

Understanding the Body from the Inside: The Role of Interoceptive
Belief Updating in Anorexia Nervosa

Alkistis Saramandi

Thesis Submitted for the Degree of
DOCTOR OF PHILOSOPHY (PhD)
UNIVERSITY COLLEGE LONDON (UCL)
February 2024

I, Alkistis Saramandi, hereby confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Anorexia Nervosa (AN; an eating disorder) has been associated with disruptions in the way patients sense, interpret, and integrate internal physiological signals (i.e., interoception). Moreover, evidence has pointed towards a compromised ability of AN patients to evaluate their own cognition and reflect on related beliefs and abilities (metacognition). Yet, the way in which metacognitive beliefs about interoception are formed and updated both retrospectively (e.g., retrospective judgements of how well one perceived their heartbeat) and prospectively (e.g., how beliefs are used in combination with new information to predict future performance) in AN has not been examined. Moreover, despite the gastrointestinal system being directly linked to the AN symptomatology, little work has been carried out in this interoceptive modality. The two-fold aim of the present thesis was to study both retrospective and prospective metacognition in two interoceptive modalities, namely cardiac and gastrointestinal interoception.

To this end, two interoceptive belief updating tasks were developed to study interoceptive belief updating in a group of women with and weight-restored from AN (AN, and AN-WR, respectively) and matched healthy controls (HCs). The experiments in the cardiac modality examined how participants formed and updated their beliefs regarding their ability to feel their own heartbeat. The experiment in the gastrointestinal modality assessed how the groups differed in predicting stomach fullness states before and after drinking certain fixed amounts of water. The results showed that interoceptive accuracy may be domain-specific (participants did not differ in their accuracy to feel their heartbeats, but the estimates of stomach fullness in AN were greater than those in the HCs). Domain-general deficits were observed in the metacognitive level, wherein AN and AN-WR participants maintained their pessimistic priors when updating their beliefs about cardiac perception abilities and when forming predictions about future stomach fullness states. This poor belief updating was partly explained by the greater reliance on pessimistic prior beliefs or the confidence regarding the accuracy of prior beliefs in light of new, interoceptive evidence.

The results have wider implications for our understanding of the role of the different levels and modalities of interoception in AN, and the findings on the precision-weighting mechanism in AN suggest it should be therapeutically targeted in order to help patients improve their metacognitive beliefs about interoceptive abilities.

Impact Statement

The findings arising from the present thesis offer valuable contributions to academic, clinical, and public health areas. Anorexia Nervosa (AN) has the highest mortality and morbidity rates among psychiatric disorders as well as significantly high relapse and treatment withdrawal rates. This work's interdisciplinary and novel approach to study interoception and metacognition allowed to show that even if the 'perceptions' about bodily states people with AN may receive from the body are 'accurate', the way they integrate them with beliefs and form future predictions may be maladaptive and this could be explored in relation to comorbid traits, and precision-weighting parameters. This work combined methods from independent domains of psychology, computational and cognitive neuroscience, and (computational) psychiatry and the preliminary scientific gains deriving from this thesis have been presented at various national and international conferences (e.g., British Affective and Cognitive Neuroscience Conference in Birmingham, 2022; 10th Mind-Body-Brain Symposium in Germany, 2023), and two of the empirical chapters have been accepted for publication in *Computational Psychiatry*, a journal which publishes original research articles which use theoretical, computational, and statistical approaches to improve the understanding of psychiatric dysfunction and mental function.

Beyond academia, the present findings have indirect clinical and public health implications. Existing treatment approaches for interventions have been put in place to target the negative and pessimistic cognitions AN patients may have, but here it was highlighted that more effort should be placed in enhancing the 'confidence' individuals may have regarding beliefs and targeting the mechanisms which underlie the development of biased beliefs in order to help patients update their beliefs about their body in light of new information they may be receiving throughout the treatment changes. Targeting the beliefs patients may have regarding their perceived fullness states over time may help improve the restrictive eating behaviours observed in this population. Additionally, although the present thesis studied interoceptive self-efficacy beliefs only, the findings could be used to advance our understanding of why AN patients hold pessimistic beliefs about their everyday activities which could in turn influence their treatment adherence and overall prognosis. Moreover, recruiting weight-restored individuals highlighted the need to consider further the premorbid behaviours and symptoms but also those which may result from a prolonged acute AN phase, often referred to as 'scar' effects.

Overall, results from this thesis, combined with further developments in research could improve the understanding of eating disorders and the development of therapy interventions.

In loving memory of my godmom Eleni. Novά, σ'αγαπώ.

Acknowledgements

My deepest gratitude and appreciation go to my primary supervisor, Prof Katerina (Aikaterini) Fotopoulou. Katerina's unwavering enthusiasm, expertise, love for research, her approach to solving problems and creating opportunities, and continuous encouragement have been inspiring and have allowed me to discover my interests in greater depth, hone my critical thinking and grow as a researcher. I will be forever grateful to the opportunities Katerina has offered me and I very much look forward to many more years of collaborations with and mentoring from Katerina.

I would also like to express my sincere thanks to the Katlab Post-docs. First, to my secondary supervisor, Dr Thanos (Athanasios) Koukoutsakis (also known as Boss²) for always being there to patiently answer my many R-related questions! I am very grateful for the support from Dr Michal Tanzer, as well as the guidance and help I received from Dr Paul Jenkinson, who while on the other side of the world has acted like my honorary supervisor and would jump into late evening calls from Australia to give me feedback on drafts.

This PhD would not have been possible without the invaluable collaborations with the clinicians, researchers and interns who have assisted with the projects and data collection. A special thank you goes to Liza Mavromara, who not only helped with recruitment and data collection but was a pillar of strength throughout. I could not have asked for a more supportive and encouraging friend than her. Liza, we are now at...0 sleeps to go...! It is over – this thesis has now been submitted! Thank you for being by my side since the very first day. Thank you to Dr Laura Crucianelli who has shared her wisdom and expertise across the many stages of this project, together with generous amounts of laughter, support, and inspiration! Thank you to my current and former lab-buddies, for being there to share frustrations, but to also share many laughs and jokes to ease up the stress which would sometimes come our way – these are none other than Marina Bobou (also my OG conference buddy!), Paulo Silva, Dr Peter Sheng Yao Hsu, Jordan Hewitt, Carolina Villegas Martinez, Oliver Singleton, and Ilia Galouzidi.

I am beyond grateful to all the participants who kindly and generously offered their time to participate in this thesis' studies as well as in the Public and Patient Involvement events – it is safe to say that without them none of this would have been made possible. I truly wish that the

insights from this thesis, together with further research breakthroughs and findings will contribute to our understanding and management of Anorexia Nervosa and help those affected.

Above ground, I am indebted to my family and friends whose love, care, and understanding encouraged me throughout this journey. First, thank you to my mom, my dad, and my sister. Thank you for giving me the courage to dream and to accomplish challenging goals, but also for reminding me of the importance of remaining patient, humble, and listening to my gut feeling (quite fitting with the research topic I chose, no?!). I am also very grateful to my friends, for their invaluable support and enthusiasm throughout. I feel beyond privileged to have found my village of friends who have become my chosen family. Thank you for choosing to become part of this journey and deciding to stick throughout it all.

Finally, I would like to thank Prof Sarah Garfinkel and Dr Jane Aspell in advance, for taking the time to read this thesis and be my viva examiners.

UCL Research Paper Declaration Form

Aspects of Chapters 2 and 3 have been preprinted and were submitted for publication prior to thesis submission and published shortly after the candidate's viva. Both preprints were edited by the candidate prior to inclusion in the present thesis to include additional information where appropriate, as well as to fit the overall style of the thesis.

Chapter 2:

1. Title of preprinted manuscript: Belief Updating About Interoception and Body Size Estimation in Anorexia Nervosa
2. Preprint server: PsyArXiv
3. Link to Preprint: <https://osf.io/preprints/psyarxiv/rntsf>
4. A summary of the methods and results presented in this preprint have been submitted (and accepted) for publication in the Journal of Computational Psychiatry together with the experiment presented in Chapter 3.
5. Author Contributions for the Preprint: LC, BD, PMJ and AF conceived and designed the experiment. AS, LC, BD, DG and VN performed data collection. OG, BD and DG supported the clinical recruitment of individuals with and weight-restored from Anorexia Nervosa. AS analysed the data under the supervision of AK, PMJ and AF. AS, LC, AB, BD, PMJ and AF wrote the manuscript. All the authors approved the manuscript before submission.

Chapter 3:

1. Title of preprinted manuscript: Updating Prospective Self-Efficacy Beliefs About Cardiac Interoception in Anorexia Nervosa: An Experimental and Computational Study.
2. Preprint server: PsyArXiv
3. Link to Preprint: <https://osf.io/preprints/psyarxiv/c8mjv>
4. The manuscript was published in the Journal of Computational Psychiatry: <https://doi.org/10.5334/cpsy.109>.
5. Author Contributions for the Preprint and Accepted Manuscript: AS, LC, BD, PMJ and AF conceived and designed the experiment. AS, LC, VN, LM, and GB performed data collection. DG, FG, BD, SB, and OG supported the clinical recruitment of individuals with and weight-restored from Anorexia Nervosa. AS analysed the data under the supervision of AK, and AF. AS, LC, AK, PMJ and AF wrote the manuscript. All the authors provided feedback and approved the manuscript before submission.

Candidate: Alkistis Saramandi

Date: 02 February 2024

Supervisor / Senior Author: Aikaterini Fotopoulou

Date: 02 February 2024

Table of Contents

Abstract.....	3
Impact Statement	4
Acknowledgements	6
UCL Research Paper Declaration Form	8
Table of Contents	9
List of Figures.....	12
List of Tables.....	13
List of Appendices	14
Publications, Preprints and Preregistrations arising from this thesis to date	15
Chapter 1: Introduction	16
1.1 <i>Overview and Aims</i>	16
1.2 <i>The Epidemiology and Aetiology of Anorexia Nervosa</i>	18
1.2.1 How is Anorexia Nervosa Diagnosed?.....	18
1.2.2 Who does Anorexia Nervosa affect?	18
1.2.3 The causes of Anorexia Nervosa	19
1.2.4 Interim Summary	21
1.3 <i>Interoception and its role in Anorexia Nervosa</i>	22
1.3.1 Interoceptive Modalities.....	23
1.3.2 Dimensions of Interoception	25
1.3.3 Interim Summary	29
1.4 <i>Metacognition and its role in Anorexia Nervosa</i>	29
1.4.1 Metacognition from an Experimental Research Standpoint.....	30
1.4.2 Metacognition from a Clinical Research and Practice Standpoint.....	37
1.5 <i>Adopting an Active Interoceptive Inference Approach to Study Belief Updating of Prospective, Metacognitive Self-Efficacy Beliefs Under a Unified Bayesian Learning Framework</i> 42	
1.5.1 The Free Energy Principle.....	42
1.5.2 The Bayesian Brain and Belief Updating.....	43
1.5.3 Predictive Processing Framework	46
1.5.4 Active (Interoceptive) Inference	49
1.5.5 A Unified Approach to Study Bayesian Interoceptive Belief Updating.....	50
1.5.6 Interim Summary	51
1.6 <i>Current Thesis</i>	51
Chapter 2: Perceptual and Metacognitive Evaluations About Cardiac Interoception in Anorexia Nervosa.....	56
2.1 <i>Introduction</i>	56
2.2 <i>Methods</i>	60
2.2.1 Participants.....	60
2.2.2 Design	64
2.2.3 Psychometric Measures.....	64

2.2.4	Experimental Measures and Procedure	67
2.2.5	Data Analysis	68
2.3	<i>Results</i>	70
2.4	<i>Discussion</i>	76
2.4.1	Conclusion.....	80
Chapter 3: Updating Prospective Self-Efficacy Beliefs About Cardiac Interoception in Anorexia Nervosa: An Experimental and Computational Study		82
3.1	<i>Introduction</i>	82
3.2	<i>Methods</i>	90
3.2.1	Participants.....	90
3.2.2	Belief Updating Task Design and Measures	94
3.2.3	Primary Hypotheses, Measures and Analyses.....	95
3.2.4	Psychometric Measures.....	100
3.2.5	Experimental Materials and Procedure	103
3.3	<i>Results</i>	105
3.3.1	Primary Hypotheses Results.....	105
3.3.2	Computational Modelling	107
3.3.3	Further control and exploratory analyses	115
3.4	<i>Discussion</i>	115
3.4.1	Conclusion.....	121
Chapter 4: Am I as full as I predicted I would be? An exploration of changes in fullness estimates before, during and after a Water-Sipping Task in women with and weight restored from Anorexia Nervosa.....		122
4.1	<i>Introduction</i>	122
4.2	<i>Methods</i>	132
4.2.1	Participants.....	132
4.2.2	Design and Data Analysis	135
4.2.3	Analyses on Primary Hypotheses.....	138
4.2.4	Secondary and Exploratory Analyses.....	145
4.2.5	Transparency and Openness.....	146
4.2.6	Materials and Measures.....	146
4.2.7	Procedure.....	147
4.3	<i>Results</i>	148
4.3.1	Primary Hypotheses Results.....	148
4.3.2	Secondary Hypotheses and Exploratory Analyses	158
4.4	<i>Discussion</i>	163
4.4.1	Conclusion.....	174
Chapter 5: The Use of Heart Rate Biofeedback Treatment to Assist the Regulation of Embodied Feelings: A Clinical Audit on the Process and Effects of Biofeedback in Clinical Practice.....		175
5.1	<i>Introduction</i>	175
5.2	<i>Methods and Materials</i>	181
5.2.1	Audit Setting and Recruitment	181
5.2.2	Development of the Heart Rate Biofeedback Practice: Feedback from the PPI Event	184
5.2.3	Summary of Process.....	185
5.2.4	Baseline Measures Obtained for Clinical Audit Purposes.....	185
5.2.5	REDU Local Practice of HRBF: Process	186

5.2.6 Part 3: Feedback on the process and the service improvements.....	189
5.2.7 Summary of Measures Used to Achieve the Audit's Aims.....	190
5.2.8 Data Management and Analysis	190
5.3 Findings and Observations.....	191
5.3.1 Assessing the Feasibility of the Process	191
5.3.2 Assessing the Service Development.....	191
5.3.3 Qualitative Feedback.....	198
5.4 Discussion.....	200
5.4.1 Primary Aim 1: Feasibility of the Heart Rate Biofeedback Process	201
5.4.2 Primary Aim 2: Assessing the Service Improvements	201
5.4.3 Action Points and Recommendations.....	202
5.4.4 Limitations	203
5.4.5 Conclusion.....	204
Chapter 6: General Discussion	206
6.1 <i>Introduction</i>	206
6.1.1 Summary of the Thesis' Main Aims.....	207
6.2 <i>Summary of Main Findings</i>	208
6.3 <i>Perceptual and Metacognitive Aspects of Interoception and their Domain-Specificity</i>	209
6.3.1 Interoceptive Accuracy.....	209
6.3.2 Interoceptive Metacognition	213
6.3.3 Updating Prospective Beliefs, Learning Rates, and the Role of Precision	217
6.3.4 The role of clinical characteristics in the formation and updating of metacognitive interoceptive beliefs	225
6.4 <i>Methodological issues</i>	228
6.4.1 Design Issues.....	228
6.4.2 Other Challenges and Limitations.....	229
6.5 <i>Clinical Impact</i>	234
6.6 <i>Directions for future research</i>	236
6.7 <i>Concluding Remarks</i>	237
References.....	239
Appendix A	314
Appendix B	322
Appendix C	346

List of Figures

Figure 2.1. Prospective Performance Estimates of the AN, AN-WR and HC Groups.....	71
Figure 2.2. Performance (IAcc) scores of the AN, AN-WR, and HC Groups.	72
Figure 2.3. Retrospective Performance Estimates between the AN, AN-WR and HC Groups.	73
Figure 2.4. Post Feedback Retrospective Performance Estimates Following False Feedback	75
Figure 3.1. Trial Timeline for the Heartbeat Counting Task (HCT).....	94
Figure 3.2. Average Performance (IAcc) Scores at the heartbeat counting task per group.....	105
Figure 3.3. Average of Posterior Prospective Self-Efficacy Beliefs per group.....	107
Figure 4.1. Trial Timeline of the Experiment and Experimental Measures.	137
Figure 4.2. Schematic illustration of the Mediation Analysis steps (Baron & Kenny, 1986).	141
Figure 4.3. Between-group differences on Prior Prospective Fullness Beliefs separated by Condition (Low and High Quantity, i.e., 2%, 5% and 12%, 15%, respectively)	149
Figure 4.4. Between-group differences on Experienced Retrospective Fullness Beliefs separated by Condition (Low and High Quantity, i.e., 2%, 5%, and 12%, 15%, respectively).....	150
Figure 4.5. Between-group differences on Posterior Prospective Fullness Beliefs separated by Condition (Low and High Quantity, i.e., 2%, 5%, and 12%, 15%, respectively).....	152
Figure 4.6. Between-group differences on Prediction Error scores separated by Condition (Low and High Quantity, i.e., 2%, 5%, and 12%, 15%, respectively).....	154
Figure 4.7. Group Differences in Updating Rates.....	156
Figure 4.8. Between-group differences in Baseline Fullness Ratings, separated by mode of rating (Perceived, i.e., 'How full are you', vs Emotional, i.e., 'How full do you feel').....	161
Figure 5.1. Interface of the Heart Rate Regulation Breathing Application and the Visual Aid for Guided Breathing.	187
Figure 5.2. Heart rate changes per patient before and after the task.....	193
Figure 5.3. Interoceptive Accuracy before and after the heart rate biofeedback.....	195
Figure 5.4. Patients' self-efficacy beliefs before and after the tasks.....	196
Figure 5.5. Beliefs about self-relaxation vs. assisted relaxation abilities.	198
Figure 5.6. Positive Affect and Negative Affect Scores.....	200

List of Tables

Table 2.1. Summary of Participant Demographics and Clinical Profiles.....	62
Table 2.2. Items of the Insight into Illness Questionnaire.....	66
Table 2.3. Hypothesis 1 and 2 Results	72
Table 2.4. Hypothesis 3 and 4 Results	74
Table 2.5. Hypothesis 5 Results	75
Table 3.1. Participant Demographics and Characteristics.....	92
Table 3.2 Main Target Models for Computational Modelling Analyses	97
Table 3.3. Insight into Illness Questionnaire.....	102
Table 3.4. Results for Analysis on Performance Group Differences.....	106
Table 3.5. Analysis for Group Differences on Posterior Prospective Self-Efficacy Beliefs.	107
Table 3.6. Main output of Bayesian Belief Updating Analyses.	109
Table 3.7. Analyses to Explore the Role of Depression and Stress on Learning Rates*	113
Table 4.1. Participant Characteristics and Clinical Profiles.	133
Table 4.2. Prior Prospective Fullness Belief Analyses.....	149
Table 4.3. Experienced Retrospective Fullness Estimate Analyses	151
Table 4.4. Group Differences on Posterior Prospective Fullness Beliefs.	153
Table 4.5. Group Differences in Prediction Error Scores.	155
Table 4.6. Group Differences in Updating Rate.....	158
Table 4.7. Group Difference on Confidence Estimates.....	159
Table 4.8. Effects of Illness Duration on Fullness.	160
Table 4.9. Effect of Fullness Estimate Mode - Between and Within Groups.....	162
Table 5.1. Clinical Profiles of the REDU in-patient users.	182
Table 5.2. Physiological Data of the Patients.....	193

