was limited support for our fifth hypothesis that autistic participants would have different distribution in interoceptive accuracy and attention scores, driven by a greater frequency of both high and low responses. Although there were significant differences in the

distributions between autistic and non-autistic participants for each of the three dimensions of interoceptions, the significance here could have been driven by the large sample size. When visually inspecting histograms, although there were slight differences in the skewness of the IAS and NIES scores, they did not show evidence of a greater proportion of high and low scores in autistic compared to non-autistic participants. This initially seems at odds with what one might expect if interoception did fit in with sensory accounts of autism that were initially modelled on exteroceptive senses (as suggested by Proff et al., 2022). These sensory accounts suggest that both hyper and hyposensitivity may be present in autism, which could manifest at a clustering of scores across each end of the distribution. However, hyper and hypo sensitivity to interoceptive information may still occur in autism, without manifesting in these differences in accuracy and attention distributions, if both hyper and hypo awareness occur in the same individual, but within different contexts or across different interoceptive domains. For example, an autistic individual may experience hypersensitivity to their heart beating in anxiety-inducing situations but feel hyposensitivity to needing to use the bathroom. The high and low scores on these individual items would cancel out in the total score. It would be interesting for future research to explore this. The same pattern of significant differences in distribution triggered mainly by differences in skewness was found when conducting exploratory comparisons between the distributions of anxious and non-anxious participants.

We did not find support for our sixth hypothesis, that autistic traits would moderate the association between anxious traits and negative interoceptive evaluation. Indeed, none of the tested interaction terms were significant, meaning that the interactions between interoceptive attention and accuracy with autistic traits also did not significantly predict anxious traits. This indicates that the effects of these three dimensions of interoception on anxious traits is consistent regardless of the degree of one's autistic traits. Similarly, none of the interactions between the three dimensions of interoception and the presence of an autism

diagnosis were significant when predicting the presence of an anxiety disorder diagnosis, suggesting that the associations between self-reported interoception and anxiety do not differ between autistic and non-autistic individuals.

There are some limitations of this study that it would be pertinent to address. Previously, caution has been advocated when interpreting results using self-reported measures of interoception, as, for example, self-reported interoceptive accuracy may not correspond to objectively measured interoceptive accuracy during a task. This is only of concern in this study if one assumes that only objective interoceptive accuracy, attention and evaluation are relevant to mental health outcomes. It is plausible, and indeed our results support, that subjectively perceived interoceptive dimensions are also relevant to mental health outcomes.

In this study, participants were not required to validate diagnoses of mental health and neurodevelopmental conditions other than anxiety disorders, autism, and eating disorders, meaning that there was a possibility of participants self-reporting conditions that they had not received a formal diagnosis for. Moreover, despite the validation checks for anxiety disorder, autism, and eating disorder diagnoses, participants may have still fabricated responses to these questions, meaning they could have passed validation checks without having a formal diagnosis. However, it is worth noting that as all participants were compensated at the same rate, there was no incentive to report a diagnosis if the participant did not have one.

There was also some concern about a degree of construct overlap in how interoceptive attention and anxious traits were measured, as the DASS-A includes questions pertaining to the noticing of bodily signals (e.g.: “I was aware of the action of my heart in the absence of physical exertion (e.g. sense of heart rate increase, heart missing a beat)”). To some degree this is unavoidable, given that physical sensations are included in definitions of anxiety.

Additionally, the DASS-A only considers the noticing of two anxiety-related and negatively connoted bodily sensations (difficulty breathing and increasing heartrate), whilst the IATS considers twenty-one signals that are not all anxiety-related. Moreover, there were no issues with multicollinearity in any of our checks, and an exploratory analysis when these items were removed from the total DASS-A scores revealed the same results as analyses including those items. However, given the unexpected discrepancy between our findings when looking at anxious traits versus anxiety diagnoses, it would be useful for further work to be done using alternative measures of anxious traits that are less conflated with interoceptive evaluation, such as the Hospital Anxiety and Depression Scale (Zigmond & Snaith, 1983).

Moreover, the group comparisons of this study (anxiety disorder versus no anxiety disorder and autistic versus non-autistic) do not reflect the considerable heterogeneity one would expect to see within groups. For example, Chapter Six of this thesis confirmed the heterogeneity of experiences of interoception in autistic adolescents. Whilst it is useful to understand how groups differ on average, it is important to understand that there is likely a high degree of variability within this.

Despite these limitations, our pre-registered study contributes to an underexplored area in the literature by examining the role of subjectively perceived interoceptive accuracy, attention and evaluation in anxiety and autism, at both the trait and diagnosis level. The study was conducted on a large sample, and several potential confounding variables, such as alexithymia, were controlled for.

In conclusion, our findings support the role of interoception in autism and anxiety. Higher self-reported interoceptive attention was related to higher anxious traits, the presence of an anxiety disorder, and higher autistic traits. Self-reported interoceptive accuracy was unrelated to either measure of anxiety or autism diagnosis, but lower interoceptive accuracy

was associated with higher autistic traits. Negative interoceptive evaluation was associated with both measures of autism, as well as anxious traits, however it showed no association with diagnosis of an anxiety disorders. No interaction effects between autism and any of the dimensions of self-reported interoception on anxiety were found.

Chapter Eight: General Discussion

The aims of this thesis were twofold: first, to illustrate the limitations of current interoception research methods and to provide some solutions, and second to enhance understanding of the role of interoception in anxiety and autism. To address the former, I began with a methodological critique of the most commonly used measures within the interoception literature, including suggestions for how their limitations may be overcome (Chapter Two). I then presented two new measures of interoception that avoid some of the limitations discussed in Chapter Two. The first was The Phase Selection Task (PST) of cardiac interoceptive accuracy (Chapter Three), which I ran a small-scale pilot study on. The second was the self-report Negative Interoceptive Evaluation Scale (NIES) questionnaire, which I conducted validity checks on in two large samples (Chapter Four).

The first study of my second section employed a meta-analysis to establish the current academic understanding of the relationship between anxiety and the most studied aspect of interoception, cardiac interoceptive accuracy (Chapter Five). Given the null but heterogenous findings, the next step was to use rich, qualitative data from interviews with autistic adolescents who experience significant anxiety to provide a guide as to which areas of interoception would most benefit from future research, and to form hypotheses that could be tested by quantitative work on larger samples (Chapter Six). Finally, these hypotheses were tested by a large-scale questionnaire study of the associations between autism, anxiety and three self-reported dimensions of interoception (interoceptive accuracy, attention, and awareness) in adults (Chapter Seven).

This General Discussion will summarise the findings of this thesis by integrating them into overall conclusions.

Conclusion one: The need to employ caution when interpreting previous interoception research

The literature review in Chapter One and the meta-analysis in Chapter Four illustrated the lack of replicability of findings across interoception research, especially when considering research into anxiety and autism. One potential reason for this, as discussed in Chapter Two, is the lack of validity of the most used measures of interoception, and the fact that numerous measures intended to assess the same construct capture very different aspects of interoception. This results in unreliable conclusions. For instance, several self-report measures of interoception claim to measure the ill-defined construct of "interoceptive awareness," although research suggests that the different questionnaires actually measure separable constructs. Similarly, both the heartbeat counting task (HCT) and heartbeat detection task (HDT) are intended to measure cardiac interoceptive accuracy, however, construct validity is impeded by the likelihood of false positives in the HCT, and false negatives in the HDT. This allows for a lower degree of concordance in the participants deemed to have cardiac interoceptive accuracy than one would expect if they truly measured the same construct, which could affect the conclusions drawn depending on which measure was used.

The reason we should be concerned about this can be illustrated with an example. If a researcher found an association between performance on the HCT and anxiety, they may conclude that reduced cardiac interoceptive accuracy leads to increased anxiety. As a result, they may decide to develop interventions to enhance cardiac interoceptive accuracy to reduce anxiety. However, given the numerous problems with the HCT, the conclusion that they are basing this intervention on could be false. It may simply be that participants with anxiety are less likely than participants without anxiety to "cheat" on the HCT by counting the number of seconds and then estimating their heart rate. If the conclusion was false, the best-case

scenario is that the development of such an intervention would be a waste of resources. The worst-case scenario is that such an intervention may have iatrogenic effects. Similarly, a clinician may attempt to establish whether or not a patient is interoceptive and therefore whether interoceptive training would be beneficial for them. The HDT may indicate the patient is not interoceptive and they may undergo training to enhance interoceptive accuracy. However, it is possible that this is a false negative, and that the training would not be beneficial as the individual can already detect their heartbeat.

The solutions here are twofold. Firstly, the field needs to move towards a unified and specific nomenclature. Despite the need for this having been recognised previously (Khalsa et al., 2018), significant progress still needs to be made. When measuring a specific dimension of interoception, authors should clearly define that dimension and address how it has been measured in different ways previously. Vague terms such as "interoceptive awareness" should be avoided unless the authors clearly define what they mean by such. Rather than describing, for instance, the Multidimensional Assessment of Interoceptive Awareness (MAIA) as a general measure of interoceptive awareness and summing a total score, each of the individual sub-scales should be used, and conclusions should be worded with reference to the *specific* dimensions of interoception that each sub-scale captures. A failure to correctly define terms allows for fishing, whereby several different ways of assessing "interoceptive awareness" or other such vague terms are used, and authors only report analyses using measures which have led to significant results. For example, if a researcher hypothesised that interoceptive awareness was associated with anxiety, and correlated anxiety with scores on three different questionnaire measures (for instance, the MAIA, the Body Perception Questionnaire (BPQ), and the Body Awareness Questionnaire (BAQ)), yet only reported a significant correlation between the BPQ and anxiety measures.

The second solution is that the field should move toward the use of more validated measures. Despite many researchers having illustrated similar concerns with commonly used measures of interoception in the past (Desmedt, Heeren, et al., 2022; Desmedt, Luminet, Maurage, et al., 2023; Ring & Brener, 2018), problematic measures continue to be used. Such a reliance on demonstrably invalid measures impedes scientific progress and wastes funding. With this thesis providing clear evidence that the HDT task results in false negatives due to the use of predefined delays, one would hope that researchers stop using this problematic measure, especially with convenient and valid alternatives becoming more readily available. One such alternative would be the PST (Chapter Three), which shows promise as a measure of cardiac interoceptive accuracy, or the Phase Adjustment Task (PAT), which has already undergone further validation checks.

Conclusion Two: Interoceptive accuracy may not be as important as previously thought

Most early theories of the role of interoception in anxiety focused on the role of interoceptive accuracy (e.g.: Ehlers, 1993), and as a result, most research has focused on the same. However, as demonstrated by the meta-analysis reported in Chapter Four, empirical evidence fails to support an association between objectively measured cardiac interoceptive accuracy and anxiety. We found no evidence for an association between performance on variants of either the HCT or HDT and state or trait anxiety (trait anxiety also included the presence of anxiety disorders). Importantly, there was no moderating effect of factors such as task instructions, anxiety measures, age or gender. We suggested several potential reasons for this, one of which was that there was simply no relationship between cardiac interoceptive accuracy and anxiety. In support of this idea is the fact that novel measures of cardiac interoceptive accuracy that do not have the same limitations as the HCT and HDT also find no association with anxiety (Plans et al., 2021).

The qualitative approach employed in Chapter Six allowed participants to bring up aspects of interoception that they found significant in explaining their experiences of autism and anxiety, meaning that the dimensions explored were not limited by what research had typically focused on in the past. Findings were interesting, as accuracy did not appear to be something that participants reflected on in isolation. Rather than discussing whether they felt their perception of certain signals was accurate, participants focused more on whether they could adaptively attend to bodily signals (i.e., noticing when there was something they needed). This concept could theoretically include interoceptive accuracy, in that detecting a homeostatic need does require an accurate detection of somatic signals. However, participants spoke more about the amount of time they spent attending to these signals and whether they could notice them, rather than whether they thought they were correct in their classification when they did notice them.

The final study in Chapter Seven supported the idea that participants' perceptions of their interoceptive accuracy might not be central to anxiety or autism. Neither anxious traits nor anxiety disorders were related to self-reported interoceptive accuracy. Although autistic traits were associated with lower self-reported interoceptive accuracy, there was no association with autism diagnoses. This suggests that neither anxiety nor autism is related to perceived, cross-domain interoceptive accuracy.

Conclusion Three: Importance of interoceptive evaluation

The assessment of the NIES in Chapter Four provided novel insights into interoceptive evaluation- a previously neglected construct. Whilst we were not able to determine a clear factor structure to establish the different components of interoceptive evaluation, there was evidence for a distinction between the evaluation of anxiety and non-anxiety-related signals, although both did appear to be related to the same underlying

construct. This suggests that interoceptive evaluation may exist as a latent variable, but one which is separable according to the interoceptive domain which is being assessed and whether or not it is related to anxiety. This distinction between anxiety-related and non-anxiety-related signals has been previously reported when considering "interoceptive awareness" (Vlemincx et al., 2023). This separability is also somewhat intuitive, as one might expect that the evaluation of signals that could indicate an anxiety response is different to the evaluation of signals that would indicate something more neutral. This separability may be especially important when considering associations between mental health and negative interoceptive evaluation. Preliminary evidence for this can be found in Chapter Seven, where the anxiety-related signals sub-scale was found to be associated with anxiety disorders, whereas total negative interoceptive evaluation was not.

In Chapter Four, an association was observed between the presence of mental health or neurodevelopmental conditions and negative interoceptive evaluation. I built upon this finding in later chapters of my thesis. Qualitative findings derived from interviews with autistic adolescents in Chapter Six highlighted the importance of negative interoceptive evaluation in participants' experiences of anxiety specifically, and quantitative findings in Chapter Seven showed that interoceptive evaluation appeared related to anxious traits, autistic traits, and the presence of an autistic diagnosis. Interestingly though, there was no association with the presence of an anxiety disorder (reasons for this were discussed in detail in Chapter Seven). As discussed in both of these chapters, our findings were consistent with the small corpus of existing empirical work.

The interviews in Chapter Six also provided a rich understanding of the way negative interoceptive evaluation would be experienced and the exact mechanisms by which it influenced anxiety. Participants discussed catastrophising about interoceptive signals, whereby the noticing of a signal would trigger anxiety as it was interpreted as a sign that

something was wrong. For some participants, this would even happen when they were able to reason that the signal was occurring for non-dangerous reasons. For others, the signals themselves were felt as aversive or overwhelming and triggered general negative affect. To speculate, this may be because of associations formed over time between certain signals and negative consequences, such as meltdowns or shutdowns occurring after participants experienced their hearts beating fast. Our interviews also suggested that different domains of interoception varied in their likelihood of being evaluated negatively, with sensations pertaining more directly to the physical sensations associated with anxiety (e.g., increasing heart rate and faster breathing) more commonly reported as being evaluated negatively. This substantiates the distinction outlined in Chapter Four of the thesis.

The mechanism outlined here is similar to that proposed by Clark and colleagues in their models of cognitive behavioural therapy (CBT; Clark, 1986), which suggests that an overly negative interpretation of somatic sensations would result in anxiety. Our findings substantiate these theories and suggest they apply to autistic populations. However, it is worth noting that this relationship is not necessarily stronger in autistic compared to non-autistic individuals, as indicated by the lack of an interaction effect identified in Chapter Seven.

Moving beyond the empirical work of this thesis to speculate further on the potential mechanisms of effect, negative interoceptive evaluation may be a moderator through which other dimensions of interoception exert their influence. This idea has been touched on briefly previously in this thesis, when considering how interoceptive attention may interact with negative interoceptive evaluation to cause anxiety. It may be that by itself, a higher-than-average degree of interoceptive attention is not enough to cause anxiety if one has adaptive interpretations of the signals that one is attending to. However, if one has a more negative interpretational bias, then the more one pays attention to these signals, the more opportunity there is for catastrophising. We found some evidence to support this in the exploratory

analyses presented in Chapter Seven, whereby the interaction between interoceptive accuracy and evaluation was predictive of trait level anxiety (but not of anxiety disorders). It also may be that other dimensions of interoception can moderate the effects of negative interoceptive evaluation. For instance, negative interoceptive evaluation may only result in anxiety in those who believe they are accurate at detecting internal sensations, whilst those who do not think they are accurately detecting signals will simply disregard them.

In summary, negative interoceptive evaluation appears to be an extremely important construct that has not yet received as much research attention as interoceptive accuracy or attention. This thesis suggests it is linked to both autism and anxiety, as well as mental health and neurodevelopmental conditions more broadly.

Conclusion Four: Relationship between anxiety and interoception

A key finding of this thesis, highlighted in particular in Chapter Six, is that there are multiple routes by which interoception can affect anxiety. The specific pathway involved will depend on both the person in question and contextual factors. Whenever interoception deviates from what is "adaptive," regardless of the direction of this deviation, anxiety may be triggered. Therefore, if one has atypically high *or* low interoceptive attention, it could be problematic.

As discussed in Chapter Six, high interoceptive attention allows for more instances where interoceptive signals could be interpreted negatively, inducing anxiety. If all signals are attended to in a heightened manner, this could also result in alexithymia, as it becomes more difficult to determine which signals are important to understanding one's emotional state amidst all of the noise. In autistic populations specifically, the signals may also be overwhelming, distracting, or difficult to switch off from, in a manner which causes anxiety. On the other hand, low interoceptive attention may result in a failure to meet homeostatic

needs and may also negatively impact emotion recognition and therefore regulation (i.e., influencing mental health through alexithymia). Theoretically, overly positive interoceptive evaluation may lead to anxiety, if individuals fail to recognise high arousal as an indication that there is something wrong in instances where there is something wrong. This may prevent them from removing themselves from stressful situations until it is too late. However, we did not find evidence of this particular route. Instead, we only found evidence that negative interoceptive evaluation caused anxiety.

Despite the potential for different routes, on average, anxious individuals appear to show different interoceptive profiles to non-anxious individuals. This suggests that although deviation from the norm in either direction can result in anxiety, deviations in one direction are more common than the other in anxious individuals. For example, greater self-reported interoceptive attention is associated both with trait anxiety and anxiety disorders, and greater self-reported negative interoceptive evaluation is associated with trait anxiety (Chapter Seven). This is consistent with prior work which has illustrated increased attention and more negative evaluation in those with anxiety disorders (Anderson & Hope, 2009; Clark, 1986; De Berardis et al., 2007; Deacon & Abramowitz, 2006; Ludewig et al., 2005; Olatunji et al., 2007). Despite the potential for both higher and lower interoception to cause anxiety, the distribution of scores within the anxiety disorder sample of Chapter Seven was still unimodal and close to normally distributed, rather than reflecting a population comprised of two subsets (one with very high scores, and one with very low scores).

In instances where higher interoceptive attention has been linked to anxiety, it may be that signal strength is amplified during interoceptive processing in anxious individuals compared to non-anxious individuals, meaning that an interoceptive signal that is objectively the same intensity may result in an amplified response. For instance, a heart rate increase of 10% may result in a greater degree of activation of interoceptive pathways in anxious versus

non-anxious individuals. According to the James Lange theory of emotion, whereby emotional experience is driven by bodily sensations (Lange & James, 1967), this may allow for the same increase in heart rate to cause a greater degree of experienced anxiety in those with anxiety disorders. There is early evidence for this, as heartbeat-evoked potential (HEP) responses are heightened in generalised anxiety disorder (GAD) patients compared to non-anxious individuals when infused with isoproterenol and saline solutions (Verdonk et al., 2024), although HEP responses did not correlate with functional magnetic resonance imaging (fMRI) signals in regions implicated in cardiac signal processing in prior fMRI studies, making results difficult to interpret.

As mentioned previously, there is significant potential for bidirectional associations between anxiety and interoception. In the case of interoceptive attention, if one is anxious, this may cause hypervigilance to any signals that might signify a threat, resulting in body monitoring. In the case of panic disorder (PD), one can imagine how, if a person had experienced a panic attack previously, they may focus on their heartbeat more in the future in anticipation of another panic attack occurring. This does not necessarily have to occur alongside negative interoceptive evaluation, as someone may be adaptively interpreting these signals (i.e., not perceiving it as a sign something is wrong in cases when there truly is nothing wrong), whilst still paying attention to the signals for a higher proportion of the time "just in case."

It might also be that those with anxiety are more likely to experience signals that do capture attention, such as rapidly increasing heart rates or laboured breathing, as symptoms of their anxiety. However, research evidence is mixed. Some studies have found that although those with GAD report a higher degree of attention to interoceptive signals than those without an anxiety disorder, their actual recordings of heart rate and respiration were not significantly different across groups (Hoehn-Saric et al., 2004; Olatunji et al., 2007). Similarly, those with

high anxiety sensitivity showed no difference in heartbeat reactivity to a stressor than those with low anxiety sensitivity (it should be noted, however, that the sample size was very small in this study; Stewart et al., 2001). Contrastingly, another study found that participants with PD and GAD (but not social anxiety disorder (SAD)) showed elevated heart rate responses to hyperventilation (Pittig et al., 2013). Given that we used self-report measures of interoceptive attention, there is the possibility that results were influenced by participants perceiving a greater degree of sudden changes in these signals. It would be interesting to conduct further work to assess the association between subjective attention and objectively measured interoceptive signals in anxious and non-anxious participants (see future directions section).

Conclusion Five: Relationship between autism and interoception

As illustrated in Chapter Six, the relationship between autism and interoception is far from straightforward. Interoception varies across autistic individuals in the same manner as in the general population, and in that sense, one should not presume that because someone is autistic, they are any more or less likely to show a certain profile of interoception. This is consistent with what we know of autism as highly heterogeneous.

Despite this variation, it does appear that autistic individuals may be more likely to have a more negative interoceptive evaluation on average than neurotypicals. Furthermore, autistic traits are associated with more negative interoceptive evaluation, reduced self-reported interoceptive accuracy, and heightened self-reported interoceptive attention (Chapter Seven).

Interoception is perhaps best understood within existing sensory accounts of autism, as suggested by Proff and colleagues (Proff et al., 2022) and outlined in Chapter One of this thesis. The findings reported in this thesis are consistent with the notion that autistic individuals can show “extremes” in interoceptive sensory processing in the same way as

exteroceptive sensory processing. Perception of interoceptive sensations may either be heightened or not felt at all. This varies according to the individual, context, and signal in question. Just as with exteroceptive sensory processing, the experience of being aware of certain signals can also be negative. For example, feelings of hunger appear to be an intensely negative experience for some autistic adolescents in the same manner that bright lights or certain smells have also been reported to be (Taels et al., 2023). One autistic individual may find themselves hyperaware of their heartbeat and sounds when they are anxious, whilst they are highly sensitive to scratchy textures all the time regardless of context, and struggle to detect when they need to use the bathroom. Another autistic individual may never be able to feel their heartbeat and find themselves unable to hear their parents calling them when they are focused, however, they are hyperaware of small visual details. This level of complexity can often be obscured in quantitative research, however, Chapter Six's qualitative study provided rich evidence for it.

Proff and colleagues reference several theories of autism that may be linked to the sensory atypicalities that have been observed in interoceptive and exteroceptive domains. One such theory is the weak central coherence theory (Happé & Frith, 2006), which suggests that autism involves increased local processing (in this instance, focusing on low-level sensory information), and reduced global processing (integration of low-level sensory information into the global context). Based on this account, autistic individuals might be expected to demonstrate heightened attention to low-level interoceptive signals like their heart rate, but a reduced ability to combine this with other information to understand what this communicates about their bodily state. Using the framework employed in this thesis, this may suggest heightened interoceptive attention and accuracy associated with autism, in addition to inaccurate interoceptive evaluation. A similar theory employs Bayesian accounts to suggest that autism is characterised by a greater weighting of bottom-up sensory

information compared to top-down prior models (Pellicano & Burr, 2012). This could in turn lead to an attentional bias towards the low-level sensory information and enhance accuracy in scenarios when top-down priors are inaccurate or reduce it when sensory information is noisy. Proff and colleagues also outline their own, integrated account, whereby atypical perceptual inference (e.g., Bayesian account) and weak stimulus binding (e.g., weak central coherence account) interact and reinforce each other, “accounting for the diversity of autistic presentations.” However, neither of these theories or the combination of the two wholly accounts for the complexity observed in this thesis, whereby autistic individuals can demonstrate both heightened and reduced attention to low-level interoceptive signals.

A theory that appears slightly more applicable to the findings reported in this thesis is the monotropism account of autism (Murray et al., 2005; see Chapter Six). This theory can account for both heightened and reduced attention and accuracy, depending on whether interoceptive information comes under an autistic individual's "spotlight" of focus at that moment. If interoceptive information is the subject of attention, the theory of monotropism would suggest that an autistic individual would experience it more intensely than a non-autistic individual and therefore may be more accurate in their detection of it. If interoceptive information is not the subject of attention, autistic individuals may struggle to notice it at all, and autistic individuals may find it harder to purposively shift attention towards it than non-autistic individuals, resulting in both reduced attention and accuracy. Whilst this theory is highly applicable to interoceptive attention and accuracy, it does little to explain why interoceptive evaluation may be more negative in autistic individuals compared to non-autistic individuals, even when anxious traits, the presence of other mental health conditions and other covariates are controlled for.