List of Appendices

Table A.1 Table for Control Analyses for Hypotheses 1 and 2 (Chapter 2).....	313
Table A.2. Analyses with Psychometric Traits as Interacting Predictors (Chapter 2).....	315
Table A.3. Effects of Illness on Prospective Performance Estimates and Performance (Chapter 2). .	317
Table A.4. Effects of Insight into Illness on Prospective Performance Estimates and Performance (Chapter 2).	319
Table B.1. Additional and Control Analyses for Performance (Chapter 3).....	325
Figure B.1. Effect of Depression on Posterior Prospective Self-Efficacy Beliefs (Chapter 3).....	328
Figure B.2. Effect of Stress on Posterior Prospective Self-Efficacy Beliefs (Chapter 3).	329
Table B.2. Additional Analyses on Posterior Prospective Self-Efficacy Beliefs (Chapter 3).	329
Figure B.3. Average of Prior Prospective Self-Efficacy Beliefs per group (Chapter 3).	332
Table B.3. Group Differences on Prior Prospective Self-Efficacy Beliefs (Chapter 3).	332
Figure B.4. Histogram of Performance Error Percentage Scores before being logarithmically transformed (Chapter 3).	333
Figure B.5. Histogram of the Logarithmically transformed Performance Error Percentage Scores (Chapter 3).	333
Figure B.6. Histogram of the difference between Posterior Retrospective Self-Efficacy Beliefs and Performance (Chapter 3).	334
Table B.4. Results from the Secondary, Behavioural Analyses (Chapter 3).	334
Table B.5. ITPE z-score Differences Between Groups (Chapter 3).....	336
Table B.6. Group Effect on Confidence Estimates (Chapter 3).	337
Table B.7. MAE and BIC Values for the Baseline Models (Chapter 3).....	338
Table B.8. Additional Linear Regressions for the Belief Updating Predictions (Chapter 3).	339
Table B.9. Control analyses for Learning Rate using WCST scores (Chapter 3).	340
Table B.10. Analysis on Actual Learning Rates (Chapter 3).	341
Table B.11. Results from Exploratory Analysis with Counterfactual Beliefs (Chapter 3).....	342
Table B.12. Correlations between Self-Efficacy Beliefs and EDI-3 Items (Chapter 3).....	343
Table B.13. Correlations with Insight into Illness Error Scores (Chapter 3).	344
Table C.1. Controlling for Effects in Stomach Fullness (Chapter 4).	345
Table C.2. Control Analyses on Post-Instructions Pre-Sipping Baseline Fullness State Rating (Chapter 4).	347
Table C.3. Mediation Analysis Step 2 (Chapter 4).....	348
Table C.4. Results from Analyses on Confidence Estimates (Chapter 4).	349
Table C.5. Results from Step 2 of the Mediation Analysis (Chapter 4).....	350

Publications, Preprints and Preregistrations arising from this thesis to date

Aspects of Chapters 2 and 3 are based on research carried out in collaboration with Dr Laura Crucianelli, Dr Athanasios Koukoutsakis, Dr Paul M Jenkinson, Liza Mavromara, Veronica Nisticò, Adam Baiza, Dr Diana Goeta, Dr Benedetta Demartini, Dr Orsola Gambini, Giovanni Boido, Dr Fragiskos Gonidakis, Prof Sara Bertelli, and Prof Aikaterini Fotopoulou and are preprinted and published in:

Saramandi, A., Crucianelli, L., Koukoutsakis, A., Nisticò, V., Baiza, A., Goeta, D., ... & Fotopoulou, A. (2022). Belief updating about Interoception and Body Size Estimation in Anorexia Nervosa. (<https://osf.io/preprints/psyarxiv/rntsfs>; PsyArXiv)

Saramandi, A., Crucianelli, L., Koukoutsakis, A., Nisticò, V., Mavromara, L., Goeta, D., ... & Fotopoulou, A. (2023). Updating Prospective Self-Efficacy Beliefs About Cardiac Interoception in Anorexia Nervosa: An Experimental and Computational Study. (<https://osf.io/preprints/psyarxiv/c8mjv>; PsyArXiv)

Saramandi, A., Crucianelli, L., Koukoutsakis, A., Nisticò, V., Mavromara, L., Goeta, D., Boido, G., Gonidakis, F., Demartini, B., Bertelli, S., Gambini, O., Jenkinson, P. M., & Fotopoulou, A. (2024). Updating Prospective Self Efficacy Beliefs About Cardiac Interoception in Anorexia Nervosa: An Experimental and Computational Study. *Computational Psychiatry*, 8(1), 92–118. DOI: <https://doi.org/10.5334/cpsy.109>

The design and analysis of the experiments presented in Chapters 3 and 4 were preregistered in OSF:

Saramandi, A., Crucianelli, L., Koukoutsakis, A., Nisticò, V., Mavromara, L., Goeta, D., ... & Fotopoulou, A. (2022). Interoceptive Belief Updating in Women With and Weight-Restored from Anorexia Nervosa. (<https://osf.io/f4zhe/>)

Chapter 1: Introduction

1.1 Overview and Aims

“I am bad at understanding my body (...) I don’t notice the anxiety build up, it just hits me.”

“I’m in hospital, so I am probably not doing well. But I feel well so I get confused about it [severity of illness].”

“I struggle with being able to give myself enough food. So, here [in the hospital ward], I need to trust the staff because it’s [eating] going to do me good.”

Anorexia Nervosa (AN) is an eating disorder which has the highest mortality and morbidity rates of all psychiatric disorders. As seen in the quotes above (collected during a clinical audit presented in full in Chapter 5), patients with AN express a difficulty in understanding internal bodily sensations (interoception), such as changes in heart rate, and show distorted insight into the severity of their condition. Interoception, broadly characterised as the process of sensing, interpreting, and integrating internal bodily signals (Critchley & Garfinkel, 2017; Garfinkel et al., 2015; Heim et al., 2022; Khalsa, Adolphs, et al., 2018), has been found to be disturbed across different levels and modalities in AN (see more in Section 1.3). For instance, some have shown that patients with AN cannot feel their heartbeats as accurately as matched healthy controls in a heartbeat perception task (e.g., Pollatos et al., 2008). Others have shown that although the accuracy during a heartbeat perception task does not significantly differ between a clinical AN and a healthy population, AN patients have poorer confidence regarding the accuracy of their answers (e.g., Kinnaird et al., 2020). Despite these indications pointing towards differences in aspects of interoception between clinical AN populations and the general population, the role of interoception as an explanatory factor to the development and maintenance of AN symptoms remains poorly understood. Moreover, the ability to ‘evaluate’ interoception and related beliefs about such abilities is closely linked to metacognition, a multifaceted concept of high clinical relevance which refers to one’s capacity and ability to reflect on, evaluate, monitor, and optimise the success of one’s own cognitive processes in various domains, such as the domains of memory and perception (Flavell, 1979; Fleming & Lau, 2014; Hoogervorst et al., 2023; Morales et al., 2018). Although experimental traditions have mainly focussed on studying metacognition using binary memory and visual perception tasks, mostly assessing the updating of isolated, ‘local’ judgements (e.g., ‘Which of the two

squares contains more dots?’) over time, clinical approaches have mainly focussed on explicit beliefs, such as negative beliefs about the controllability and danger of mental events (see more in Section 1.4). Given the initial findings on metacognitive deficits in neuropsychiatric conditions, more work should be carried out to bridge the gaps across the experimental and clinical research traditions. For example, little attention has been placed on how individuals with AN form and metacognitively update beliefs about interoceptive abilities and their role in AN symptomatology, despite evidence from interoception, experimental, and clinical metacognition research traditions pointing towards poor interoception and metacognitive deficits in AN. This was the key aim of the present thesis. Another key aim was to study this belief formation and updating across different modalities of interoception, namely the cardiac and gastrointestinal systems, in order to characterise the domain-specificity or generality of the metacognitive and interoceptive deficits in AN. These novel research aims were addressed using methods from experimental psychology and neuroscience and tools from computational neuroscience and computational psychiatry. Next, insights from a variety of research traditions, including psychiatry, neuroscience, and psychology, were combined to interpret the findings in an interdisciplinary manner.

This first chapter of the thesis will introduce and discuss the background to the above aims and the related hypotheses of the studies. Specifically, Section 1.2 will cover the epidemiology and known causes of AN. Section 1.2 also describes the need to adopt an interdisciplinary approach to understand the mechanisms underlying symptom development and maintenance and explains how these may influence remission and prognosis overall. Then, Sections 1.3 and 1.4 will bring forward the understanding of AN not as a disorder that is merely caused and maintained by food restriction, starvation, and emaciation, but by a dynamic interplay between how one is in tune with embodied feelings (see review of interoception and related theories and findings in Section 1.3) and distortions in the updating of cognitions and beliefs (metacognition; see Section 1.4). Next, Section 1.5 introduces a unified Bayesian framework under which one can study and attempt to characterise the continuity between perceptual (interoceptive) and metacognitive beliefs in AN. The chapter ends with Section 1.6 which provides an overview to the present thesis’ aims and core hypotheses for each empirical chapter separately (Chapters 2 – 5), as well as a brief overview of what will then be covered in the General Discussion (Chapter 6).

1.2 The Epidemiology and Aetiology of Anorexia Nervosa

1.2.1 How is Anorexia Nervosa Diagnosed?

AN is an eating disorder, which according to psychiatric, categorical classifications is characterised by (i) restricted caloric intake relative to requirements, leading to a significant low body weight in the context of one's age, sex, developmental trajectory, and physical health, (ii) an intense fear of gaining weight, becoming fat, or a persistent behaviour that interferes with weight gain, and (iii) a disturbed view of one's body weight or shape, self-worth influenced by body weight or shape, or persistent lack of recognition of seriousness of low bodyweight (Diagnostic and Statistical Manual of Mental Disorders 5 (DSM-5); American Psychiatric Association, 2013). AN can be further subcategorised as restricting type (r-AN) if the individual has not engaged in binge-eating or purging for at least three months during their current episode, or as binge-eating/purging type (bp-AN) if the individual has regularly engaged in such behaviours (American Psychiatric Association, 2013). The DSM-5 characterisation of AN is similar to the characterisation of AN provided in the 11th revision of the International Classification of Diseases (ICD-11), which also acknowledges that the significant low body weight is not due to another health condition or to the unavailability of food; instead, the low body weight is accompanied by behaviours that prevent restoration of normal weight and fears around weight gain, and one's self-evaluation being dictated by one's body weight and shape (World Health Organization, 2022). Importantly, however, while the above DSM-5 classification is frequently used for the conceptualisation and diagnosis of AN, the exact definition of AN and labelling can vary (for review see Broomfield et al., 2017) and the DSM-5 conceptualisation of AN been criticised for its restrictive nature and neglect of including perceptual processes relating to internal bodily processes and body image in the signs and symptoms of AN, despite these being involved in AN (Stinson, 2019). Although such conceptualisation and classification differences can influence the direction of research and more appropriate frameworks for organising and classifying psychiatric disorders should be developed, this discussion falls beyond the scope of the present thesis. For consistency purposes, all patients recruited in the present study had received a diagnosis of AN according to the DSM-5 diagnostic criteria which were the criteria used at the collaborating clinical sites.

1.2.2 Who does Anorexia Nervosa affect?

AN can affect any individual irrespective of sex, gender, age, or cultural background, and the average length of illness can vary between 6.5 to 14 years (Fernández-Aranda et al., 2021;

Zerwas et al., 2013). However, AN primarily affects adolescent girls and young adult women (Hoek & van Hoeken, 2003; Kjelsås et al., 2004; Smink et al., 2012; Zipfel et al., 2015) – a 10:1 female-to-male ratio has been found in clinical populations with AN (American Psychiatric Association, 2013; Watson et al., 2019) and the lifetime prevalence of AN among women can be up to 4% (Smink et al., 2013). Incidence studies of AN on a community level are scarce, but a community study of female Finnish twins (1975-1979 birth cohort) found the incidence of AN to be 490 per 100,000 person years among 15- and 19-year-old girls (Keski-Rahkonen et al., 2007). Even more strikingly, the crude mortality rate (CMR; defined as the number of deaths within the study population over a specified period) of AN is 5.1 deaths per 1,000 person-years (see meta-analysis of 36 studies on AN mortality: Arcelus et al., 2011), making AN the disorder with the highest mortality rate among psychiatric disorders (Harris & Barraclough, 1998; Smink et al., 2012). Deaths due to AN are primarily a direct consequence of starvation-related medical complications (e.g., cardiac complications and severe infections), and about one in five deaths are from suicide (Smink et al., 2012). Unfortunately, remission and recovery rates are low – these rates vary but include 18% to 52% at the 6-year follow up, 60.3% at the 13-year follow-up, and 62.8% at the 22-year follow up (Castellini et al., 2011; Franko et al., 2018; Rigaud et al., 2011; Zerwas et al., 2013). Notwithstanding, it is worth noting that the variety of reported outcomes may not only be due to the diverse study protocols and type of treatment patients receive, but also and most importantly to the lack of a standardised definition of relapse and recovery (for more see: Khalsa et al., 2017; Miskovic-Wheatley et al., 2023). Similarly, the relapse rates of AN can be highly variable – depending on the way relapse is defined (Khalsa et al., 2017) – ranging between 9-52% following treatment, although most studies report relapse rates greater than 25% (e.g., Carter et al., 2004, 2012; Fichter & Quadflieg, 1999; Helverskov et al., 2010; Kordy et al., 2002; Löwe et al., 2001; Strober et al., 1997). Moreover, almost three quarters of AN patients also report a lifetime mood disorder (most commonly major depressive disorder (Fernández-Aranda et al., 2007), but also anxiety disorder (Swinbourne & Touyz, 2007), obsessive-compulsive disorder (Halmi et al., 2003; Salbach-Andrae et al., 2008), among other. Patients can also have physical comorbidities, with high rates of osteoporosis (21% of patients) and osteopenia of the lumbar spine (54%) being observed in patients who have had a chronic AN course (Misra & Klibanski, 2014; Zipfel et al., 2001). These reports also make AN the disorder with the highest morbidity rates among psychiatric disorders.

1.2.3 The causes of Anorexia Nervosa

Initial attempts to understand the underlying mechanisms and causes involved in the development and maintenance of AN, focussed on isolated components and processes (e.g., biological, psychological, or social causes). The strong genetic contribution to AN has been well established, with heritability estimates ranging between 28% to 74% (Raevuori et al., 2014; Strober et al., 2000; Trace et al., 2013), but there is limited evidence on the role of epigenetic mechanisms (Kim et al., 2014). Neurobiological models have also focussed on (i) disturbances in the emotional processing at a non-conscious level, affecting one's thinking, feeling, and self-regulation (Hatch et al., 2010), (ii) a noradrenergic dysregulation, which results to raised anxiety levels, and hypoperfusion of the insula which is likely to be influencing one's body image (Nunn et al., 2012), (iii) a general dysfunction in key limbic modulatory structures, such as the insula and the subcallosal cingulate, both believed to contribute to the dysregulation of emotional processes (Lipsman et al., 2015). Some have also looked at dysfunctions in the dopaminergic, serotonergic, and noradrenergic systems in AN, and more recently also at the role of the oxytocinergic system (Kaye et al., 2013; Maguire et al., 2013).

Moving to the psychological causes, most work has focussed on the cognitive-behavioural and inter-personal causes. For instance, some models have addressed the role of transdiagnostic traits in the development of AN, suggesting that factors such as perfectionism, low self-esteem, mood intolerance, and interpersonal difficulties are key in the development and maintenance of all eating disorders including AN (Fairburn et al., 2003), while cognitive-interpersonal models have underpinned the role of perfectionism, experiential avoidance, beliefs about the value of the illness and other interpersonal difficulties as key maintaining factors of AN (Schmidt & Treasure, 2006). Next, a role of comorbid traits has also been highlighted, with suggestions that poor set-shifting (Tchanturia, Davies, Roberts, et al., 2012) and central coherence (i.e., preference of local (detail-focussed) over global (bigger picture) processing; Lang et al., 2014) together with obsessive-compulsive traits, difficulties in socio-emotional processing which may also cause anxiety and avoidance, as well as interpersonal difficulties may also all perpetuate the illness (Treasure & Schmidt, 2013). Moreover, some have shown that people with AN might struggle with tasks which require the understanding of implicit thoughts (rather than explicit words, i.e., theory of mind; Gillberg et al., 2010; Russell et al., 2009), while some also suggest that people with AN might have a difficulty in recognising facial emotions, interpreting emotional meaning from the voice, body movement, or even from watching a film (Kucharska-Pietura et al., 2004; Oldershaw et al., 2010, 2011; Zucker et al., 2013). Yet, although these traits and vulnerability factors have been identified and studied

extensively, they do not explain how certain mechanisms relating the ability to perceive and interpret bodily signals may be compromised in AN.

Others have also focussed on the developmental and environmental factors which may nurture the development of AN symptoms. For instance, some have looked at the role of feeding and sleeping difficulties in infancy (Jacobi, Hayward, et al., 2004) while others have also tapped on the role of hormonal changes and dysregulations during puberty as underlying risk factors (Herpertz-Dahlmann et al., 2011). Unfortunately, the social environment can also be a risk factor for AN. It is believed that peers, parents and the media are the three socio-cultural mechanisms with the highest environmental influence (Keery et al., 2004). Concomitantly, an increase of AN in low-income and middle-income countries may also be viewed as the influence of a so-called western lifestyle (including nutritional habits and thin ideal internalisation; Smink et al., 2012), further highlighting the role of sociocultural pressures in AN; although the role of the ‘western lifestyle’ as a vulnerability factor for AN needs to be further investigated (see Shekriladze & Tchaturia, 2017 for the role of acculturation to Western culture in eating disorders; and see Pike et al., 2021 for a comparison of risk factors for the development of AN in Japan vs. the US).

Although the above are some of the known biopsychosocial causes and risk factors for AN, they have not been studied in an integrative approach manner which allows to simultaneously combine our understanding of relevant biopsychosocial components and processes, contributing to our understanding of the complex and multifaceted nature of AN (for reviews see Khalsa et al., 2017; Munro et al., 2017). Additionally, these approaches do not encompass the role of other processes, such as the interpretation and integration of internal bodily signals, in the development of mental health symptoms, including those observed in AN.

1.2.4 Interim Summary

To sum up, this section introduced AN, the diagnostic criteria, epidemiological background of and the risk factors for AN. AN is a severe, multifaceted disorder which mainly affects adolescent girls and young adult women, and although most work has been carried out on the biopsychosocial causes of AN, there is a need to further understand the role of other mechanisms relying at the integration and interpretation of perceptual signals of bodily states

(for example, at the integration of bottom-up afferent signals and top-down predictions of bodily states) in symptom development and maintenance.