This thesis has also provided evidence that associations between autism and interoception exist even when controlling for alexithymia. Due to a historical conflation between autism

and alexithymia (see Chapter One), research in the last decade has demonstrated that several traits previously attributed to autism may actually be linked to alexithymia, which creates an artificial association due to the high co-occurrence of autism and alexithymia. These traits include recognition of emotional facial expressions and empathy (Bird et al., 2010; Cook et al., 2013). Some have suggested that interoception may be another example of this, given the known association between recognition of emotions and experiences of bodily sensations (Shah et al., 2016). However, despite alexithymia (as measured by the TAS-20) showing significant associations with self-reported interoceptive attention, accuracy and evaluation, this did not completely account for the association between autism and negative interoceptive evaluation, or autistic traits and all three dimensions, suggesting that relationships exist between certain aspects of autism and interoception above and beyond that which is attributable to alexithymia.

The findings of this thesis also provide further support that interoception might be one of the many factors that can increase anxiety in autistic individuals. As discussed in Chapter One, this is best conceptualised within a wider developmental psychopathology framework, whereby interoception is one, internal factor that can interact with many other internal and external factors to heighten levels of anxiety in certain autistic individuals (Rutter & Sroufe, 2000). The potential for these interactions is illustrated in the following example: an autistic individual may find social settings anxiety-inducing, which results in an increase in heart rate which is then interpreted negatively, causing further anxiety. Chapter Seven suggests the associations between anxiety and interoception are the same between autistic and non-autistic individuals, however, as autistic individuals tended to have more negative interoceptive evaluation, this may cause greater anxiety than in the non-autistic group.

General implications

Interventions

The main implication of the findings presented in this thesis is the potential for interoception-based interventions to alleviate anxiety. The potential for interventions may be especially relevant for autistic populations. This is not necessarily because interoception has a greater relationship to anxiety (as we failed to find evidence for this in Chapter Seven), but more so because autistic individuals appear to have a higher likelihood of having an interoceptive profile that may confer a greater risk of anxiety compared to non-autistics. Additionally conventional therapies for anxiety, such as CBT, have been demonstrated to be slightly less efficacious in autistic compared to non-autistic patients (Spain et al., 2023; van Steensel & Bögels, 2015) and autistic individuals report that existing therapies do not meet their unique and personalised needs (Camm-Crosbie et al., 2019). This may be because interventions such as CBT and talking therapy to a degree rely on patients being able to talk about their feelings. With such high rates of alexithymia in autism (Kinnaird et al., 2019), this might prove a barrier to the therapy's efficacy. However, therapies focused on interoception may provide a double benefit, in that they could address anxiety, and also help reduce alexithymia (Bornemann & Singer, 2017), thus rendering more conventional types of therapies more effective if used in tandem.

Of the existing research on interoceptive interventions, many focus on enhancing interoceptive accuracy by delivering tasks used to measure interoceptive accuracy and combining this with feedback or slightly adapting the task to increase the difficulty gradually over time by reducing the signal. Indeed, one study conducted on autistic populations reported that their intervention focused on improving performance on HCT and HDT tasks through feedback and manipulations to transiently increase cardiac arousal enhanced cardiac interoceptive accuracy over time and reduced anxiety (Quadt et al., 2021). However, this was

potentially confounded by the fact that participants in the interoception-based intervention group also received leaflets with CBT-based material that the control group did not receive, meaning that reductions in anxiety may have been attributable to the leaflet, rather than improvements in interoceptive accuracy. Other studies have demonstrated conflicting results. Although an increase in interoceptive accuracy over time was found, only state and not trait anxiety were reduced in an adult sample (Sugawara et al., 2020). Moreover, interoceptive accuracy might not be the exact mechanism by which anxiety is being impacted in these studies. The interventions in question may have downstream effects on other dimensions of interoception. For example, these tasks may act as a form of interoceptive exposure therapy and thus reduce negative interoceptive evaluation. There is some support for these downstream effects, as participants also demonstrate changes in self-reported interoceptive "awareness" that capture a mix of meta-awareness, attention and evaluation (Quadt et al., 2021). We must explore this further, to understand which specific interoceptive changes might underlie the results that have been observed on anxious traits.

As interoceptive evaluation appears to be a very important construct, it might be that some interventions involve a mix of conditioning and psychoeducation, both of which could reduce overly negative interoceptive evaluation. Conditioning would involve guided exposure therapy to interoceptive signals, so participants learn to decouple the experience of these bodily sensations from anticipated negative consequences. Psychoeducation would involve making participants aware that certain fluctuations in bodily signals are normal and do not always indicate danger. For example, a degree of heart rate variability is normal and nothing to be concerned about. Mindfulness approaches could also be employed, that encourage a noticing of bodily sensations with a neutral interpretation. These interventions could be integrated into existing therapeutic approaches. In particular, given that CBT already involves elements of this approach, it would be straightforward to expand on what is already there.

Early research findings have been promising, illustrating that interoceptive evaluation-based interventions seem to reduce anxiety (see review in Heim et al., 2023). However, these have mainly relied on small samples, with short follow-up times precluding conclusions being made about the longevity of results.

The research included within this thesis draws attention to the fact that a "one-size-fits-all" approach to interoceptive interventions may not only be ineffective, but it could be iatrogenic. This can be illustrated with the example of anxiety. It may be that for Person A their anxiety is exacerbated by a lack of attention to their bodily signals. Person A is unable to recognise when they are becoming stressed until the emotion becomes overwhelming, and it is hard for them to regulate their emotions in adaptive ways. For Person A, enhancing interoceptive attention by encouraging a mindful focus on what they are feeling at a given time may be helpful for them. However, Person B finds themselves focusing on interoceptive signals a lot of the time, and they negatively evaluate these signals. Their experience of signals is often overwhelming and anxiety-inducing. For Person B, an intervention that aimed to enhance attention to interoceptive signals may be anxiogenic and a more effective intervention may be the combined exposure therapy and psychoeducation discussed above. When developing interoceptive interventions, it is of fundamental importance that clinicians or therapists understand exactly what the interoceptive profile of the individual is, and how they feel interoception plays into their experiences with mental health.

Psychoeducation

Beyond professional intervention, if these findings are effectively communicated to those for whom interoception affects their mental health, then they may be able to make more informed decisions to benefit their mental health. For example, one of the participants in the qualitative study in Chapter Six who struggled to recognise interoceptive signals discussed

how after participating in the study, she bought a smartwatch to allow her to better track bodily signals such as her heart rate, allowing her to recognise anxiety at an earlier stage.

Another example of how a knowledge of interoception can help individuals make better choices would be if an individual who struggles to detect when they are hungry starts scheduling their eating at certain points in the day, rather than relying on bodily signals to know when they are hungry. This may lead to their bodily needs being met better, resulting in enhanced mental and physical health. The same example can be applied to thirst.

For those who struggle with negatively interpreting interoceptive signals, they might try gradually exposing themselves to these signals in controlled environments, so over time they can learn that most of the time when they feel these sensations, they are not in any danger. One example of how to do this would be slightly raising one's heart rate by doing some exercise, then gradually building the amount of exercise done over time, so the signal of a raised heart rate becomes more salient.

In the same way that certain interventions may be iatrogenic if employed using a "one-size-fits-all" approach, psychoeducation that suggests that teaching oneself to focus more on interoceptive signals will improve mental health could be problematic for those who already are attending to these signals a disproportionate amount of the time. Communicating to people how interoception may vary, how both non-adaptively high and low interoceptive attention can cause issues, and how to appraise signals more neutrally can avoid this and empower individuals to use interoception to improve their mental well-being.

Strengths of the work reported in this thesis

This thesis sought to address several outstanding questions, and in doing so made novel contributions to the field. Firstly, it explained the heterogenous, and often contradictory, results obtained so far when exploring associations between interoception,

mental health and neurodevelopmental profiles. I argue that the conflicting findings can at least be partially explained through the inconsistent definitions and operationalisations of various dimensions of interoception, as well as the use of measures that lack validity. This provides a necessary degree of clarity and sets the groundwork for future research of a higher quality.

Chapter Two provided the first empirical demonstration of the main limitation of the frequently used yet often critiqued Heartbeat Detection Task in a large sample. Findings across two methods of analyses (plotting the selected delays across the cardiac cycle and calculating whether the most selected delays for each participant were closer to the asynchronous or synchronous conditions of the HDT) were triangulated, thus accounting for the limitations of each. Moreover, an extremely conservative approach was taken, in terms of allowing for tones to either precede or follow the heartbeat, meaning that our results are relevant to even the most generous interpretations of the HDT.

The PST, an alternative measure of cardiac interoceptive accuracy, was then piloted in Chapter Three. The pilot demonstrated that with a few amendments to the delivery of the app, this could be a promising measure of cardiac interoceptive accuracy that could be used as an alternative to the HDT. Feedback was collected for a variety of questions, and the use of in-person or videocall-based design allowed me to watch participants' performances, ensuring no cheating and reducing the likelihood that they encountered problems with the app that, for any reason, they did not report on.

The NIES introduced in Chapter Four addresses the formerly neglected construct of interoceptive evaluation, facilitating further research using two large samples into an area that is already deemed important to mental health at a theoretical level. In developing this measure, interoceptive evaluation was operationalised more clearly than it had been

previously, and its relationship to other dimensions of interoception and physical and mental health was explored. Our findings further substantiated the relationship between interoceptive evaluation and mental health and neurodevelopmental conditions. Moreover, a thorough approach was taken to the assessment of this questionnaire, enhancing confidence in its use.

The meta-analysis in Chapter Five provided a clear picture of current research into the most researched aspect of interoception (cardiac interoceptive accuracy) and anxiety and also allowed for several potential moderators to be explored.

Chapter Six provided an in-depth perspective on the complex relationships between autism, anxiety, and interoception, whereby both heightened and reduced interoceptive accuracy and attention could result in heightened anxiety, through differing mechanisms. It also allowed for research directions to be driven by participants with lived experiences of autism. The study design was driven by thorough public and patient involvement work and much thought was put into making the study as accessible as possible.

The final chapter of this thesis provided further insight into how subjective facets of interoception relate to anxiety and autism when controlling for relevant confounders. Of particular note is the fact that this study controlled for alexithymia, given concerns that the association between alexithymia and interoception may lead to artificial associations being observed between autism and interoception. The approach taken in this study was novel, as most studies employing self-report measures of interoception do not disambiguate between interoceptive accuracy, attention and evaluation, despite research illustrating the separability of these dimensions. The analysis was comprehensive, and some explanations for unexpected findings were explored in further exploratory analyses.

Limitations

There are some limitations of this thesis that it would be pertinent to address. First and foremost, the operationalisation of various clinical groups within this body of work is constrained by conventions in the field. Perhaps the clearest example of this is with autism. Much concern has been expressed about the operationalisation of autism within the field and how it might be gender-biased. Both commonly used measures of autistic traits (Belcher et al., 2023), as well as the diagnostic criteria of autism (Cruz et al., 2024; Lai et al., 2015), have been critiqued for not being representative of how autism may manifest in females or non-binary individuals due to their development based on primarily male samples. Indeed, there is evidence for gender biases impacting the likelihood of a child receiving an autism diagnosis, whereby girls need to present more pronounced autistic traits than boys to receive a diagnosis (Duvekot et al., 2017; Dworzynski et al., 2012). Clinicians may miss signs of autism in females, as they are less educated in how to recognise them, and because females are more likely to employ camouflaging strategies like masking (Lai et al., 2015). I tried to address this in my thesis, by using the Comprehensive Autistic Trait Inventory- a measure of autism which preliminary evidence suggests is less susceptible to gender bias than the autism quotient, and by over-sampling autistic females in my qualitative study. However, in Chapter Seven, participants reported whether or not they had received an autism diagnosis by a clinician- a metric which may be susceptible to gender biases.

Similarly, our results about autism may not be generalisable to everyone within the autistic community. Although there were no explicit requirements about IQ levels when participating in our research, due to some of the complex demands (for example, completing a one-hour questionnaire, or an hour-long interview), the studies were likely less accessible for those with intellectual disability. As those with intellectual disability form between 40 to

70% of autistic individuals (McKenzie et al., 2023), it is important that they are also represented in research. We hope to find ways to address this in our future research.

Another limitation was that due to concerns expressed by the Ethics Committee, we were not able to verify mental health or neurodevelopmental conditions with official letters of diagnosis, and therefore, mental health conditions were self-reported. Although in our final study, participants were able to validate this with information about their diagnosis, it is still possible that they populated this with false information. However, given that participation in this study was not dependent on the presence of a mental health or neurodevelopmental condition, nor were those with mental health or neurodevelopmental conditions differentially compensated, there was little incentive for participants to falsely report a diagnosis. In a similar vein, the gold standard within autism research would be for the researcher to administer the ADOS (Gotham et al., 2007) to validate participants' self-reported diagnoses of autism. However, given the time constraints of this DPhil, this was not feasible. Moreover, in both Chapter Four and Chapter Seven, we did not collect data on biological sex, only on gender. Given there is evidence for sex differences in interoception (Prentice & Murphy, 2022), it would have been interesting to include this as a variable in analyses.

Another limitation is the use of the NIES in the final study, despite the study in Chapter Four indicating that the NIES does not have a replicable factor structure. However, as there are currently no better-validated measures to assess the construct of interoceptive evaluation and all other validation checks suggested the NIES measures what it is intended to, I chose to include this questionnaire in Chapter Seven, despite a lack of certainty of its factor structure. By including the NIES in this study, it would provide a third sample on whom the factor structure of the NIES could be tested, and importantly, there would be sub-samples within which we would anticipate experience overly negative interoceptive evaluation (namely, autistics and participants with anxiety and eating disorders). This allows

for future testing of measure invariance by assessing the factor structure within these groups, thus testing whether the unexplained variability in the initial models could be attributed to this. Moreover, despite issues with the factor structure, the questionnaire appeared to be valid according to other metrics. As such, it is possible that the questionnaire does capture negative interoceptive evaluation, but that the questionnaire can be factorable in different ways and that there is variability among participants in terms of how they answer.

Finally, all of the work presented in this thesis is cross-sectional, precluding a developmental approach, or formal assessment of directionality and mediation. This was due to time constraints. The possibility of future longitudinal work and suggestions for this will be discussed in the following section.

Future directions

As a priority, future research should aim to take a longitudinal approach, measuring interoception in the same individuals across their lifespan. This would allow for a developmental approach to interoception and provide a further understanding of how interoception might change over time, and how these changes might influence mental health. Whilst there is some limited evidence that interoception might change according to one's life stage (Murphy et al., 2017), most of this relies on cross-sectional data, and there is a lack of research in age groups such as the elderly. Longitudinal studies can also provide further insight into the degree to which interoception is state-like, as well as the factors which influence it to vary.

It would also be interesting to explore how subjective, self-reported interoception relates to objective measures. In the case of interoceptive accuracy, this could be done by comparing scores on the IAS with scores on objective measures of interoception, such as the PAT and PST in the cardiac domain, occlusion tasks in the respiratory domain, and water-loading tasks

in the gastric domain. For interoceptive attention, ecological momentary assessment, whereby participants used a smartphone or similar to record what they are attending to at a given moment at multiple time points throughout the day, could be used (Murphy, Catmur, et al., 2019; Poerio et al., 2024). For interoceptive evaluation, one could measure participants' changes in self-reported state anxiety before and after engaging in tasks which either encourage focus on interoceptive sensations or make certain interoceptive sensations more salient (e.g., raising heart rate through exercise, making participants breathe through a small straw). The correspondence between these objective and subjective measures would provide measures of meta-cognition, which may also be clinically relevant (Suksasilp & Garfinkel, 2022).

There is also a need for further qualitative work into interoception in different populations. Within Chapter Six, we found that a qualitative approach allowed for an understanding of the multiple different ways by which interoception might influence anxiety, and how both hyper and hypoawareness may have an effect. However, our findings only considered autistic adolescents. It would be interesting to explore whether similar complex associations can be observed in non-autistic participants, as well as autistic adults.

It would also be beneficial to explore in more detail the degree to which the different dimensions of interoception are separable. Chapter Four provided some evidence for the separability of self-reported interoceptive evaluation from interoceptive accuracy and attention through correlation analysis. This separability is consistent with what we would expect given previous findings (Desmedt et al., 2022; Garfinkel et al., 2015; Herbert et al., 2013; Pollatos et al., 2016; Yang et al., 2023; see Chapter One), however, our approach is novel in terms of the degree of standardisation in how each dimension was measured (i.e., all three were measured by subjectively measured self-report, meaning differences were not confounded by the distinction between subjective and objective measures). However, this

analysis relied on the correlations between the total scores, rather than assessing whether the individual questions from each of the three questionnaires demonstrated the expected factor structure that would indicate that accuracy, attention and evaluation are separate constructs. This could be assessed through exploratory and confirmatory factor analyses.

In a similar vein, future research would also benefit from a more thorough investigation of how different dimensions of interoception interact to affect mental health outcomes. Some potential mechanisms have been outlined within this thesis (i.e., greater interoceptive attention only causing increased anxiety within the context of negative interoceptive evaluation) and a preliminary, exploratory analysis has been conducted (Chapter Seven). The next step would be to empirically test more potential interactions in pre-registered analyses.

Conclusion

To conclude, at present the field of interoception research is constrained by a lack of reliable and valid findings, which is partially a result of poor methodology. The introduction of novel measures of interoception and standardised definitions may be able to ameliorate these problems. This thesis introduced a new measure of cardiac interoceptive accuracy and a new measure of negative interoceptive evaluation. It is important for future research of a high quality to be conducted, as certain dimensions of self-reported interoception appear linked to both anxiety and autism. In particular, interoceptive evaluation and attention appear to be associated with autism and may also have bidirectional relationships with anxiety.

References

Abma, I. L., Rovers, M., & Van Der Wees, P. J. (2016). Appraising convergent validity of patient-reported outcome measures in systematic reviews: Constructing hypotheses and interpreting outcomes. *BMC Research Notes*, *9*(1), 226. https://doi.org/10.1186/s13104-016-2034-2

Abrams, K., Cieslowski, K., Johnson, S., Krimmel, S., La Rosa, G. B.-D., Barton, K., & Silverman, P. (2018). The effects of alcohol on heartbeat perception: Implications for anxiety. *Addictive Behaviors*, *79*, 151–158. https://doi.org/10.1016/j.addbeh.2017.12.023

Acker, L., Knight, M., & Knott, F. (2018). 'Are they just gonna reject me?' Male adolescents with autism making sense of anxiety: An Interpretative Phenomenological Analysis. *Research in Autism Spectrum Disorders*, *56*, 9–20. https://doi.org/10.1016/j.rasd.2018.07.005

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 and Biobehavioral Reviews*, *140*, 104754. https://doi.org/10.1016/j.neubiorev.2022.104754

Adams, K. L., Mandy, W., Catmur, C., & Bird, G. (2024). Potential mechanisms underlying the association between feeding and eating disorders and autism. *Neuroscience & Biobehavioral Reviews*, *162*, 105717. https://doi.org/10.1016/j.neubiorev.2024.105717

Adams, R. A., Friston, K. J., & Bastos, A. M. (2015). Active Inference, Predictive Coding and Cortical Architecture. In M. F. Casanova & I. Opris (Eds.), *Recent Advances on the*

Modular Organization of the Cortex (pp. 97–121). Springer Netherlands. https://doi.org/10.1007/978-94-017-9900-3_7

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th edition). American Psychiatric Association.

Anderson, E. R., & Hope, D. A. (2009). The relationship among social phobia, objective and perceived physiological reactivity, and anxiety sensitivity in an adolescent population. *Journal of Anxiety Disorders*, *23*(1), 18–26. https://doi.org/10.1016/j.janxdis.2008.03.011

Antony, M. M., Brown, T. A., Craske, M. G., Barlow, D. H., Mitchell, W. B., & Meadows, E. A. (1995). Accuracy of heartbeat perception in panic disorder, social phobia, and nonanxious subjects. *Journal of Anxiety Disorders*, *9*(5), 355–371. https://doi.org/10.1016/0887-6185(95)00017-I

Ardizzi, M., & Ferri, F. (2018). Interoceptive influences on peripersonal space boundary. *Cognition*, *177*, 79–86. https://doi.org/10.1016/j.cognition.2018.04.001

Auerbach, R. P., Alonso, J., Axinn, W. G., Cuijpers, P., Ebert, D. D., Green, J. G., Hwang, I., Kessler, R. C., Liu, H., Mortier, P., Nock, M. K., Pinder-Amaker, S., Sampson, N. A., Aguilar-Gaxiola, S., Al-Hamzawi, A., Andrade, L. H., Benjet, C., Caldas-de-Almeida, J. M., Demyttenaere, K., … Bruffaerts, R. (2016). Mental disorders among college students in the World Health Organization World Mental Health Surveys. *Psychological Medicine*, *46*(14), 2955–2970. https://doi.org/10.1017/S0033291716001665

Austin, D. W., Richards, J. C., & Klein, B. (2006). Modification of the Body Sensations Interpretation Questionnaire (BSIQ-M): Validity and reliability. *Journal of Anxiety Disorders*, *20*(2), 237–251. https://doi.org/10.1016/j.janxdis.2005.01.001

Avery, J. A., Kerr, K. L., Ingeholm, J. E., Burrows, K., Bodurka, J., & Simmons, W. K. (2015). A common gustatory and interoceptive representation in the human mid-insula. *Human Brain Mapping*, *36*(8), 2996–3006. https://doi.org/10.1002/hbm.22823

Bagby, R. M., Taylor, G. J., & Parker, J. D. A. (1994). The twenty-item Toronto Alexithymia Scale: II. Convergent, discriminant, and concurrent validity. *Journal of Psychosomatic Research*, *38*(1), 33–40. https://doi.org/10.1016/0022-3999(94)90006-X

Baiano, C., Raimo, G., Zappullo, I., Cecere, R., Rauso, B., Positano, M., Conson, M., The LabNPEE Group, Aversana, L. D., Di Rosa, A., Esposito, G., Milo, R., Polito, F., Raimondo, C., & Turco, A. (2022). Anxiety Sensitivity Domains are Differently Affected by Social and Non-social Autistic Traits. *Journal of Autism and Developmental Disorders*, *52*(8), 3486–3495. https://doi.org/10.1007/s10803-021-05228-w

Baker, A., Simon, N., Keshaviah, A., Farabaugh, A., Deckersbach, T., Worthington, J. J., Hoge, E., Fava, M., & Pollack, M. P. (2019). Anxiety Symptoms Questionnaire (ASQ): Development and validation. *General Psychiatry*, *32*(6), e100144. https://doi.org/10.1136/gpsych-2019-100144

Balduzzi, S., Rücker, G., & Schwarzer, G. (2019). How to perform a meta-analysis with R: A practical tutorial. *Evidence Based Mental Health*, *22*(4), 153–160. https://doi.org/10.1136/ebmental-2019-300117

Bandelow, B., & Michaelis, S. (2015). Epidemiology of anxiety disorders in the 21st century. *Dialogues in Clinical Neuroscience*, *17*(3), 327–335. https://doi.org/10.31887/DCNS.2015.17.3/bbandelow

Banellis, L., & Cruse, D. (2020). Skipping a Beat: Heartbeat-Evoked Potentials Reflect Predictions during Interoceptive-Exteroceptive Integration. *Cerebral Cortex Communications*, *1*(1), tgaa060. https://doi.org/10.1093/texcom/tgaa060

Barlow, D. H. (1988). *Anxiety and its disorders: The nature and treatment of anxiety and panic* (pp. xvii, 698). The Guilford Press.