1.3 Interoception and its role in Anorexia Nervosa

The concept of interoception was firstly introduced as a way to describe the processes that were happening entoperipherally, i.e., on the internal surface of the body, as opposed to processes that were happening exteroceptively, or with the body's external surface being in direct contact with the environment (Ceunen et al., 2016; Munro et al., 2017; for the first known usage of the concept of interoception see Sherrington, 1906). Ever since, several definitions of the concept of interoception have been introduced, and though there are several variations regarding what entails interoception, most scholars generally agree that interoception is concerned with one's perception of the body's state (for reviews see: Ceunen et al., 2016; Khalsa, Adolphs, et al., 2018). In this thesis, I will use the definition used by Khalsa and colleagues during the first Interoception Summit in 2016 (Khalsa, Adolphs, et al., 2018). Specifically, the definition used by Khalsa, Adolphs and colleagues (2018) states that "interoception refers to the process by which the nervous system senses, interprets, and integrates signals originating from within the body, providing a moment-by-moment mapping of the body's internal landscape across conscious and unconscious levels". Additionally, interoception is a multifaceted process, and "the act of sensing, interpreting, and integrating information about the state of the inner body systems can be related to different elements such as interoceptive attention, detection, discrimination, accuracy, insight, sensibility, and self-report" (Khalsa, Adolphs et al., 2018). These (interoceptive), physiological body signals arise from major visceral systems such as the gastrointestinal, respiratory, and cardiovascular systems (Khalsa, Adolphs, et al., 2018). In some definitions interoception is also used to describe certain cutaneous sensations such as affective touch, pain and temperature (Craig, 2003; Crucianelli & Ehrsson, 2023; von Mohr & Fotopoulou, 2018). Interoception has been related to a vast range of human bodily functions, processes, symptoms, abilities such as pain, awareness and regulation of cognitive and emotional states at a behavioural (Füstös et al., 2013; Schandry, 1981) and neural level (Craig, 2003; Pollatos, Schandry, et al., 2007), medically unexplained symptoms (Bogaerts et al., 2010; Schaefer et al., 2012), anxiety, anxiety disorders, and affective disorders (Barlow et al., 2004; Domschke et al., 2010; Dunn, Stefanovitch, et al., 2010; Paulus & Stein, 2010; Stern, 2014), decision making (Bechara et al., 1997; Clark et al., 2008; Damasio, 1994; Dunn, Galton, et al., 2010; Dunn et al., 2012; Paulus, 2007), subjective (self)-awareness and consciousness (Craig, 2004; Seth et al., 2012), food and water intake (Berthoud, 2006; Brannigan et al., 2015; Herbert

et al., 2012), eating disorders (Crucianelli et al., 2016, 2021; Herbert & Pollatos, 2014; Khalsa et al., 2015; Pollatos et al., 2008), and addiction (Naqvi & Bechara, 2010; Paulus et al., 2009; Verdejo-Garcia et al., 2012), among others. Moreover, disruptions of interoceptive processing and damage, or functional alterations to a related neural network centred around the mid-dorsal insula have been proposed as a ‘transdiagnostic’ mechanism conferring a common vulnerability across multiple mental health disorders (Nord et al., 2021) and particularly those at the interface between physical and mental health, such as eating disorders (Berner et al., 2019; Brewer et al., 2021; Frank et al., 2013; Gordon et al., 2001; Kerr et al., 2016; Khalsa, Adolphs, et al., 2018; Kojima et al., 2005; Nunn et al., 2008, 2011; Uher et al., 2003, 2004; Wagner et al., 2008; Zucker et al., 2017).

Taking a step back, however, it is important to review and understand that interoception is a multimodal, multidimensional concept, and although its domain-specificity is not always clear, it is acknowledged that interoceptive processes are crucial in helping the body maintain its homeostasis (Berntson et al., 1993; Strigo & Craig, 2016; Tsakiris & Critchley, 2016). Although the study of interoception has gained increased popularity, hitherto most interoception research adopts a single-modality and -dimension approach, and thus the understanding of interoception and related deficits across multiple organ systems and across the different (hierarchical) levels of processing remains unclear (see below). For instance studies on interoception might only look at the cardiac modality and not directly compare patterns and mechanisms in the cardiac modality with those which may be seen in the respiratory modality. Moreover, although interoception is multidimensional, approaches to date typically either focus on the accuracy of one’s ability to attend to interoceptive signals in the cardiac modality (e.g., how many heartbeats did one feel correctly) or on self-report questionnaires on one’s awareness regarding such interoceptive feelings and abilities (e.g., how well one thinks they can feel their arm is going to bruise), failing to disentangle the complexity of interoception. Notwithstanding, while initial empirical paradigms had focussed on the study of interoceptive dimensions and modalities in an isolated manner, incoming support not only points towards the influence between modalities and dimensions themselves, but also on the role of each modality and dimension in the development and maintenance of mental health symptoms and related, maladaptive behaviours. Subsequently a better understanding of these interactions and mechanisms is needed.

1.3.1 Interoceptive Modalities

As mentioned above, interoception is a multi-modal concept. This means that the interoceptive processing occurs across all the major organ systems that are involved in maintaining the body's homeostasis, including the cardiovascular (Oppenheimer & Cechetto, 2016; Shivkumar et al., 2016), pulmonary (Von Leupoldt et al., 2013), gastrointestinal (Kaye et al., 2009; Mayer et al., 2006), nociceptive (Simons et al., 2014), thermoregulatory (Crucianelli & Ehrsson, 2023; Fealey, 2013), and immune systems (Capuron & Miller, 2011; Irwin & Cole, 2011), among others (see Khalsa, Adolphs, et al., 2018). Impairments in modalities and/or dimensions of interoception may lead to the manifestation of neuropsychiatric symptoms, including symptoms seen in AN as briefly reviewed below. In this thesis the focus was placed on two major organ systems – one which is frequently used to study interoception and one which is linked to the primary symptomatology of AN, namely the cardiovascular and gastrointestinal systems, respectively.

Various self-report and behavioural studies in AN have shown deficits in non-gastrointestinal systems, such as in the cutaneous, cardiac, and respiratory modalities (e.g., Berner et al., 2018; Crucianelli et al., 2021; Kerr et al., 2016; Lapidus et al., 2020; Pollatos et al., 2008, 2016) which may interact with gastrointestinal-related interoception to maintain eating disorder symptoms. For instance, the anticipation of eating is associated with heightened sensations of heart palpitations, dyspnoea, and anxiety in AN, all of which decrease after meal completion (Khalsa, Hassanpour, et al., 2018) suggesting abnormal multisensory integration of gastrointestinal and cardiorespiratory afferent signalling. In another exploration of interoceptive differences between AN patients and controls, Crucianelli and colleagues showed that both patients with and recovered from AN anticipated and perceived affective touch as less pleasant than matched healthy controls (Crucianelli et al., 2021), offering support for long-standing findings regarding general anhedonia in AN. Previous studies, for example, have shown anhedonia regarding food (Davis & Woodside, 2002), amusing stimuli and sexual activity (Tchanturia, Davies, Harrison, et al., 2012), while others have suggested that the increased cognitive and behavioural control tendencies (including control of energy intake) may be correlated with the reduced reward sensitivity in AN (Kaye et al., 2013; Wagner et al., 2008), all offering support for an aberrant reward processing mechanism in AN which may lead to anhedonia across several domains. Others have shown that patients with AN report increased and prolonged satiety after a meal in comparison to healthy controls (Bluemel et al., 2017; Robinson, 1989), and this could be attributed to reduced interoceptive sensitivity to gastric distention as seen by the failure of patients to detect small volumes of liquid milkshake

delivered directly into their stomach (Coddington & Bruch, 1970). However, the findings remain inconclusive and warrant further investigation given the low power and methodological limitations. Additionally, there are indications that deficits in gastrointestinal interoception in AN may be due to the perception and processing of gastrointestinal signals, rather than actual differences in visceral sensitivity, as noted by the reports of fullness and no hunger even when on an empty stomach (Bluemel et al., 2017) for more on gastrointestinal interoception see Chapter 4). Such findings show the interwoven relationship between interoceptive parameters and AN symptomatology, but also the bidirectional relationship between underlying risk factors in AN and symptom manifestation.

An example that supports the multimodal properties of interoception in the processing of interoceptive afferents is offered by imaging data collected during a breathing load task (i.e., to assess respiratory interoception; Berner et al., 2018). Women in remission from AN and matched healthy controls underwent functional magnetic resonance imaging (fMRI) while wearing a nose clip and breathing through a hose which intermittently restricted their breathing. Participants rated the pleasantness, unpleasantness, and intensity of the experience before and right after the fMRI scan. Although the two groups rated the experience similarly, the clinical group showed altered activation in the insula (a key neural region for interoception) before, during, and after the breathing load (Berner et al., 2018). These findings suggest that there can be a discrepancy between how one experiences an interoceptive state and how the brain may predict and adapt to such an interoceptive state and changes (e.g., anticipating for breathing restriction, meal consumption, and returning to a post-breathing restriction or post-meal consumption state, respectively). Yet, such studies have not drawn direct comparisons between the organ systems that are directly linked with the AN symptomatology and those that are indirectly linked. Specifically, questions regarding the generalisability and domain-specificity of key AN deficits in the cardiac and respiratory modalities in the gastrointestinal modality remain unanswered.

1.3.2 Dimensions of Interoception

The model by Garfinkel and colleagues proposes three main dimensions of interoception, namely (1) interoceptive accuracy, or sensitivity (i.e., performance on objective behavioural tests, such as heartbeat detection tasks), (2) interoceptive sensibility (i.e., self-report assessment of subjective interoception, typically measured through interviews or questionnaires), and (3)

interoceptive awareness, or interoceptive insight (i.e., metacognitive awareness of interoceptive accuracy, typically measured via the confidence-accuracy correspondence; Garfinkel et al., 2015 but for a description of interoceptive dimensions beyond interoceptive accuracy, sensibility and awareness see: Khalsa, Adolphs, et al., 2018; Suksasilp & Garfinkel, 2022). In the present thesis, the focus was predominantly placed on interoceptive accuracy, interoceptive awareness, and self-report and interoceptive beliefs (which encompass both interoceptive sensibility and confidence ratings about interoceptive abilities; see more below). The following subsections provide a more detailed explanation of each dimension and how each dimension is typically assessed, before presenting existing evidence and literature gaps about each respective dimension in mental health and particularly in AN.

1.3.2.1 Interoceptive Accuracy

Attending to one's emotions and related embodied feelings could rely, to an extent, to interoceptive accuracy. Interoceptive accuracy is most often assessed in the cardiac modality; cardiac interoception tasks usually include a heartbeat discrimination task (e.g., Katkin et al., 1982), or the heartbeat tracking task (HCT; Schandry, 1981). In the former, participants receive an external stimulus (e.g., tones or lights), and need to state whether this external stimulus is synchronous or asynchronous with their own heart, while in the latter, participants need to count how many heartbeats they feel within a given time frame without attempting a physical manipulation to facilitate detection. A score is then calculated to assess how accurate one was in feeling or detecting their heartbeats. Both tasks have been frequently employed to study interoceptive accuracy due to their non-invasive nature, their time-efficiency, and the fact that they are easy to understand by both clinical and subclinical populations. Findings to date point towards a positive relationship between interoceptive accuracy and the intensity of feelings, and clinical symptoms manifesting in light of distortions in one's interoceptive accuracy (Herbert et al., 2007; Katkin et al., 2001; Pollatos, Gramann, et al., 2007; Pollatos, Herbert, et al., 2007; Schandry, 1981; Wiens, 2005). For instance, while individuals with heightened cardiac interoceptive accuracy may be better at self-regulating emotions in response to negative affect (Füstös et al., 2013; Kever et al., 2015; Pollatos et al., 2015), individuals with asthma show reduced respiratory interoceptive accuracy (Jacobi, Paul, et al., 2004; Meyer et al., 1990). Poor interoceptive accuracy has also been demonstrated in patients with moderate depression (Dunn et al., 2007). In AN, most initial studies suggested that patients may have poorer cardiac interoceptive accuracy as measured using traditional cardiac interoception tasks (e.g., Pollatos et al., 2008, 2016). However, recent evidence from cardiac interoceptive accuracy assessments

in AN contradicts those earlier accounts, and points towards the role of the other interoceptive dimensions in the anorectic symptomatology (e.g., Kinnaird et al., 2020; Lutz et al., 2019; Richard et al., 2019).

1.3.2.2 Self-report and Interoceptive Beliefs: Interoceptive Sensibility and Confidence Judgements

This dimension of interoception includes two main measures which offer insight into one's self-reported experiences and beliefs about interoceptive sensations and beliefs (also previously referred to as interoceptive sensibility; Garfinkel et al., 2015) and confidence ratings, or judgements which are a type of self-report ratings typically measuring one's belief in a state-like or trial-by-trial manner. Typically, interoceptive sensibility is assessed using questionnaires in which individuals self-report the degree to which they may feel/notice an interoceptive signal (e.g., Body Perception Questionnaire, Multidimensional Assessment of Interoceptive Awareness; Mehling et al., 2012, 2018; Porges, 1993). For instance, individuals need to assess the frequency in which they are aware of bodily states and sensations, such as *"My heart often beats irregularly"*, *"When I breathe, I feel like I cannot get enough oxygen"*, or *"After eating I have digestion problems"*, among others. These self-report questionnaire measures usually target global, interoceptive beliefs which are usually trait constructs and may not necessarily reflect one's beliefs on how well or accurately they think they have performed in an interoceptive task (Suksasilp & Garfinkel, 2022). This latter aspect of interoceptive beliefs is usually measured by assessing one's subjective uncertainty (or, confidence) regarding their interoceptive accuracy during their performance in a specific interoceptive task (Mehling et al., 2012, 2018). For example, during the aforementioned HCT, participants are asked to report how many heartbeats they felt (to measure their interoceptive accuracy) and following that they may be asked *"How confident are you in the accuracy of your answer?"* as a measure of their state-like belief in the moment. This dimension of interoception can highlight the relationship between one's subjective (or, perceived; interoceptive beliefs) and objective (or, actual; interoceptive accuracy) interoceptive ability (Garfinkel et al., 2015; Khalsa et al., 2008, 2015). Recently, network analysis approaches have been used to identify the central features of the AN psychopathology and their connections to previously recognised vulnerabilities (e.g., Brown et al., 2020; Vervaeke et al., 2021). A strong connection has been observed between perfectionistic evaluative concerns and one's need for control over one's self and body, suggesting that a perceived inability to attend to internal bodily sensations, could be contributing to the maladaptive behaviours observed in AN, such as restricted food intake

(Barca & Pezzulo, 2020; Martini et al., 2021). When assessing group differences in subjective ratings of performance in the HCT, AN patients reported lower confidence estimates in comparison to a matched healthy control group, despite no evidence of significant differences in interoceptive accuracy (Kinnaird et al., 2020). Interestingly, lower confidence ratings in the AN group were also correlated with higher eating disorder symptomatology. The authors proposed that this reduced confidence in AN regarding interoceptive abilities could result in reliance on other cues, such as previously existing (or, prior) beliefs individuals may have regarding such abilities (Kinnaird et al., 2020), or it could also reflect the generalised low self-esteem seen in AN (Jacobi, Paul, et al., 2004; Karpowicz et al., 2009; Kästner et al., 2019). In the present thesis, one of the aims was to also delineate the role of interoceptive beliefs further in AN symptomatology, by studying the changes in interoceptive beliefs (both global, trait-like measures of interoceptive abilities and trial-by-trial confidence estimates regarding the accuracy of such abilities) during cardiac and gastrointestinal interoception tasks and study whether the predicted poor interoceptive beliefs in AN will be observed only in the acute AN state, or if they persists beyond remission and during the weight-restoration phase. For these purposes, the term ‘interoceptive self-efficacy’ was conceptualised and coined for this line of work to refer to the (retrospective and prospective) beliefs one has regarding their abilities to attend to interoceptive signals and perform well in respective interoceptive tasks. Interoceptive self-efficacy includes one’s prior prospective beliefs (e.g., ‘*How well will you do?*’), one’s posterior retrospective beliefs (e.g., ‘*How well did you do?*’), and one’s posterior prospective beliefs (e.g., ‘*How well will you do in the future?*’) regarding performance in an interoceptive task, allowing one to study how the brain can combine top-down beliefs with bottom-up afferent sensory signals to predict a bodily state and also update such predictions over time (see more in Section 1.6).

1.3.2.3 Interoceptive Awareness, or Interoceptive Insight

The correspondence between how accurately, or well one performed in an interoceptive task (i.e., interoceptive accuracy) and subjective uncertainty (or, confidence) ratings regarding the performance, offer insight into one’s metacognitive beliefs regarding their performance and interoceptive abilities (Garfinkel et al., 2015). In general, good interoceptive awareness allows an individual to know when they have done well or poorly on something. For instance, one could explain good/poor post-task meta-awareness by using a science test example. If a student who has performed poorly during a science test is asked to rate how well they think they did and report they did very well, this would be suggestive of poor meta-awareness. On the other

hand, if they said they did very poorly, this would be suggestive of good meta-awareness. With regards to interoception, high levels of interoceptive awareness would be suggestive that the individual knows or is aware of their good or bad interoceptive decisions in terms of interoceptive accuracy, whereas low levels of interoceptive awareness would be suggestive of the opposite (Garfinkel et al., 2015). Poor interoceptive awareness has been observed in both neurological and psychiatric disorders – overestimations of performance and poor interoceptive awareness were found in patients with frontotemporal dementia and Alzheimer’s disease (García-Cordero et al., 2016), and poor sleep has been found to impair interoceptive awareness in individuals with depression and mixed diagnoses (e.g., depression and anxiety; Ewing et al., 2017). These examples highlight two potential implications: (1) the potential bidirectional relation between structural and functional neuroanatomical changes and interoceptive awareness, and (2) the clinical relevance of interoceptive awareness, and its role as a transdiagnostic feature in psychiatric symptoms. In AN, patients exhibit poor interoceptive awareness as indexed by their low ratings of confidence following above-average performance in a cardiac interoception task (Kinnaird et al., 2020). However, such empirical paradigms, usually assess interoceptive awareness in isolation [i.e., correspondence between how well participants did (interoceptive accuracy) and how well they think they did (trial-by-trial confidence estimates)] and do not investigate how such beliefs may update over time and also be related to global, interoceptive self-efficacy beliefs about one’s interoceptive abilities (e.g., “*How well will you do in the future?*”). Thus, studying the formation and updating of beliefs regarding interoceptive self-efficacy was a key aim of the present thesis.

1.3.3 Interim Summary

This section introduced the modalities and dimensions of interoception and presented current findings on the role of interoception in mental health and particularly in AN. Hitherto, the extent to which relevant modalities and/or dimensions of interoception contribute to state and/or trait mechanisms of symptom manifestation and maintenance in AN remains unclear. Adopting a multidimensional and cross-modal comparison for the study of interoception and further delineation of such state and trait mechanisms were thus two key aims of the present thesis as outlined above.

1.4 Metacognition and its role in Anorexia Nervosa

The deficits discussed in the interoceptive awareness dimension of interoception are closely linked to the study of metacognitive deficits by both experimental and clinical research

traditions. Metacognition is a multifaceted concept of high clinical relevance which refers to one's capacity and ability to reflect on, evaluate, monitor, and optimise the success of one's own cognitive processes in various domains, such as memory and perception (Flavell, 1979; Fleming & Lau, 2014; Hoogervorst et al., 2023; Morales et al., 2018). An overview of what metacognition is, how it can be measured and studied from an experimental standpoint, and the clinical relevance of metacognition are provided below. The links between AN symptoms and characteristics, namely clinical and cognitive insight and self-efficacy, and metacognition are drawn, with a focus on why metacognitive deficits in the interoceptive domain should be studied to improve the understanding of neuropsychiatric symptomatology and processes and what gaps exist in current literature.

1.4.1 Metacognition from an Experimental Research Standpoint

Metacognition can be studied as a class of mechanisms which allow one to form beliefs about other mental operations (for review see Fleming, 2024). Psychophysical assessments using a staircase approach are typically employed when assessing metacognition. For instance, participants may be shown a square consisting of red and blue dots and asked to indicate which of the two stimulus classes occurred more frequently (the dot stimuli are presented in a non-sequential and randomised order on a 7 x 7 grid, i.e., the position of the red and blue dots is not fixed), followed by a confidence rating (ranging between 1 (*low*) to 4 (*high*)) on the accuracy of their response (e.g., Haddara & Rahnev, 2022), and this is repeated over several trials across multiple blocks (e.g., five blocks with 20 trials each). Confidence (or, subjective uncertainty) ratings can be used to estimate individuals' metacognition and differentiate between two different aspects of metacognition, namely metacognitive bias and metacognitive sensitivity. The former describes an individual's average confidence levels in their task performance, irrespective of how well they actually did (i.e., if the judgements in the binary task were correct or incorrect), overlapping conceptually with the notion of global metacognition (broadly referring to self-performance estimates about overall performance; see more below; Bhome et al., 2022; Fleming & Lau, 2014). The latter – metacognitive sensitivity – offers insight into one's ability to distinguish between correct and incorrect decisions using the trial-by-trial confidence ratings. Metacognitive sensitivity is typically quantified using a combination of task-based measures and Type 2 Signal Detection Theory (SDT; Fleming & Lau, 2014; Galvin et al., 2003; Maniscalco & Lau, 2012), which allows the investigation of how one makes decisions under uncertainty (Peters et al., 2023). However, to control for the first-order

cognitive performance and bias regarding positive or negative confidence rating confounds on metacognitive sensitivity, a third parameter, known as metacognitive efficiency, has also been introduced to allow the exploration of metacognitive sensitivity relative to a specific level of task performance (Maniscalco & Lau, 2012). Although these aspects of metacognition have been studied extensively in healthy populations and in domains of visual perception and memory, literature on metacognition and its relevance in AN remains scarce. To date there exists one pilot study looking at neuropsychological aspects of anorexia and the links with (meta)cognitive performance (Arbel et al., 2013). Patients with AN and matched healthy controls completed a computerised body-size discrimination task and rated the level of confidence regarding how correct their estimation of the body size stimuli (smaller, equal to, or larger than the body to which the stimuli were compared to), and the results showed that poor metacognition mediated poor insight (e.g., severity of illness; see more on insight below) more than cognitive deficits *per se*. These preliminary findings highlight the need to consider the role of metacognitive deficits in the maintenance of (incorrect) beliefs and subsequent decisions. Given that daily decisions are frequently made in the face of uncertainty and without immediate feedback (Bang & Fleming, 2018), understanding how these metacognitive processes and deficits may influence interoception (including beliefs, and awareness of interoceptive states and subsequent decisions) and AN psychopathology is key and of high interest.

1.4.1.1 Domain-Specific Metacognition vs. Domain-General Metacognition

Although there is anatomical (Fleming et al., 2010, 2012; Maniscalco et al., 2017; McCurdy et al., 2013), functional (Baird et al., 2013; Fleck et al., 2006; Hilgenstock et al., 2014; Yokoyama et al., 2010), and neuropsychological (Fleming et al., 2014; Schnyer et al., 2004; Shimamura & Squire, 1986) evidence of neural substrates (such as frontolateral, frontomedial and parietal regions) contributing to metacognition across domains, it remains unclear if metacognition relies on one common, domain-general resource for performance evaluation across different task domains, or if it relies on domain-specific components. Thus, a key to understanding the role of metacognition as a higher-order process, is delineating whether it is a domain-general or domain-specific process (e.g., do metacognitive processes share resources across several tasks and situations or not). For example, if metacognition is domain-general, this could mean that individuals use a common resource when evaluating their performance across different types of tasks (Mazancieux et al., 2020; Rouault, McWilliams, et al., 2018). An attempt to delineate between the domain-generality and domain-specificity of metacognition has been

facilitated by progresses in our understanding of the neural basis of metacognition (for review see Fleming & Dolan, 2012).

A recent cross-domain (episodic memory and visual perception) examination of metacognitive bias (measured by confidence) and efficiency offered support on common resources shared across domains for both metacognitive aspects as seen by the correlation between the tendency to report high confidence in one task extending to the tendency to report high confidence in another task of a different domain (Mazancieux et al., 2020). Similarly, another recent study showed that the updating of metacognitive beliefs across the memory, visual, and general knowledge domains (including nutrition and socioeconomic judgements) was domain-general, as indexed by lower local confidence estimates across domains (Hoogervorst et al., 2023). Interestingly, global, metacognitive beliefs showed high domain-specificity (Hoogervorst et al., 2023). These findings are consistent with indications of both domain-specific and domain-general resources being involved in metacognitive efficiency, as also shown by previous neuroimaging data (Mazancieux et al., 2020; Morales et al., 2018), and models of confidence formation (Fleming & Daw, 2017). Thus, two main suggestions can be formed: (i) usage of global resources may differ between domains (e.g., memory vs perception), but (ii) some components of metacognitive efficiency in retrospective confidence (i.e., when one rates decision using trial-by-trial confidence estimates) may be domain general, while aspects of metacognitive belief updating may be domain-specific (Hoogervorst et al., 2023; Mazancieux et al., 2020). On the other hand, results from neuroimaging studies have pointed towards a domain-specificity process in metacognition. In one study in individuals with frontotemporal cortex (but not precuneus) lesions, impairments in the metacognitive efficiency in the perceptual but not the memory domain were noted (Fleming et al., 2014), while the reverse dissociation was seen in theta burst TMS in the precuneus (a brain stimulation process which leads to strong changes in brain activity; Ye et al., 2018). Recent studies have thus come to suggest that both domain-specific and domain-general processes of metacognition coexist and the extent to which tasks rely on a common, shared circuitry of metacognition may depend on the context or type of judgement/task one may have to metacognitively evaluate (e.g., De Gardelle et al., 2016; Fleming & Daw, 2017; Fu et al., 2022; Lund et al., 2023; Mazancieux et al., 2020; Pouget et al., 2016; West et al., 2023; for review see: Morales et al., 2018; Rouault, McWilliams, et al., 2018).

As seen in more detail below, although the domain-general and/or domain-specificity of metacognition has been frequently studied in the perceptual, decision making, and memory domains and psychopathologies, such research in other modalities of perception, beyond visual perception, such as interoception and comparisons between modalities of interoception (e.g., metacognition about cardiac vs gastrointestinal interoception) and especially in AN remain scarce (but see Nikolova et al., 2022 for a psychophysical task to study respiratory interoception and metacognition trialled in a sample of healthy adults). Yet, the study of metacognition offers an opportunity to experimentally assess one's subjective estimates (e.g., confidence and beliefs) about internal states and insights into the metacognitive mechanisms of neuropsychiatric symptom pathogenesis and maintenance and thus guide assessment and management (Rouault, McWilliams, et al., 2018; Seow et al., 2021; Wells et al., 2012).

1.4.1.2 Local Metacognition vs. Global Metacognition

It is also important to note the differences across the different hierarchical metacognitive levels. The two levels of interest are 'local' and 'global' metacognition (Fleming & Lau, 2014; Garfinkel et al., 2015; Rouault, McWilliams, et al., 2018; Rouault & Fleming, 2020; Seow et al., 2021). Local measures of metacognition can include trial-by-trial (retrospective) confidence ratings and can be regarded as isolated decisions or judgements of performance and can also be characterised by the two aforementioned parameters, namely metacognitive bias and metacognitive sensitivity (Bhome et al., 2022; Fleming, 2024; Rouault & Fleming, 2020). Most research has focussed on the psychological and neural bases of local metacognition and has shown that local confidence ratings can be involved in behaviours and processes such as the regulation of subsequent decisions (Haddara & Rahnev, 2022) and information gathering (Desender et al., 2018). Others have also explored the role of external trial-by-trial feedback on performance and its impact on local metacognition and have shown that such feedback may not affect perceptual or metacognitive sensitivity but can reduce both perceptual and metacognitive bias as measured via local confidence estimates (Haddara & Rahnev, 2022; e.g., for effect of different feedback types on confidence ratings see: Benson & Önköl, 1992; Frömer et al., 2021). On the other hand, global metacognition refers to one's overall sense of confidence about our abilities (sometimes also referred to as global self-performance estimates; Bandura, 1977; Rouault et al., 2019; Stephan et al., 2016) and may be more closely related to daily functioning, extending to the subjective experience of mental health symptoms (Seow et al., 2021). These global metacognitive beliefs can also include explicit beliefs about one's cognition and related (dis)abilities, or its consequences (termed clinical insight; David, 2004a),

and beliefs about one's capacity to succeed in, or cope with different tasks, situations and contexts (or, self-efficacy; Bandura, 1977), and are of high clinical relevance. Global metacognition is thus highly related to future behaviour and can influence one's adaptive behaviour (e.g., knowing when to seek help; Fleming, 2024).

Relevant studies have shown that key brain networks supporting the formation of local confidence (Heereman et al., 2015; Kepecs et al., 2008; Kiani & Shadlen, 2009; McCurdy et al., 2013; Morales et al., 2018) may be implicated together with brain regions implicated in self-reflection (Christoff et al., 2003; Fleming et al., 2010) in the support of a range of metacognitive abilities (Rouault & Fleming, 2020). Further neuroimaging work has also noted brain activity in networks in which local confidence encoding was modulated that global self-performance (e.g., strong local confidence corresponded to lower global self-performance estimate during a visual perception task), whereas shifts in overall global self-performance estimates was identified in other brain networks irrespective of local confidence (Rouault & Fleming, 2020). These findings offer initial support for the coexistence of local and global metacognition in the brain and how performance in one can modulate the other (Rouault et al., 2019), and are aligned with theories of learning suggesting that the brain accumulates evidence (e.g., local confidence) to update (posterior) beliefs over predicted global self-performance estimates (Donoso et al., 2014; Rouault et al., 2019). These findings also tap onto the idea of a hierarchical, and nested nature of local (subjective) confidence, and global beliefs (Allen et al., 2020; Fleming, 2020; Fleming & Daw, 2017). Yet, little is known regarding the formation and updating of retrospective and prospective, metacognitive beliefs across these two levels and their role in psychopathology. Thus, a key aim of the present thesis was to experimentally and computationally study how patients with or weight-restored from AN, in comparison to matched healthy controls use nested, local and global retrospective interoceptive beliefs about performance in traditional interoceptive tasks to update their explicit, (global) prospective beliefs about such abilities. This was done in an attempt to apply the initial experimental insights on the importance of global (mostly retrospective) metacognition in mental health, and to study how they pertain to eating disorders, and interoception research (for a first theoretical proposal and model see Stephan et al., 2016).

1.4.1.3 Metacognition in Mental Health: Empirical Insights

Impairments in metacognition have been identified in a range of neuropsychiatric symptoms and conditions, such as Alzheimer's disease, brain injury, schizophrenia, drug addiction, and

eating disorders (e.g., Arbel et al., 2013; Hoven et al., 2019; Kirsch et al., 2021; Palmer-Cooper et al., 2023; Rouault, Seow, et al., 2018; for reviews see: David et al., 2012; Goldstein et al., 2009; Seow et al., 2021). Although determining the domain generality and/or specificity in clinical metacognition deficits warrants further transdiagnostic and interdisciplinary investigations, there are indications that disruptions in the mapping of local confidence and global self-beliefs can impact mental health and cognitive function, such as mood disorder symptom severity (Benwell et al., 2022; Rouault, Seow, et al., 2018; Seow et al., 2021). Recent evidence has also pointed towards an interplay between chronic underconfidence in subclinical anxiety and depression, stemming from a distortion in interactions between local and global metacognition (Katyál et al., 2023) but what remains elusive is whether the underconfidence resulting from metacognitive deficits is a premorbid risk factor of depression and anxiety traits, or the other way around. The latter proposal is in line with a recent theoretical framework which suggests that depression and fatigue symptoms could be the response to a chronic ‘surprise’ (or, dyshomeostasis) about internal (or, interoceptive) bodily signals which influence the brain’s self-efficacy beliefs, and thus impacting one’s domain-general metacognitive abilities (Stephan et al., 2016).

Notwithstanding, it is also likely that the domain-generality or specificity of metacognitive deficits differs between symptoms and conditions. For instance, while some have shown that domain-general self-beliefs are systematically lowered in individuals with depression and anxiety (Wells et al., 2012), others have suggested that domain-specific metacognitive deficits in perception may be involved in the formation of hallucinations in psychosis (Klein et al., 2013; Moritz et al., 2014) and deficits in the domain of memory may explain symptoms of functional memory loss (Stone et al., 2015), or the lack of insight into memory deficits in patients with Alzheimer’s disease (Chapman et al., 2018). Overall, several lines of work have shown that metacognitive distortions and the influence of domain-specific and domain-general mechanisms of metacognition may be associated with specific personality traits (Buratti et al., 2013) and belief systems (Rollwage et al., 2018), both in the general population and (sub)clinical populations (Gillan & Seow, 2020; Hoven et al., 2019; Rouault, Seow, et al., 2018). For example, individuals with high anxious-depression traits (hereafter referred to as AD dimension) may show low overall confidence (Rouault, Seow, et al., 2018; as also seen in clinical samples, e.g., Wells et al., 2012), while individuals with high compulsive behaviour and intrusive thought traits (hereafter referred to as CIT dimension) show greater confidence in perceptual decisions, but reduced metacognitive insight (Hoven et al., 2023; Rouault, Seow,

et al., 2018; Seow & Gillan, 2020). Individuals scoring higher in the CIT dimension, a dimension which features links to impulsivity, obsessive-compulsive disorder, schizotypy, and eating disorders, showed metacognitive deficits across both perception and general knowledge tasks, suggestive of a domain-general mechanism of metacognition (Benwell et al., 2022). On the other hand, the individuals scoring higher on the AD dimension, a dimension which is mainly linked to features of apathy, anxiety, and depression, showed greater underconfidence across cognitive domains (in tasks of perception and general knowledge), although this underconfidence was not associated with their ability to reliably dissociate correct from incorrect task responses (i.e., hinting towards an absent relationship between metacognitive efficiency and anxious-depressive symptoms; Benwell et al., 2022; Hoven et al., 2023; Moses-Payne et al., 2019). There are two important points to consider. Firstly, these findings across both clinical, subclinical and general population samples reveal that domain-specific and domain-general mechanisms may relate to distinct psychological symptom manifestations and psychopathology. Secondly, there is a need to study these domain-specific and domain-general mechanisms of metacognition across eating disorders, especially in domains which play a significant role in the AN symptomatology, such as interoception. Although CIT traits are often observed in AN, AN is also comorbid with anxiety and depression, and thus there is a need to better understand which domain-specific and domain-general patterns of metacognition observed in these comorbid traits and conditions which are also present in AN.

At the same time, work in neurological conditions such as Alzheimer's and Parkinson's disease has further shown a dissociation between the local and global metacognitive levels in some cognitive domains (Chapman et al., 2018; Cosentino et al., 2011; Souchay et al., 2006). For example, local memory and motor metacognition was found to be relatively intact in the participants with Alzheimer's disease (Chapman et al., 2018; Cosentino et al., 2011), but a specific deficit in the global metacognition in the memory and not the motor domain was noted (Chapman et al., 2018). Differences in local assessments of recognition memory were also observed between participants with Parkinson's disease and healthy controls, yet the two groups did not differ in their global prediction of accuracy (feeling of knowing how accurately they will perform in the recognition memory test; Souchay et al., 2006). All these observations, across both psychological and neurological conditions, could also be suggestive of a symptom-severity influence in metacognitive levels, with different symptoms or levels of severity manifesting differently across the hierarchical levels of metacognition and higher-order processes in general (Seow et al., 2021). Yet, all these proposals and theories remain highly

speculative as research to date has not studied these concepts in unison. Given the high relevance of global self-beliefs and metacognition to daily life functions, understanding their underlying function in global aspects of one's subjective experience, including interoceptive signals, and their relation to shaping abstract self-beliefs, such as self-esteem and self-efficacy is key. Addressing this important gap in experimental traditions that has direct clinical implications was another aim of the present thesis.

1.4.2 Metacognition from a Clinical Research and Practice Standpoint

Approaches towards uncovering associations between metacognition and mental health symptomatology can contribute towards our understanding of mechanistic underpinnings of (low) self-beliefs (e.g., low interoceptive self-efficacy beliefs; see above) and their impact on key psychological constructs and neuropsychiatric symptoms. The subsections below focus on present insights from clinical research on self-beliefs and mental health, zooming in on self-efficacy, and insight into illness, and their relevance to the study of metacognition.

1.4.2.1 Assessing Metacognition and Self-Beliefs in Clinical Settings

Global metacognition and self-beliefs in clinical psychology studies have mainly been studied using questionnaires to sample explicit beliefs. However, unlike experimental traditions, the examination of metacognition and psychopathology in clinical psychology approaches relies on the assessment of metacognitive beliefs under the Self-Regulatory Executive (S-REF) model (Wells & Matthews, 1994, 1996). The S-REF model proposes that the metacognitive beliefs that are linked to the activation and maintenance of coping strategies - including worry, rumination, and thought suppression – can result in the endurance of psychological distress (Wells & Matthews, 1994, 1996). These metacognitive beliefs can be classified into three subtypes, namely (i) generic metacognitive beliefs relating to internal cognitive-affective experiences and their importance, (ii) positive metacognitive beliefs about the usefulness of coping strategies, which are also linked to the activation of these strategies (also known as Cognitive Attentional Strategies; (Wells, 2001, 2013), and (iii) negative metacognitive beliefs about the controllability and danger of mental events (Spada et al., 2015). Most studies on metacognitive beliefs as conceptualised in the S-REF model thus use self-report questionnaires (e.g., Metacognitions Questionnaire (MCQ); Cartwright-Hatton & Wells, 1997; Wells & Cartwright-Hatton, 2004) to assess five dimensions of generic metacognitive beliefs (Cartwright-Hatton & Wells, 1997; Wells & Cartwright-Hatton, 2004). These dimensions are (i) positive beliefs about worry, (ii) negative beliefs about the uncontrollability and danger of

thoughts, (iii) cognitive confidence to assess the level of confidence in one's mnemonic and attentional capabilities, (iv) beliefs about the need to control thoughts, and (v) cognitive self-consciousness, to assess one's monitoring of thinking processes. In turn, relevant studies have consistently found metacognitive belief differences between AN patients and the general population, such as higher levels of negative beliefs about worry (i.e., beliefs that thoughts are uncontrollable and dangerous; Cooper et al., 2007; Davenport et al., 2015; McDermott & Rushford, 2011; Olstad et al., 2015; Palomba et al., 2017) and higher positive beliefs about worry (e.g., beliefs that perseverative thinking is useful; Davenport et al., 2015; McDermott & Rushford, 2011; for review on metacognition and eating disorders see: Palmieri et al., 2021; Sun et al., 2017). Moreover, these metacognitive beliefs, and specifically positive beliefs about worry and negative beliefs about the uncontrollability have been found to predict drive for thinness in AN (Davenport et al., 2015), whilst metacognitive dysfunctions manifesting in ruminations over distorted cognitions pertaining to food, weight and shape hinder the ability of individuals to engage in helpful cognitive processes such as problem solving (Safdari et al., 2013; Tchanturia et al., 2013). These findings further highlighting the role of aberrant metacognitive beliefs in the AN psychopathology.

Interestingly, in an attempt to characterise metacognitive impairments and brain morphometry in AN, and whether these constitute reversible or stable markers of AN (i.e., whether deficits in metacognitive beliefs are a characteristic of the acute AN state resulting from the restrictive energy intake or if they persist beyond recovery), a recent study combined psychometric assessment (using the MCQ; Cartwright-Hatton & Wells, 1997; Wells & Cartwright-Hatton, 2004) and assessments of comorbid AN traits, such as alexithymia) with Voxel-Based Morphometry in a sample of women recovered from AN and matched healthy controls (Oliva et al., 2020). Although previous studies in clinical, acute AN populations have shown differences in the brain volumes and functions, often located in regions involved in metacognitive and cognitive functioning, in comparison to health controls (e.g., Brooks et al., 2011; Friederich et al., 2012; Fuglset et al., 2016; Fujisawa et al., 2015; Gaudio et al., 2016; Seitz et al., 2016; Steward et al., 2018; Titova et al., 2013), no significant differences in the psychometric profiles of the two groups or their brain morphometry were noted (Oliva et al., 2020). Although such findings offer initial contributions to the understanding of functions and brain structures which may be sensitive to recovery from AN, the role of metacognitive processes and their manifestation through clinical symptoms and behaviours during the different stages of the disorder remain largely unexplored.