Barlow, D. H., Gorman, J. M., Shear, M. K., & Woods, S. W. (2000). Cognitive-behavioral therapy, imipramine, or their combination for panic disorder: A randomized controlled trial. *JAMA*, *283*(19), 2529–2536. https://doi.org/10.1001/jama.283.19.2529

Barnes, L. L. B., Harp, D., & Jung, W. S. (2002). Reliability Generalization of Scores on the Spielberger State-Trait Anxiety Inventory. *Educational and Psychological Measurement*, *62*(4), 603–618. https://doi.org/10.1177/0013164402062004005

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, *31*(1), 5–17. https://doi.org/10.1023/A:1005653411471

Barsky, A. J., Ahern, D. K., Brener, J., Surman, O. S., Ring, C., & Dec, G. W. (1998). Palpitations and cardiac awareness after heart transplantation. *Psychosomatic Medicine*, *60*(5), 557–562. https://doi.org/10.1097/00006842-199809000-00007

Barsky, A. J., Cleary, P. D., Sarnie, M. K., & Ruskin, J. N. (1994). Panic Disorder, Palpitations, and the Awareness of Cardiac Activity: *The Journal of Nervous and Mental Disease*, *182*(2), 63–71. https://doi.org/10.1097/00005053-199402000-00001

Baxter, A. J., Scott, K. M., Vos, T., & Whiteford, H. A. (2013). Global prevalence of anxiety disorders: A systematic review and meta-regression. *Psychological Medicine*, *43*(5), 897–910. https://doi.org/10.1017/S003329171200147X

Beck, A., Emery, G., & Greenberg, R. (1985). *Anxiety Disorders and Phobias. A Cognitive Perspective*. Basic Books.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology*, *56*(6), 893–897. https://doi.org/10.1037/0022-006X.56.6.893

Behrens, B., Swetlitz, C., Pine, D. S., & Pagliaccio, D. (2019). The Screen for Child Anxiety Related Emotional Disorders (SCARED): Informant Discrepancy, Measurement Invariance, and Test–Retest Reliability. *Child Psychiatry & Human Development*, *50*(3), 473–482. https://doi.org/10.1007/s10578-018-0854-0

Belcher, H. L., Uglik-Marucha, N., Vitoratou, S., Ford, R. M., & Morein-Zamir, S. (2023). Gender bias in autism screening: Measurement invariance of different model frameworks of the Autism Spectrum Quotient. *BJPsych Open*, *9*(5), e173. https://doi.org/10.1192/bjo.2023.562

Ben Hassen, N., Molins, F., Garrote-Petisco, D., & Serrano, M. Á. (2023). Emotional regulation deficits in autism spectrum disorder: The role of alexithymia and interoception. *Research in Developmental Disabilities*, *132*, 104378. https://doi.org/10.1016/j.ridd.2022.104378

Bendig, A. W. (1966). Reliability of and Intercorrelations Between Cattell's IPAT Anxiety and Neuroticism Scales. *The Journal of General Psychology*, *75*(1), 1–7. https://doi.org/10.1080/00221309.1966.9710344

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300.

Berardis, D., Campanella, D., Nicola, S., Gianna, S., Alessandro, C., Chiara, C., Valchera, A., Marilde, C., Salerno, R., & Ferro, F. (2008). The Impact of Alexithymia on Anxiety Disorders: A Review of the Literature. *Current Psychiatry Reviews*, *4*(2), 80–86. https://doi.org/10.2174/157340008784529287

Berdica, E., Gerdes, A. B. M., Bublatzky, F., White, A. J., & Alpers, G. W. (2018). Threat vs. Threat: Attention to Fear-Related Animals and Threatening Faces. *Frontiers in Psychology*, *9*, 1154. https://doi.org/10.3389/fpsyg.2018.01154

Berthoz, S., & Hill, E. L. (2005). The validity of using self-reports to assess emotion regulation abilities in adults with autism spectrum disorder. *European Psychiatry*, *20*(3), 291–298. https://doi.org/10.1016/j.eurpsy.2004.06.013

Betka, S., Łukowska, M., Silva, M., King, J., Garfinkel, S., & Critchley, H. (2020). *Feeling the beat: Temporal and spatial perception of heartbeat sensations*. Neuroscience. https://doi.org/10.1101/2020.07.27.222810

Betka, S., Pfeifer, G., Garfinkel, S., Prins, H., Bond, R., Sequeira, H., Duka, T., & Critchley, H. (2018). How Do Self-Assessment of Alexithymia and Sensitivity to Bodily Sensations Relate to Alcohol Consumption? *Alcoholism: Clinical and Experimental Research*, *42*(1), 81–88. https://doi.org/10.1111/acer.13542

Bird, G., & Cook, R. (2013). Mixed emotions: The contribution of alexithymia to the emotional symptoms of autism. *Translational Psychiatry*, *3*(7), e285–e285. https://doi.org/10.1038/tp.2013.61

Bird, G., Silani, G., Brindley, R., White, S., Frith, U., & Singer, T. (2010). Empathic brain responses in insula are modulated by levels of alexithymia but not autism. *Brain*, *133*(5), 1515–1525. https://doi.org/10.1093/brain/awq060

Birmaher, B., Khetarpal, S., Brent, D., Cully, M., Balach, L., Kaufman, J., & Neer, S. M. (1997). The Screen for Child Anxiety Related Emotional Disorders (SCARED): Scale Construction and Psychometric Characteristics. *Journal of the American Academy of Child & Adolescent Psychiatry*, *36*(4), 545–553. https://doi.org/10.1097/00004583-199704000-00018

Bishop-Fitzpatrick, L., Mazefsky, C. A., & Eack, S. M. (2018). The combined impact of social support and perceived stress on quality of life in adults with autism spectrum disorder and without intellectual disability. *Autism*, *22*(6), 703–711. https://doi.org/10.1177/1362361317703090

Boettcher, H., Brake, C. A., & Barlow, D. H. (2016). Origins and outlook of interoceptive exposure. *Journal of Behavior Therapy and Experimental Psychiatry*, *53*, 41–51. https://doi.org/10.1016/j.jbtep.2015.10.009

Bogaerts, K., Notebaert, K., Van Diest, I., Devriese, S., De Peuter, S., & Van Den Bergh, O. (2005). Accuracy of respiratory symptom perception in different affective contexts. *Journal of Psychosomatic Research*, *58*(6), 537–543. https://doi.org/10.1016/j.jpsychores.2004.12.005

Bogaerts, K., Walentynowicz, M., Van Den Houte, M., Constantinou, E., & Van den Bergh, O. (2022). The Interoceptive Sensitivity and Attention Questionnaire: Evaluating Aspects of Self-Reported Interoception in Patients With Persistent Somatic Symptoms, Stress-Related Syndromes, and Healthy Controls. *Psychosomatic Medicine*, *84*(2), 251–260. https://doi.org/10.1097/PSY.0000000000001038

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to Meta-Analysis* (1st ed.). Wiley. https://doi.org/10.1002/9780470743386

Borg, C., Chouchou, F., Dayot-Gorlero, J., Zimmermann, P., Maudoux, D., Laurent, B., & Michael, G. (2018). Pain and emotion as predictive factors of interoception in fibromyalgia. *Journal of Pain Research*, *Volume 11*, 823–835. https://doi.org/10.2147/JPR.S152012

Bornemann, B., & Singer, T. (2017). Taking time to feel our body: Steady increases in heartbeat perception accuracy and decreases in alexithymia over 9 months of

contemplative mental training. *Psychophysiology*, *54*(3), 469–482. https://doi.org/10.1111/psyp.12790

Boswell, J. F., Farchione, T. J., Sauer-Zavala, S., Murray, H. W., Fortune, M. R., & Barlow, D. H. (2013). Anxiety Sensitivity and Interoceptive Exposure: A Transdiagnostic Construct and Change Strategy. *Behavior Therapy*, *44*(3), 417–431. https://doi.org/10.1016/j.beth.2013.03.006

Brand, S., Meis, A. C., Tünte, M. R., Murphy, J., Woller, J. P., Jungmann, S. M., Witthöft, M., Hoehl, S., Weymar, M., Hermann, C., & Ventura-Bort, C. (2023). A multi-site German validation of the Interoceptive Accuracy Scale and its relation to psychopathological symptom burden. *Communications Psychology*, *1*(1), 14. https://doi.org/10.1038/s44271-023-00016-x

Braun, V., & Clarke, V. (2013). *Successful Qualitative Research: A Practical Guide for Beginners.* SAGE Publications, Ltd.

Braun, V., & Clarke, V. (2021). To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. *Qualitative Research in Sport, Exercise and Health*, *13*(2), 201–216. https://doi.org/10.1080/2159676X.2019.1704846

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brener, J., & Kluvitse, C. (1988). Heartbeat Detection: Judgments of the Simultaneity of External Stimuli and Heartbeats. *Psychophysiology*, *25*(5), 554–561. https://doi.org/10.1111/j.1469-8986.1988.tb01891.x

Brener, J., Liu, X., & Ring, C. (1993). A method of constant stimuli for examining heartbeat detection: Comparison with the Brener-Kluvitse and Whitehead methods.

Psychophysiology, *30*(6), 657–665. https://doi.org/10.1111/j.1469-8986.1993.tb02091.x

Brener, J., & Ring, C. (2016). Towards a psychophysics of interoceptive processes: The measurement of heartbeat detection. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1708), 20160015. https://doi.org/10.1098/rstb.2016.0015

Brener, J., Ring, C., & Liu, X. (1994). Effects of data limitations on heartbeat detection in the method of constant stimuli. *Psychophysiology*, *31*(3), 309–312. https://doi.org/10.1111/j.1469-8986.1994.tb02219.x

Brewer, R., Cook, R., & Bird, G. (2016). Alexithymia: A general deficit of interoception. *Royal Society Open Science*, *3*(10), 150664. https://doi.org/10.1098/rsos.150664

Brewer, R., Happé, F., Cook, R., & Bird, G. (2015). Commentary on "Autism, oxytocin and interoception": Alexithymia, not Autism Spectrum Disorders, is the consequence of interoceptive failure. *Neuroscience & Biobehavioral Reviews*, *56*, 348–353. https://doi.org/10.1016/j.neubiorev.2015.07.006

Brewer, R., Murphy, J., & Bird, G. (2021). Atypical interoception as a common risk factor for psychopathology: A review. *Neuroscience and Biobehavioral Reviews*, *130*, 470–508. https://doi.org/10.1016/j.neubiorev.2021.07.036

Brown, T. A., Berner, L. A., Jones, M. D., Reilly, E. E., Cusack, A., Anderson, L. K., Kaye, W. H., & Wierenga, C. E. (2017). Psychometric Evaluation and Norms for the Multidimensional Assessment of Interoceptive Awareness (MAIA) in a Clinical Eating Disorders Sample. *European Eating Disorders Review: The Journal of the Eating Disorders Association*, *25*(5), 411–416. https://doi.org/10.1002/erv.2532

Brown, T. A., Chorpita, B. F., Korotitsch, W., & Barlow, D. H. (1997). Psychometric properties of the Depression Anxiety Stress Scales (DASS) in clinical samples.

Behaviour Research and Therapy, *35*(1), 79–89. https://doi.org/10.1016/S0005-7967(96)00068-X

Bruffaerts, R., Mortier, P., Kiekens, G., Auerbach, R. P., Cuijpers, P., Demyttenaere, K., Green, J. G., Nock, M. K., & Kessler, R. C. (2018). Mental health problems in college freshmen: Prevalence and academic functioning. *Journal of Affective Disorders*, *225*, 97–103. https://doi.org/10.1016/j.jad.2017.07.044

Budiman, H., Wakita, R., Ito, T., & Maeda, S. (2022). Factors Associated with Variability in Pulse Wave Transit Time Using Pulse Oximetry: A Retrospective Study. *Journal of Clinical Medicine*, *11*(14), Article 14. https://doi.org/10.3390/jcm11143963

Butera, C., Kaplan, J., Kilroy, E., Harrison, L., Jayashankar, A., Loureiro, F., & Aziz-Zadeh, L. (2023). The relationship between alexithymia, interoception, and neural functional connectivity during facial expression processing in autism spectrum disorder. *Neuropsychologia*, *180*, 108469. https://doi.org/10.1016/j.neuropsychologia.2023.108469

Camilleri, M., Coulie, B., & Tack, J. F. (2001). Visceral hypersensitivity: Facts, speculations, and challenges. *Gut*, *48*(1), 125–131. https://doi.org/10.1136/gut.48.1.125

Camm-Crosbie, L., Bradley, L., Shaw, R., Baron-Cohen, S., & Cassidy, S. (2019). 'People like me don't get support': Autistic adults' experiences of support and treatment for mental health difficulties, self-injury and suicidality. *Autism*, *23*(6), 1431–1441. https://doi.org/10.1177/1362361318816053

Campbell, D. B., Datta, D., Jones, S. T., Batey Lee, E., Sutcliffe, J. S., Hammock, E. A. D., & Levitt, P. (2011). Association of oxytocin receptor (OXTR) gene variants with multiple phenotype domains of autism spectrum disorder. *Journal of Neurodevelopmental Disorders*, *3*(2), 101–112. https://doi.org/10.1007/s11689-010-9071-2

Carter, M. M., Sbrocco, T., & Ayati, F. (2009). Predicting anxious response to a social challenge and hyperventilation: Comparison of the ASI and ASI-3. *Journal of Behavior Therapy and Experimental Psychiatry*, *40*(3), 434–442. https://doi.org/10.1016/j.jbtep.2009.05.001

Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., Meier, M. H., Ramrakha, S., Shalev, I., Poulton, R., & Moffitt, T. E. (2014). The p Factor: One General Psychopathology Factor in the Structure of Psychiatric Disorders? *Clinical Psychological Science*, *2*(2), 119–137. https://doi.org/10.1177/2167702613497473

Cattell, R. B. (1969). *Scala di ansieta*. Organizzazioni Speciali.

Ceunen, E., Vlaeyen, J. W. S., & Van Diest, I. (2016). On the Origin of Interoception. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00743

Chan, G. S. H., Middleton, P. M., Celler, B. G., Wang, L., & Lovell, N. H. (2007). Change in pulse transit time and pre-ejection period during head-up tilt-induced progressive central hypovolaemia. *Journal of Clinical Monitoring and Computing*, *21*(5), 283–293. https://doi.org/10.1007/s10877-007-9086-8

Chapman, L., Rose, K., Hull, L., & Mandy, W. (2022). "I want to fit in… but I don't want to change myself fundamentally": A qualitative exploration of the relationship between masking and mental health for autistic teenagers. *Research in Autism Spectrum Disorders*, *99*, 102069. https://doi.org/10.1016/j.rasd.2022.102069

Chen, W. G., Schloesser, D., Arensdorf, A. M., Simmons, J. M., Cui, C., Valentino, R., Gnadt, J. W., Nielsen, L., Hillaire-Clarke, C. S., Spruance, V., Horowitz, T. S., Vallejo, Y. F., & Langevin, H. M. (2021). The Emerging Science of Interoception: Sensing, Integrating, Interpreting, and Regulating Signals within the Self. *Trends in Neurosciences*, *44*(1), 3–16. https://doi.org/10.1016/j.tins.2020.10.007

Chorpita, B. F., Moffitt, C. E., & Gray, J. (2005). Psychometric properties of the Revised Child Anxiety and Depression Scale in a clinical sample. *Behaviour Research and Therapy*, *43*(3), 309–322. https://doi.org/10.1016/j.brat.2004.02.004

Chorpita, B. F., Yim, L., Moffitt, C., Umemoto, L. A., & Francis, S. E. (2000). Assessment of symptoms of DSM-IV anxiety and depression in children: A revised child anxiety and depression scale. *Behaviour Research and Therapy*, *38*(8), 835–855. https://doi.org/10.1016/S0005-7967(99)00130-8

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*(4), 284–290. https://doi.org/10.1037/1040-3590.6.4.284

Clark, D. M. (1986). A cognitive approach to panic. *Behaviour Research and Therapy*, *24*(4), 461–470. https://doi.org/10.1016/0005-7967(86)90011-2

Clark, D. M., Salkovskis, P. M., Hackmann, A., Middleton, H., Anastasiades, P., & Gelder, M. (1994). A Comparison of Cognitive Therapy, Applied Relaxation and Imipramine in the Treatment of Panic Disorder. *British Journal of Psychiatry*, *164*(6), 759–769. https://doi.org/10.1192/bjp.164.6.759

Clark, D. M., Salkovskis, P. M., Öst, L.-G., Breitholtz, E., Koehler, K. A., Westling, B. E., Jeavons, A., & Gelder, M. (1997). Misinterpretation of body sensations in panic disorder. *Journal of Consulting and Clinical Psychology*, *65*(2), 203–213. https://doi.org/10.1037/0022-006X.65.2.203

Clark, L. A., & Watson, D. (1991). Theoretical and empirical issues in differentiating depression from anxiety. In J. Becker & A. Kleinman (Eds.), *Psychosocial aspects of depression* (pp. 39–65). Lawrence Erlbaum Associates, Inc.

Clemens, W. J. (1984). Temporal Arrangement of Signals in Heartbeat Discrimination Procedures. *Psychophysiology*, *21*(2), 187–190. https://doi.org/10.1111/j.1469-8986.1984.tb00202.x

Clemente, R., Murphy, A., & Murphy, J. (2024). The relationship between self-reported interoception and anxiety: A systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews*, *167*, 105923. https://doi.org/10.1016/j.neubiorev.2024.105923

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences (2nd ed.)*.

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

Condon, D. M., & Revelle, W. (2014). The international cognitive ability resource: Development and initial validation of a public-domain measure. *Intelligence*, *43*, 52–64. https://doi.org/10.1016/j.intell.2014.01.004

Cook, J., Hull, L., Crane, L., & Mandy, W. (2021). Camouflaging in autism: A systematic review. *Clinical Psychology Review*, *89*, 102080. https://doi.org/10.1016/j.cpr.2021.102080

Cook, R., Brewer, R., Shah, P., & Bird, G. (2013). Alexithymia, Not Autism, Predicts Poor Recognition of Emotional Facial Expressions. *Psychological Science*, *24*(5), 723–732. https://doi.org/10.1177/0956797612463582

Costello, E. J., Egger, H. L., & Angold, A. (2005). The Developmental Epidemiology of Anxiety Disorders: Phenomenology, Prevalence, and Comorbidity. *Child and Adolescent Psychiatric Clinics of North America*, *14*(4), 631–648. https://doi.org/10.1016/j.chc.2005.06.003

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

Craig, A. D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nature Reviews. Neuroscience*, *3*(8), 655–666. https://doi.org/10.1038/nrn894

Crane, L., Goddard, L., & Pring, L. (2009). Sensory processing in adults with autism spectrum disorders. *Autism*, *13*(3), 215–228. https://doi.org/10.1177/1362361309103794

Craske, M. G., Rowe, M., Lewin, M., & Noriega-Dimitri, R. (1997). Interoceptive exposure versus breathing retraining within cognitive-behavioural therapy for panic disorder with agoraphobia [1]. *British Journal of Clinical Psychology*, *36*(1), 85–99. https://doi.org/10.1111/j.2044-8260.1997.tb01233.x

Critchley, H. D., Wiens, S., Rotshtein, P., Ohman, A., & Dolan, R. J. (2004). Neural systems supporting interoceptive awareness. *Nature Neuroscience*, *7*(2), 189–195. https://doi.org/10.1038/nn1176

Cropley, M., Plans, D., Morelli, D., Sütterlin, S., Inceoglu, I., Thomas, G., & Chu, C. (2017). The Association between Work-Related Rumination and Heart Rate Variability: A Field Study. *Frontiers in Human Neuroscience*, *11*. https://doi.org/10.3389/fnhum.2017.00027

Cruz, S., Zubizarreta, S. C.-P., Costa, A. D., Araújo, R., Martinho, J., Tubío-Fungueiriño, M., Sampaio, A., Cruz, R., Carracedo, A., & Fernández-Prieto, M. (2024). Is There a Bias

Towards Males in the Diagnosis of Autism? A Systematic Review and Meta-Analysis. *Neuropsychology Review*. https://doi.org/10.1007/s11065-023-09630-2

Cusack, J. P., & Sterry, R. (2016). *Your Questions Shaping Future Autism Research*. Autistica.

Damasio, A., & Damasio, H. (2022). Homeostatic feelings and the biology of consciousness. *Brain*, *145*(7), 2231–2235. https://doi.org/10.1093/brain/awac194

Damasio, A. R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. Putnam.

De Berardis, D., Campanella, D., Gambi, F., La Rovere, R., Sepede, G., Core, L., Canfora, G., Santilli, E., Valchera, A., Mancini, E., Salerno, R. M., Moschetta, F. S., & Ferro, F. M. (2007). Alexithymia, Fear of Bodily Sensations, and Somatosensory Amplification in Young Outpatients With Panic Disorder. *Psychosomatics*, *48*(3), 239–246. https://doi.org/10.1176/appi.psy.48.3.239

De Pascalis, V., Alberti, M. L., & Pandolfo, R. (1984). Anxiety, Perception, and Control of Heart Rate. *Perceptual and Motor Skills*, *59*(1), 203–211. https://doi.org/10.2466/pms.1984.59.1.203

De Pascalis, V., Palumbo, G., & Ronchitelli, V. (1991). Heartbeat perception, instructions, and biofeedback in the control of heart rate. *International Journal of Psychophysiology*, *11*(2), 179–193. https://doi.org/10.1016/0167-8760(91)90010-U

Deacon, B., & Abramowitz, J. (2006). Anxiety sensitivity and its dimensions across the anxiety disorders. *Journal of Anxiety Disorders*, *20*(7), 837–857. https://doi.org/10.1016/j.janxdis.2006.01.003

Desmedt, O., Corneille, O., Luminet, O., Murphy, J., Bird, G., & Maurage, P. (2020). Contribution of Time Estimation and Knowledge to Heartbeat Counting Task Performance under Original and Adapted Instructions. *Biological Psychology*, *154*, 107904. https://doi.org/10.1016/j.biopsycho.2020.107904

Desmedt, O., Heeren, A., Corneille, O., & Luminet, O. (2022). What do measures of self-report interoception measure? Insights from a systematic review, latent factor analysis, and network approach. *Biological Psychology*, *169*, 108289. https://doi.org/10.1016/j.biopsycho.2022.108289

Desmedt, O., Luminet, O., & Corneille, O. (2018). The heartbeat counting task largely involves non-interoceptive processes: Evidence from both the original and an adapted counting task. *Biological Psychology*, *138*, 185–188. https://doi.org/10.1016/j.biopsycho.2018.09.004

Desmedt, O., Luminet, O., Maurage, P., & Corneille, O. (2023). Discrepancies in the Definition and Measurement of Human Interoception: A Comprehensive Discussion and Suggested Ways Forward. *Perspectives on Psychological Science*, 17456916231191537. https://doi.org/10.1177/17456916231191537

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

Desmedt, O., Van Den Houte, M., Walentynowicz, M., Dekeyser, S., Luminet, O., & Corneille, O. (2022). How Does Heartbeat Counting Task Performance Relate to Theoretically-Relevant Mental Health Outcomes? A Meta-Analysis. *Collabra: Psychology*, *8*(1), 33271. https://doi.org/10.1525/collabra.33271

Di Lernia, D., Lacerenza, M., Ainley, V., & Riva, G. (2020). Altered Interoceptive Perception and the Effects of Interoceptive Analgesia in Musculoskeletal, Primary, and Neuropathic Chronic Pain Conditions. *Journal of Personalized Medicine*, *10*(4), 201. https://doi.org/10.3390/jpm10040201

Di Lernia, D., Serino, S., Pezzulo, G., Pedroli, E., Cipresso, P., & Riva, G. (2018). Feel the Time. Time Perception as a Function of Interoceptive Processing. *Frontiers in Human Neuroscience*, *12*, 74. https://doi.org/10.3389/fnhum.2018.00074

Dobson, K. S. (1985). The relationship between anxiety and depression. *Clinical Psychology Review*, *5*(4), 307–324. https://doi.org/10.1016/0272-7358(85)90010-8

Domschke, K., Stevens, S., Pfleiderer, B., & Gerlach, A. L. (2010). Interoceptive sensitivity in anxiety and anxiety disorders: An overview and integration of neurobiological findings. *Clinical Psychology Review*, *30*(1), 1–11. https://doi.org/10.1016/j.cpr.2009.08.008

Dunn, B. D., Galton, H. C., Morgan, R., Evans, D., Oliver, C., Meyer, M., Cusack, R., Lawrence, A. D., & Dalgleish, T. (2010). *Listening to Your Heart: How Interoception Shapes Emotion Experience and Intuitive Decision Making*. https://journals.sagepub.com/doi/full/10.1177/0956797610389191?casa_token=Egd-uLI5-psAAAAA%3AbOBGvj9qN9s6EgI9eHesGugMiqWM3B-z4l0nM743QR7n_Q3qTIxUgM7ZC_gRVnDvO6KXaChxkkSP-g

Dunn, B. D., Stefanovitch, I., Evans, D., Oliver, C., Hawkins, A., & Dalgleish, T. (2010). Can you feel the beat? Interoceptive awareness is an interactive function of anxiety- and depression-specific symptom dimensions. *Behaviour Research and Therapy*, *48*(11), 1133–1138. https://doi.org/10.1016/j.brat.2010.07.006

Durlik, C., Brown, G., & Tsakiris, M. (2014). Enhanced interoceptive awareness during anticipation of public speaking is associated with fear of negative evaluation. *Cognition and Emotion*, *28*(3), 530–540. https://doi.org/10.1080/02699931.2013.832654