The metacognitive beliefs under the R-SEF model, and particularly negative beliefs that thoughts are uncontrollable and dangerous can extend to the elevated rumination and negative affect observed in AN patients (e.g., Cowdrey & Park, 2012; Seidel et al., 2016; Smith et al., 2018; Startup et al., 2013) which may in part be triggered by heightened feelings or beliefs of inefficiency (or, low self-efficacy; Bruch, 1978; Jacobi, Paul, et al., 2004; Surgenor et al., 2007). Self-efficacy beliefs, defined as one's beliefs in their capacity to succeed in a certain situation, influencing how one thinks, behaves, and feels (Bandura, 1977), can also be studied under the two-dimensional concept of self-esteem which broadly refers to one's personal judgement of one's own worth (Bardone et al., 2003; Rosenberg, 2015). Individuals with high self-efficacy are more confident in their ability to perform at a certain level or control the outcome of a task (Weerdmeester et al., 2020; Zulkosky, 2009). Contrastingly, poor self-efficacy beliefs are believed to be involved in the development of mental health symptoms (Seow et al., 2021). For instance, in eating disorders, self-efficacy has been found to influence drive for thinness, body dissatisfaction, and treatment outcomes (Keshen et al., 2017; Toray & Cooley, 1997), with some also suggesting that self-efficacy to change eating and weight control behaviours is necessary to sustain behaviour change in and recovery from AN (Pinto et al., 2006, 2008; Schebendach et al., 2008). Yet, although self-efficacy in experimental traditions has been studied with respect to 'global' aspects of metacognition (such as overall prospective judgements about one's capabilities in different contexts), clinical traditions have studied self-efficacy in relation to beliefs about one's ability to regulate eating, other bodily domains and cognition (e.g., Berman, 2006; Goodrick et al., 1999; Linardon et al., 2017; O'Leary, 1985). For example, studies have assessed patients' confidence in what they can eat in a restaurant with friends, or confidence in their ability to succeed if they decide to change a behaviour (e.g., Goddard et al., 2013; Pinto et al., 2008; Steele et al., 2011). As such, one of the aims of this present thesis was to assess both local and global (self-efficacy) beliefs about one's interoceptive ability in different interoceptive modalities.

1.4.2.2 Associations between Metacognition and Insight into Illness

Metacognitive deficits and low self-efficacy beliefs may also be related to the ego-syntonic nature of AN and impaired insight (Gregertsen et al., 2017; Serpell et al., 1999). The term ego-syntonicity is used in this context to describe the fact that patients with AN may regard their restrictive eating as adaptive to their psychological needs and their identity, while disregarding the illness' negative consequences on their body and health (Serpell et al., 1999). Ego-

syntonicity may thus be linked with impaired insight. Insight impairments can be multidimensional, affecting one's appreciation of one's illness and symptom severity, symptom causes, and treatment needs (Beck et al., 2004; David, 2004b; Greenfeld et al., 1991). Insight impairments in AN can vary in severity and specificity, from a general illness unawareness to more specific distortions in how the individual experiences and perceives their body, extending to delusional body image beliefs and overvalued beliefs about one's body weight and size (Hartmann et al., 2013; Konstantakopoulos et al., 2012; Steinglass et al., 2007).

Insight deficits in AN have been found to be a maintaining factor in AN, hindering treatment adherence and recovery, and are especially noted in patients with the restricting type of AN (r-AN) in comparison to patients with the binge-purge type of the disorder (bp-AN) or patients with other eating disorders (Konstantakopoulos et al., 2011, 2012). These insight deficits are irreducible to other clinical, cognitive, or emotional symptoms in AN (Konstantakopoulos et al., 2012), and instead appear to have a specific metacognitive basis (Arbel et al., 2013). Unfortunately, the explanatory mechanisms for the lack of insight have mainly been studied in other psychiatric disorders, and particularly in psychosis and schizophrenia (Bleuler, 1950; Davis et al., 2020; Lysaker et al., 2005, 2019, 2021; Vohs et al., 2015), with evidence for links between lack of insight and poor treatment compliance, severity of psychopathology, and poor global functioning (Mintz et al., 2003; Pyne et al., 2001). Metacognitive deficits may limit one's ability to recognise changes in mental states, disruptions in the course of their life, and prevent them from constructing a coherent account of one's experience of mental illness (Lysaker et al., 2005, 2015; Nicolò et al., 2012; Vohs et al., 2016), and thus compromising one's clinical insight into illness (for links between metacognition and clinical insight in schizophrenia and first episode psychosis see: Lysaker et al., 2021; Wright et al., 2023). Instead, in eating disorders, assessment of clinical insight is usually achieved by coding clinical notes, scoring self-report measures of eating disorder symptoms, vignettes, or direct measures of insight [see Fatt et al., 2022 for a scoping review on existing methods to assess insight in eating disorders; and, see Konstantakopoulos et al., 2020 for the first validated questionnaire to assess insight in eating disorders; Schedule for the Assessment of Insight in Eating Disorders (SAI-ED)]. Unfortunately, these aforementioned methods do not help extend our understanding of the (potential) associations between poor insight, metacognitive deficits and poor self-efficacy in AN. Importantly, work in other conditions (e.g., addiction) has shown that impaired insight into one's addictive behaviours (Goldstein et al., 2009) is consistent with impairments of metacognitive sensitivity in this population (Moeller et al., 2016), also documented in

pathological gambling (Brevers et al., 2014), brain injury (Fleming et al., 2014; Ham et al., 2014; Pannu & Kaszniak, 2005), and anosognosia (Kirsch et al., 2021). Thus, exploring the relation between insight and metacognition in AN may be useful in further exploring the basis of deficits in either or both of these aspects.

1.4.2.3 Associations between the Experimental and Clinical Research Insights, Identifying Current Gaps and Section Summary

Initial findings on specific metacognitive deficits in neuropsychiatric conditions and symptoms have offered insight into the multiple facets of metacognition (e.g., domain-general, domain-specific metacognition), and hierarchical levels (e.g., local metacognition, global metacognition). Whilst traditional experimental approaches had largely focussed on the study of metacognitive domains at a local level only (e.g., assessing trial-by-trial confidence in a memory or visual perception task), clinical traditions have increasingly highlighted the associations between self-beliefs (e.g., self-efficacy) and treatment adherence and prognosis. Thus, interdisciplinary efforts should be made to study both local and global metacognition and their association with mental health symptoms in order to also better understand how aberrant or distorted metacognitive beliefs could be therapeutically targeted. Another effort should be made towards a better understanding of how (clinical) metacognition deficits can influence belief updating about inner bodily states and their contributions to symptom development and maintenance. This would come after recent efforts in studying mental health under a Bayesian framework of active interoceptive inference (see Section 1.5) and understanding how aberrant predictions about the body can contribute to a chronic, maladaptive predictive state as manifested through a range of clinical, neuropsychiatric symptoms. Although the role of interoception has been acknowledged in mental health, most studies in metacognition have not looked at metacognitive interoceptive deficits. For instance, in a sample of participants with functional cognitive disorder, a condition in which interoceptive abnormalities have been noted, metacognitive processing was studied in the visual perception domain (for local metacognition) and global metacognition was assessed using a questionnaire (Bhome et al., 2022). As a result, insights from such studies cannot characterise the associations between higher-order mechanisms of metacognition and dimensions or levels of interoception (e.g., *How well will I be able to notice my hunger/ How well will I be able to feel my heartbeat*). Addressing these above gaps in the literature with a focus on Anorexia Nervosa (AN) was thus the key, overall aim of this present thesis.

1.5 Adopting an Active Interoceptive Inference Approach to Study Belief Updating of Prospective, Metacognitive Self-Efficacy Beliefs Under a Unified Bayesian Learning Framework

Work on interoception has revealed that instead of addressing interoception as a unidimensional entity, one should consider the many layers and dimensions involved, ranging from assessments of interoceptive accuracy (performance on objective behavioural tests) and self-report measures of interoceptive sensibility to ‘higher’ level dimensions that may extend to static and dynamic beliefs and models of one’s bodily state (see more on interoception in Section 1.3; Critchley & Garfinkel, 2017; Seth, 2013; Tsakiris & Critchley, 2016) and the accuracy of one’s interoceptive beliefs (or, metacognition, see more in Section 1.4; Petzschner et al., 2017; Stephan et al., 2016). Moving beyond the traditional, experimental approaches to study interoception and metacognition, recent computational methods have been adopted as they can complement existing experimental methods by capturing the complex dynamics of cognitive processes, offering a more nuanced understanding of their underlying mechanisms. Such computational approaches may offer unique insights into how the body continually makes and updates predictions about future body states and adjusts behaviour accordingly in order to maintain homeostasis (i.e., autonomic nervous system regulation), thus extending our understanding of affected processes in neuropsychiatric disorders (Owens, Allen, et al., 2018). These computational approaches rely on neurocomputational theories and frameworks which broadly aim to characterise how the brain creates representations of its environment. These are introduced below.

1.5.1 The Free Energy Principle

An influential neurocomputational theory, the ‘Free Energy Principle’ (FEP), is a principle that can be applied to perception, learning, and action in biological organisms (Friston et al., 2006; Friston, 2010; Gershman, 2019). The FEP relies on the premise that individuals are biological, self-organising organisms who in order to remain in (homeostatic) states they can sustain, they need to minimise the surprise (or, free energy) of sensory inputs (Friston et al., 2006; Friston, 2010). In other words, under the FEP, homeostasis is seen as a process that restricts the organism to visiting only a given number of states that are conducive to its ongoing existence (Corcoran & Hohwy, 2017; Friston et al., 2006; Friston, 2009, 2010). The FEP can serve as a foundational and unifying perspective underlying broader frameworks (e.g., the Bayesian Brain Hypothesis; see below) which attempt to explain how the brain achieves the minimisation of free energy by updating predictions and internal models of bodily states, combining prior

knowledge and beliefs of sensory states with actual sensory input (Friston, 2010). In interoceptive metacognition this would mean that the brain minimises the uncertainty regarding interoceptive states and can update its internal model in order to notify the organism about its deviation from homeostatic states. The FEP subsumes three hypotheses and frameworks. First, it can be studied in parallel to the theory that the brain has a model of the world and uses sensory inputs to construct a representation of it (the Bayesian Brain; Knill & Pouget, 2004; see more below). Secondly, in order to minimise the free energy, the brain needs to reduce the surprise or uncertainty; this aligns with the central concept of the Predictive Processing Framework (see more below), which states that to update one's internal model, the discrepancy between predictions and actual sensory input needs to be lessened (Aitchison & Lengyel, 2017; Gershman, 2019). Finally, active inference is rooted in the FEP, positing that an organism will actively minimise the free energy between internal models and the sensory input from the environment to fulfil their predictions (see more below; Friston, 2009; Friston et al., 2016; Pezzulo et al., 2015). Combining these three offers a unified view on how the brain may operate to combine top-down predictions about bodily interoceptive states with bottom-up afferent sensory input to generate a model of its internal state and also update this based on the continuous feedback it is receiving.

1.5.2 The Bayesian Brain and Belief Updating

One way to computationally study interoception is by using the influential Bayesian models of perception and inference (a method of statistical inference in which new observations are used to update or infer the probability that a hypothesis was true; see below) to characterise neurocomputational abnormalities in interoceptive and non-interoceptive perceptual and metacognitive inferences across psychopathologies (for interoceptive examples see: Fogelson et al., 2014; Gu, Eilam-Stock, et al., 2015; Gu, Lohrenz, et al., 2015; Owens, Allen, et al., 2018; Owens, Friston, et al., 2018; Petzschner et al., 2021; Stephan et al., 2016; for non-interoceptive examples see: Adams, 2018; Adams et al., 2016; Fletcher & Frith, 2009; Huys et al., 2016, 2021; Stephan et al., 2017; Sterzer et al., 2018). The Bayesian models of perception and inference rely on the theoretical framework of the Bayesian brain – the hypothesis of the Bayesian brain assumes that the brain has actively learned statistical patterns by combining prior knowledge with incoming evidence to develop predictions and improve the way in which information is processed and the way in which sensory input is interpreted (Knill & Pouget, 2004; Yon & Frith, 2021). Such process of forming and adjusting predictions based on a

combination of prior knowledge or beliefs and new sensory information is necessary to make judgements and guide action in both the external and internal world, thus playing a crucial role in perception, including interoception, and metacognition (Knill & Pouget, 2004). For instance, Bayesian mechanisms allow the brain to make inferences about emotions, visceral sensations, and physiological states, helping the body maintain its homeostatic balance. Similarly, Bayesian principles underlie metacognitive processes wherein an agent integrates self-generated predictions with actual experiences via monitoring and updating beliefs about one's own mental states, allowing to reflect on one's own cognitive processes.

Difficulty or impairment in an agent's ability to integrate new information into their existing beliefs and probabilistically infer a future state (or, prediction) has been linked with neuropsychiatric symptoms. For example, a recent study examined how individuals with AN and matched healthy controls rated the possible interpretations of given scenarios (e.g., 'A girl cannot fall asleep') and then adjusted their interpretation (e.g., why can the girl not fall asleep) after receiving additional information to specify the context of each scenario which became increasingly more specific over the trials (e.g., it gradually becomes clear that the girl cannot fall asleep due to the excitement over the impending Christmas morning; 'Bias against Disconfirmatory Evidence (BADE)' Task; Tenconi et al., 2023; Woodward et al., 2006). The results showed that the AN group (both acute and weight-restored individuals) had difficulties in disconfirming their prior beliefs (plausibility ratings for the presented scenarios) in light of new information (Tenconi et al., 2023). This finding of poor belief updating was particularly marked in the early stages of receiving additional information, suggesting that when the uncertainty levels regarding the new information are high, individuals with AN are even more unwilling to use the new information and update their own prior beliefs. The authors further suggested this finding might be relevant to the understanding of why patients with AN who have high levels of rigidity, perfectionism and obstinacy, may not discontinue maladaptive behaviours (e.g., eating restriction, purging, hyperactivity; Tenconi et al., 2023). Similar difficulties in belief updating have been observed in other domains and symptoms. For instance, a recent study in patients who lack awareness of their paralysis following right hemisphere stroke (or, anosognosia in hemiplegia) showed that patients were not able to update their beliefs about motor abilities prospectively (Kirsch et al., 2021). Specifically, at a behavioural level, patients were able to retrospectively recognise that they had failed to perform an attempted motor action (e.g., cutting a steak with both hands), despite having overestimated such motor ability prior to performing the task (e.g., when asked '*How well would you be able*

to cut a steak with both hands?”). Yet, patients seemed unable to ‘transfer’ this retrospective insight to update their prospective beliefs about motor abilities when asked to rate how well they would be able to perform this task in the future. Although these two examples differ in that they were carried out in different clinical populations, assessed belief updating in distinct domains, and also did not use the same kind of ratings, they both show that the way in which beliefs were updated could also rely on other parameters, such as the uncertainty regarding the new information versus the uncertainty regarding the prior.

In fact, when generating a prediction or updating a belief, an agent needs to balance prior knowledge with new, incoming evidence based on the ‘precision’ of these different sources (Knill & Pouget, 2004; Yon & Frith, 2021). Mathematically, precisions are the inverse dispersion or variance, i.e., the inverse of uncertainty, and play a key role both at the perceptual and metacognitive level. For instance, when walking in the forest and hearing a sound, the individual needs to use prior knowledge and new evidence to predict where the sound is coming from – the resulting percept relies on the brain’s estimation of the precision of incoming signals by combining data across the senses, or by combining inputs and expectations (Yon & Frith, 2021). This percept can be updated over time; for example, if the individual then sees where the sound came from (e.g., a rabbit), then this new evidence can be used in the future, should the agent hear a similar sound in the same forest. Estimating the precision of a prior belief against that of incoming sensory information can also be seen as a high-level mental ability, such as metacognition, or the ability to subjectively monitor one’s own mind’s reliability (Yon & Frith, 2021). A recent MRI study showed that when individuals had to make perceptual judgements about tilted visual patterns, the neural precision (as indexed by responses evoked by stimuli across the early visual cortex) predicted the participants’ subjective confidence rating on the respective trial, offering support for a previous hypothesis that the precision of other mental states is directly tracked by metacognitive mechanisms (Geurts et al., 2021). Furthermore, the anterior insular cortex has been found to have a central role in the integration of prior beliefs during tactile perception (Allen et al., 2016). The observed aberrant insular processes that have been observed in AN together with aberrant functioning in other brain regions could be contributing to the aberrant encoding of precision-weighted prediction errors, in turn explaining the observed belief updating difficulties in AN.

Although this Bayesian view of the brain offers a foundation to ways in which the self can anticipate the sensory input it will experience and subsequently generate representations,

modify its bodily (interoceptive) states, and update pre-existing representations in response to unexpected outcomes or evidence, it mainly describes the behaviour resulting from that brain computation, rather than the underlying mechanism(s) involved.

1.5.3 Predictive Processing Framework

Our beliefs about what constitutes our external and internal world do not reflect what this world actually is but are rather a construction of the world based on a template individuals have created to perceive oneself, others, and the world in general (Kube & Rozenkrantz, 2021). This notion thus suggests that our beliefs are continuously tested and challenged against the background of new information and experiences which can either confirm or disconfirm our existing prior beliefs and influence our updated, posterior beliefs. The roots of how we make predictions and update beliefs date back to von Helmholtz, who introduced the idea of unconscious inference, i.e., the idea that in order to make sense of a visual scene, the human brain fills in visual information (von Helmholtz, 1867). This early theory formed the basis for the predictive processing framework (PPF), a dominant theoretical framework in neuroscience which assumes that perception is linked to prediction and can be regarded as a formulation of the Bayesian Brain Hypothesis: prior predictions will bias our perception and thus shape posterior predictions (Kube & Rozenkrantz, 2021; Owens, Allen, et al., 2018; Owens, Friston, et al., 2018; Seth, 2013). Importantly, however, the PPF assumes that the processing of incoming sensory (bottom-up) signals is not passive, and instead the brain constantly generates (top-down) predictions about expected sensory input which is tested against prior predictions (Clark, 2013; Huang & Rao, 2011; O'Reilly et al., 2012). The sensory input contains statistical regularities which are captured in what are known as generative models – these models aim to use a probabilistic format to represent the causal structure of the world and are constantly updated taking into account new evidence (Barron et al., 2020; Clark, 2013; Friston, 2005; Harkness & Keshava, 2017). Under the PPF, the brain then uses the actual sensory input against predicted states (e.g., if a face that is lit from the top turns to the left, the brain will need to predict how the lighting on the face will change given that the effects of the light are also dependent on what has been perceived in the past) in an attempt to minimise the discrepancy between an actual and a predicted state, i.e., to minimise prediction errors (Friston, 2009; Harkness & Keshava, 2017).

Prediction errors can be used to provide corrective feedback for future predictions (or belief updating), and this can occur both by updating generative models (perceptual inference and learning; changing the model to fit the world), or by performing actions to align sensory states with predictions (active inference; changing the world to fit the model; Friston et al., 2009; Seth, 2013). In order to minimise prediction errors there needs to be a reciprocal exchange of bottom-up (low level) and top-down (high level) signals between the brain's hierarchical levels; the prediction error (feedforward loop), propagated up the hierarchical levels, is then used to update the original prediction (feedback loop) at that level (Aitchison & Lengyel, 2017; Barrett & Simmons, 2015; Owens, Allen, et al., 2018; Seth, 2013). This hierarchical structure of predictive coding and processing frameworks adds on to the Bayesian brain hypothesis by emphasising the role of prediction errors in the updating of internal models, and can also be compared to the multifaceted aspects of metacognition: low-level predictions can be more targeted predictions about the exact form of a sensory input, similar to the 'local' aspects of metacognition (e.g., an isolated confidence estimate of how well one performed in a trial), while higher-level predictions which are more abstract and require multisensory integration across domains or information about the context can be juxtaposed with the 'global' aspects of metacognition (e.g., an overall confidence estimate of how well one may perform in a future task; Friston, 2008; Pezzulo et al., 2015; Williams, 2020).