Duvekot, J., Van Der Ende, J., Verhulst, F. C., Slappendel, G., Van Daalen, E., Maras, A., & Greaves-Lord, K. (2017). Factors influencing the probability of a diagnosis of autism

spectrum disorder in girls versus boys. *Autism*, *21*(6), 646–658. https://doi.org/10.1177/1362361316672178

Dworzynski, K., Ronald, A., Bolton, P., & Happé, F. (2012). How Different Are Girls and Boys Above and Below the Diagnostic Threshold for Autism Spectrum Disorders? *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(8), 788–797. https://doi.org/10.1016/j.jaac.2012.05.018

Ebisch, S. J. H., Gallese, V., Willems, R. M., Mantini, D., Groen, W. B., Romani, G. L., Buitelaar, J. K., & Bekkering, H. (2011). Altered intrinsic functional connectivity of anterior and posterior insula regions in high-functioning participants with autism spectrum disorder. *Human Brain Mapping*, *32*(7), 1013–1028. https://doi.org/10.1002/hbm.21085

Ehlers, A. (1993). Interoception and panic disorder. *Advances in Behaviour Research and Therapy*, *15*(1), 3–21. https://doi.org/10.1016/0146-6402(93)90001-I

Ehlers, A. (1995). A 1-year prospective study of panic attacks: Clinical course and factors associated with maintenance. *Journal of Abnormal Psychology*, *104*(1), 164–172. https://doi.org/10.1037/0021-843X.104.1.164

Ehlers, A., & Breuer, P. (1992). Increased cardiac awareness in panic disorder. *Journal of Abnormal Psychology*, *101*(3), 371–382. https://doi.org/10.1037/0021-843X.101.3.371

Ehlers, A., Breuer, P., Dohn, D., & Fiegenbaum, W. (1995). Heartbeat perception and panic disorder: Possible explanations for discrepant findings. *Behaviour Research and Therapy*, *33*(1), 69–76. https://doi.org/10.1016/0005-7967(94)E0002-Z

Ehlers, A., Mayou, R. A., Sprigings, D. C., & Birkhead, J. (2000). Psychological and Perceptual Factors Associated With Arrhythmias and Benign Palpitations:

Psychosomatic Medicine, *62*(5), 693–702. https://doi.org/10.1097/00006842-200009000-00014

Eley, T. C., Gregory, A. M., Clark, D. M., & Ehlers, A. (2007). Feeling anxious: A twin study of panic/somatic ratings, anxiety sensitivity and heartbeat perception in children. *Journal of Child Psychology and Psychiatry*, *48*(12), 1184–1191. https://doi.org/10.1111/j.1469-7610.2007.01838.x

Eley, T. C., Stirling, L., Ehlers, A., Gregory, A. M., & Clark, D. M. (2004). Heart-beat perception, panic/somatic symptoms and anxiety sensitivity in children. *Behaviour Research and Therapy*, *42*(4), 439–448. https://doi.org/10.1016/S0005-7967(03)00152-9

Elwin, M., Ek, L., Schröder, A., & Kjellin, L. (2012). Autobiographical Accounts of Sensing in Asperger Syndrome and High-Functioning Autism. *Archives of Psychiatric Nursing*, *26*(5), 420–429. https://doi.org/10.1016/j.apnu.2011.10.003

English, M. C. W., Gignac, G. E., Visser, T. A. W., Whitehouse, A. J. O., Enns, J. T., & Maybery, M. T. (2021). The Comprehensive Autistic Trait Inventory (CATI): Development and validation of a new measure of autistic traits in the general population. *Molecular Autism*, *12*(1), 37. https://doi.org/10.1186/s13229-021-00445-7

Ernst, J., Böker, H., Hättenschwiler, J., Schüpbach, D., Northoff, G., Seifritz, E., & Grimm, S. (2014). The association of interoceptive awareness and alexithymia with neurotransmitter concentrations in insula and anterior cingulate. *Social Cognitive and Affective Neuroscience*, *9*(6), 857–863. https://doi.org/10.1093/scan/nst058

Etkin, A., & Wager, T. D. (2007). Functional neuroimaging of anxiety: A meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. *The American Journal of Psychiatry*, *164*(10), 1476–1488. https://doi.org/10.1176/appi.ajp.2007.07030504

Ewing, D. L., Manassei, M., Gould Van Praag, C., Philippides, A. O., Critchley, H. D., & Garfinkel, S. N. (2017). Sleep and the heart: Interoceptive differences linked to poor experiential sleep quality in anxiety and depression. *Biological Psychology*, *127*, 163–172. https://doi.org/10.1016/j.biopsycho.2017.05.011

Ewing, D. L., Manassei, M., van Praag, C. G., Philippides, A. O., Critchley, H. D., & Garfinkel, S. N. (2017). Sleep and the heart: Interoceptive differences linked to poor experiential sleep quality in anxiety and depression. *Biological Psychology*, *127*, 163–172. https://doi.org/10.1016/j.biopsycho.2017.05.011

Fabriger, L. R., & Wegener, D. T. (2012). *Exploratory Factor Analysis*. Oxford University Press.

Failla, M. D., Bryant, L. K., Heflin, B. H., Mash, L. E., Schauder, K., Davis, S., Gerdes, M. B., Weitlauf, A., Rogers, B. P., & Cascio, C. J. (2020). Neural Correlates of Cardiac Interoceptive Focus Across Development: Implications for Social Symptoms in Autism Spectrum Disorder. *Autism Research*, *13*(6), 908–920. https://doi.org/10.1002/aur.2289

Farb, N., Daubenmier, J., Price, C. J., Gard, T., Kerr, C., Dunn, B. D., Klein, A. C., Paulus, M. P., & Mehling, W. E. (2015). Interoception, contemplative practice, and health. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00763

Feinstein, J. S., Khalsa, S. S., Yeh, H., Al Zoubi, O., Arevian, A. C., Wohlrab, C., Pantino, M. K., Cartmell, L. J., Simmons, W. K., Stein, M. B., & Paulus, M. P. (2018). The Elicitation of Relaxation and Interoceptive Awareness Using Floatation Therapy in Individuals With High Anxiety Sensitivity. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *3*(6), 555–562. https://doi.org/10.1016/j.bpsc.2018.02.005

Fenigstein, A., Scheier, M. F., & Buss, A. H. (1975). Public and private self-consciousness: Assessment and theory. *Journal of Consulting and Clinical Psychology*, *43*(4), 522–527. https://doi.org/10.1037/h0076760

Ferentzi, E., Bogdány, T., Szabolcs, Z., Csala, B., Horváth, Á., & Köteles, F. (2018). Multichannel Investigation of Interoception: Sensitivity Is Not a Generalizable Feature. *Frontiers in Human Neuroscience*, *12*, 223. https://doi.org/10.3389/fnhum.2018.00223

Field, A. P. (2001). Meta-analysis of correlation coefficients: A Monte Carlo comparison of fixed- and random-effects methods. *Psychological Methods*, *6*(2), 161–180. https://doi.org/10.1037/1082-989X.6.2.161

Fiene, L., & Brownlow, C. (2015). Investigating interoception and body awareness in adults with and without autism spectrum disorder. *Autism Research*, *8*(6), 709–716. https://doi.org/10.1002/aur.1486

Fiene, L., Ireland, M. J., & Brownlow, C. (2018). The Interoception Sensory Questionnaire (ISQ): A Scale to Measure Interoceptive Challenges in Adults. *Journal of Autism and Developmental Disorders*, *48*(10), 3354–3366. https://doi.org/10.1007/s10803-018-3600-3

Fietz, J., Valencia, N., & Silani, G. (2018). Alexithymia and autistic traits as possible predictors for traits related to depression, anxiety, and stress: A multivariate statistical approach. *Journal of Evaluation in Clinical Practice*, *24*(4), 901–908. https://doi.org/10.1111/jep.12961

Fissler, M., Winnebeck, E., Schroeter, T., Gummersbach, M., Huntenburg, J. M., Gaertner, M., & Barnhofer, T. (2016). An investigation of the effects of brief mindfulness training on self-reported interoceptive awareness, the ability to decenter, and their role

in the reduction of depressive symptoms. *Mindfulness*, *7*(5), 1170–1181. https://doi.org/10.1007/s12671-016-0559-z

Flasinski, T., Dierolf, A. M., Rost, S., Lutz, A. P. C., Voderholzer, U., Koch, S., Bach, M., Asenstorfer, C., Münch, E. E., Mertens, V.-C., Vögele, C., & Schulz, A. (2020). Altered Interoceptive Awareness in High Habitual Symptom Reporters and Patients With Somatoform Disorders. *Frontiers in Psychology*, *11*, 1859. https://doi.org/10.3389/fpsyg.2020.01859

Fletcher-Watson, S., & Happé, F. (2019). *Autism: A New Introduction to Psychological Theory and Current Debate* (1st ed.). Routledge. https://doi.org/10.4324/9781315101699

Francis, S. M., Camchong, J., Brickman, L., Goelkel-Garcia, L., Mueller, B. A., Tseng, A., Lim, K. O., & Jacob, S. (2019). Hypoconnectivity of insular resting-state networks in adolescents with Autism Spectrum Disorder. *Psychiatry Research: Neuroimaging*, *283*, 104–112. https://doi.org/10.1016/j.pscychresns.2018.12.003

Furman, D. J., Waugh, C. E., Bhattacharjee, K., Thompson, R. J., & Gotlib, I. H. (2013). Interoceptive awareness, positive affect, and decision making in Major Depressive Disorder. *Journal of Affective Disorders*, *151*(2), 780–785. https://doi.org/10.1016/j.jad.2013.06.044

Fydrich, T., Dowdall, D., & Chambless, D. L. (1992). Reliability and validity of the beck anxiety inventory. *Journal of Anxiety Disorders*, *6*(1), 55–61. https://doi.org/10.1016/0887-6185(92)90026-4

Gabriele, E., Spooner, R., Brewer, R., & Murphy, J. (2022). Dissociations between self-reported interoceptive accuracy and attention: Evidence from the Interoceptive Attention Scale. *Biological Psychology*, *168*, 108243. https://doi.org/10.1016/j.biopsycho.2021.108243

Gaebler, M., Daniels, J. K., Lamke, J.-P., Fydrich, T., & Walter, H. (2013). Heart rate variability and its neural correlates during emotional face processing in social anxiety disorder. *Biological Psychology*, *94*(2), 319–330. https://doi.org/10.1016/j.biopsycho.2013.06.009

Gaggero, G., Bizzego, A., Dellantonio, S., Pastore, L., Lim, M., & Esposito, G. (2021). Clarifying the relationship between alexithymia and subjective interoception. *PLOS ONE*, *16*(12), e0261126. https://doi.org/10.1371/journal.pone.0261126

Gardner, R. M., Morrell, J. A., Watson, D. N., & Sandoval, S. L. (1990). Cardiac self-perception in obese and normal persons. *Perceptual and Motor Skills*, *70*(3 Pt 2), 1179–1186.

Garfinkel, S. N., & Critchley, H. D. (2013). Interoception, emotion and brain: New insights link internal physiology to social behaviour. Commentary on: *Social Cognitive and Affective Neuroscience*, *8*(3), 231–234. https://doi.org/10.1093/scan/nss140

Garfinkel, S. N., Manassei, M. F., Hamilton-Fletcher, G., In Den Bosch, Y., Critchley, H. D., & Engels, M. (2016). Interoceptive dimensions across cardiac and respiratory axes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1708), 20160014. https://doi.org/10.1098/rstb.2016.0014

Garfinkel, S. N., Seth, A. K., Barrett, A. B., Suzuki, K., & Critchley, H. D. (2015). Knowing your own heart: Distinguishing interoceptive accuracy from interoceptive awareness. *Biological Psychology*, *104*, 65–74. https://doi.org/10.1016/j.biopsycho.2014.11.004

Garfinkel, S. N., Tiley, C., O'Keeffe, S., Harrison, N. A., Seth, A. K., & Critchley, H. D. (2016). Discrepancies between dimensions of interoception in autism: Implications for emotion and anxiety. *Biological Psychology*, *114*, 117–126. https://doi.org/10.1016/j.biopsycho.2015.12.003

Garner, D. M., Olmstead, M. P., & Polivy, J. (1983). Development and validation of a multidimensional eating disorder inventory for anorexia nervosa and bulimia. *International Journal of Eating Disorders*, *2*(2), 15–34. https://doi.org/10.1002/1098-108X(198321)2:2<15::AID-EAT2260020203>3.0.CO;2-6

Ghisi, M., Bottesi, G., Altoè, G., Razzetti, E., Melli, G., & Sica, C. (2016). Factor Structure and Psychometric Properties of the Anxiety Sensitivity Index-3 in an Italian Community Sample. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00160

Gotham, K., Risi, S., Pickles, A., & Lord, C. (2007). The Autism Diagnostic Observation Schedule: Revised Algorithms for Improved Diagnostic Validity. *Journal of Autism and Developmental Disorders*, *37*(4), 613–627. https://doi.org/10.1007/s10803-006-0280-1

Grabauskaitė, A., Baranauskas, M., & Griškova-Bulanova, I. (2017). Interoception and gender: What aspects should we pay attention to? *Consciousness and Cognition*, *48*, 129–137. https://doi.org/10.1016/j.concog.2016.11.002

Green, L., Fein, D., Modahl, C., Feinstein, C., Waterhouse, L., & Morris, M. (2001). Oxytocin and autistic disorder: Alterations in peptide forms. *Biological Psychiatry*, *50*(8), 609–613. https://doi.org/10.1016/S0006-3223(01)01139-8

Green, S. A., & Ben-Sasson, A. (2010). Anxiety Disorders and Sensory Over-Responsivity in Children with Autism Spectrum Disorders: Is There a Causal Relationship? *Journal of Autism and Developmental Disorders*, *40*(12), 1495–1504. https://doi.org/10.1007/s10803-010-1007-x

Gregory, S. G., Connelly, J. J., Towers, A. J., Johnson, J., Biscocho, D., Markunas, C. A., Lintas, C., Abramson, R. K., Wright, H. H., Ellis, P., Langford, C. F., Worley, G., Delong, G. R., Murphy, S. K., Cuccaro, M. L., Persico, A., & Pericak-Vance, M. A.

(2009). Genomic and epigenetic evidence for oxytocin receptor deficiency in autism. *BMC Medicine*, *7*(1), 62. https://doi.org/10.1186/1741-7015-7-62

Groskurth, K., Bluemke, M., & Lechner, C. M. (2024). Why we need to abandon fixed cutoffs for goodness-of-fit indices: An extensive simulation and possible solutions. *Behavior Research Methods*, *56*(4), 3891–3914. https://doi.org/10.3758/s13428-023-02193-3

Gurrieri, F., & Neri, G. (2009). Defective oxytocin function: A clue to understanding the cause of autism? *BMC Medicine*, *7*(1), 63. https://doi.org/10.1186/1741-7015-7-63

Hale, W. W., Crocetti, E., Raaijmakers, Q. A. W., & Meeus, W. H. J. (2011). A meta-analysis of the cross-cultural psychometric properties of the Screen for Child Anxiety Related Emotional Disorders (SCARED): Meta-analysis of the SCARED. *Journal of Child Psychology and Psychiatry*, *52*(1), 80–90. https://doi.org/10.1111/j.1469-7610.2010.02285.x

Halim, A. T., Richdale, A. L., & Uljarević, M. (2018). Exploring the nature of anxiety in young adults on the autism spectrum: A qualitative study. *Research in Autism Spectrum Disorders*, *55*, 25–37. https://doi.org/10.1016/j.rasd.2018.07.006

Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *36*(1), 5–25. https://doi.org/10.1007/s10803-005-0039-0

Harrer, M., Cuijpers, P., Furukawa, T., & Ebert, D. D. (2019). *Dmetar: Companion R Package For The Guide 'Doing Meta-Analysis in R'. R package version 0.0.9000.* [Computer software]. http://dmetar.protectlab.org

Harrison, O. K., Köchli, L., Marino, S., Luechinger, R., Hennel, F., Brand, K., Hess, A. J., Frässle, S., Iglesias, S., Vinckier, F., Petzschner, F. H., Harrison, S. J., & Stephan, K.

E. (2021). Interoception of breathing and its relationship with anxiety. *Neuron*, *109*(24), 4080-4093.e8. https://doi.org/10.1016/j.neuron.2021.09.045

Hart, N., McGowan, J., Minati, L., & Critchley, H. D. (2013). Emotional Regulation and Bodily Sensation: Interoceptive Awareness Is Intact in Borderline Personality Disorder. *Journal of Personality Disorders*, *27*(4), 506–518. https://doi.org/10.1521/pedi_2012_26_049

Haruki, Y., & Ogawa, K. (2023). Cardiac and Gastric Interoceptive Awareness Have Distinct Neural Substrates. *eNeuro*, *10*(1), ENEURO.0157-22.2023. https://doi.org/10.1523/ENEURO.0157-22.2023

Harver, A., Balrd, J. C., Mcgovern, J. F., & Daubenspeck, J. A. (1988). Grouping and multidimensional organization of respiratory sensations. *Perception & Psychophysics*, *44*(3), 285–292. https://doi.org/10.3758/BF03206297

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

Hazen, E. P., Stornelli, J. L., O'Rourke, J. A., Koesterer, K., & McDougle, C. J. (2014). Sensory Symptoms in Autism Spectrum Disorders. *Harvard Review of Psychiatry*, *22*(2), 112–124. https://doi.org/10.1097/01.HRP.0000445143.08773.58

Heim, N., Bobou, M., Tanzer, M., Jenkinson, P. M., Steinert, C., & Fotopoulou, A. (2023). Psychological interventions for interoception in mental health disorders: A systematic review of randomized-controlled trials. *Psychiatry and Clinical Neurosciences*, *77*(10), 530–540. https://doi.org/10.1111/pcn.13576

Herbert, B. M. (2020). Interoception and its role for eating, obesity, and eating disorders: Empirical findings and conceptual conclusions. *European Journal of Health Psychology*, *27*(4), 188–205. https://doi.org/10.1027/2512-8442/a000062

Herbert, B. M., Blechert, J., Hautzinger, M., Matthias, E., & Herbert, C. (2013). Intuitive eating is associated with interoceptive sensitivity. Effects on body mass index. *Appetite*, *70*, 22–30. https://doi.org/10.1016/j.appet.2013.06.082

Herbert, B. M., Muth, E. R., Pollatos, O., & Herbert, C. (2012). Interoception across modalities: On the relationship between cardiac awareness and the sensitivity for gastric functions. *PloS One*, *7*(5), e36646. https://doi.org/10.1371/journal.pone.0036646

Herbert, B. M., Pollatos, O., Flor, H., Enck, P., & Schandry, R. (2010). Cardiac awareness and autonomic cardiac reactivity during emotional picture viewing and mental stress. *Psychophysiology*, *47*(2), 342–354. https://doi.org/10.1111/j.1469-8986.2009.00931.x

Herman, A. M., Rae, C. L., Critchley, H. D., & Duka, T. (2019). Interoceptive accuracy predicts nonplanning trait impulsivity. *Psychophysiology*, *56*(6), e13339. https://doi.org/10.1111/psyp.13339

Hernandez, R., Bassett, S. M., Boughton, S. W., Schuette, S. A., Shiu, E. W., & Moskowitz, J. T. (2018). Psychological Well-being and Physical Health: Associations, Mechanisms, and Future Directions. *Emotion Review: Journal of the International Society for Research on Emotion*, *10*(1), 18–29. https://doi.org/10.1177/1754073917697824

Hickman, L., Seyedsalehi, A., Cook, J. L., Bird, G., & Murphy, J. (2020). The relationship between heartbeat counting and heartbeat discrimination: A meta-analysis. *Biological Psychology*, *156*, 107949. https://doi.org/10.1016/j.biopsycho.2020.107949

Hina, F., & Aspell, J. E. (2019). Altered interoceptive processing in smokers: Evidence from the heartbeat tracking task. *International Journal of Psychophysiology*, *142*, 10–16. https://doi.org/10.1016/j.ijpsycho.2019.05.012

Hoehn-Saric, R., & McLeod, D. R. (2000). Anxiety and arousal: Physiological changes and their perception. *Journal of Affective Disorders*, *61*(3), 217–224. https://doi.org/10.1016/S0165-0327(00)00339-6

Hoehn-Saric, R., McLeod, D. R., Funderburk, F., & Kowalski, P. (2004). Somatic Symptoms and Physiologic Responses in Generalized Anxiety Disorderand Panic Disorder: An Ambulatory Monitor Study. *Archives of General Psychiatry*, *61*(9), 913. https://doi.org/10.1001/archpsyc.61.9.913

Hogeveen, J., Bird, G., Chau, A., Krueger, F., & Grafman, J. (2016). Acquired alexithymia following damage to the anterior insula. *Neuropsychologia*, *82*, 142–148. https://doi.org/10.1016/j.neuropsychologia.2016.01.021

Holzschneider, K., & Mulert, C. (2011). Neuroimaging in anxiety disorders. *Dialogues in Clinical Neuroscience*, *13*(4), 453–461. https://doi.org/10.31887/DCNS.2011.13.4/kholzschneider

Hothorn, T., Zeileis, A., Farebrother, R. W., Cummins, C., Millo, G., & Mitchell, D. (2022). *Testing Linear Regression Models* (Version 0.9-40) [R].

Hunter, J. E., & Schmidt, F. L. (2000). Fixed Effects vs. Random Effects Meta-Analysis Models: Implications for Cumulative Research Knowledge. *International Journal of Selection and Assessment*, *8*(4), 275–292. https://doi.org/10.1111/1468-2389.00156

Ironside, M., DeVille, D. C., Kuplicki, R. T., Burrows, K. P., Smith, R., Teed, A. R., Paulus, M. P., & Khalsa, S. S. (2023). The unique face of comorbid anxiety and depression: Increased interoceptive fearfulness and reactivity. *Frontiers in Behavioral Neuroscience*, *16*, 1083357. https://doi.org/10.3389/fnbeh.2022.1083357

Itti, L., & Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision Research*, *49*(10), 1295–1306. https://doi.org/10.1016/j.visres.2008.09.007

Jackson, L., Keville, S., & Ludlow, A. (2022). Anxiety in female adolescents with autism spectrum disorder: Lessons for healthcare professionals. *Qualitative Health Communication*, *1*(2), 80–98. https://doi.org/10.7146/qhc.v1i2.128871

Jacob, S., Brune, C. W., Carter, C. S., Leventhal, B. L., Lord, C., & Cook, E. H. (2007). Association of the oxytocin receptor gene (OXTR) in Caucasian children and adolescents with autism. *Neuroscience Letters*, *417*(1), 6–9. https://doi.org/10.1016/j.neulet.2007.02.001

Jakobson, L. S., & Rigby, S. N. (2021). Alexithymia and Sensory Processing Sensitivity: Areas of Overlap and Links to Sensory Processing Styles. *Frontiers in Psychology*, *12*, 583786. https://doi.org/10.3389/fpsyg.2021.583786

James, R. J. A., Irons, D. W., Holmes, C., Charlton, A. L., Drewett, R. F., & Baylis, P. H. (1995). Thirst induced by a suckling episode during breast feeding and its relation with plasma vasopressin, oxytocin and osmoregulation. *Clinical Endocrinology*, *43*(3), 277–282. https://doi.org/10.1111/j.1365-2265.1995.tb02032.x

Jenkinson, R., Milne, E., & Thompson, A. (2020). The relationship between intolerance of uncertainty and anxiety in autism: A systematic literature review and meta-analysis. *Autism*, *24*(8), 1933–1944. https://doi.org/10.1177/1362361320932437

Jeon, H., Hur, A., Lee, H., Shin, Y.-W., Lee, S.-I., Shin, C.-J., Kim, S., Ju, G., Lee, J., Jung, J. H., Chung, S., & Son, J.-W. (2024). The Relationship Between Brain Activation for Taking Others' Perspective and Interoceptive Abilities in Autism Spectrum Disorder: An fMRI Study. *Journal of the Korean Academy of Child and Adolescent Psychiatry*, *35*(3), 197–209. https://doi.org/10.5765/jkacap.240008

Jolliffe, R., Adams, D., & Simpson, K. (2023). Trait Anxiety in Individuals on the Autism Spectrum: A Systematic Review. *Review Journal of Autism and Developmental Disorders*, *10*(3), 523–545. https://doi.org/10.1007/s40489-022-00308-8

Jones, G. E., & Hollandsworth, J. G. (1981). Heart rate discrimination before and after exercise-induced augmented cardiac activity. *Psychophysiology*, *18*(3), 252–257. https://doi.org/10.1111/j.1469-8986.1981.tb03029.x

Kaiser, H. F., & Rice, J. (1974). Little Jiffy, Mark Iv. *Educational and Psychological Measurement*, *34*(1), 111–117. https://doi.org/10.1177/001316447403400115

Kandasamy, N., Garfinkel, S. N., Page, L., Hardy, B., Critchley, H. D., Gurnell, M., & Coates, J. M. (2016). Interoceptive Ability Predicts Survival on a London Trading Floor. *Scientific Reports*, *6*(1), 32986. https://doi.org/10.1038/srep32986

Karukivi, M., Hautala, L., Kaleva, O., Haapasalo-Pesu, K.-M., Liuksila, P.-R., Joukamaa, M., & Saarijärvi, S. (2010). Alexithymia is associated with anxiety among adolescents. *Journal of Affective Disorders*, *125*(1–3), 383–387. https://doi.org/10.1016/j.jad.2010.02.126

Katkin, E. S. (1985). Blood, Sweat, and Tears: Individual Differences in Autonomic Self-Perception. *Psychophysiology*, *22*(2), 125–137. https://doi.org/10.1111/j.1469-8986.1985.tb01573.x

Kaufman, J., & Charney, D. (2000). Comorbidity of mood and anxiety disorders. *Depression and Anxiety*, *12*(S1), 69–76. https://doi.org/10.1002/1520-6394(2000)12:1+<69::AID-DA9>3.0.CO;2-K

Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C., & Pellicano, E. (2016). Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism: The International Journal of Research and Practice*, *20*(4), 442–462. https://doi.org/10.1177/1362361315588200

Kerr, K. L., Moseman, S. E., Avery, J. A., Bodurka, J., Zucker, N. L., & Simmons, W. K. (2016). Altered Insula Activity during Visceral Interoception in Weight-Restored

Patients with Anorexia Nervosa. *Neuropsychopharmacology*, *41*(2), 521–528. https://doi.org/10.1038/npp.2015.174

Kessler, R. C., Angermeyer, M., Anthony, J. C., DE Graaf, R., Demyttenaere, K., Gasquet, I., DE Girolamo, G., Gluzman, S., Gureje, O., Haro, J. M., Kawakami, N., Karam, A., Levinson, D., Medina Mora, M. E., Oakley Browne, M. A., Posada-Villa, J., Stein, D. J., Adley Tsang, C. H., Aguilar-Gaxiola, S., … Ustün, T. B. (2007). Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry: Official Journal of the World Psychiatric Association (WPA)*, *6*(3), 168–176.

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

Khalsa, S. S., Craske, M. G., Li, W., Vangala, S., Strober, M., & Feusner, J. D. (2015). Altered interoceptive awareness in anorexia nervosa: Effects of meal anticipation, consumption and bodily arousal. *International Journal of Eating Disorders*, *48*(7), 889–897. https://doi.org/10.1002/eat.22387

Khalsa, S. S., & Lapidus, R. C. (2016). Can Interoception Improve the Pragmatic Search for Biomarkers in Psychiatry? *Frontiers in Psychiatry*, *7*, 121. https://doi.org/10.3389/fpsyt.2016.00121

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

Kinnaird, E., Stewart, C., & Tchanturia, K. (2019). Investigating alexithymia in autism: A systematic review and meta-analysis. *European Psychiatry: The Journal of the Association of European Psychiatrists*, *55*, 80–89. https://doi.org/10.1016/j.eurpsy.2018.09.004

Kinnaird, E., & Tchanturia, K. (2021). Looking beneath the surface: Distinguishing between common features in autism and anorexia nervosa. *Journal of Behavioral and Cognitive Therapy*, *31*(1), 3–13. https://doi.org/10.1016/j.jbct.2020.09.001

Kline, P. A. (1986). *A handbook of test construction*. Methuen.

Knapp, K., Ring, C., & Brener, J. (1997). Sensitivity to mechanical stimuli and the role of general sensory and perceptual processes in heartbeat detection. *Psychophysiology*, *34*(4), 467–473. https://doi.org/10.1111/j.1469-8986.1997.tb02391.x

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

Knapp-Kline, K., Ring, C., Emmerich, D., & Brener, J. (2021). The effects of vibrotactile masking on heartbeat detection: Evidence that somatosensory mechanoreceptors transduce heartbeat sensations. *Psychophysiology*, *58*(6), e13817. https://doi.org/10.1111/psyp.13817

Knoll, J. F., & Hodapp, V. (1992). A Comparison between Two Methods for Assessing Heartbeat Perception. *Psychophysiology*, *29*(2), 218–222. https://doi.org/10.1111/j.1469-8986.1992.tb01689.x

Konting, M. M., Kamaruddin, N., & Man, N. A. (2009). Quality Assurance in Higher Education Institutions: Exist Survey among Universiti Putra Malaysia Graduating

Students. *International Education Studies*, *2*(1), p25. https://doi.org/10.5539/ies.v2n1p25

Koreki, A., Goeta, D., Ricciardi, L., Eilon, T., Chen, J., Critchley, H. D., Garfinkel, S. N., Edwards, M., & Yogarajah, M. (2022). The relationship between interoception and agency and its modulation by heartbeats: An exploratory study. *Scientific Reports*, *12*(1), 13624. https://doi.org/10.1038/s41598-022-16569-6

Körmendi, J., & Ferentzi, E. (2024). Heart activity perception: Narrative review on the measures of the cardiac perceptual ability. *Biologia Futura*, *75*(1), 3–15. https://doi.org/10.1007/s42977-023-00181-4

Körmendi, J., Ferentzi, E., & Köteles, F. (2022). A heartbeat away from a valid tracking task. An empirical comparison of the mental and the motor tracking task. *Biological Psychology*, *171*, 1–8. https://doi.org/10.1016/j.biopsycho.2022.108328

Körmendi, J., Ferentzi, E., Petzke, T., Gál, V., & Köteles, F. (2023). Do we need to accurately perceive our heartbeats? Cardioceptive accuracy and sensibility are independent from indicators of negative affectivity, body awareness, body image dissatisfaction, and alexithymia. *PLOS ONE*, *18*(7), e0287898. https://doi.org/10.1371/journal.pone.0287898

Krautwurst, S., Gerlach, A. L., Gomille, L., Hiller, W., & Witthöft, M. (2014). Health anxiety – An indicator of higher interoceptive sensitivity? *Journal of Behavior Therapy and Experimental Psychiatry*, *45*(2), 303–309. https://doi.org/10.1016/j.jbtep.2014.02.001

Krautwurst, S., Gerlach, A. L., & Witthöft, M. (2016). Interoception in pathological health anxiety. *Journal of Abnormal Psychology*, *125*(8), 1179–1184. https://doi.org/10.1037/abn0000210

Kroenke, K., Yu, Z., Wu, J., Kean, J., & Monahan, P. O. (2014). Operating Characteristics of PROMIS Four-Item Depression and Anxiety Scales in Primary Care Patients with Chronic Pain. *Pain Medicine*, *15*(11), 1892–1901. https://doi.org/10.1111/pme.12537

Kumsta, R., Hummel, E., Chen, F. S., & Heinrichs, M. (2013). Epigenetic regulation of the oxytocin receptor gene: Implications for behavioral neuroscience. *Frontiers in Neuroscience*, *7*. https://doi.org/10.3389/fnins.2013.00083

La Greca, A. M., & Stone, W. L. (1993). Social Anxiety Scale for Children-Revised: Factor Structure and Concurrent Validity. *Journal of Clinical Child Psychology*, *22*(1), 17–27. https://doi.org/10.1207/s15374424jccp2201_2

Lackner, R. J., & Fresco, D. M. (2016). Interaction effect of brooding rumination and interoceptive awareness on depression and anxiety symptoms. *Behaviour Research and Therapy*, *85*, 43–52. https://doi.org/10.1016/j.brat.2016.08.007

Lai, M.-C., Kassee, C., Besney, R., Bonato, S., Hull, L., Mandy, W., Szatmari, P., & Ameis, S. H. (2019). Prevalence of co-occurring mental health diagnoses in the autism population: A systematic review and meta-analysis. *The Lancet. Psychiatry*, *6*(10), 819–829. https://doi.org/10.1016/S2215-0366(19)30289-5

Lai, M.-C., Lombardo, M. V., Auyeung, B., Chakrabarti, B., & Baron-Cohen, S. (2015). Sex/Gender Differences and Autism: Setting the Scene for Future Research. *Journal of the American Academy of Child & Adolescent Psychiatry*, *54*(1), 11–24. https://doi.org/10.1016/j.jaac.2014.10.003

Lamers, F., Van Oppen, P., Comijs, H. C., Smit, J. H., Spinhoven, P., Van Balkom, A. J. L. M., Nolen, W. A., Zitman, F. G., Beekman, A. T. F., & Penninx, B. W. J. H. (2011). Comorbidity Patterns of Anxiety and Depressive Disorders in a Large Cohort Study: The Netherlands Study of Depression and Anxiety (NESDA). *The Journal of Clinical Psychiatry*, *72*(03), 341–348. https://doi.org/10.4088/JCP.10m06176blu

Lange, C. G., & James, W. (1967). *The Emotions*. Hafner.

Larsson, D. E. O., Esposito, G., Critchley, H. D., Dienes, Z., & Garfinkel, S. N. (2021). Sensitivity to changes in rate of heartbeats as a measure of interoceptive ability. *Journal of Neurophysiology*, *126*(5), 1799–1813. https://doi.org/10.1152/jn.00059.2021

Lawson, E. A., Holsen, L. M., Santin, M., Meenaghan, E., Eddy, K. T., Becker, A. E., Herzog, D. B., Goldstein, J. M., & Klibanski, A. (2012). Oxytocin Secretion Is Associated with Severity of Disordered Eating Psychopathology and Insular Cortex Hypoactivation in Anorexia Nervosa. *The Journal of Clinical Endocrinology & Metabolism*, *97*(10), E1898–E1908. https://doi.org/10.1210/jc.2012-1702

Leganes-Fonteneau, M., Cheang, Y., Lam, Y., Garfinkel, S., & Duka, T. (2019). Interoceptive awareness is associated with acute alcohol-induced changes in subjective effects. *Pharmacology Biochemistry and Behavior*, *181*, 69–76. https://doi.org/10.1016/j.pbb.2019.03.007

Legrand, N., Nikolova, N., Correa, C., Brændholt, M., Stuckert, A., Kildahl, N., Vejlø, M., Fardo, F., & Allen, M. (2022). The heart rate discrimination task: A psychophysical method to estimate the accuracy and precision of interoceptive beliefs. *Biological Psychology*, *168*, 108239. https://doi.org/10.1016/j.biopsycho.2021.108239

Leigh, E., & Clark, D. M. (2023). Internet-delivered therapist-assisted cognitive therapy for adolescent social anxiety disorder (OSCA): A randomised controlled trial addressing preliminary efficacy and mechanisms of action. *Journal of Child Psychology and Psychiatry*, *64*(1), 145–155. https://doi.org/10.1111/jcpp.13680

Leonidou, C., Pollatos, O., & Panayiotou, G. (2020). Emotional responses to illness imagery in young adults: Effects of attention to somatic sensations and levels of illness

anxiety. *Biological Psychology*, *149*, 107812. https://doi.org/10.1016/j.biopsycho.2019.107812

Levitt, E. E., & Persky, H. (1962). Experimental evidence for the validity of the IPAT anxiety scale. *Journal of Clinical Psychology*, *18*(4), 458–461. https://doi.org/10.1002/1097-4679(196210)18:4<458::AID-JCLP2270180418>3.0.CO;2-Z

Leweke, F., Leichsenring, F., Kruse, J., & Hermes, S. (2012). Is Alexithymia Associated with Specific Mental Disorders. *Psychopathology*, *45*(1), 22–28. https://doi.org/10.1159/000325170

Li, H., Zhang, B., Hu, Q., Zhang, L., Jin, Y., Wang, J., Cui, H., Pang, J., & Li, C. (2020). Altered heartbeat perception sensitivity associated with brain structural alterations in generalised anxiety disorder. *General Psychiatry*, *33*(1), e100057. https://doi.org/10.1136/gpsych-2019-100057

Lischke, A., Pahnke, R., Mau-Moeller, A., Jacksteit, R., & Weippert, M. (2020). Sex-Specific Relationships Between Interoceptive Accuracy and Emotion Regulation. *Frontiers in Behavioral Neuroscience*, *14*, 67. https://doi.org/10.3389/fnbeh.2020.00067

Liu, X., Kawamura, Y., Shimada, T., Otowa, T., Koishi, S., Sugiyama, T., Nishida, H., Hashimoto, O., Nakagami, R., Tochigi, M., Umekage, T., Kano, Y., Miyagawa, T., Kato, N., Tokunaga, K., & Sasaki, T. (2010). Association of the oxytocin receptor (OXTR) gene polymorphisms with autism spectrum disorder (ASD) in the Japanese population. *Journal of Human Genetics*, *55*(3), 137–141. https://doi.org/10.1038/jhg.2009.140

Livermore, J. J. A., Holmes, C. L., Moga, G., Adamatzky, K., Critchley, H. D., Garfinkel, S. N., & Campbell-Meiklejohn, D. (2022). A single oral dose of citalopram increases interoceptive insight in healthy volunteers. *Psychopharmacology*, *239*(7), 2289–2298. https://doi.org/10.1007/s00213-022-06115-7

Lombardo, T. W., & Epstein, L. H. (1986). The nicotine paradox: Effect of smoking on autonomic discrimination. *Addictive Behaviors*, *11*(3), 341–344. https://doi.org/10.1016/0306-4603(86)90061-4

Longarzo, M., D'Olimpio, F., Chiavazzo, A., Santangelo, G., Trojano, L., & Grossi, D. (2015). The relationships between interoception and alexithymic trait. The Self-Awareness Questionnaire in healthy subjects. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01149

Longarzo, M., D'Olimpio, F., Chiavazzo, A., Santangelo, G., Trojano, L., & Grossi, D. (2016). *Self-Awareness Questionnaire* [Dataset]. https://doi.org/10.1037/t45816-000

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy*, *33*(3), 335–343. https://doi.org/10.1016/0005-7967(94)00075-U

Low, J. (2019). A Pragmatic Definition of the Concept of Theoretical Saturation. *Sociological Focus*, *52*(2), 131–139. https://doi.org/10.1080/00380237.2018.1544514

Ludewig, S., Geyer, M. A., Ramseier, M., Vollenweider, F. X., Rechsteiner, E., & Cattapan-Ludewig, K. (2005). Information-processing deficits and cognitive dysfunction in panic disorder. *Journal of Psychiatry and Neuroscience*, *30*(1), 37–43.

Lyyra, P., & Parviainen, T. (2018). Behavioral Inhibition Underlies the Link Between Interoceptive Sensitivity and Anxiety-Related Temperamental Traits. *Frontiers in Psychology*, *9*, 1026. https://doi.org/10.3389/fpsyg.2018.01026

Maenner, M. J., Warren, Z., Williams, A. R., Amoakohene, E., Bakian, A. V., Bilder, D. A., Durkin, M. S., Fitzgerald, R. T., Furnier, S. M., Hughes, M. M., Ladd-Acosta, C. M., McArthur, D., Pas, E. T., Salinas, A., Vehorn, A., Williams, S., Esler, A., Grzybowski, A., Hall-Lande, J., … Shaw, K. A. (2023). Prevalence and Characteristics of Autism

Spectrum Disorder Among Children Aged 8 Years—Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2020. *MMWR. Surveillance Summaries*, *72*(2), 1–14. https://doi.org/10.15585/mmwr.ss7202a1

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

Maisel, M. E., Stephenson, K. G., South, M., Rodgers, J., Freeston, M. H., & Gaigg, S. B. (2016). Modeling the cognitive mechanisms linking autism symptoms and anxiety in adults. *Journal of Abnormal Psychology*, *125*(5), 692–703. https://doi.org/10.1037/abn0000168

Mallorquí--Bagué, N., Garfinkel, S. N., Engels, M., Eccles, J. A., Pailhez, G., Bulbena, A., & Critchley, H. D. (2014). Neuroimaging and psychophysiological investigation of the link between anxiety, enhanced affective reactivity and interoception in people with joint hypermobility. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.01162

Malterud, K., Siersma, V. D., & Guassora, A. D. (2016). Sample Size in Qualitative Interview Studies: Guided by Information Power. *Qualitative Health Research*, *26*(13), 1753–1760. https://doi.org/10.1177/1049732315617444

Martínez, K., Martínez-García, M., Marcos-Vidal, L., Janssen, J., Castellanos, F. X., Pretus, C., Villarroya, Ó., Pina-Camacho, L., Díaz-Caneja, C. M., Parellada, M., Arango, C., Desco, M., Sepulcre, J., & Carmona, S. (2020). Sensory-to-Cognitive Systems Integration Is Associated With Clinical Severity in Autism Spectrum Disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *59*(3), 422–433. https://doi.org/10.1016/j.jaac.2019.05.033

Masdrakis, V. G., Papakostas, Y. G., Vaidakis, N., Papageorgiou, C., & Pehlivanidis, A. (2008). Caffeine challenge in patients with panic disorder: Baseline differences between those who panic and those who do not. *Depression and Anxiety*, *25*(9), E72–E79. https://doi.org/10.1002/da.20333

Mash, L. E., Schauder, K. B., Cochran, C., Park, S., & Cascio, C. J. (2017). Associations Between Interoceptive Cognition and Age in Autism Spectrum Disorder and Typical Development. *Journal of Cognitive Education and Psychology*, *16*(1), 23–37. https://doi.org/10.1891/1945-8959.16.1.23

Mayeli, A., Al Zoubi, O., White, E. J., Chappelle, S., Kuplicki, R., Morton, A., Bruce, J., Smith, R., Feinstein, J. S., Bodurka, J., Paulus, M. P., & Khalsa, S. S. (2023). Parieto-occipital ERP indicators of gut mechanosensation in humans. *Nature Communications*, *14*(1), 3398. https://doi.org/10.1038/s41467-023-39058-4

McKenzie, K., Metcalfe, D., & Murray, A. L. (2023). Screening for intellectual disability in autistic people: A brief report. *Research in Autism Spectrum Disorders*, *100*, 102076. https://doi.org/10.1016/j.rasd.2022.102076

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. https://doi.org/10.1371/journal.pone.0208034

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

Meza-Concha, N., Arancibia, M., Salas, F., Behar, R., Salas, G., Silva, H., & Escobar, R. (2017). Towards a neurobiological understanding of alexithymia. *Medwave*, *17*(04), e6960–e6960. https://doi.org/10.5867/medwave.2017.04.6960

Michal, M., Reuchlein, B., Adler, J., Reiner, I., Beutel, M. E., Vögele, C., Schächinger, H., & Schulz, A. (2014). Striking Discrepancy of Anomalous Body Experiences with Normal Interoceptive Accuracy in Depersonalization-Derealization Disorder. *PLoS ONE*, *9*(2), e89823. https://doi.org/10.1371/journal.pone.0089823

Miller, D., Rees, J., & Pearson, A. (2021). "Masking Is Life": Experiences of Masking in Autistic and Nonautistic Adults. *Autism in Adulthood*, *3*(4), 330–338. https://doi.org/10.1089/aut.2020.0083

Miller, L. C., Murphy, R., & Buss, A. H. (1981). Consciousness of body: Private and public. *Journal of Personality and Social Psychology*, *41*(2), 397–406. https://doi.org/10.1037/0022-3514.41.2.397

Milosavljevic, B., Carter Leno, V., Simonoff, E., Baird, G., Pickles, A., Jones, C. R. G., Erskine, C., Charman, T., & Happé, F. (2016). Alexithymia in Adolescents with Autism Spectrum Disorder: Its Relationship to Internalising Difficulties, Sensory Modulation and Social Cognition. *Journal of Autism and Developmental Disorders*, *46*(4), 1354–1367. https://doi.org/10.1007/s10803-015-2670-8

Mimura, C., Nishioka, M., Sato, N., Hasegawa, R., Horikoshi, R., & Watanabe, Y. (2011). A Japanese Scale to Assess Anxiety Severity: Development and Psychometric Evaluation. *The International Journal of Psychiatry in Medicine*, *41*(1), 29–45. https://doi.org/10.2190/PM.41.1.d

Modahl, C., Green, L. A., Fein, D., Morris, M., Waterhouse, L., Feinstein, C., & Levin, H. (1998). Plasma oxytocin levels in autistic children. *Biological Psychiatry*, *43*(4), 270–277. https://doi.org/10.1016/S0006-3223(97)00439-3

Mojtabai, R., Stuart, E. A., Hwang, I., Susukida, R., Eaton, W. W., Sampson, N., & Kessler, R. C. (2015). Long-term effects of mental disorders on employment in the National

Comorbidity Survey ten-year follow-up. *Social Psychiatry and Psychiatric Epidemiology*, *50*(11), 1657–1668. https://doi.org/10.1007/s00127-015-1097-z

Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). Multisensory auditory–visual interactions during early sensory processing in humans: A high-density electrical mapping study. *Cognitive Brain Research*, *14*(1), 115–128. https://doi.org/10.1016/S0926-6410(02)00066-6

Montoya-Hurtado, O. L., Gómez-Jaramillo, N., Criado-Gutiérrez, J. M., Pérez, J., Sancho-Sánchez, C., Sánchez-Barba, M., Tejada-Garrido, C. I., Criado-Pérez, L., Sánchez-González, J. L., Santolalla-Arnedo, I., Del Pozo-Herce, P., & Juárez-Vela, R. (2023). Exploring the Link between Interoceptive Body Awareness and Suicidal Orientation in University Students: A Cross-Sectional Study. *Behavioral Sciences (Basel, Switzerland)*, *13*(11), 945. https://doi.org/10.3390/bs13110945

Morelli, D., Bartoloni, L., Colombo, M., Plans, D., & Clifton, D. A. (2018). Profiling the propagation of error from PPG to HRV features in a wearable physiological-monitoring device. *Healthcare Technology Letters*, *5*(2), 59–64. https://doi.org/10.1049/htl.2017.0039

Mul, C., Stagg, S. D., Herbelin, B., & Aspell, J. E. (2018). The Feeling of Me Feeling for You: Interoception, Alexithymia and Empathy in Autism. *Journal of Autism and Developmental Disorders*, *48*(9), 2953–2967. https://doi.org/10.1007/s10803-018-3564-3

Murphy, J., Brewer, R., Catmur, C., & Bird, G. (2017). Interoception and psychopathology: A developmental neuroscience perspective. *Developmental Cognitive Neuroscience*, *23*, 45–56. https://doi.org/10.1016/j.dcn.2016.12.006

Murphy, J., Brewer, R., Coll, M.-P., Plans, D., Hall, M., Shiu, S. S., Catmur, C., & Bird, G. (2019). I feel it in my finger: Measurement device affects cardiac interoceptive

accuracy. *Biological Psychology*, *148*, 107765. https://doi.org/10.1016/j.biopsycho.2019.107765

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

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

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

Murphy, J., Catmur, C., & Bird, G. (2019). Classifying individual differences in interoception: Implications for the measurement of interoceptive awareness. *Psychonomic Bulletin & Review*, *26*(5), 1467–1471. https://doi.org/10.3758/s13423-019-01632-7

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

Murray, D., Lesser, M., & Lawson, W. (2005). Attention, monotropism and the diagnostic criteria for autism. *Autism*, *9*(2), 139–156. https://doi.org/10.1177/1362361305051398

Nagai, M., Kishi, K., & Kato, S. (2007). Insular cortex and neuropsychiatric disorders: A review of recent literature. *European Psychiatry*, *22*(6), 387–394. https://doi.org/10.1016/j.eurpsy.2007.02.006

Nakao, M., Barsky, A. J., Kumano, H., & Kuboki, T. (2002). Relationship Between Somatosensory Amplification and Alexithymia in a Japanese Psychosomatic Clinic. *Psychosomatics*, *43*(1), 55–60. https://doi.org/10.1176/appi.psy.43.1.55

Näring, G. W. B., & Van Der Staak, C. P. F. (1995). Perception of Heart Rate and Blood Pressure: The Role of Alexithymia and Anxiety. *Psychotherapy and Psychosomatics*, *63*(3–4), 193–200. https://doi.org/10.1159/000288959

National Autistic Society. (2021). *Good practice guide For professionals delivering talking therapies for autistic adults and children*. National Autistic Society. https://s2.chorus-mk.thirdlight.com/file/24/asDKIN9as.klK7easFDsalAzTC/NAS-Good-Practice-Guide-A4.pdf

Nemiah, J. C., Freyberger, H., & Sifneos, P. E. (1976). Alexithymia: A view of the psychosomatic process. In O. W. Hill (Ed.), *Modern trends in psychosomatic medicine* (Vol. 3, pp. 430–439). Butterworths.