Prediction errors and the way they can be minimised also depend on the assigned precision to priors and new evidence. Specifically, precisions significantly influence subsequent hierarchical processing, since they can modulate the extent to which predictions are reduced by updating generative models or by performing actions (Friston et al., 2010). Understanding how prediction errors about interoception may arise from a mismatch between predictions regarding bodily afferents (i.e., being aware of an interoceptive bodily state) and the actual experienced bodily afferents (i.e., what the central nervous system actually experiences) is key in computational psychiatry since disruptions at level(s) of the predictive processing have been associated with a disruption of one's ability to maintain physiological homeostasis and experience one's (meta)cognitive self, resulting in neuropsychiatric symptoms and disorders, such as AN (Seth et al., 2012; Smith, Badcock, et al., 2021). Similarly, precision estimation in metacognition allows the generation of strong predictions, but aberrant precision-weighting of prior beliefs in light of new, sensory information will lead to prediction errors (Yon & Frith, 2021). If this persists over time, then the updated predictions one makes over time will be based on a misrepresentation of one's actual state, such as a misrepresentation of one's interoceptive,

embodied self (see Saini et al., 2022 for a model of depersonalisation disorder as a downregulation of interoceptive prediction errors, giving rise to a disembodied self; and see Paulus & Stein, 2006 for a view on anxiety as the result of a discrepancy between predicted and experienced interoceptive signals). A recent study in individuals in a transdiagnostic clinical sample of individuals with psychiatric disorders (such as substance use disorders (SUDs), and also a small eating disorders sample, N=14) and matched healthy controls used a Bayesian computational approach to identify which parameters best explained participants' behaviour in a cardiac perception task (Smith et al., 2020). Participants had to tap a button each time they felt a heartbeat and were then also asked to hold their breath (to amplify the ascending heart signals). Both control and SUDs groups had similar prior beliefs and precision estimates (assessed by rating the accuracy of their performance at the end of each condition) in the condition at rest (i.e., when not holding their breath to alter the perception of the cardiac signal). Yet, in the hold-breath condition, the precision of the control group's estimates increased, but the SUDs group showed no change, suggesting an inability to update beliefs about the reliability of the sensory signals arising from their body (Smith et al., 2020). These findings show the importance in precision-weighting mechanisms in interoceptive processing, suggesting that differences in the precision attributed to sensory input rather than prior beliefs may best account for interoceptive differences in certain clinical populations (Smith et al., 2020). This novel approach to model behaviour offers insights into the distinct computational mechanisms which may be involved in belief updating with implications in interoceptive processing in psychiatric disorders (Smith et al., 2020). But given methodological limitations it was not able to capture and model how these precision-weighting mechanisms and both prior beliefs as well as 'local' metacognitive beliefs (retrospective evaluations of performance after participants completed the task conditions) could influence the formation of more 'global' metacognitive beliefs. This approach was also not able to disentangle which computational mechanisms and parameters contribute to the formation of posterior prospective, interoceptive beliefs (a key aim here).

The PPF assumptions are closely linked to a similar framework, known as active inference, which relies on the concept of Free Energy Principle (see Section 1.5.2). Under this framework, the brain is believed to adjust (future state) predictions to minimise prediction errors (or, the free energy) and use incoming sensory information in the most optimal way in order to create the most optimal and accurate representation of oneself, others, and the world in general, in turn maximising the evidence for its generative model and maintaining homeostasis (Barrett &

Simmons, 2015; Friston, 2010; Friston et al., 2015; Friston & Kiebel, 2009; Moran et al., 2013; Owens, Allen, et al., 2018; Pezzulo et al., 2015; Seth & Friston, 2016).

1.5.4 Active (Interoceptive) Inference

Under the aforementioned FEP, it is believed that action can minimise the free energy through what is known as active inference (Friston et al., 2009). Under the theory of active inference, interoceptive perceptions are believed to be dynamically constructed by the brain, i.e., interoception is the outcome of a repetitive process whereby the brain compares its expectation of sensation with incoming sensations (Barrett & Simmons, 2015; Paulus & Stein, 2010; Pezzulo et al., 2015; Stephan et al., 2016). The goal of the active inference process is to form the most accurate representation of the world in order to facilitate the most adaptive behaviour despite the uncertainties pertaining to what an expected sensation will actually be like (Paulus et al., 2019). Thus, active inference extends beyond only providing homeostatic control for the body, and also encompasses allostatic control (where allostasis broadly refers to the active process that leads to adaptation to a stressor; McEwen & Wingfield, 2003; Romero et al., 2009) to explain how the organism may proactively prepare for homeostatic disturbances before they occur (e.g., increasing blood pressure in anticipation to the hypotension that is induced by postural change when one stands up from a chair; Corcoran & Hohwy, 2017) and how such allostatic, anticipatory beliefs may be used to predict a bodily state (see Stephan et al., 2016 for a formalised Bayesian implementation of hierarchical allostatic control and the effects of chronic dysregulation on fatigue and depression). Moreover, it is hypothesised that perceptual processes of active inference occur hierarchically and bidirectionally (Friston, 2008). In other words, the brain has various models which exist at different levels, and the models at the lower levels of the hierarchy are used as evidence for models at higher levels, while the models at the higher levels are used to modify the expectations of those at lower levels (Friston, 2008; Paulus et al., 2019). For instance, a higher heartbeat might be used to increase one's expectation of arousal, which could then provide evidence for fear or excitement – depending on contextual factors levels (Friston, 2008; Paulus et al., 2019). By weighing the model's accuracy in representing the world, the active inference process allows the brain to settle on a perception which has the least discrepancy between incoming sensory information (from the inside or outside world) and the expected model (Friston et al., 2013; Paulus et al., 2019). In short, active inference reduces prediction errors either by changing the prediction (or, perception), or by altering the predicted sensation (or, action; Owens, Allen, et al., 2018).

Recent efforts to understand the mechanisms underlying neuropsychiatric symptoms have thus suggested that symptoms may occur following the brain's difficulty in using afferent bodily signals (and prediction errors) to make an interoceptive inference of the body (Paulus et al., 2019; Smith et al., 2020). For example, individuals with hypochondriasis and somatisation disorders often report with somatic hypervigilance and interoceptive sensitivity which could be indicators of impaired interoceptive precision (Anderson & Hope, 2009; Barsky, 1992; Owens, Allen, et al., 2018; Rief et al., 1998). In another example, hypervigilant responses in anxiety were explained by increased postsynaptic gain and modulation of feedforward pathways, i.e., aberrant precision-weighting of bottom-up prediction errors (Cornwell et al., 2017). Evidence in AN has shown that during a cardiac interoception task patients with AN experienced cardiorespiratory visceral illusions (i.e., they were reporting heightened sensations of palpitations and dyspnoea, and did not accurately localise the cardiac sensations) when receiving saline infusions which did not change the observed state of the body; this was particularly observed at the premeal stage, a particularly anxiogenic time period for individuals with AN (Khalsa, Hassanpour, et al., 2018). Combined with another finding of reduced activation in the insular cortex during the anticipation and resolution of restricted breathing in recovered AN individuals (Berner et al., 2018), these results support the idea that individuals with AN may have hyperprecise priors about context-specific scenarios (e.g., hunger state) which in turn create strong expectations about visceral sensations (Paulus et al., 2019). In turn, such strong priors with imprecise weighting could result in maladaptive behaviours which are put in place to reduce uncertainty about interoceptive states, such as food restriction and starvation (see Barca & Pezzulo, 2020 for a recent theoretical framework of active interoceptive inference in AN).

1.5.5 A Unified Approach to Study Bayesian Interoceptive Belief Updating

Using insights from the above frameworks and theories, the aim of the present thesis was to use a unifying Bayesian, computational approach (Friston, 2010; Petzschner et al., 2017; Smith et al., 2020; Smith, Mayeli, et al., 2021; Stephan et al., 2016) to study under one continuous framework how perception and local metacognition influence the updating of explicit, global prospective beliefs about interoception in AN (see Kirsch et al., 2021 for a recent application of this approach in a different neuropsychiatric context). In particular, the focus was placed on the beliefs that individuals may have about interoceptive abilities (e.g., detecting heartbeats or

predicting stomach fullness states), and how they are generated and updated across the cardiac and gastrointestinal modalities. To my knowledge, this is the first attempt to characterise cross-modal dysfunctions in interoception and across levels and domains of metacognition (local and global metacognition) in an AN sample (acute and weight-restored) in comparison to a control sample. These aims come after proposals that the aberrant belief updating in mental health could result from a dysfunctional ‘precision-weighting’ mechanism which in turn influences the perception and processing of interoceptive signals (Smith et al., 2020). This thesis’ overarching goal was to further extend our understanding of how the anorectic self may in part be the result of belief updating and interoceptive inference distortions across modalities and hierarchical levels of metacognitive and learning processes, e.g., the selective attenuation of the processing of incoming sensory input and iterative use of noisy sensory data to update interoceptive inferences.

1.5.6 Interim Summary

In brief, this section introduced the unified Bayesian framework which was adopted in this thesis to study the formation and updating of beliefs about interoception in AN. The background to key frameworks and theories which explain how the brain makes predictions and updates its internal models was introduced, also providing evidence of the role of deficits in those mechanisms in mental health. The section concluded with the present thesis’ overarching aim in studying the updating of interoceptive, metacognitive beliefs in AN.

1.6 Current Thesis

Based on the above review, in a series of empirical studies and one clinical audit, this thesis focussed on experimentally and computationally investigating how patients with and weight-restored from AN differ from HCs in the formation and updating of beliefs regarding their cardiac interoceptive abilities. In parallel to this, a second key aim was assessing the domain specificity of the findings in the cardiac modality, and whether patterns and deficits observed in a cardiac modality would replicate in the gastrointestinal modality, which is directly linked to the primary AN symptomatology. Moreover, patients with AN present with severe structural and functional modifications of brain and peripheral-organ function and well-documented, related cognitive and mood effects (e.g., Bomba et al., 2015; Castro-Fornieles et al., 2009; Kaye et al., 2009; Mainz et al., 2012; but see Seitz et al., 2016 for review on brain morphological changes in adolescent and adult AN patients) which are sequelae of the prolonged malnutrition and starvation. However, despite the observed consequences from the

acute AN state, several traits are thought to constitute a vulnerability factor to the development of AN and may sometimes endure beyond the acute state – these include harm avoidance, perfectionism, and obsessive-compulsive personality traits among others (e.g., Anderluh et al., 2003; Bulik et al., 2003; Crane et al., 2007; Lilenfeld et al., 2006; but see Gárriz et al., 2021 for a different view). Although interoception and metacognitive deficits have been observed in AN, the role of state- and/or trait-related disturbances across relevant interoceptive and metacognitive domains have not been studied extensively. By recruiting both an acute and a weight-restored group in addition to the matched control group, the present thesis aimed to further delineate whether the hypothesised deficits in the retrospective and prospective metacognitive updating of interoceptive beliefs are secondary to the acute AN state or stable premorbid and enduring traits. Adopting this sampling approach across both the cardiac and gastrointestinal interoception studies of this thesis allowed qualitative comparisons of the domain-specificity and generality of these state-trait related disturbances. At the same time, the role of individual differences including clinical and cognitive traits which could explain part of the predicted differences between the three groups was considered. Finally, this thesis also presents the observations from a clinical audit which assessed whether using a biofeedback-enhanced intervention to target self-efficacy beliefs regarding self-regulatory abilities (e.g., regulating one's own heart rate) brought improvements to the practices of a local NHS eating disorders unit (in comparison to the existing standard practice of mindfulness). These aims were assessed in acute and weight-restored patients with Anorexia Nervosa (AN and AN-WR, respectively), and matched healthy controls (HCs). The studies and specific aims and hypotheses are outlined below:

Chapter 2, Study 1: Perceptual and Metacognitive Evaluations About Cardiac Interoception in Anorexia Nervosa

As reviewed in Section 1.4, most experimental psychiatric research to date has focussed on local metacognition, despite recent accounts showing that global aspects of metacognition relate more closely to everyday neuropsychiatric symptoms and their awareness (Kirsch et al., 2021; Seow et al., 2021). Yet, to my knowledge, these aspects of global metacognition and their relation to key interoceptive impairments noted in AN (as seen in Section 1.3) have yet to be assessed in AN. Thus, in the first pilot, empirical study of this thesis, the local and global beliefs about one's perceptual abilities in the cardiac interoception modality were assessed. Specifically, in this study a revised version of the Heartbeat Counting Task (HCT; Schandry, 1981) was used to obtain a measure of belief updating, by assessing participants' belief about

their performance abilities (i.e., how well they can feel and count heartbeats) before and after completing the task. Despite its known methodological limitations (e.g., Brener & Ring, 2016; Desmedt et al., 2018; Murphy et al., 2018; Richard et al., 2019), the HCT is a feasible task to do in an acute hospital setting and the requirement to form beliefs about one's performance can be intuitively and meaningfully understood by patients (e.g., '*How accurately will you be able to count your own heartbeats?*'). In line with previous evidence, more pessimistic beliefs about performance were predicted in the AN group in comparison to HCs, both prior prospectively (i.e., before even engaging in the task) and posterior retrospectively (i.e., when evaluating the task after completion).

Chapter 3, Study 2: Updating Prospective Self-Efficacy Beliefs About Cardiac Interoception in Anorexia Nervosa: An Experimental and Computational Study

Building on the findings of Chapter 2, patients with AN showed a marked pessimism regarding their interoceptive abilities to feel their heartbeat, despite no evidence of group differences on performance (or, interoceptive accuracy) in comparison to HCs. Moreover, although Experiment 1 (Chapter 2) did not employ a full belief updating paradigm, it was observed that after participants received arbitrary feedback from the experimenter regarding their performance (half were told they did better than the rest of the participants, and the other half were told they did worse), irrespective of group, all participants gave more optimistic performance ratings if the social feedback (even false, as in this case) was positive, and more pessimistic ones if it was negative. This finding suggests that the belief updating in AN may be particularly compromised regarding interoceptive signals and may be more domain-specific rather than domain-general. To further explore belief updating difficulties in AN, the preregistered Experiment 2 of the present Chapter further focussed on why AN patients struggle to update their 'pessimistic' prospective self-efficacy beliefs despite this comparable performance to HCs on formal interoceptive tasks, such as the HCT (Schandry, 1981). Using a behavioural and also computational modelling approach under a unifying Bayesian Learning Framework (see Section 1.5 for more), the aim of Experiment 2 was to study both global retrospective and prospective self-efficacy beliefs about interoceptive abilities. To date, experimental research on metacognition has mainly focussed on local and retrospective measures, while clinical traditions have typically focussed on global, retrospective and prospective metacognitive beliefs (see Section 1.4 for more). Yet, despite the interwoven relation between the two, there have been no attempts to combine the theoretical insights generated by empirical approaches with the insights generated through clinical studies. The

main hypotheses were that AN patients would have pessimistic prior and posterior prospective interoceptive self-efficacy beliefs, and they would show less belief updating and poorer precision-weighted Learning Rates than HCs. The role of comorbid traits and cognitive differences was also explored to examine if it explained the observed group differences in self-efficacy beliefs and Learning Rates.

Chapter 4, Study 3: Am I as full as I predicted I would be? An exploration of changes in fullness estimates during and after a Water-Sipping Task in women with and weight restored from Anorexia Nervosa

The primary symptomatology of AN that is used to diagnose the disorder (e.g., eating restriction) could extend beyond the desire to not eat in order to achieve one's ideal body weight or shape. Instead, it has been theorised to also be influenced by distortions in the way patients with AN experience fullness, satiety, and hunger. However, what remains unknown is whether the reported and noted distortions in fullness, satiety, and hunger are premorbid risk factors that may entertain the disorder and worsen during the acute states, and endure beyond the remission and recovery state, or if these distortions result from a prolonged period of starvation and emaciation, for instance as a secondary consequence to the eating restriction. This brings up some of the following questions which were addressed in Chapter 4: (1) How does a prolonged period of starvation (i.e., acute AN state) influence one's ability to feel hunger and fullness? (2) Do individuals who are weight-restored from AN have difficulty in attending to hunger and fullness signals? (3) How does an AN population differ from healthy controls regarding their ability to predict and update beliefs regarding fullness states? These are some of the questions that remain unanswered, partly because most clinical studies on interoception to date have focussed on the cardiac modality, and not on the gastrointestinal modality. The aim of Experiment 3 was to thus study the differences in gastrointestinal interoception in a clinical sample of AN and AN-WR patients in comparison to HCs. Specifically, using a water-sipping task, participants had to estimate how full drinking certain quantities of water (shown to them in a bottle) would make them, and then provide a fullness estimate after having consumed that indicated amount of water. Comparisons between the three groups' predicted and reported fullness estimates were made, expecting AN and AN-WR patients to expect to be fuller, and show a lower rate at which they updated their posterior prospective fullness estimates.

Chapter 5, Study 4: The Use of Heart Rate Biofeedback Treatment to Assist the Regulation of Embodied Feelings: A Clinical Audit on the Process and Effects of Biofeedback in Clinical Practice

Chapter 5 reports the results of an eating disorder service improvement evaluation (clinical audit), which was designed using the theoretical framework under examination. Specifically, following evidence on the effects of mindfulness and other third-wave CBT practices, a local eating disorders unit added a biofeedback with guided breathing element in their practice to assist patients with AN become aware of embodied feelings, such as heart rate, and learn how to self-regulate in order to help them target the overall pessimism patients may feel regarding such self-regulatory abilities (typically observed in the AN population – see Chapters 1-4). Shortly after this biofeedback-enhanced practice to target the pessimistic self-efficacy beliefs about interoceptive abilities was introduced, a clinical audit was conducted to evaluate the new process. The aims of this clinical audit were to evaluate the service improvements at a local ward, by assessing if the process in which this biofeedback-enhanced practice was feasible and whether it successfully met the local standards of care.

Overall, the present thesis aimed to provide insights into the role of interoception in AN and how perceptions and beliefs about interoceptive (cardiac and gastric) abilities are updated retrospectively and prospectively, while also taking into account the role of individual differences and comorbid cognitive and psychological traits. This thesis also presents the observations from a clinical audit which evaluated a biofeedback-enhanced practice against existing practices of mindfulness for AN patients receiving inpatient treatment. A general discussion about the pattern of findings arising from this thesis is presented in Chapter 6. The findings and their importance in further understanding the role of interoception in AN as a mechanism involved in the development and maintenance of symptoms will be discussed, and the benefits of therapeutically targeting interoceptive metacognitive beliefs will also be considered. Finally, the general limitations and implications will be discussed in detail.

Chapter 2: Perceptual and Metacognitive Evaluations About Cardiac Interoception in Anorexia Nervosa

2.1 Introduction

Despite recent advances in treatments for Anorexia Nervosa (AN), studies report low adherence to treatment and high relapse rates (Carter et al., 2004; Halmi et al., 2005; Roux et al., 2016). As such, AN has the highest mortality and morbidity rates among psychiatric disorders (Arcelus et al., 2011). Poor compliance to AN treatment has been associated with resistance to treatment and the disorder's 'egosyntonic' nature (Gregertsen et al., 2017; Serpell et al., 1999). The term egosyntonicity is used in this context to describe the fact that patients with AN may regard their restrictive eating as adaptive to their psychological needs and identity, while disregarding the illness' negative impact on their health (Serpell et al., 1999). Egosyntonicity may be linked with impaired 'insight', i.e., a lack of appreciation of one's illness. Insight deficits have been found to be a maintaining factor in AN, hindering treatment adherence and recovery (Abbate-Daga et al., 2013; Konstantakopoulos et al., 2012) and are more frequently observed in restrictive AN (r-AN) compared to the binge-eating/purging subtype and in patients with Bulimia Nervosa (Greenfeld et al., 1991; Konstantakopoulos et al., 2011; see Sections 1.2 and 1.4, Chapter 1 for more details on the epidemiology of AN and insight deficits).