Nicholson, T. M., Williams, D., Carpenter, K., & Kallitsounaki, A. (2019). Interoception is Impaired in Children, But Not Adults, with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *49*(9), 3625–3637. https://doi.org/10.1007/s10803-019-04079-w

Nicholson, T. M., Williams, D. M., Grainger, C., Christensen, J. F., Calvo-Merino, B., & Gaigg, S. B. (2018). Interoceptive impairments do not lie at the heart of autism or alexithymia. *Journal of Abnormal Psychology*, *127*(6), 612–622. https://doi.org/10.1037/abn0000370

Noda, M., Sato, Y., Suetsugu, Y., & Morokuma, S. (2022). Interoception is associated with anxiety and depression in pregnant women: A pilot study. *PLOS ONE*, *17*(5), e0267507. https://doi.org/10.1371/journal.pone.0267507

Nord, C. L., & Garfinkel, S. N. (2022). Interoceptive pathways to understand and treat mental health conditions. *Trends in Cognitive Sciences*, *26*(6), 499–513. https://doi.org/10.1016/j.tics.2022.03.004

Oakley, B. F. M., Jones, E. J. H., Crawley, D., Charman, T., Buitelaar, J., Tillmann, J., Murphy, D. G., Loth, E., & The EU-AIMS LEAP Group. (2022). Alexithymia in autism: Cross-sectional and longitudinal associations with social-communication difficulties, anxiety and depression symptoms. *Psychological Medicine*, *52*(8), 1458–1470. https://doi.org/10.1017/S0033291720003244

Ohst, B., & Tuschen-Caffier, B. (2018). Catastrophic misinterpretation of bodily sensations and external events in panic disorder, other anxiety disorders, and healthy subjects: A systematic review and meta-analysis. *PloS One*, *13*(3), e0194493. https://doi.org/10.1371/journal.pone.0194493

Olatunji, B. O., Deacon, B. J., Abramowitz, J. S., & Valentiner, D. P. (2007). Body Vigilance in Nonclinical and Anxiety Disorder Samples: Structure, Correlates, and Prediction of Health Concerns. *Behavior Therapy*, *38*(4), 392–401. https://doi.org/10.1016/j.beth.2006.09.002

Olatunji, B. O., & Wolitzky-Taylor, K. B. (2009). Anxiety sensitivity and the anxiety disorders: A meta-analytic review and synthesis. *Psychological Bulletin*, *135*(6), 974–999. https://doi.org/10.1037/a0017428

Olthuis, J. V., Watt, M. C., & Stewart, S. H. (2014). Anxiety Sensitivity Index (ASI-3) subscales predict unique variance in anxiety and depressive symptoms. *Journal of Anxiety Disorders*, *28*(2), 115–124. https://doi.org/10.1016/j.janxdis.2013.04.009

Ortmann, J., Lutz, A. P. C., Rose, G., Happ, C., Vögele, C., Schulz, A., & van Dyck, Z. (2024). Development and initial validation of a self-report measure to assess eating disorder-specific interoceptive perception. *Psychological Assessment*, *36*(2), 162–174. https://doi.org/10.1037/pas0001283

Owens, A. P., Allen, M., Ondobaka, S., & Friston, K. J. (2018). Interoceptive inference: From computational neuroscience to clinic. *Neuroscience and Biobehavioral Reviews*, *90*, 174–183. https://doi.org/10.1016/j.neubiorev.2018.04.017

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ*, n160. https://doi.org/10.1136/bmj.n160

Palser, E. R., Fotopoulou, A., Pellicano, E., & Kilner, J. M. (2018). The link between interoceptive processing and anxiety in children diagnosed with autism spectrum disorder: Extending adult findings into a developmental sample. *Biological Psychology*, *136*, 13–21. https://doi.org/10.1016/j.biopsycho.2018.05.003

Palser, E. R., Palmer, C. E., Galvez-Pol, A., Hannah, R., Fotopoulou, A., & Kilner, J. M. (2018). Alexithymia mediates the relationship between interoceptive sensibility and anxiety. *PLoS ONE*, *13*(9), e0203212. https://doi.org/10.1371/journal.pone.0203212

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

Paton, J. F., Li, Y. W., Deuchars, J., & Kasparov, S. (2000). Properties of solitary tract neurons receiving inputs from the sub-diaphragmatic vagus nerve. *Neuroscience*, *95*(1), 141–153. https://doi.org/10.1016/s0306-4522(99)00416-9

Paulus, M. P. (2013). The breathing conundrum-interoceptive sensitivity and anxiety. *Depression and Anxiety*, *30*(4), 315–320. https://doi.org/10.1002/da.22076

Paulus, M. P., Feinstein, J. S., & Khalsa, S. S. (2019). An Active Inference Approach to Interoceptive Psychopathology. *Annual Review of Clinical Psychology*, *15*(1), 97–122. https://doi.org/10.1146/annurev-clinpsy-050718-095617

Paulus, M. P., & Stein, M. B. (2006). An Insular View of Anxiety. *Biological Psychiatry*, *60*(4), 383–387. https://doi.org/10.1016/j.biopsych.2006.03.042

Paulus, M. P., & Stein, M. B. (2010). Interoception in anxiety and depression. *Brain Structure & Function*, *214*(5–6), 451–463. https://doi.org/10.1007/s00429-010-0258-9

Pearson, A., & Rose, K. (2021). A Conceptual Analysis of Autistic Masking: Understanding the Narrative of Stigma and the Illusion of Choice. *Autism in Adulthood*, *3*(1), 52–60. https://doi.org/10.1089/aut.2020.0043

Pellicano, E., & Burr, D. (2012). When the world becomes 'too real': A Bayesian explanation of autistic perception. *Trends in Cognitive Sciences*, *16*(10), 504–510. https://doi.org/10.1016/j.tics.2012.08.009

Petersen, S., Schroijen, M., Mölders, C., Zenker, S., & Van Den Bergh, O. (2014). Categorical Interoception: Perceptual Organization of Sensations From Inside. *Psychological Science*, *25*(5), 1059–1066. https://doi.org/10.1177/0956797613519110

Petersen, S., Van Staeyen, K., Vögele, C., von Leupoldt, A., & Van den Bergh, O. (2015). Interoception and symptom reporting: Disentangling accuracy and bias. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00732

Pfeifer, J. H., & Allen, N. B. (2021). Puberty Initiates Cascading Relationships Between Neurodevelopmental, Social, and Internalizing Processes Across Adolescence. *Biological Psychiatry*, *89*(2), 99–108. https://doi.org/10.1016/j.biopsych.2020.09.002

Pfleiderer, B., Berse, T., Stroux, D., Ewert, A., Konrad, C., & Gerlach, A. L. (2014). Internal focus of attention in anxiety-sensitive females up-regulates amygdale activity: An fMRI study. *Journal of Neural Transmission*, *121*(11), 1417–1428. https://doi.org/10.1007/s00702-014-1248-5

Pickard, H., Hirsch, C., Simonoff, E., & Happé, F. (2020). Exploring the cognitive, emotional and sensory correlates of social anxiety in autistic and neurotypical adolescents. *Journal of Child Psychology and Psychiatry*, *61*(12), 1317–1327. https://doi.org/10.1111/jcpp.13214

Pile, V., Lau, J. Y. F., Topor, M., Hedderly, T., & Robinson, S. (2018). Interoceptive Accuracy in Youth with Tic Disorders: Exploring Links with Premonitory Urge, Anxiety and Quality of Life. *Journal of Autism and Developmental Disorders*, *48*(10), 3474–3482. https://doi.org/10.1007/s10803-018-3608-8

Pilkonis, P. A., Choi, S. W., Reise, S. P., Stover, A. M., Riley, W. T., Cella, D., & PROMIS Cooperative Group. (2011). Item Banks for Measuring Emotional Distress From the Patient-Reported Outcomes Measurement Information System (PROMIS®): Depression, Anxiety, and Anger. *Assessment*, *18*(3), 263–283. https://doi.org/10.1177/1073191111411667

Pilowsky, I. (1967). Dimensions of Hypochondriasis. *British Journal of Psychiatry*, *113*(494), 89–93. https://doi.org/10.1192/bjp.113.494.89

Pittig, A., Arch, J. J., Lam, C. W. R., & Craske, M. G. (2013). Heart rate and heart rate variability in panic, social anxiety, obsessive–compulsive, and generalized anxiety disorders at baseline and in response to relaxation and hyperventilation. *International*

Journal of Psychophysiology, *87*(1), 19–27. https://doi.org/10.1016/j.ijpsycho.2012.10.012

Plans, D., Ponzo, S., Morelli, D., Cairo, M., Ring, C., Keating, C. T., Cunningham, A. C., Catmur, C., Murphy, J., & Bird, G. (2021). Measuring interoception: The phase adjustment task. *Biological Psychology*, *165*, 108171. https://doi.org/10.1016/j.biopsycho.2021.108171

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: Biological Sciences*, *379*(1908), 20230256. https://doi.org/10.1098/rstb.2023.0256

Pollatos, O., & Georgiou, E. (2016). Normal interoceptive accuracy in women with bulimia nervosa. *Psychiatry Research*, *240*, 328–332. https://doi.org/10.1016/j.psychres.2016.04.072

Pollatos, O., Herbert, B., Kaufmann, C., Auer, D., & Schandry, R. (2007). Interoceptive awareness, anxiety and cardiovascular reactivity to isometric exercise. *International Journal of Psychophysiology*, *65*(2), 167–173. https://doi.org/10.1016/j.ijpsycho.2007.03.005

Pollatos, O., Herbert, B. M., Mai, S., & Kammer, T. (2016). Changes in interoceptive processes following brain stimulation. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *371*(1708), 20160016. https://doi.org/10.1098/rstb.2016.0016

Pollatos, O., Matthias, E., & Schandry, R. (2007). Heartbeat perception and P300 amplitude in a visual oddball paradigm. *Clinical Neurophysiology*, *118*(10), 2248–2253. https://doi.org/10.1016/j.clinph.2007.06.057

Pollatos, O., Schandry, R., Auer, D., & Kaufmann, C. (2007). Brain structures mediating cardiovascular arousal and interoceptive awareness. *Brain Research*, *1141*, 178–187. https://doi.org/10.1016/j.brainres.2007.01.026

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

Pollatos, O., Traut-Mattausch, E., Schroeder, H., & Schandry, R. (2007). Interoceptive awareness mediates the relationship between anxiety and the intensity of unpleasant feelings. *Journal of Anxiety Disorders*, *21*(7), 931–943. https://doi.org/10.1016/j.janxdis.2006.12.004

Porges. (1993). *Body Perception Questionnaire.*

Postorino, V., Kerns, C. M., Vivanti, G., Bradshaw, J., Siracusano, M., & Mazzone, L. (2017). Anxiety Disorders and Obsessive-Compulsive Disorder in Individuals with Autism Spectrum Disorder. *Current Psychiatry Reports*, *19*(12), 92. https://doi.org/10.1007/s11920-017-0846-y

Preece, D. A., Mehta, A., Becerra, R., Chen, W., Allan, A., Robinson, K., Boyes, M., Hasking, P., & Gross, J. J. (2022). Why is alexithymia a risk factor for affective disorder symptoms? The role of emotion regulation. *Journal of Affective Disorders*, *296*, 337–341. https://doi.org/10.1016/j.jad.2021.09.085

Prentice, F., & Murphy, J. (2022). Sex differences in interoceptive accuracy: A meta-analysis. *Neuroscience and Biobehavioral Reviews*, *132*, 497–518. https://doi.org/10.1016/j.neubiorev.2021.11.030

Proff, I., Williams, G. L., Quadt, L., & Garfinkel, S. N. (2022). Sensory processing in autism across exteroceptive and interoceptive domains. *Psychology & Neuroscience*, *15*(2), 105–130. https://doi.org/10.1037/pne0000262

Quadt, L., Garfinkel, S. N., Mulcahy, J. S., Larsson, D. E., Silva, M., Jones, A.-M., Strauss, C., & Critchley, H. D. (2021). Interoceptive training to target anxiety in autistic adults (ADIE): A single-center, superiority randomized controlled trial. *EClinicalMedicine*, *39*, 101042. https://doi.org/10.1016/j.eclinm.2021.101042

Quattrocki, E., & Friston, K. (2014). Autism, oxytocin and interoception. *Neuroscience and Biobehavioral Reviews*, *47*, 410–430. https://doi.org/10.1016/j.neubiorev.2014.09.012

Reichenberger, J., Schwarz, M., König, D., Wilhelm, F. H., Voderholzer, U., Hillert, A., & Blechert, J. (2016). Angst vor negativer sozialer Bewertung: Übersetzung und Validierung der Furcht vor negativer Evaluation–Kurzskala (FNE-K). *Diagnostica*, *62*(3), 169–181. https://doi.org/10.1026/0012-1924/a000148

Reiss, S. (1991). Expectancy model of fear, anxiety, and panic. *Clinical Psychology Review*, *11*(2), 141–153. https://doi.org/10.1016/0272-7358(91)90092-9

Reiss, S., Peterson, R. A., Gursky, D. M., & McNally, R. J. (1986). Anxiety sensitivity, anxiety frequency and the prediction of fearfulness. *Behaviour Research and Therapy*, *24*(1), 1–8. https://doi.org/10.1016/0005-7967(86)90143-9

Riem, M. M. E., Bakermans-Kranenburg, M. J., Pieper, S., Tops, M., Boksem, M. A. S., Vermeiren, R. R. J. M., Van IJzendoorn, M. H., & Rombouts, S. A. R. B. (2011). Oxytocin Modulates Amygdala, Insula, and Inferior Frontal Gyrus Responses to Infant Crying: A Randomized Controlled Trial. *Biological Psychiatry*, *70*(3), 291–297. https://doi.org/10.1016/j.biopsych.2011.02.006

Ring, C., & Brener, J. (1996). Influence of beliefs about heart rate and actual heart rate on heartbeat counting. *Psychophysiology*, *33*(5), 541–546. https://doi.org/10.1111/j.1469-8986.1996.tb02430.x

Ring, C., & Brener, J. (2018). Heartbeat counting is unrelated to heartbeat detection: A comparison of methods to quantify interoception. *Psychophysiology*, *55*(9), e13084. https://doi.org/10.1111/psyp.13084

Ring, C., Brener, J., Knapp, K., & Mailloux, J. (2015). Effects of heartbeat feedback on beliefs about heart rate and heartbeat counting: A cautionary tale about interoceptive awareness. *Biological Psychology*, *104*, 193–198. https://doi.org/10.1016/j.biopsycho.2014.12.010

Robertson, A. E., Stanfield, A. C., Watt, J., Barry, F., Day, M., Cormack, M., & Melville, C. (2018). The experience and impact of anxiety in autistic adults: A thematic analysis. *Research in Autism Spectrum Disorders*, *46*, 8–18. https://doi.org/10.1016/j.rasd.2017.11.006

Rodgers, J., & Ofield, A. (2018). Understanding, Recognising and Treating Co-occurring Anxiety in Autism. *Current Developmental Disorders Reports*, *5*(1), 58–64. https://doi.org/10.1007/s40474-018-0132-7

Rodriguez, B. F., Bruce, S. E., Pagano, M. E., Spencer, M. A., & Keller, M. B. (2004). Factor structure and stability of the Anxiety Sensitivity Index in a longitudinal study of anxiety disorder patients. *Behaviour Research and Therapy*, *42*(1), 79–91. https://doi.org/10.1016/S0005-7967(03)00074-3

Rosseel, Y. (2012). lavaan: An *R* Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2). https://doi.org/10.18637/jss.v048.i02

Rosso, I. M., Makris, N., Britton, J. C., Price, L. M., Gold, A. L., Zai, D., Bruyere, J., Deckersbach, T., Killgore, W. D. S., & Rauch, S. L. (2010). Anxiety sensitivity correlates with two indices of right anterior insula structure in specific animal phobia. *Depression and Anxiety*, *27*(12), 1104–1110. https://doi.org/10.1002/da.20765

Rost, S., Van Ryckeghem, D. M. L., Schulz, A., Crombez, G., & Vögele, C. (2017). Generalized hypervigilance in fibromyalgia: Normal interoceptive accuracy, but reduced self-regulatory capacity. *Journal of Psychosomatic Research*, *93*, 48–54. https://doi.org/10.1016/j.jpsychores.2016.12.003

Rouse, C. H., Jones, G. E., & Jones, K. R. (1988). The Effect of Body Composition and Gender on Cardiac Awareness. *Psychophysiology*, *25*(4), 400–407. https://doi.org/10.1111/j.1469-8986.1988.tb01876.x

Rutter, M., & Sroufe, L. A. (2000). Developmental psychopathology: Concepts and challenges. *Development and Psychopathology*, *12*(3), 265–296. https://doi.org/10.1017/S0954579400003023

Salkovskis, P. M., Rimes, K. A., Warwick, H. M. C., & Clark, D. M. (2002). The Health Anxiety Inventory: Development and validation of scales for the measurement of health anxiety and hypochondriasis. *Psychological Medicine*, *32*(05). https://doi.org/10.1017/S0033291702005822

Salomon, R., Ronchi, R., Dönz, J., Bello-Ruiz, J., Herbelin, B., Martet, R., Faivre, N., Schaller, K., & Blanke, O. (2016). The Insula Mediates Access to Awareness of Visual Stimuli Presented Synchronously to the Heartbeat. *The Journal of Neuroscience*, *36*(18), 5115–5127. https://doi.org/10.1523/JNEUROSCI.4262-15.2016

Santangelo, G., Vitale, C., Baiano, C., D'Iorio, A., Longo, K., Barone, P., Amboni, M., & Conson, M. (2018). Interoceptive processing deficit: A behavioral marker for subtyping Parkinson's disease. *Parkinsonism & Related Disorders*, *53*, 64–69. https://doi.org/10.1016/j.parkreldis.2018.05.001

Scarpazza, C., Zangrossi, A., Huang, Y.-C., Sartori, G., & Massaro, S. (2022). Disentangling interoceptive abilities in alexithymia. *Psychological Research*, *86*(3), 844–857. https://doi.org/10.1007/s00426-021-01538-x

Schaefer, M., Egloff, B., Gerlach, A. L., & Witthöft, M. (2014). Improving heartbeat perception in patients with medically unexplained symptoms reduces symptom distress. *Biological Psychology*, *101*, 69–76. https://doi.org/10.1016/j.biopsycho.2014.05.012

Schandry, R. (1981). Heart beat perception and emotional experience. *Psychophysiology*, *18*(4), 483–488. https://doi.org/10.1111/j.1469-8986.1981.tb02486.x

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

Schandry, R., & Specht, G. (1981). The influence of psychological and physical stress on the perception of heartbeats [Abstract]. *Psychophysiology*, *18*(2), 137–205. https://doi.org/10.1111/j.1469-8986.1981.tb02928.x

Schauder, K. B., Mash, L. E., Bryant, L. K., & Cascio, C. J. (2015). Interoceptive ability and body awareness in autism spectrum disorder. *Journal of Experimental Child Psychology*, *131*, 193–200. https://doi.org/10.1016/j.jecp.2014.11.002

Schirmer-Mokwa, K. L., Fard, P. R., Zamorano, A. M., Finkel, S., Birbaumer, N., & Kleber, B. A. (2015). Evidence for Enhanced Interoceptive Accuracy in Professional Musicians. *Frontiers in Behavioral Neuroscience*, *9*. https://doi.org/10.3389/fnbeh.2015.00349

Schmitz, J., Blechert, J., Krämer, M., Asbrand, J., & Tuschen-Caffier, B. (2012). Biased Perception and Interpretation of Bodily Anxiety Symptoms in Childhood Social Anxiety. *Journal of Clinical Child & Adolescent Psychology*, *41*(1), 92–102. https://doi.org/10.1080/15374416.2012.632349

Schultchen, D., Zaudig, M., Krauseneck, T., Berberich, G., & Pollatos, O. (2019). Interoceptive deficits in patients with obsessive-compulsive disorder in the time

course of cognitive-behavioral therapy. *PLOS ONE*, *14*(5), e0217237. https://doi.org/10.1371/journal.pone.0217237

Schulz, A., Lass-Hennemann, J., Sütterlin, S., Schächinger, H., & Vögele, C. (2013). Cold pressor stress induces opposite effects on cardioceptive accuracy dependent on assessment paradigm. *Biological Psychology*, *93*(1), 167–174. https://doi.org/10.1016/j.biopsycho.2013.01.007

Schulz, A., & Vögele, C. (2015). Interoception and stress. *Frontiers in Psychology*, *6*, 993. https://doi.org/10.3389/fpsyg.2015.00993

Schwarz, D., Witthöft, M., & Bailer, J. (2007). Kritische Überprüfung der Faktorenstruktur und Konstruktvalidität des Whiteley-Index—Einem etablierten Screening-Instrument zur Erfassung von Hypochondrie. *Zeitschrift Für Klinische Psychologie Und Psychotherapie*, *36*(2), 128–136. https://doi.org/10.1026/1616-3443.36.2.128

Schwarzer, G., Carpenter, J. R., & Rücker, G. (2015). *Meta-Analysis with r*. Springer.

Scott, K. M., Bruffaerts, R., Tsang, A., Ormel, J., Alonso, J., Angermeyer, M. C., Benjet, C., Bromet, E., de Girolamo, G., de Graaf, R., Gasquet, I., Gureje, O., Haro, J. M., He, Y., Kessler, R. C., Levinson, D., Mneimneh, Z. N., Oakley Browne, M. A., Posada-Villa, J., … Von Korff, M. (2007). Depression-anxiety relationships with chronic physical conditions: Results from the World Mental Health Surveys. *Journal of Affective Disorders*, *103*(1–3), 113–120. https://doi.org/10.1016/j.jad.2007.01.015

Scott, K. M., Lim, C., Al-Hamzawi, A., Alonso, J., Bruffaerts, R., Caldas-de-Almeida, J. M., Florescu, S., De Girolamo, G., Hu, C., De Jonge, P., Kawakami, N., Medina-Mora, M. E., Moskalewicz, J., Navarro-Mateu, F., O'Neill, S., Piazza, M., Posada-Villa, J., Torres, Y., & Kessler, R. C. (2016). Association of Mental Disorders With Subsequent Chronic Physical Conditions: World Mental Health Surveys From 17 Countries. *JAMA Psychiatry*, *73*(2), 150. https://doi.org/10.1001/jamapsychiatry.2015.2688

Seth, A. K. (2013). Interoceptive inference, emotion, and the embodied self. *Trends in Cognitive Sciences*, *17*(11), 565–573. https://doi.org/10.1016/j.tics.2013.09.007

Shah, P., Hall, R., Catmur, C., & Bird, G. (2016). Alexithymia, not autism, is associated with impaired interoception. *Cortex*, *81*, 215–220. https://doi.org/10.1016/j.cortex.2016.03.021

Sherrington, C. S. (1906). *The integrative action of the nervous system* (pp. xvi, 411). Yale University Press. https://doi.org/10.1037/13798-000

Shields, S. A., Mallory, M. E., & Simon, A. (1989). The Body Awareness Questionnaire: Reliability and validity. *Journal of Personality Assessment*, *53*(4), 802–815. https://doi.org/10.1207/s15327752jpa5304_16

Sidik, K., & Jonkman, J. N. (2007). A comparison of heterogeneity variance estimators in combining results of studies. *Statistics in Medicine*, *26*(9), 1964–1981. https://doi.org/10.1002/sim.2688

Silverman, W. K., Fleisig, W., Rabian, B., & Peterson, R. A. (1991). Childhood Anxiety Sensitivity Index. *Journal of Clinical Child Psychology*, *20*(2), 162–168. https://doi.org/10.1207/s15374424jccp2002_7

Skuse, D. H., Lori, A., Cubells, J. F., Lee, I., Conneely, K. N., Puura, K., Lehtimäki, T., Binder, E. B., & Young, L. J. (2014). Common polymorphism in the oxytocin receptor gene (*OXTR*) is associated with human social recognition skills. *Proceedings of the National Academy of Sciences*, *111*(5), 1987–1992. https://doi.org/10.1073/pnas.1302985111

Slotta, T., Witthöft, M., Gerlach, A. L., & Pohl, A. (2021). The interplay of interoceptive accuracy, facets of interoceptive sensibility, and trait anxiety: A network analysis. *Personality and Individual Differences*, *183*, 111133. https://doi.org/10.1016/j.paid.2021.111133

Slotta, T., Wolters, C., Marx, Z., Witthöft, M., Gerlach, A. L., & Pohl, A. (2023). Respiratory Interoception and Pathological Illness Anxiety: Disentangling Bias. *Psychosomatic Medicine*, *85*(9), 778–784. https://doi.org/10.1097/PSY.0000000000001244

Smári, J., Erlendsdóttir, G., Björgvinsdóttir, A., & Ágústsdóttir, V. R. (2003). ANXIETY SENSITIVITY AND TRAIT-SYMPTOM MEASURES OF ANXIETY AND DEPRESSION. *Anxiety, Stress & Coping*, *16*(4), 375–386. https://doi.org/10.1080/1061580031000107791

Smith, R., Feinstein, J. S., Kuplicki, R., Forthman, K. L., Stewart, J. L., Paulus, M. P., Tulsa 1000 Investigators, Aupperle, R. L., Bodurka, J., Savitz, J. B., Victor, T. A., & Khalsa, S. S. (2021). Perceptual insensitivity to the modulation of interoceptive signals in depression, anxiety, and substance use disorders. *Scientific Reports*, *11*(1), 2108. https://doi.org/10.1038/s41598-021-81307-3

Smith, R., & Lane, R. D. (2015). The neural basis of one's own conscious and unconscious emotional states. *Neuroscience and Biobehavioral Reviews*, *57*, 1–29. https://doi.org/10.1016/j.neubiorev.2015.08.003

Smith, R., Mayeli, A., Taylor, S., Al Zoubi, O., Naegele, J., & Khalsa, S. S. (2021). Gut inference: A computational modelling approach. *Biological Psychology*, *164*, 108152. https://doi.org/10.1016/j.biopsycho.2021.108152

Solano López, A. L., & Moore, S. (2019). Dimensions of Body-Awareness and Depressed Mood and Anxiety. *Western Journal of Nursing Research*, *41*(6), 834–853. https://doi.org/10.1177/0193945918798374

Spain, D., Milner, V., Mason, D., Iannelli, H., Attoe, C., Ampegama, R., Kenny, L., Saunders, A., Happé, F., & Marshall-Tate, K. (2023). Improving Cognitive Behaviour Therapy for Autistic Individuals: A Delphi Survey with Practitioners. *Journal of Rational-*

Emotive & Cognitive-Behavior Therapy, *41*(1), 45–63. https://doi.org/10.1007/s10942-022-00452-4

Spain, D., Zıvralı Yarar, E., & Happé, F. (2020). Social anxiety in adults with autism: A qualitative study. *International Journal of Qualitative Studies on Health and Well-Being*, *15*(1), 1803669. https://doi.org/10.1080/17482631.2020.1803669

Speckens, A. E. M., Spinhoven, P., Sloekers, P. P. A., Bolk, J. H., & Van Hemert, A. M. (1996). A validation study of the Whitely Index, the Illness Attitude Scales, and the Somatosensory Amplification Scale in general medical and general practice patients. *Journal of Psychosomatic Research*, *40*(1), 95–104. https://doi.org/10.1016/0022-3999(95)00561-7

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory.* Consulting Psychologists Press.