However, to date, insight disturbances in AN have mainly been studied in relation to treatment outcomes rather than explanatory mechanisms. To my knowledge, there is only one pilot study on insight in AN that suggests a potential association between insight deficits and impairments in other abilities, such as metacognition – the ability to reflect upon and monitor one's own cognition (Arbel et al., 2013; Flavell, 1979). Further experimental studies and mechanism-based accounts of insight deficits are lacking in AN. However, work on other conditions (e.g., in schizophrenia; Beck et al., 2004; Lysaker et al., 2021) suggests that clinical insight could be dependent on other higher-order abilities, such as metacognition, including for example, patients' ability to judge the veracity of their own perceptions or beliefs. This metacognitive dimension of insight is also termed 'cognitive insight' (Beck et al., 2004; Lysaker et al., 2021) and several psychiatric studies have explored the relation between clinical and cognitive insight (for review see Van Camp et al., 2017). However, cognitive insight and metacognition have not been experimentally addressed in AN research - one of the aims of the present experiment.

Moreover, as discussed in Section 1.4 (Chapter 1), metacognition itself is a multifaceted ability. One dimension of interest here is the difference between ‘local’ (e.g., retrospective judgements on specific, isolated cognitions, such as confidence judgements on the accuracy of a percept) versus more abstract and ‘global’ aspects of metacognition, such as overall retrospective beliefs about one’s performance across many tasks, or prospective judgements about one’s capabilities in different contexts. The latter aspect also relates to the concept of self-efficacy beliefs, defined by Bandura as one’s beliefs in their capacity to succeed in a certain situation, influencing how one thinks, behaves, and feels (Bandura, 1977). Individuals with high self-efficacy are more confident in their ability to perform at a certain level or control the outcome of a task (Weerdmeester et al., 2020; Zulkosky, 2009). On the other hand, poor self-efficacy may be involved in the development of mental health symptoms (Seow et al., 2021). In eating disorders, self-efficacy influences drive for thinness, body dissatisfaction, and treatment outcomes (Keshen et al., 2017; Toray & Cooley, 1997). However, self-efficacy in AN has only been assessed with questionnaires. Indeed, hitherto most experimental psychiatric research on metacognition has focused on local metacognition. It is only recently that global aspects of metacognition have been found to relate more closely to everyday neuropsychiatric symptoms and their awareness (Kirsch et al., 2021; Seow et al., 2021). Aspects of global metacognition, and their relation to key interoceptive and body perception impairments have not been assessed in AN. Thus, assessing both local and global beliefs about one’s perceptual ability was also one of the aims of the present experiment.

The main domain of interest in the present experiment was interoception. Interoception refers to the act of sensing, integrating, and interpreting information about the state of inner body systems (Craig, 2003; Khalsa, Adolphs, et al., 2018; see Section 1.3, Chapter 1 for more details). Sensing and interpreting the physiological state of the body at least partly depends on the ability to sense afferent signals, including respiratory, gastrointestinal, and cardiac signals (Cameron, 2001; Critchley & Garfinkel, 2017). In addition to sensation, interoception can be examined at different levels of signal processing, such as the accuracy at which one perceives interoceptive signals (typically referred to as interoceptive accuracy or interoceptive sensitivity), the self-reported beliefs regarding one’s interoceptive abilities (self-report and interoceptive beliefs, which include both measures collected using questionnaires and trial-by-trial confidence ratings), and the metacognitive evaluation of one’s interoceptive performance, or ability (e.g., measured via the confidence-accuracy correspondence between interoceptive accuracy in an interoceptive task and subjective interoceptive measures, such as one’s

confidence regarding the accuracy of their performance; Critchley & Garfinkel, 2017; Garfinkel et al., 2015; Khalsa et al., 2018; Suksasilp & Garfinkel, 2022; see Section 1.3, Chapter 1 for more).

Impairments in these different dimensions of interoception have been associated with several mental health conditions, including AN (e.g., Fassino et al., 2004; Fisher et al., 2016; Herbert & Pollatos, 2014; Jenkinson et al., 2018; Khalsa et al., 2018; Pollatos et al., 2008; Pollatos et al., 2009). Although difficulties with interoception in AN have been noted and studied across cardiac and other domains (e.g., Ambrosecchia et al., 2017; Crucianelli et al., 2021; Khalsa et al., 2015; Kinnaird et al., 2020; Lutz et al., 2019), the methodological differences or the lack of rigorous methods to measure interoception, and the lack of a consensus on the characterisation of its dimensions has contributed to several inconsistencies in the results (Arnaud et al., 2023). For example, studies on interoceptive accuracy have offered mixed findings – while some groups have found that AN patients do not differ from matched controls in their ability to accurately feel their heartbeat in the absence of any physical manipulation to facilitate the detection of heartbeats (e.g., Kinnaird et al., 2020; Lutz et al., 2019; Richard et al., 2019), others have observed poorer cardiac interoceptive accuracy in AN patients in comparison to HCs (e.g., Pollatos et al., 2008, 2016). On the other hand, preliminary pilot study results in the gastrointestinal modality of interoception point towards differences in the perception of fullness between a group of AN patients and a group of control participants (Brown et al., 2022). Taken together, these results from studies assessing one's ability to use 'objective' standards to represent one's own bodily sensations points towards the hypothesis that difficulties in interoceptive accuracy may be domain-specific, particularly relating to difficulties assessing gastrointestinal feelings, rather than extending to a generalised deficit across all interoceptive modalities (see more on gastrointestinal interoception in Chapter 4). Moreover, another study found that AN patients often report prolonged fullness, and one in three individuals with AN will report satiety and lack of hunger cues even when on an empty stomach (Bluemel et al., 2017). Another study tested changes in the perception of interoception before, during and after meal consumption and found that AN patients had dysfunctional interoceptive responses during meal anticipation, compared to controls (Khalsa et al., 2015). These dysfunctional interoceptive responses could be explained by pre-meal anxiety and related emotional responses (Khalsa et al., 2018), metacognitive or interoceptive awareness biases (Crucianelli et al., 2021), indexed by confidence ratings or questionnaires (Kinnaird et al., 2020). Yet, while AN patients have shown poor interoception across dimensions and

modalities, no study has tested directly the role of anticipatory metacognitive beliefs about interoception in AN – one of the present aims in this Chapter’s experiment.

Specifically, in the present experiment, certain theoretical approaches to interoception and mental health (e.g., Stephan et al., 2016) that take into account both retrospective and prospective aspects of interoceptive metacognition, such as prospective, self-efficacy beliefs about interoceptive abilities and retrospective beliefs about one’s interoceptive accuracy in a specific task were followed. Thus, this allowed the investigation of how acute and weight-restored AN patients (AN and AN-WR, respectively) in comparison to HCs update their prospective beliefs about their ability to detect their heartbeats after engaging in a modified heartbeat counting task (HCT; Schandry, 1981) and receiving feedback about their performance in the HCT. Under a Bayesian Learning Framework (BLF; Friston, 2003; Mathys et al., 2011), the following measures were regarded: individuals’ (1) initial, prospective beliefs about how well they will do in the task as ‘prior beliefs’, (2) their subsequent estimates after the task as ‘posterior retrospective beliefs’ and (3) their final, retrospective beliefs on how well they did following false performance feedback from the experimenter (i.e., they were told they performed better or worse than the other participants, irrespective of their actual interoceptive accuracy scores, hereafter referred to as Performance) as ‘post false-feedback retrospective’ beliefs. In this context, the term ‘prior beliefs’ refers to the beliefs (or, probability distributions) the individual has before considering some new information. After the integration of this new information, these beliefs are referred to as ‘posterior beliefs’, and this change from prior beliefs to posterior beliefs is known as belief updating.

It was predicted that the AN group would anticipate performing worse in the HCT, i.e., have low prior beliefs in their ability to perform well in the HCT in comparison to the HCs (*Hypothesis 1*), given previous findings of affected interoceptive anticipation (Khalsa et al., 2015). However, it was expected that Performance (i.e., interoceptive accuracy scores) of the AN and AN-WR groups would be similar to that of the HCs, i.e., the accuracy of reporting felt heartbeats would not differ between the groups (*Hypothesis 2*), given previous related findings on group-differences in interoceptive accuracy (Eshkevari et al., 2014; Lutz et al., 2019; Richard et al., 2019). Moreover, it was predicted that the AN and AN-WR groups would provide lower ratings on their retrospective beliefs in comparison to HCs when asked to evaluate their Performance on the HCT (as in some previous studies measuring confidence following HCT; Kinnaird et al., 2020), despite an otherwise successful Performance

(*Hypothesis 3*). Next, it was also hypothesised that the AN and AN-WR groups, in comparison to HCs, would rely more on their prior beliefs rather than their Performance when providing a rating on their retrospective beliefs (given the hypotheses seen in the aforementioned literature regarding low confidence and self-esteem in AN and relation to interoception; *Hypothesis 4*). Hence, it was predicted that a greater discrepancy between Performance and prior beliefs (Prediction Error) in the two clinical groups compared to HCs would explain the patients' lower retrospective beliefs regarding their Performance. Additionally, the influence of false feedback (within each group) on post false-feedback retrospective beliefs was examined (*Hypothesis 5*). Specifically, this was done to examine if individuals during the acute and weight-restored stages of AN have difficulty changing their prior beliefs based on all kinds of feedback (here, indexed by external false feedback), or whether these difficulties were specific to (interoceptive) performance. Hence, all participants were provided with false feedback on their performance (i.e., each group was arbitrarily divided into two subgroups and receive false feedback of high vs low performance). It was expected that participants in the negative false feedback subgroup would provide lower false-feedback retrospective performance estimates, in comparison to those in the positive false feedback subgroup, and this would be observed in all three groups. This would suggest that any difficulties in the retrospective beliefs in the AN group (i.e., *Hypotheses 3* and *4* above) would be specific to embodied performance and not any kind of feedback.

Finally, self-report measures were used to assess whether relevant key traits previously associated with AN would predict group differences in prior beliefs and Performance, separately. Specifically, the effects of trait depression, anxiety, stress, alexithymia, and body awareness were examined separately. It was predicted that greater trait depression, anxiety, stress and alexithymia, and poorer body awareness would predict pessimistic prior beliefs and Performance across all groups, and in particular in the two clinical groups. In the AN group only, dimensions of an exploratory measure of insight (awareness on illness severity, consequences and prognosis) were used to explore their role on the predicted pessimistic prior beliefs and Performance. Finally, the effects of illness duration and severity in the AN and AN-WR groups' prior beliefs and Performance were also explored separately.

2.2 Methods

2.2.1 Participants

The convenience sample included a total of $N=161$ female participants ($N_{AN}=51$, $N_{AN-WR}=47$, and $N_{HC}=63$). Patients with AN and AN-WR were recruited from the Psychiatric Daycare Unit of Sao Paolo General Hospital (ASST Santi Paolo e Carlo) in Milan, from Comunità Villa Miralago in Cuasso al Monte (VI) and from Casa di Cura Villa Margherita in Arcugnano, Italy. HCs were Italian volunteers recruited from the University of Milan and University College London (UCL). Patients with AN met DSM-5 criteria (American Psychiatric Association, 2013) for restrictive subtype AN, as diagnosed by an experienced clinician, using standard clinical measures (patient history and psychometric questionnaires and a physical assessment), and had a BMI <18.5 . The delineation of remission and recovery criteria for AN remain a challenge for the field (see also Section 1.2, Chapter 1) and thus a combination of at least two of the following was chosen: (i) BMI >16.5 , (ii) clinical and behavioural signs of AN recovery (e.g., no restrictive eating patterns) for at least 6 months, or (iii) a global Eating Disorders Examination Questionnaire (EDE-Q; Fairburn & Beglin, 1994) score <4 (similar to the inclusion criteria of (Jenkinson et al., 2023)). This combination of behavioural and clinical criteria was decided given that BMI cannot account for other body composition complexities (e.g., race, muscularity, or age), and does not capture the complicated biopsychological determinants of AN and symptom severity, rendering BMI-only criteria outdated for the characterisation of AN and recovery (Harrop et al., 2021; Khalsa et al., 2017; Lebow et al., 2018; Machado et al., 2017; Ralph et al., 2022). Additionally, if an AN-WR participant had a BMI <18.5 , the clinical status of the individual was further confirmed by an experienced clinician. All HCs had a healthy BMI between 18.5 and 25. HCs were not included if they had a history of an eating disorder, or a first-degree relative with a history of an eating disorder. All participants were aged between 18 and 45 years, with no history of neurological disease, brain damage, drug dependence or severe psychiatric disease (e.g., psychosis; see Table 2.1 for other clinical characteristics including comorbidities and pharmacological treatment). Other exclusion criteria included being male, being pregnant, substance and/or alcohol abuse.

Institutional ethics approval was granted and written informed consent was obtained from all participants prior to their participation. All procedures were conducted in accordance with the World Medical Association Declaration of Helsinki. HCs and AN-WR participants received either monetary compensation at a fixed rate of £7.5/hour, or university course credits. The AN group received £10 Amazon vouchers.

Table 2.1. Summary of Participant Demographics and Clinical Profiles

	AN	AN-WR	HC	Test Statistic [β , (SE)], p^1	
	Mean (SD)	Mean (SD)	Mean (SD)	AN vs HC	AN-WR vs HC
N	51	47	63	-	-
Age	25.51 (9.48)	25.38 (6.23)	25.14 (4.77)	1.37(1.31), 0.299	0.24(1.34), 0.858
BMI	15.80 (1.53)	19.89 (2.46)	21.17 (2.98)	-5.37(0.47), <0.001	-1.29(0.47), 0.008
Average heartbeat *	45.80(8.60)	47.59(11.45)	66.40(18.24)	-20.60(3.10), <0.001	-18.81(3.18), <0.001
Illness Duration (years)	8.21 (8.45)	4.36 (3.85)	-	-	-
Age of Onset	16.65 (3.54)	17.59 (4.24)	-	-	-
Years from Weight Restoration	-	3.47 (5.40)	-	-	-
DASS-21 Depression	12.85(5.47)	7.65(6.00)	3.31(3.69)	9.54(1.28), <0.001	4.34(1.34), 0.002
DASS-21 Anxiety	7.33(4.52)	6.04(4.99)	2.29(2.35)	5.05(1.01), <0.001	3.76(1.05), 0.001
DASS-21 Stress	13.04(4.59)	8.61(5.29)	6.34(4.27)	6.69(1.19), <0.001	2.27(1.25), 0.074
BAQ Total	72.08(18.00)	74.21(17.86)	81.50(14.82)	-9.42(4.77), 0.053	-7.29(4.88), 0.140
TAS-20 Total	61.53(16.17)	48.24(16.08)	39.66(9.88)	21.87(2.66), <0.001	8.58(2.74), 0.002

Psychiatric Comorbidities	5 OCD 7 Mood Disorder 3 Panic/Anxiety Disorders 4 Personality Disorder 15 Multiple Comorbidities	2 OCD 8 Mood Disorder 10 Panic/Anxiety Disorder 2 Borderline Personality Disorder 1 Trichotillomania 9 Multiple Comorbidities	4 Subclinical Anxiety Symptoms
Psychiatric Treatment during the acute stage	16 Antidepressants 3 Mood Stabilisers 13 Antipsychotics 11 Sedatives	21 Antidepressants 3 Mood Stabilisers 4 Antipsychotics 5 Sedatives	1 Sedatives (<i>4y prior to participation in the study</i>)

N.B. Dashes indicate measure was not taken. Abbreviations: AN (Anorexia Nervosa); AN-WR (Anorexia Nervosa Weight-Restored); HC (Healthy Control); DASS-21 (21-Item Depression, Anxiety and Stress Scale; Lovibond & Lovibond, 1995); BAQ (Body Awareness Questionnaire; Shields et al., 1989); TAS-20 (20-Item Toronto Alexithymia Scale; Bagby et al., 1994); OCD (Obsessive Compulsive Disorder). Multiple comorbidities include at least two of the following: OCD, Mood Disorder, Borderline Personality Disorder, Panic/Anxiety Disorder. * Heart rate was recorded at the beginning of the session (for 5 minutes) and was then divided by 5 to obtain the average number of heartbeats per minute. ¹Linear regressions were run to examine group differences, with HC as the intercept. As expected, there were significant between group differences on BMI, consistent with the inclusion criteria, but the BMI differences between the AN-WR and HC groups were unexpected and thus the effect of BMI was controlled for where necessary. Given some of the expected group differences in psychometric traits, these were taken into account in the exploratory analyses (see Appendix A, Table A.2). Bolded values denote statistical significance ($p < 0.05$).

2.2.2 Design

An existing cardiac perception task (HCT; Schandry, 1981) was adapted to create a well-controlled experimental task for belief updating. The adapted task included the traditional measure of interoceptive accuracy (IAcc), hereafter referred to as Performance, and three additional measures. These measures include participants' (1) Prospective Performance Estimates (i.e., prior beliefs on the task they will be asked to complete), (2) Retrospective Performance Estimates (i.e., posterior beliefs about performance on the task they completed), and their (3) Post false-feedback Retrospective Performance Estimates (i.e., posterior beliefs about performance following false feedback regarding their performance). Prospective Performance Estimates were assessed by asking participants '*How well do you think you will do [on the task] on a scale of 0 (very poorly) to 100 (very well)?*' For Retrospective Performance Estimates the experimenter asked the participant '*How well did you do on the task on a scale of 0 (very poorly) to 100 (very well)?*'. Participants gave their answer verbally. The experimenter gave false feedback to each participant by telling them they performed better or worse than the other participants and Post False-feedback Retrospective Performance Estimates were obtained by asking participants: '*Now, if I were to ask you again, how well do you think you did on the task on a 0 (very poorly) to 100 (very well) scale?*'.

The independent variables (IVs) were Group (AN, AN-WR, and HC) and false feedback (positive and negative). The Dependent Variables (DVs) were the Prospective Performance Estimates, Performance, Prediction Error (difference between Performance and Prospective Performance Estimates), Retrospective Performance Estimates, and Post False-feedback Retrospective Performance Estimates.

2.2.3 Psychometric Measures

Depression, Anxiety, Stress Scale (DASS-21; Lovibond & Lovibond, 1995): The DASS-21 was only used in a subset of the sample ($N_{AN}=27$, $N_{AN-WR}= 23$, $N_{HC}=33$) due to clinical time constraints. The DASS-21 is a reliable scale (Cronbach's α can range between 0.86 and 0.90; (Gloster et al., 2008) and has 21 items used to assess any symptoms relating to negative mood and emotions on a 4-point Likert scale (0 – *Did not apply to me to all* to 3 – *Applied to me very much*). It has three subscales to assess depression, anxiety, and stress, with higher scores relating to higher levels of depression, anxiety, and stress, with cut-offs being 5, 4, and 8, respectively. The DASS-21 has positive psychometric properties in samples of adults with

anxiety and depression. The present sample showed an excellent internal consistency, Cronbach's $\alpha = 0.958$.