Spooner, R., Bird, J. M., Clemente, R., Izagirre, N. I., Fueyo, E. F., Watling, D., Plans, D., Brewer, R., Bird, G., & Murphy, J. (2024). No differences between remote and laboratory-based testing of cardiac interoceptive accuracy using the phase adjustment task. *Scientific Reports*, *14*(1), 28524. https://doi.org/10.1038/s41598-024-79125-4

Spooner, R., Bird, J. M., Irigoras Izagirre, N., Clemente, R., Fernandez Fueyo, E., Budworth, G., Cocirla, D., Todd, J., Aspell, J., Leganes, M., Watling, D., Plans, D., Brewer, R., & Murphy, J. (2024). Exploring sex differences in cardiac interoceptive accuracy using the phase adjustment task. *Psychophysiology*, *61*(12), e14689. https://doi.org/10.1111/psyp.14689

Stark, E., Stacey, J., Mandy, W., Kringelbach, M. L., & Happé, F. (2021). Autistic Cognition: Charting Routes to Anxiety. *Trends in Cognitive Sciences*, *25*(7), 571–581. https://doi.org/10.1016/j.tics.2021.03.014

Stekhoven, D. J. (2022). *missForest: Nonparametric Missing Value Imputation using Random Forest.* [R package version 1.5].

Steptoe, A., & Vögele, C. (1992). Individual differences in the perception of bodily sensations: The role of trait anxiety and coping style. *Behaviour Research and Therapy*, *30*(6), 597–607. https://doi.org/10.1016/0005-7967(92)90005-2

Stern, Y., Koren, D., Moebus, R., Panishev, G., & Salomon, R. (2020). Assessing the Relationship between Sense of Agency, the Bodily-Self and Stress: Four Virtual-Reality Experiments in Healthy Individuals. *Journal of Clinical Medicine*, *9*(9), 2931. https://doi.org/10.3390/jcm9092931

Stevens, S., Gerlach, A. L., Cludius, B., Silkens, A., Craske, M. G., & Hermann, C. (2011). Heartbeat perception in social anxiety before and during speech anticipation. *Behaviour Research and Therapy*, *49*(2), 138–143. https://doi.org/10.1016/j.brat.2010.11.009

Stewart, S. H., Buffett-Jerrott, S. E., & Kokaram, R. (2001). Heartbeat awareness and heart rate reactivity in anxiety sensitivity: A further investigation. *Journal of Anxiety Disorders*, *15*(6), 535–553. https://doi.org/10.1016/S0887-6185(01)00080-9

Stewart, S. H., Taylor, S., & Baker, J. M. (1997). Gender differences in dimensions of anxiety sensitivity. *Journal of Anxiety Disorders*, *11*(2), 179–200. https://doi.org/10.1016/s0887-6185(97)00005-4

Storch, E. A., Eisenberg, P. S., Roberti, J. W., & Barlas, M. E. (2003). Reliability and Validity of the Social Anxiety Scale for Children—Revised for Hispanic Children. *Hispanic Journal of Behavioral Sciences*, *25*(3), 410–422. https://doi.org/10.1177/0739986303256915

Striepens, N., Scheele, D., Kendrick, K. M., Becker, B., Schäfer, L., Schwalba, K., Reul, J., Maier, W., & Hurlemann, R. (2012). Oxytocin facilitates protective responses to

aversive social stimuli in males. *Proceedings of the National Academy of Sciences*, *109*(44), 18144–18149. https://doi.org/10.1073/pnas.1208852109

Sturges, L. V., Goetsch, V. L., Ridley, J., & Whittal, M. (1998). Anxiety Sensitivity and Response to Hyperventilation Challenge. *Journal of Anxiety Disorders*, *12*(2), 103–115. https://doi.org/10.1016/S0887-6185(98)00004-8

Sugawara, A., Terasawa, Y., Katsunuma, R., & Sekiguchi, A. (2020). Effects of interoceptive training on decision making, anxiety, and somatic symptoms. *BioPsychoSocial Medicine*, *14*(1), 7. https://doi.org/10.1186/s13030-020-00179-7

Suksasilp, C., & Garfinkel, S. N. (2022). Towards a comprehensive assessment of interoception in a multi-dimensional framework. *Biological Psychology*, *168*, 108262. https://doi.org/10.1016/j.biopsycho.2022.108262

Suschinsky, K. D., & Lalumière, M. L. (2014). The Relationship Between Sexual Concordance and Interoception in Anxious and Nonanxious Women. *The Journal of Sexual Medicine*, *11*(4), 942–955. https://doi.org/10.1111/jsm.12250

Tabor, A., Vollaard, N., Keogh, E., & Eccleston, C. (2019). Predicting the consequences of physical activity: An investigation into the relationship between anxiety sensitivity, interoceptive accuracy and action. *PLOS ONE*, *14*(3), e0210853. https://doi.org/10.1371/journal.pone.0210853

Taels, L., Feyaerts, J., Lizon, M., De Smet, M., & Vanheule, S. (2023). 'I felt like my senses were under attack': An interpretative phenomenological analysis of experiences of hypersensitivity in autistic individuals. *Autism*, *27*(8), 2269–2280. https://doi.org/10.1177/13623613231158182

Tan, Y., Wei, D., Zhang, M., Yang, J., Jelinčić, V., & Qiu, J. (2018). The role of mid-insula in the relationship between cardiac interoceptive attention and anxiety: Evidence from

an fMRI study. *Scientific Reports*, *8*, 17280. https://doi.org/10.1038/s41598-018-35635-6

Taylor, G. J., & Bagby, R. M. (2004). New Trends in Alexithymia Research. *Psychotherapy and Psychosomatics*, *73*(2), 68–77. https://doi.org/10.1159/000075537

Taylor, S., Koch, W. J., & Crockett, D. J. (1991). Anxiety sensitivity, trait anxiety, and the anxiety disorders. *Journal of Anxiety Disorders*, *5*(4), 293–311. https://doi.org/10.1016/0887-6185(91)90030-W

Taylor, S., Zvolensky, M. J., Cox, B. J., Deacon, B., Heimberg, R. G., Ledley, D. R., Abramowitz, J. S., Holaway, R. M., Sandin, B., Stewart, S. H., Coles, M., Eng, W., Daly, E. S., Arrindell, W. A., Bouvard, M., & Cardenas, S. J. (2007). Robust dimensions of anxiety sensitivity: Development and initial validation of the Anxiety Sensitivity Index-3. *Psychological Assessment*, *19*(2), 176–188. https://doi.org/10.1037/1040-3590.19.2.176

Teed, A. R., Feinstein, J. S., Puhl, M., Lapidus, R. C., Upshaw, V., Kuplicki, R. T., Bodurka, J., Ajijola, O. A., Kaye, W. H., Thompson, W. K., Paulus, M. P., & Khalsa, S. S. (2022). Association of Generalized Anxiety Disorder With Autonomic Hypersensitivity and Blunted Ventromedial Prefrontal Cortex Activity During Peripheral Adrenergic Stimulation: A Randomized Clinical Trial. *JAMA Psychiatry*, *79*(4), 323. https://doi.org/10.1001/jamapsychiatry.2021.4225

Tiller, J., Pain, M., & Biddle, N. (1987). Anxiety Disorder and Perception of Inspiratory Resistive Loads. *Chest*, *91*(4), 547–551. https://doi.org/10.1378/chest.91.4.547

Todd, J., Plans, D., Lee, M. C., Bird, J. M., Morelli, D., Cunningham, A., Ponzo, S., Murphy, J., Bird, G., & Aspell, J. E. (2024). Heightened interoception in adults with fibromyalgia. *Biological Psychology*, *186*, 108761. https://doi.org/10.1016/j.biopsycho.2024.108761

Trembath, D., Germano, C., Johanson, G., & Dissanayake, C. (2012). The Experience of Anxiety in Young Adults With Autism Spectrum Disorders. *Focus on Autism and Other Developmental Disabilities*, *27*(4), 213–224. https://doi.org/10.1177/1088357612454916

Trevisan, D. A., Mehling, W. E., & McPartland, J. C. (2021). Adaptive and Maladaptive Bodily Awareness: Distinguishing Interoceptive Sensibility and Interoceptive Attention from Anxiety-Induced Somatization in Autism and Alexithymia. *Autism Research*, *14*(2), 240–247. https://doi.org/10.1002/aur.2458

Trevisan, D. A., Parker, T., & McPartland, J. C. (2021). First-Hand Accounts of Interoceptive Difficulties in Autistic Adults. *Journal of Autism and Developmental Disorders*, *51*(10), 3483–3491. https://doi.org/10.1007/s10803-020-04811-x

Vabba, A., Porciello, G., Monti, A., Panasiti, M. S., & Aglioti, S. M. (2023). A longitudinal study of interoception changes in the times of COVID-19: Effects on psychophysiological health and well-being. *Heliyon*, *9*(4), e14951. https://doi.org/10.1016/j.heliyon.2023.e14951

Van Bael, K., Ball, M., Scarfo, J., & Suleyman, E. (2023). Assessment of the mind-body connection: Preliminary psychometric evidence for a new self-report questionnaire. *BMC Psychology*, *11*(1), 309. https://doi.org/10.1186/s40359-023-01302-3

Van Den Bergh, O., Winters, W., Devriese, S., Van Diest, I., Vos, G., & De Peuter, S. (2004). Accuracy of respiratory symptom perception in persons with high and low negative affectivity. *Psychology & Health*, *19*(2), 213–222. https://doi.org/10.1080/08870440410001675627

Van Den Houte, M., Vlemincx, E., Franssen, M., Van Diest, I., Van Oudenhove, L., & Luminet, O. (2021). The respiratory occlusion discrimination task: A new paradigm to

measure respiratory interoceptive accuracy. *Psychophysiology*, *58*(4), e13760. https://doi.org/10.1111/psyp.13760

Van Der Cruijsen, R., Murphy, J., & Bird, G. (2019). Alexithymic traits can explain the association between puberty and symptoms of depression and anxiety in adolescent females. *PLOS ONE*, *14*(1), e0210519. https://doi.org/10.1371/journal.pone.0210519

van Dyck, Z., Vögele, C., Blechert, J., Lutz, A. P. C., Schulz, A., & Herbert, B. M. (2016). The water load test as a measure of gastric interoception: Development of a two-stage protocol and application to a healthy female population. *PLoS ONE*, *11*(9). https://doi.org/10.1371/journal.pone.0163574

van Steensel, F. J. A., & Bögels, S. M. (2015). CBT for anxiety disorders in children with and without autism spectrum disorders. *Journal of Consulting and Clinical Psychology*, *83*(3), 512–523. https://doi.org/10.1037/a0039108

van Steensel, F. J. A., Bögels, S. M., & Perrin, S. (2011). Anxiety disorders in children and adolescents with autistic spectrum disorders: A meta-analysis. *Clinical Child and Family Psychology Review*, *14*(3), 302–317. https://doi.org/10.1007/s10567-011-0097-0

van Steensel, F. J. A., & Heeman, E. J. (2017). Anxiety Levels in Children with Autism Spectrum Disorder: A Meta-Analysis. *Journal of Child and Family Studies*, *26*(7), 1753–1767. https://doi.org/10.1007/s10826-017-0687-7

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

Vig, L., Köteles, F., & Ferentzi, E. (2022). Questionnaires of interoception do not assess the same construct. *PLoS ONE*, *17*(8). https://doi.org/10.1371/journal.pone.0273299

Vlemincx, E., Walentynowicz, M., Zamariola, G., Van Oudenhove, L., & Luminet, O. (2023). A novel self-report scale of interoception: The three-domain interoceptive sensations questionnaire (THISQ). *Psychology & Health*, *38*(9), 1234–1253. https://doi.org/10.1080/08870446.2021.2009479

von Leupoldt, A., & Dahme, B. (2013). The impact of emotions on symptom perception in patients with asthma and healthy controls. *Psychophysiology*, *50*(1), 1–4. https://doi.org/10.1111/j.1469-8986.2012.01480.x

Vormbrock, F., & Neuser, J. (1983). Konstruktion zweier spezifischer Trait-Fragebögen zur Erfassung von Angst in sozialen Situationen (SANB und SVSS). *Diagnostica*, *29*, 165–182.

Warwick, H. M. C., & Salkovskis, P. M. (1989). Cognitive and Behavioural Characteristics of Primary Hypochondriasis. *Scandinavian Journal of Behaviour Therapy*, *18*(2), 85–92. https://doi.org/10.1080/16506078909455848

Watkins, M. (2018). Exploratory Factor Analysis: A Guide to Best Practice. *Journal of Black Psychology*, *44*, 009579841877180. https://doi.org/10.1177/0095798418771807

Watson, D., & Clark, L. A. (1991). The mood and anxiety symptom questionnaire. *(Unpublished Manuscript), Department of Psychology, University of Iowa*.

Watson, D., & Friend, R. (1969). Measurement of social-evaluative anxiety. *Journal of Consulting and Clinical Psychology*, *33*(4), 448–457. https://doi.org/10.1037/h0027806

Watson, D. R., Garfinkel, S. N., Gould Van Praag, C., Willmott, D., Wong, K., Meeten, F., & Critchley, H. D. (2019). Computerized Exposure Therapy for Spider Phobia: Effects of Cardiac Timing and Interoceptive Ability on Subjective and Behavioral Outcomes. *Psychosomatic Medicine*, *81*(1), 90–99. https://doi.org/10.1097/PSY.0000000000000646

Webster, K. E., & Colrain, I. M. (2000). The relationship between respiratory-related evoked potentials and the perception of inspiratory resistive loads. *Psychophysiology*, *37*(6), 831–841.

Wechsler, D. (2012). *Wechsler Adult Intelligence Scale—Fourth Edition* [Dataset]. https://doi.org/10.1037/t15169-000

Wermter, A., Kamp-Becker, I., Hesse, P., Schulte-Körne, G., Strauch, K., & Remschmidt, H. (2010). Evidence for the involvement of genetic variation in the *oxytocin receptor* gene (*OXTR*) in the etiology of autistic disorders on high-functioning level. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *153B*(2), 629–639. https://doi.org/10.1002/ajmg.b.31032

Werner, N. S., Duschek, S., Mattern, M., & Schandry, R. (2009). Interoceptive sensitivity modulates anxiety during public speaking. *Journal of Psychophysiology*, *23*(2), 85–94. https://doi.org/10.1027/0269-8803.23.2.85

Wheaton, M. G., Deacon, B. J., McGrath, P. B., Berman, N. C., & Abramowitz, J. S. (2012). Dimensions of anxiety sensitivity in the anxiety disorders: Evaluation of the ASI-3. *Journal of Anxiety Disorders*, *26*(3), 401–408. https://doi.org/10.1016/j.janxdis.2012.01.002

White, S. W., Albano, A. M., Johnson, C. R., Kasari, C., Ollendick, T., Klin, A., Oswald, D., & Scahill, L. (2010). Development of a Cognitive-Behavioral Intervention Program to Treat Anxiety and Social Deficits in Teens with High-Functioning Autism. *Clinical Child and Family Psychology Review*, *13*(1), 77–90. https://doi.org/10.1007/s10567-009-0062-3

White, S. W., Oswald, D., Ollendick, T., & Scahill, L. (2009). Anxiety in children and adolescents with autism spectrum disorders. *Clinical Psychology Review*, *29*(3), 216–229. https://doi.org/10.1016/j.cpr.2009.01.003

Whitehead, W. E., & Drescher, V. M. (1980). Perception of gastric contractions and self-control of gastric motility. *Psychophysiology*, *17*(6), 552–558. https://doi.org/10.1111/j.1469-8986.1980.tb02296.x

Whitehead, W. E., Drescher, V. M., & Blackwell, B. (1976). Lack of relationship between autonomic perception questionnaire scores and actual sensitivity for perceiving one's heart beat. *Psychophysiology*, *13*, 177.

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

Wibmer, T., Doering, K., Kropf-Sanchen, C., Rüdiger, S., Blanta, I., Stoiber, K. M., Rottbauer, W., & Schumann, C. (2014). Pulse Transit Time and Blood Pressure During Cardiopulmonary Exercise Tests. *Physiological Research*, 287–296. https://doi.org/10.33549/physiolres.932581

Wickman, H., Hester, J., Chang, W., & Bryan, J. (2021). *devtools: Tools to Make Developing R Packages Easier* [Computer software]. https://CRAN.R-project.org/package=devtools

Wiens, S., & Palmer, S. N. (2001). Quadratic trend analysis and heartbeat detection. *Biological Psychology*, *58*(2), 159–175. https://doi.org/10.1016/S0301-0511(01)00110-7

Willem Van der Does, A. J., Antony, M. M., Ehlers, A., & Barsky, A. J. (2000). Heartbeat perception in panic disorder: A reanalysis. *Behaviour Research and Therapy*, *38*(1), 47–62. https://doi.org/10.1016/s0005-7967(98)00184-3

Williams, Z. J., Suzman, E., Bordman, S. L., Markfeld, J. E., Kaiser, S. M., Dunham, K. A., Zoltowski, A. R., Failla, M. D., Cascio, C. J., & Woynaroski, T. G. (2023). Characterizing Interoceptive Differences in Autism: A Systematic Review and Meta-

analysis of Case–control Studies. *Journal of Autism and Developmental Disorders*, *53*(3), 947–962. https://doi.org/10.1007/s10803-022-05656-2

Windmann, S., Schonecke, O. W., Fröhlig, G., & Maldener, G. (1999). Dissociating beliefs about heart rates and actual heart rates in patients with cardiac pacemakers. *Psychophysiology*, *36*(3), 339–342. https://doi.org/10.1017/S0048577299980381

Wiśniewski, P., Jakubczyk, A., Trucco, E. M., Kobyliński, P., Suszek, H., Zaorska, J., Rydzewska, M., & Kopera, M. (2023). Interoception, alexithymia, and anxiety among individuals with alcohol use disorder. *Frontiers in Psychiatry*, *14*, 1229985. https://doi.org/10.3389/fpsyt.2023.1229985

Wittkamp, M. F., Bertsch, K., Vögele, C., & Schulz, A. (2018). A latent state-trait analysis of interoceptive accuracy. *Psychophysiology*, *55*(6), e13055. https://doi.org/10.1111/psyp.13055

Wolff, N., Stroth, S., Kamp-Becker, I., Roepke, S., & Roessner, V. (2022). Autism Spectrum Disorder and IQ – A Complex Interplay. *Frontiers in Psychiatry*, *13*, 856084. https://doi.org/10.3389/fpsyt.2022.856084

Wolters, C., Gerlach, A. L., & Pohl, A. (2022). Interoceptive accuracy and bias in somatic symptom disorder, illness anxiety disorder, and functional syndromes: A systematic review and meta-analysis. *PLOS ONE*, *17*(8), e0271717. https://doi.org/10.1371/journal.pone.0271717

Wood, H., Rusbridge, S., Lei, J., Lomax, C., Elliston, J., & Russell, A. (2022). Exploring the Cognitive Model of Social Anxiety in Autistic Young People—The Central Role of Bodily Symptoms. *Journal of Autism and Developmental Disorders*, *52*(12), 5500–5514. https://doi.org/10.1007/s10803-021-05359-0

Wood, J. J., & Gadow, K. D. (2010). Exploring the nature and function of anxiety in youth with autism spectrum disorders. *Clinical Psychology: Science and Practice*, *17*(4), 281–292. https://doi.org/10.1111/j.1468-2850.2010.01220.x

World Health Organization. (2022). *ICD-11: International classification of diseases (11th revision)*.

World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. (2013). *JAMA*, *310*(20), 2191. https://doi.org/10.1001/jama.2013.281053

Wu, S., Jia, M., Ruan, Y., Liu, J., Guo, Y., Shuang, M., Gong, X., Zhang, Y., Yang, X., & Zhang, D. (2005). Positive Association of the Oxytocin Receptor Gene (OXTR) with Autism in the Chinese Han Population. *Biological Psychiatry*, *58*(1), 74–77. https://doi.org/10.1016/j.biopsych.2005.03.013

Yang, H., Zhang, Y., hu, H., Wang, L., Yan, Y., Lui, S., Wang, Y., & Chan, R. (2023). Relationship Between Interoception and Autistic Traits: A Resting-State Functional Connectivity Study. *Journal of Autism and Developmental Disorders*, 1–11. https://doi.org/10.1007/s10803-023-06050-2

Yang, H.-X., Hu, H.-X., Zhang, Y.-J., Wang, Y., Lui, S. S. Y., & Chan, R. C. K. (2022). A network analysis of interoception, self-awareness, empathy, alexithymia, and autistic traits. *European Archives of Psychiatry and Clinical Neuroscience*, *272*(2), 199–209. https://doi.org/10.1007/s00406-021-01274-8

Yang, H.-X., Zhou, H.-Y., Zheng, H., Wang, Y., Wang, Y.-Y., Lui, S. S. Y., & Chan, R. C. K. (2022). Individuals with Autistic Traits Exhibit Heightened Alexithymia But Intact Interoceptive-Exteroceptive Sensory Integration. *Journal of Autism and Developmental Disorders*, *52*(7), 3142–3152. https://doi.org/10.1007/s10803-021-05199-y

Yang, X., Fang, Y., Chen, H., Zhang, T., Yin, X., Man, J., Yang, L., & Lu, M. (2021). Global, regional and national burden of anxiety disorders from 1990 to 2019: Results from the Global Burden of Disease Study 2019. *Epidemiology and Psychiatric Sciences*, *30*, e36. https://doi.org/10.1017/S2045796021000275

Yates, A. J., Jones, K. E., Marie, G. V., & Hogben, J. H. (1985). Detection of the Heartbeat and Events in the Cardiac Cycle. *Psychophysiology*, *22*(5), 561–567. https://doi.org/10.1111/j.1469-8986.1985.tb01651.x

Yoris, A., Esteves, S., Couto, B., Melloni, M., Kichic, R., Cetkovich, M., Favaloro, R., Moser, J., Manes, F., Ibanez, A., & Sedeño, L. (2015). The roles of interoceptive sensitivity and metacognitive interoception in panic. *Behavioral and Brain Functions*, *11*(1), 14. https://doi.org/10.1186/s12993-015-0058-8

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

Young, H. A., Gaylor, C. M., De Kerckhove, D., Watkins, H., & Benton, D. (2019). Interoceptive accuracy moderates the response to a glucose load: A test of the predictive coding framework. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1898), 20190244. https://doi.org/10.1098/rspb.2019.0244

Young, S. R. (2020). *Valid Cognitive Ability Measures in the Public Domain: A Convergent Validity Study of the ICAR16 Using the WAIS-IV - ProQuest*. https://www.proquest.com/openview/0fe64e7aaafbba408ea394b483486f5f/1?cbl=18750&diss=y&pq-origsite=gscholar

Zamariola, G., Maurage, P., Luminet, O., & Corneille, O. (2018). Interoceptive accuracy scores from the heartbeat counting task are problematic: Evidence from simple bivariate correlations. *Biological Psychology*, *137*, 12–17. https://doi.org/10.1016/j.biopsycho.2018.06.006

Zdankiewicz-Ścigała, E., Ścigała, D., Sikora, J., Kwaterniak, W., & Longobardi, C. (2021). Relationship between interoceptive sensibility and somatoform disorders in adults with autism spectrum traits. The mediating role of alexithymia and emotional dysregulation. *PLOS ONE*, *16*(8), e0255460. https://doi.org/10.1371/journal.pone.0255460

Zeidan, J., Fombonne, E., Scorah, J., Ibrahim, A., Durkin, M. S., Saxena, S., Yusuf, A., Shih, A., & Elsabbagh, M. (2022). Global prevalence of autism: A systematic review update. *Autism Research: Official Journal of the International Society for Autism Research*, *15*(5), 778–790. https://doi.org/10.1002/aur.2696

Zhao, W., Martin, A. D., & Davenport, P. W. (2002). Detection of inspiratory resistive loads in double-lung transplant recipients. *Journal of Applied Physiology (Bethesda, Md.: 1985)*, *93*(5), 1779–1785. https://doi.org/10.1152/japplphysiol.00210.2002

Zhou, H., Gao, Q., Chen, W., & Wei, Q. (2022). Action Understanding Promoted by Interoception in Children: A Developmental Model. *Frontiers in Psychology*, *13*, 724677. https://doi.org/10.3389/fpsyg.2022.724677

Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, *67*(6), 361–370. https://doi.org/10.1111/j.1600-0447.1983.tb09716.x

Zoellner, L. A., & Craske, M. G. (1999). Interoceptive Accuracy and Panic. *Behaviour Research and Therapy*, *37*(12), 1141–1158. https://doi.org/10.1016/S0005-7967(98)00202-2

Appendix

Chapter Two

Appendix 1

Demographics for all participants before only interoceptive participants were selected

Study	Sex	Age in years (average / S.D.)	Presence of mental health condition	Presence of physical health condition	Missing data (N)	Interoceptive at BF30 (N / %)
Study 1 (Todd et al., 2024 Pain free sample) N= 84	Male: 0.0% Female: 100.0% Other: 0.0% *	40.1 (8.3)	13.3%	28.9%	1	19 (22.6%)
Study 2 N= 57	Male: 28.1% Female: 71.9%	27.2 (12.9)	0%	Data not collected	0	14 (24.6%)
Study 3 N= 91	Male: 38.5% Female: 61.5%	23.3 (7.6)	19.8%	13.2%	0	23 (25.3%)
Study 4 N= 10	Male: 40% Female: 60%	23.5 (3.3)	Data not collected	Data not collected	0	3 (30.0%)

Other studies *N= 115*	Male: 46.9% Female: 53.1%	30.9 (9.4)	0.9% (data not collected for 10.4%)	Data not collected	2	27 (23.5%)

Note. This table reflects the demographics of the initial samples, before only those participants who were interoceptive at Bayes Factor 30 were selected. For the demographics of the interoceptive sample, see Table 3 in the main manuscript

* These studies collected information on gender rather than sex

Appendix 2

Full instructions for the PAT

Screen 1: “How well can you match a sound with your heartbeat? Let’s find out!”

Screen 2: “First, find a quiet place where you can sit comfortably upright with your earphones on for around 10 minutes.”

Screen 3: “You will be asked to place your finger on the phone camera (on the back) so that the app can read your heartbeat. Once your finger is in position, you will hear a series of sounds. Each sound actually represents one of your own heartbeats!”

Screen 4: “It might seem like there is a delay between the sounds and the heartbeats you feel. Play the video below to hear an example! [VIDEO 1]”

Screen 5: “In order to rectify the delay, you will be asked to move a dial until the sounds are in sync with your heartbeats. Play the video below to hear an example! [VIDEO 2]”

Screen 6: “If you move the dial to the right, the delay between the heartbeat and the sound will get longer; if you move it to the left, the delay will get shorter.”

Screen 7: “Want to know how this might look like? Press “continue” to watch a short tutorial.”

Screen 8: [VIDEO 3]

Screen 9: “After you have matched the sound with your heartbeat, you will be asked how sure you are about the answer you gave. Press “confirm” then “continue” to start the following trial. In this task, there will be 20 trials in total. [EXAMPLE OF THE CONFIDENCE SCALE]”

Screen 10: “You can feel your heartbeat in different places in your body, such as your chest or yours fingers. You will be asked to indicate where you felt your heartbeat on a body map (like the one below) once every 5 trials. You can choose any of the highlighted body parts or you can select “nowhere” if you haven’t felt your heartbeat in any particular place. [EXAMPLE OF THE BODY MAP]”

Screen 11: “For the duration of this task, please do not actively try to feel your pulse with your hand; we are only interested in what you feel! When you are ready to start, please sit comfortably upright with your earphones on and press “continue”.”

Screen 12: [HEART RATE BASELINE READING]

Screen 13: “You will now get a chance to do two practice trials. Focus on feeling your heartbeat and try to match the sounds to your own heartbeat.”

Chapter Three

Appendix 3

Full instructions for the PST

Experimental task.

1. Make sure you are sitting comfortably upright. This task will take around 10 minutes.
2. The phone camera will measure your pulse.
3. In this task, you will be presented with a series of tones. Your task is to decide which series of tones seems the most in sync with your heartbeat.
4. Each series of tones will be presented with a coloured shape. The delay between your heartbeat and the tone will be different for each shape. When asked, please rank each of the shapes in terms of how synchronous they were with your heartbeat by tapping on them. The shape you tap on first should be the one you perceive as most synchronous with your heartbeat, and the one you tap on last should be the least synchronous.
5. [VIDEO OF EXPERIMENTER DEMONSTRATING HOW TO DO THE TASK]
6. Each tone will be triggered by your heartbeat, meaning that AFTER we detect your heartbeat, the tone will play.
7. If you want to start again, tap “reset choices.” If you are happy with your choices, tap “continue”.

8. Please do not attempt to take your pulse or use any other methods to detect your heart beats. You should try and focus on feeling the sensation of your heart beats.
9. This task may be difficult for some people, and there is no reason to worry if you cannot feel your heartbeat.

10. Sometimes paying attention to your heart rate might make you feel anxious. If you feel yourself begin to panic, please stop the task immediately.
11. Do you understand? If not, please ask the experimenter.
12. First, we will do two practice trials…

Control task.

1. In this task, you will be presented with a series of tones. Your task is to decide which flashing shape is most in sync with the tones that you hear.
2. When asked, please rank each of the shapes in terms of how synchronous they are by clicking on them. The shape you tap on first should be the one you perceive as most synchronous with the tones, and the one you tap on last should be the least synchronous.
3. [VIDEO OF EXPERIMENTER DEMONSTRATING HOW TO DO THE TASK]
4. Each tone will be triggered by the flash of the shape, meaning that you should hear the tone AFTER the flash.
5. If you want to start again, tap “reset choices.” If you are happy with your choices, tap “continue”.
6. Do you understand?

Chapter Four

Appendix 4

The mental health or neurodevelopmental conditions self-reported by participants in Sample One

Mental health/ neurodevelopmental condition	Number of participants
Anxiety (unspecified)	168
Depression	140
Attention Deficit Hyperactivity Disorder	72

Generalised Anxiety Disorder	39
Autism	35
Obsessive Compulsive Disorder	32
Post traumatic stress disorder (complex or not)	30
Bipolar disorder	12
Eating disorder (unspecified)	11
Panic disorder	10
Social Anxiety	10
Dyslexia	4
Emotionally unstable personality disorder	3
Agoraphobia	3
Anorexia nervosa	3
Seasonal Affective Disorder	2
Bulimia nervosa	2
Tic disorder	2
Body dysmorphic disorder	2
Borderline personality disorder	3
Cyclothymia	1
ARFID	1
Psychosis	1
Premenstrual dysphoric disorder	1
Dyscalculia	1
Insomnia	1
Dissociative identity disorder	1
Affective disorder	1
Eating Disorder not otherwise specified	1
Schizoaffective	1
Auditory hallucinations	1
Somatoform disorder	1
Dermatillomania	1
Dysthymia	1

Note. This shows the frequency each mental health or neurodevelopmental condition was reported in Sample One. Participants may have reported more than one mental health or neurodevelopmental condition. This only includes conditions that were not reported to be self-diagnosed.

Appendix 5

The physical health conditions reported by participants in Sample One

Physical health condition	Number of participants
Asthma	35
Irritable bowel syndrome	23
Hypermobility	14
Migraine	12
Underactive thyroid	12

Polycystic ovary syndrome	11
Ehlers Danlos	11
Eczema	10
Anemia	9
Postural orthostatic tachycardia syndrome	9
Diabetes	8
Scoliosis	8
Endometriosis	7
Chronic fatigue syndrome	7
Hypertension	6
Hearing loss	5
Fibromyalgia	5
High blood pressure	5
Long COVID	5
Allergies	5
Epilepsy	4
Hashimoto’s	4
Chronic pain	3
Crohn’s disease	3
Gastroesophageal Reflux Disease	3
Psoriasis	3
Sleep apnea	3
Thalassaemia	3
High cholesterol	3
Coeliac disease	2
Obesity	2
Brain tumour	2
Alopecia	2
Adenomyosis	2
Hiatus hernia	2
Multiple sclerosis	2
Mitral valve prolapse	2
Raynaud’s	2
Pre-diabetes	2
Kidney issues	1
Hormonal imbalance	1
Arthritis	1
Juvenile idiopathic arthritis	1
Human Immunodeficiency Virus	1
Cancer	1
Tachycardia	1
Urticaria	1
Haemochromatosis	1
Low blood pressure	1
Urinary tract infection	1
Cataracts	1
Trigeminal neuralgia	1
Laryngopharyngeal reflux	1
Acoustic neuroma	1

Mast cell activation syndrome	1
Gastroparesis	1
Temporomandibular Joint Disorder	1
Hand amputation	1
Chronic headaches	1
Interstitial cystitis	1
Hidradenitis suppurativa	1
Complex regional pain syndrome	1
Back pain	1
Syringomyelia	1
Chiara malformation	1
Osteoporosis	1
Klinefelter syndrome	1
Essential tremor	1
Sciatica	1
Cystic fibrosis	1
Renal stones	1
Uterine issues	1
Heart palpitations	1
Septum deviation	1
Leg injury	1
Haemorrhoids	1
Heart condition	1
Partially blind	1
Glaucoma	1
Traumatic brain injury	1
Gastritis	1
Heart murmur	1
Shoulder injury	1
Gastrointestinal issue	1
Narcolepsy	1
Knee injury	1
Hay fever	1
Skin condition	1
Systemic Lupus Erythematosus	1
Blood clotting disorder	1
Hand deformation	1
Sjogren’s syndrome	1
Dysautonomia	1
Spinal degradation	1
Slipped disc	1
Post viral fatigue	1
Duane syndrome	1
Anti-phospholipid syndrome	1
Bell’s palsy	1
Hemiplegia	1
Back injury	1
Ménière's disease	1
Constipation	1

Ectopic heartbeat	1
Ichthyosis	1
Tetralogy of Fallot	1
Immune thrombocytopenic purpura	1
Neuropathic pain	1
Drop foot	1
Functional neurological disorder	1
Shoulder injury	1
Sarcoidosis	1
Occipital neuralgia	1
Neurogenic bladder and bowel	1
Extreme respiratory problems	1
G6PD deficiency	1

Note. This shows the frequency each physical health condition was reported in Sample One.

Participants may have reported more than one physical health condition.

Appendix 6

The mental health or neurodevelopmental conditions self-reported by participants in Sample Two

Mental health/ neurodevelopmental condition	Number of participants
Anxiety (unspecified)	90
Depression	79
Attention Deficit Hyperactivity Disorder	36
Autism	19
Post traumatic stress disorder (complex or not)	14
Generalised Anxiety Disorder	9
Obsessive Compulsive Disorder	6
Bipolar disorder	6
Social Anxiety	6
Panic disorder	3
Anorexia nervosa	3
Borderline personality disorder	3
Dysthymia	3
Trauma	2
Binge eating disorder	1
Trichotillomania	1
Postpartum anxiety	1
Depersonalisation derealisation	1
Bulimia nervosa	1
Addiction	1
Body dysmorphic disorder	1
Dyspraxia	1
Health Anxiety	1
Obsessive Compulsive Personality Disorder	1
Eating Disorder not otherwise specified	1

Auditory processing disorder	1
Tourette's	1

Note. This shows the frequency each mental health or neurodevelopmental condition was reported in Sample Two. Participants may have reported more than one mental health or neurodevelopmental condition. This only includes conditions that were not reported to be self-diagnosed.

Appendix 7

The physical health conditions reported by participants in Sample Two

Physical health condition	Number of participants
Asthma	27
Polycystic ovary syndrome	10
Migraine	8
Eczema	7
Diabetes	7
Irritable bowel syndrome	7
Hypermobility	4
Hearing loss	4
Chronic pain	3
Endometriosis	3
Fibromyalgia	3
Scoliosis	3
High blood pressure	3
Long COVID	3
Crohn's disease	3
Ehlers Danlos	3
Kidney issues	2
Underactive thyroid	2
Knee injury	2
Hypogonadism	2
Chronic fatigue syndrome	2
Anemia	2
Lipedema	2
Epilepsy	2
Gastroesophageal Reflux Disease	2
Psoriasis	2
Coeliac disease	2
Obesity	2
Peripheral neuropathy	1
Respiratory condition	1
Recto colitis	1
Sleep apnea	1

Spinal injury	1
Amputation	1
Hashimoto’s	1
Pancreatitis	1
Epicysts	1
Psoriatic arthritis	1
Joint issues	1
Chronic prostatitis	1
Juvenile idiopathic arthritis	1
Heart palpitations	1
Insulin resistance	1
Achalasia	1
Rheumatoid arthritis	1
Uterine malformation	1
Premature ovarian insufficiency	1
Human Immunodeficiency Virus	1
Hiatal hernia	1
Cancer	1
Supraventricular tachycardia	1
Deep vein thrombosis	1
Urticaria	1
Visual impairment	1
Rosacea	1
Haemochromatosis	1
Postural orthostatic tachycardia syndrome	1
Heart disease	1
Migralepsy	1
Allergic rhinitis	1
Juvenile myoclonic epilepsy	1
Dermatitis	1
Ulcerative colitis	1
Secondary acquired immunosuppression	1
Keratoconus	1
Low blood pressure	1
Fracture	1
Polycystic kidney	1
Chronic inflammation response syndrome	1
Urinary tract infection	1
Repetitive strain injury	1
Hypertension	1
Benign prostatic hyperplasia	1

Note. This shows the frequency each physical health condition was reported in Sample Two.

Participants may have reported more than one physical health condition.

Appendix 8

What do you think this questionnaire was trying to measure?	Was this questionnaire easy to understand? If not, what made it difficult to understand?	This questionnaire was intended to measure the degree to which we negatively evaluate signals coming from our body. Do you have any comments that you think would help us improve this questionnaire?
It didn’t let me answer the questions there but I think it was to measure how people feel their internal body actions and how that makes us feel	It didn’t let me answer the questions, but I would’ve gone slightly agree for most	Didn’t let me click the boxes
Interoception and how this relates to anxiety around bodily sensations	I didn't find it difficult to understand	Maybe the option for people to explain why certain bodily sensations make them feel nervous? e.g. someone might know they they are susceptible to panic attacks or be aware of a family history of heart attacks etc etc
Anxiety levels in response to physical symptoms	Yes	I wasn’t sure if I interpreted the pain question correctly
Anxiety / Worrying	Yes	No,
Body pain and emotions	Moderate, keep it short and sweet	It's more generalized
How people experience bodily sensations and the amount of distress that they cause.	Mostly	Maybe include something about the degree to which we are able to understand our bodily sensations/recognise them as what they are and whether this contributes to the negative evaluation?
The questionnaire was trying to measure how in tune we are with the physical responses our body makes to certain situations	It was easy to understand	The questionnaire is somewhat repetitive in it's wordings, so towards the end people might tend to skip reading the whole sentence and not answer correctly
Anxiety/Mental Illness	Yes	No

The degree to which perception of bodily needs induces a negative emotional reaction.	It was easy to understand due to its consistent phrasing and accessible wording.	I think evaluating different forms of negative reactions would be interesting and provide more insight. For example, annoyance could be a potential reaction, possibly dismissal. Otherwise, the formatting is very clear, I just think if possible that the study could dive into the negative evaluation piece more.
What being aware of throughout are body leads to us feeling anxious, seeing if there's a link	It was easy to understand	Nope
People's perception of and feelings on our bodies involuntary processes and sensations	Not terrifically difficult but not completely easy either. Some questions where there's an element of a double negative, such as there was a question along the lines of 'I don't like the sensation of being aware of my heartbeat' where you'd answer disagree if you actually do happen to like it. Also, trying to separate my feelings towards the event or action causing the sensation from my feeling just towards the sensation required a bit of thought/introspection (although this might be me utterly overthinking things).	Perhaps include an example of an answer to a prompt and why answer was picked (if that wouldn't bias people's answers and the sample)? To a much smaller degree, maybe minor rephrasing of the questions and prompts to make them a bit easier to grasp quicker.
the way we think about things we feel inside our body and oru needs and whether it makes us anxious	yes	the question about not liking pain seemed a bit pointless because surely noone likes pain
People’s perceptions and reactions to the feelings they have in their bodies.	Yes, I found it simple to understand and answer.	No.

I think it was trying to measure how we feel things and then how they make us feel/how we process them etc. I understand it but its hard to put into exact words what it was trying to measure.	I think it was easy to understand.	No, I hink it was a good survey, questions were cleary worded and easy to understand.
it was trying to measure people's perceptions of their feelings towards what goes on inside their body, and how they respond to these stimuli	It was very easy to understand. All of the prompts were given in seperate short sentences	
How I perceive internal signals, and as they relate to magnitude and emotional response	Fairly straightforward, no major difficulty in comprehension	No comments
My comfortability in how satisfied my body is	Yes it was easy to understand	I would the questions are self explanationatory and precised
How accurately I could assess internal cues in my body about how I was feeling and the impact of this on my emotions and mood	Yes I think it was generally easy to understand but sometimes it can be confusing to know the difference between a physical sensation and an emotion	To make it a bit clearer language to distinguish between bodily feelings and other feelings such as more emotional/psychological feelings

Chapter Six

Appendix 9

Topic guide

Exploring experiences of anxiety and interoception in autistic adolescents: A reflexive thematic analysis
Ethics reference: R85616/RE001

Section One: *Welcome and introduction!*

- Thank for participating
- Introduce myself and my research
- Tell them I will start recording and take notes
- Show them the information powerpoint again
- Ask them if they have questions
- Reiterate the right to withdraw, take breaks, or skip questions
- Make it clear there are no right answers
- Get verbal consent
- Show them tissues, pens, sensory toys etc.
- Re-establish communication preferences

Section 2: *General info*

I'm going to first ask you some questions about yourself so I can get to know you a little bit, and then I am going to ask a few more questions about your experiences of anxiety.

- Could you please tell me a little about yourself?

Prompts: experience in relation to their autism

- Could you tell me a little bit about your experiences of anxiety?

Prompts: diagnoses, treatment

Tip: use info and interests discussed here throughout to build rapport.

Section 3: *Anxiety*

Now, we're going to move on to some deeper questions about your anxiety. Would you like a break before we start?

- If you were going to describe anxiety to someone that has never felt it before, how would you describe it?

Prompts: hot cross bun model from CBT. Ask about thoughts, emotions, bodily sensations and behaviour.

- How do you think your autism affects your anxiety?

Tip: be aware of any stress in participant.

Section 4: *Alexithymia*

We're going to talk a bit now about emotions other than anxiety. These are things like happiness, sadness, or anger [show picture prompts]. As well as recognising emotions in other people, we can also recognise our own emotions.

- How would you tell if you were feeling very happy? What if you were feeling sad?

Prompt: how easy/hard it is to tell how you're feeling?

Section 5: *Interoception*

As I said, I research our ability to feel signals coming from inside our body. Some examples are how fast our heart is beating, how hungry we are, or how badly we need the toilet. [[Shows pulse on wrist]]. For example, some people can feel their heart beating - even when they're not taking their pulse!

- Are you aware of these type of signals in your everyday life?

Prompts: which are you aware of? Which are you not aware of? Specific signals relating to different areas of the body

- Are there times you feel more or less aware of these signals?

Prompts: what about when you are anxious?

- How do these signals feel to you? Are they neutral, positive, negative?

Prompts: for example, if you feel your heart beating really fast, how does that make you feel?

- Ok, now we're going to focus on a specific example. Think about one signal that you feel is really obvious to you, or easy to talk about. How do you experience this?

Prompt: If they can't think of one, I can suggest something (e.g.: hunger/going to the bathroom/heartbeat)

- How does that impact your everyday life?

Section 6: *Cool down!*

So- we're coming to the end of my questions now. I just wanted to check if there was anything else you would like to add, or anything else you would like to talk about that we haven't yet?

- Tell them they will be sent transcripts when they are written, and they have the choice to edit these slightly but don't have to
- They will be sent the initial themes at the end of analysis, and can also provide feedback but don't have to
- Prior to publication, they will be sent a powerpoint with our findings
- Ask them to refer any friends they think would be interested in participating
- Explain how they will receive their voucher

Thank you so much for participating! This interview will really help us in our research.

Appendix 10

Screening questionnaire

Start of Block: Introduction

Q1 Thank you so much for agreeing to take part. The next section involves questions to find out a bit more about you and to get your contact information.

We want to make sure our sample is as mixed as possible, so we can hear the experiences of people from different backgrounds (e.g., of different ages). This means that we are not able to interview everyone as the amount of data would be unmanageable. So, we will select people on the basis of the following information to get a good mix of young people in the study.

Remember, if you are not selected this time round, you can leave your contact details and we will get in touch with you about future studies. You can also opt in to a prize draw to win one of two £25 vouchers.

End of Block: Introduction

Start of Block: For the adolescent

Q2.1 These questions concern the **young person** who wishes to participate. You may answer these by yourself, or with your parent/guardian

Q2.2 What is your age in years and months?

Q2.3 Years
▼ 14 (1) ... 17

Q2.4 Months

▼ 1 (1) ... 12

Q2.5 What is your gender?

- Male
- Female
- Non-binary or other gender (please write how you identify below)
- Prefer not to say
- _____

Page Break

Q2.7 What is your ethnic group? You may select multiple answers if you are mixed/multiple ethnic groups.

- White
- Asian or Asian British
- Black/African/Caribbean/Black British
- Other ethnic group (please enter) _____
- Prefer not to state (5)

Q2.8 What type of anxiety do you feel? You can select more than one option

- Social anxiety (fear of social situations)
- Specific phobia (extreme fear of a certain thing)
- Panic attacks
- Separation anxiety (fear about being away from home, or from a caregiver)
- General anxiety (persistent worry about everyday situations)
- Selective mutism (not being able to talk in certain situations)
- Agoraphobia (fear and avoidance of situations where you could feel trapped or panicked)
- Other (please explain) _____

End of Block: For the adolescent

Start of Block: For parent/carer 1

Q3.1 These questions concern **parent/caregiver 1**.

Q3.2 Who is this about?

- Mum
- Dad

- Guardian
- Other (please specify) _____

Q3.3 Do you work?

- No
- Yes

Display This Question:
If Do you work? = Yes

Q3.4 What is your job?

_____

Q3.5 How old was your child when they received an autism diagnosis?

_____

Q3.6 If you can remember, please can you give us details of where your child was diagnosed with autism?

_____

Q3.7 Does your child have any of the following diagnosed anxiety disorders?

- Social anxiety disorder
- Specific phobia
- Panic disorder
- Separation anxiety disorder
- General anxiety disorder
- Selective mutism (not being able to talk in certain situations)
- Agoraphobia
- Other (please explain) _____

Q3.8 Is there a second parent/caregiver?

- No
- Yes

End of Block: For parent/carer 1

Start of Block: For parent/carer 2

Q4.1 These questions concern **parent/caregiver 2**.

Q4.2 Who is this about?

- ○ Mum
- ○ Dad
- ○ Guardian
- ○ Other (please specify) _____

Q4.3 Do you work?

- ○ No
- ○ Yes

Display This Question:
If Do you work? = Yes

Q4.4 What is your job?

_____

End of Block: For parent/carer 2