Toronto Alexithymia Scale (TAS-20; Bagby et al., 1994): The TAS-20 was used as a measure of alexithymia as it is the most commonly used self-report questionnaire, with demonstrated good reliability and factorial validity in both clinical and non-clinical populations (González-Arias et al., 2018). All participants, except for two AN and one AN-WR participants, completed the TAS-20, which consists of 20 items rated on a 5-point likert scale of 1 (*strongly disagree*) to 5 (*strongly agree*). Total scores may range from 20 to 100 (sum of the three subscale scores), and a total score of 61 or above is generally indicative of alexithymia (Franz et al., 2008; Honkalampi et al., 2001). The total TAS-20 score was used in the present experiment's analysis, instead of each subscale score separately due to the relatively small sample sizes and subsequent power-related issues. In the current experiment, the TAS-20 showed an excellent internal consistency, Cronbach's $\alpha = 0.913$.

Body Awareness Questionnaire (BAQ; Shields et al., 1989): The BAQ was used to measure self-reported interoceptive sensibility. Due to clinical time constraints the BAQ was only completed by a subset of the sample ($N_{AN}=24$, $N_{AN-WR}=22$, $N_{HC}=26$). The BAQ is an 18-item scale measuring attentiveness to normal, internal bodily processes and sensations and has four factors: the ability to note responses in body processes, and predict bodily sensations, sleep-wake cycle, and onset of illness. All responses are measured on a 7-point likert scale of 0 (*Not at all true about me*) to 7 (*Very true about me*), with higher scores indicating greater body awareness. The internal consistency of the BAQ is good, as shown both in an undergraduate student sample, Cronbach's $\alpha = .82$ (Shields et al., 1989) and a clinical sample (Cronbach's $\alpha = 0.786$; Unal et al., 2021). In this experiment, the BAQ demonstrated a high internal consistency, with Cronbach's $\alpha = 0.851$ in the present sample.

Insight into Illness: Finally, a subset of the AN and AN-WR participants completed an insight into illness questionnaire developed by the research team (see Table 2.2), given that to the best of my knowledge, there is no questionnaire that focusses only on insight in eating disorders. Specifically, there is one self-report measure of insight into illness in eating disorders but not specific to AN only (Konstantakopoulos et al., 2011; Konstantakopoulos, Georgantopoulos, et al., 2020) which was developed largely for clinical, psychiatric use. However, at the time of study design and recruitment it had not yet been validated, its content was too general for the

current experiment's scope, and its completion also did not fit within the time limits of the present experiment. Instead, participants from the two clinical groups answered five questions on different dimensions of restrictive, anorexic symptom awareness. The focus was placed on three dimensions, namely on the (i) General, Allocentric Perspective of Self, (ii) Health Consequences, and (iii) Future Perspectives: Hopes and Fears. The self-report items measure patients' awareness of their emaciated appearance and illness, severity, the negative consequences of their insight disordered eating and fears and beliefs regarding behaviours which may hinder prognosis. Each question was rated on a 0 (*not at all*) to 10 (*very much*) likert scale, and each participant's clinician (both for AN and AN-WR groups) also rated symptom and illness severity on these questions based on their professional evaluation of the individual's symptoms and the individual's medical files. Individual scores were averaged to obtain a 'subjective' global clinical insight score, with higher scores suggesting more insight into illness, while clinician scores were averaged to obtain an objective score, with higher scores indicating greater illness severity.

Table 2.2. Items of the Insight into Illness Questionnaire.

Item		Dimension of insight
		<i>General, Allocentric</i>
1	<i>How severe is your condition?</i>	<i>Perspective of Self</i>
<i>Do you think that your eating behaviour could have negative</i>		
2	<i>consequences for your health?</i>	
<i>Do you think that your body weight and shape could have</i>		<i>Health Consequences</i>
3	<i>negative consequences for your health?</i>	
<i>Are you afraid of losing control of your eating behaviour?</i>		
<i>Do you believe that the severity of your condition will</i>		<i>Future Perspectives:</i>
5	<i>improve in the future?</i>	<i>Hopes and Fears</i>

Participants at the UK site were administered the validated questionnaires in English. The participants at the Italian site were given the Italian-validated versions of the DASS-21 and TAS-20. The BAQ has not been validated in Italian, and the Insight into Illness Questionnaire, which was developed for the present experimental purposes is also not validated. Therefore, two members of the research team (including the PhD student) who are fluent in English and

Italian (separately) translated the BAQ and Insight into Illness Questionnaire, and another two members from the Italian site, fluent in both languages as well, backtranslated the questionnaires. Good agreement was shown across all items and between the two translations and backtranslations.

2.2.4 Experimental Measures and Procedure

All participants provided informed consent and were given written task instructions prior to completing the HCT and were also given the opportunity to ask the experimenter clarification questions. Demographic data and self-report, standardised questionnaire measures were completed prior to the HCT. Heart rate (both for baseline measurements and the counting phases of the HCT) were recorded using the Polar wrist worn heart rate monitor (model RS 800CX, see Emanuelsen et al., 2015; Fischer et al., 2016). As also described in detail below and in Appendix A, participants were encouraged to remain still and with their arms resting on the desk in front of them during the recording of their heart rate and the heartbeat counting trials so that no movement artifacts or noise is generated. The heart rate activity was programmed to be shown – in real time – on a computer screen (using the Biopac software; www.biopac.com). The computer screen was only visible to the experimenter, and not to the participant, and the experimenter was able to mark the beginning and end of each recording by clicking on a ‘recording start/stop’ button on the computer interface. After the experimental session was over, the experimenter and another member of the team independently checked the heart rate activity output to screen for potential artefacts and noise, as well as to count the number of actual heartbeats during the separate trials and then compared observations. The time of the recording was displayed on the computer interface, and was also noted manually to check for potential inconsistencies. In case of disagreement between the two research team members (e.g., in the number of actual heartbeats counted during a certain trial), a third member was asked to check the recording as well. However, no major disagreements were noted.

Initially, all participants were asked to remove any accessories from their wrists, roll up their sleeves, and to sit down with their wrist gently resting on the band of the heart rate monitor which was located on the table in front of them. Participants were allowed to keep their eyes opened or closed, whichever felt more comfortable, but with their legs uncrossed. Before completing the HCT, a baseline measurement for heart rate was obtained. Participants were asked to sit and relax, while the experimenter recorded their heart rate for 5-minutes using the

Polar monitor. Then, participants were given task instructions. Specifically, participants were told that they would have to silently count their heartbeats without attempting any physical manipulation or trying to take their pulse to facilitate detection. Participants were also told that they would have to only count the heartbeats they felt and not guess (during specific time periods as instructed by the experimenter) and verbally report the number of felt heartbeats when prompted by the experimenter. Next, they were given the opportunity to ask questions regarding the instructions (detailed instructions script is available in Appendix A). Based on the instructions they had just received, participants were then asked to provide a Prospective Performance Estimate (i.e., rate how well they thought they would do in the task). Next, participants were asked to silently count their heartbeats to measure their Performance. Performance was measured as one's ability to detect their own heartbeats, without feeling their chest or taking their pulse (Schandry, 1981). The participant was prompted with 'Go' and 'Stop' signals at the start and end of each counting phase, respectively. After each 'Stop', participants verbally reported the number of felt heartbeats. This was repeated for the three counting phases of 25s, 45s, and 65s, with 30s rest breaks in between. The order of the counting phases was randomised between all participants. Information about the length of the counting phases or participant performance was not given at this stage.

Upon completion of the three counting phases, participants were asked to give a Retrospective Performance Estimate. Finally, participants were told they performed either better or worse than the rest of sample and were asked to provide a Post False-feedback Retrospective Performance Estimate. Once participants gave a verbal answer, they were debriefed and told that this feedback was for experimental purposes only, and in no way was a reflection of their actual performance during the task. Due to clinical time constraints, false-feedback was not given to N=32 participants, hence these participants' Post False-feedback Retrospective Performance Estimates are missing.

2.2.5 Data Analysis

First, a linear regression was conducted to assess whether group (IV) predicted Prospective Performance Estimates (DV; *Hypothesis 1*). Control analyses for the effects of age and baseline heart rate were conducted to examine their effects on Prospective Performance Estimates given that significant differences between the groups' age and average heartbeat counts (per minute).

However, these two covariates did not significantly affect the main findings (see Appendix A, Table A.1).

Next, a linear regression was conducted to assess whether group (IV) had an effect on Performance (DV; *Hypothesis 2*). Performance (IAcc) was calculated using the Schandry (1981) transformation:

$$Performance = \left[1 - \frac{1}{3} \sum \frac{|Recorded\ Heartbeats - Counted\ Heartbeats|}{Recorded\ Heartbeats} \right] * 100 \quad 2.1$$

As seen in the equation above (2.1), the Performance scores were multiplied by 100 in order to maintain consistency with participants' Performance Estimates. Performance scores varied between 0 and 100, with a score of 100 implying best performance, i.e., a smaller difference between the recorded and counted (felt heartbeats), and a score closer to 0 indicating worst performance. Control analyses for Performance in each time interval separately were also conducted to ensure that overall Performance results were not exclusively driven by any one of the intervals. Control analyses with age and baseline heart rate as predictors were conducted to control for the effects of these factors, however they did not affect the main findings (see Appendix A, Table A.1).

Next, for *Hypothesis 3* a linear regression was run to assess whether the three Groups (IV) differed in their Retrospective Performance Estimates (DV) when controlling for their Prospective Performance Estimates and Performance (covariates). For *Hypothesis 4*, a separate linear regression was conducted to assess whether the difference between Performance and Prospective Performance Estimates (Prediction Error, PE; IV) of each group predicted their Retrospective Performance Estimates (DV).

Furthermore, three separate linear regressions were conducted to assess whether the false feedback (IV) participants received predicted their Post False-feedback Retrospective Performance Estimates (DV), within each Group separately (*Hypothesis 5*). However, given that false feedback was randomly given to participants, irrespective of their Performance in the task, three control linear regressions with Performance as IV and false feedback as DV were also carried out. These latter three analyses were conducted to look for possible within-group

differences in Performance and to control that positive false feedback was not randomly given to participants with higher Performance scores and negative false feedback to participants with low Performance scores.

Then, exploratory linear regressions were run for *Hypotheses 1* and *2* with the aforementioned psychometric measures as interacting IVs with group, as well as with the illness duration and severity (in the AN and AN-WR groups, separately) as covariates. However, none of these variables significantly influenced the main findings (see Appendix A, Tables A.2, and A.3, respectively). Finally, in the AN group exploratory analyses were conducted to explore the relationship between their Prospective Performance Estimates and Performance and their clinical insight scores (calculated as the difference between the clinician- and patient-ratings on each item, Table 2.2). The output is available in the Appendix A, Table A.4 since the results were not significant and this analysis was not part of a main hypothesis in the present experiment.

All data were analysed performed using R and figures were generated using ggplot2 (Wickham, 2016). This Chapter has also been preprinted and is available here: <https://osf.io/preprints/psyarxiv/rntsf>.

2.3 Results

Women with AN, but not AN-WR women, significantly underestimated their performance abilities prior to completing the HCT.

Results from the linear regression showed that group was a significant predictor of Prospective Performance Estimates, with the Prospective Performance Estimates of the AN group being significantly lower than those of the HCs (Figure 2.1; Table 2.3). That is, women with AN anticipated that they would perform worse in the HCT. The difference in Prospective Performance Estimates was small and not statistically significant between the AN-WR and HC groups.

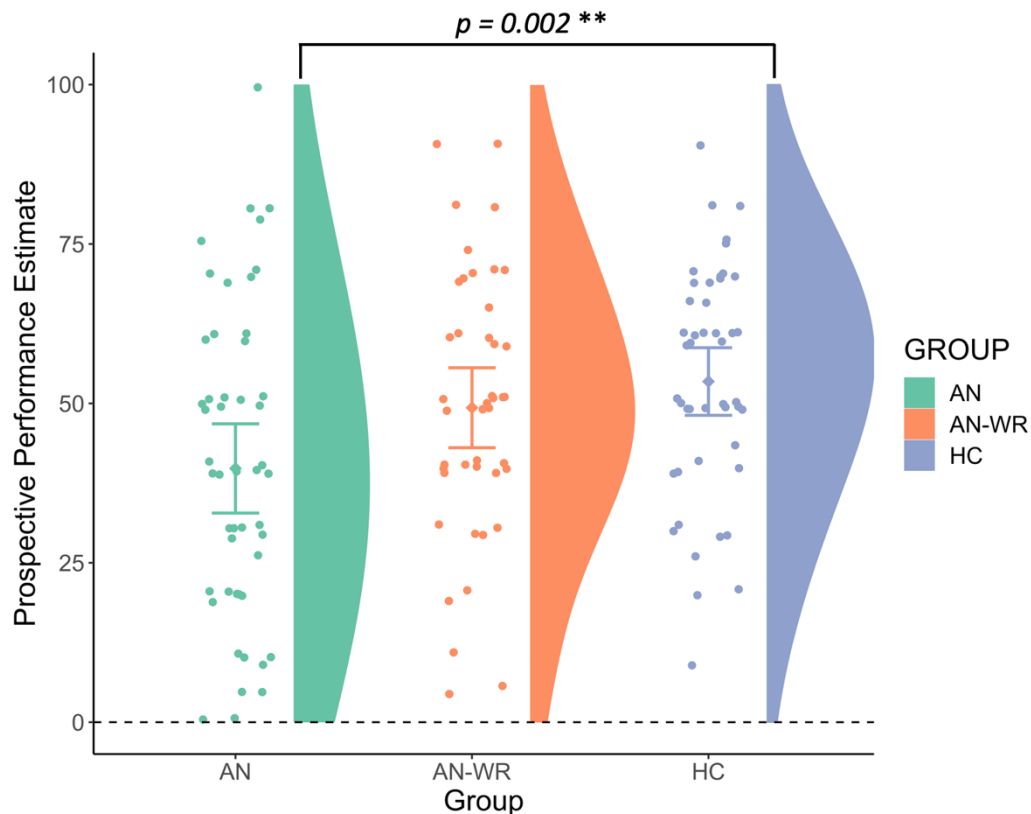


Figure 2.1. Prospective Performance Estimates of the AN, AN-WR and HC Groups.

HCT performance did not differ significantly between women with AN, AN-WR and healthy women.

The three groups did not differ significantly in their actual HCT Performance (or, IAcc; Figure 2.2; Table 2.3). That is, all groups were able to discern felt heartbeats at a similar level of accuracy, with only very small and non-significant differences between groups. BMI was not added as an IV in the linear regression given that it was a criterion for the three groups, but three separate regressions were conducted to examine whether BMI affected Performance within each group, since recent criticisms of the heartbeat counting task suggest that BMI may affect sensitivity to heartbeat detection (Richard et al., 2019), and in turn Performance. However, none of the three within-group analyses yielded significant results, suggesting that in the present cohort, Performance was not significantly affected by BMI within each group.

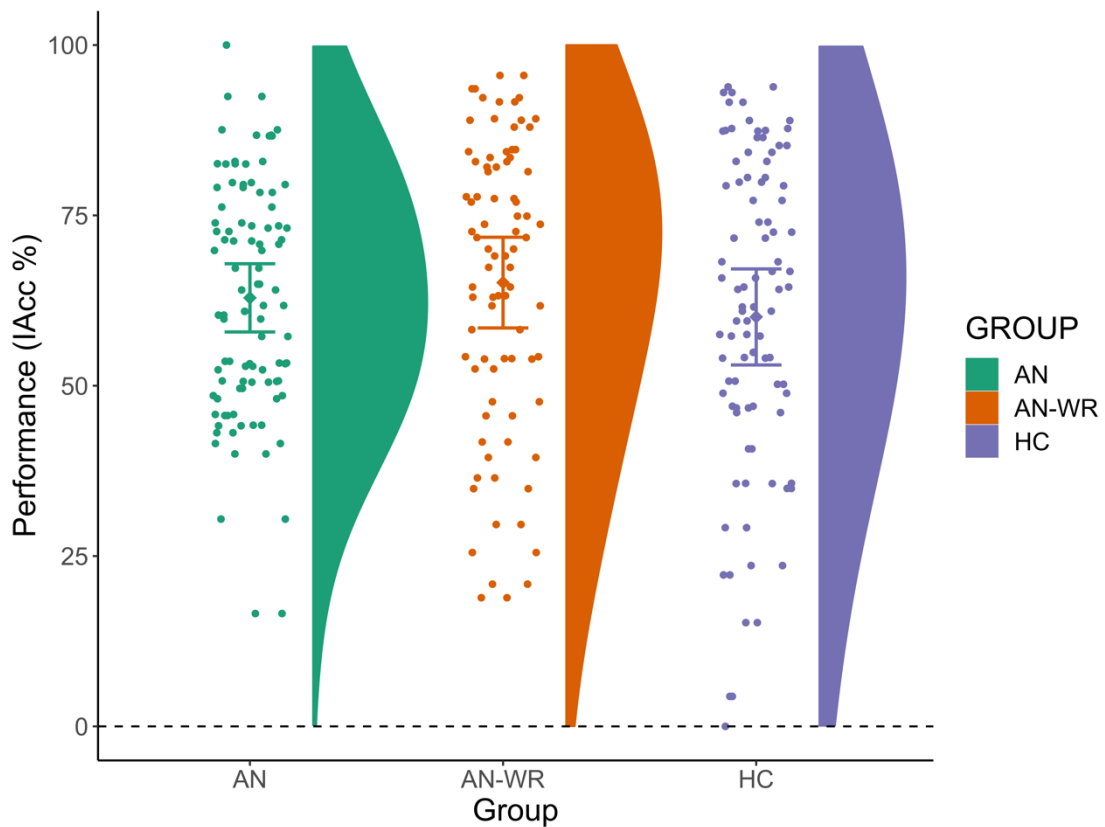


Figure 2.2. Performance (IAcc) scores of the AN, AN-WR, and HC Groups.

Table 2.3. Hypothesis 1 and 2 Results

Predictors	β (SE)	t	95% CI	p
Prospective Performance Estimates Between Groups				
Intercept	53.435(3.14)	16.99	47.22 – 59.65	<0.001
AN	-13.635(4.36)	-3.23	-22.25 - -5.02	0.002
AN-WR	-4.12(4.50)	-0.92	-13.01 – 4.78	0.362
R^2 / Adj R^2 : 0.07 / 0.06				
Performance Between Groups				
Intercept	60.10(3.05)	19.70	54.06 – 66.13	<0.001
AN	2.98(4.25)	0.70	-5.42 – 11.38	0.484
AN-WR	5.18(4.41)	1.17	-3.55 – 13.90	0.243
R^2 / Adj R^2 : 0.01 / -0.004				
Control Analyses for BMI effects on Performance (within each Group separately)				
Intercept ¹	83.20(26.21)	3.18	30.45 – 135.96	0.002
BMI	-1.24(1.65)	-0.75	-4.56 – 2.08	0.456
R^2 / Adj R^2 : 0.012 / -0.009; ¹ in AN Group				
Intercept ²	61.95(27.12)	2.29	7.19 – 116.72	0.028

BMI	0.17(1.36)	0.12	-2.58 – 2.91	0.903
R^2 / Adj R^2 : 0.00 / -0.024; ² in the AN-WR Group				
Intercept ³	33.42(24.32)	1.37	-15.57 – 82.40	0.176
BMI	1.24(1.12)	1.11	-1.01 – 3.49	0.274
R^2 / Adj R^2 : 0.027 / 0.005; ³ in the HC Group				
Abbreviations: AN (Acute Anorexia Nervosa Group); AN-WR (Group Weight-Restored from Anorexia Nervosa); HC (Healthy Controls Group). Bolded values denote statistical significance ($p < 0.05$).				

Women with AN, but not AN-WR women, gave lower retrospective ratings on their HCT performance compared HCs.

The effect of group on Retrospective Performance Estimates when controlling for Prospective Performance Estimates and Performance was examined using a linear regression. The AN group gave lower Retrospective Performance Estimates compared to HCs (Figure 2.3; Table 2.4). The difference between the Retrospective Performance Estimates of the AN-WR group in comparison to the HCs was small and not statistically significant.

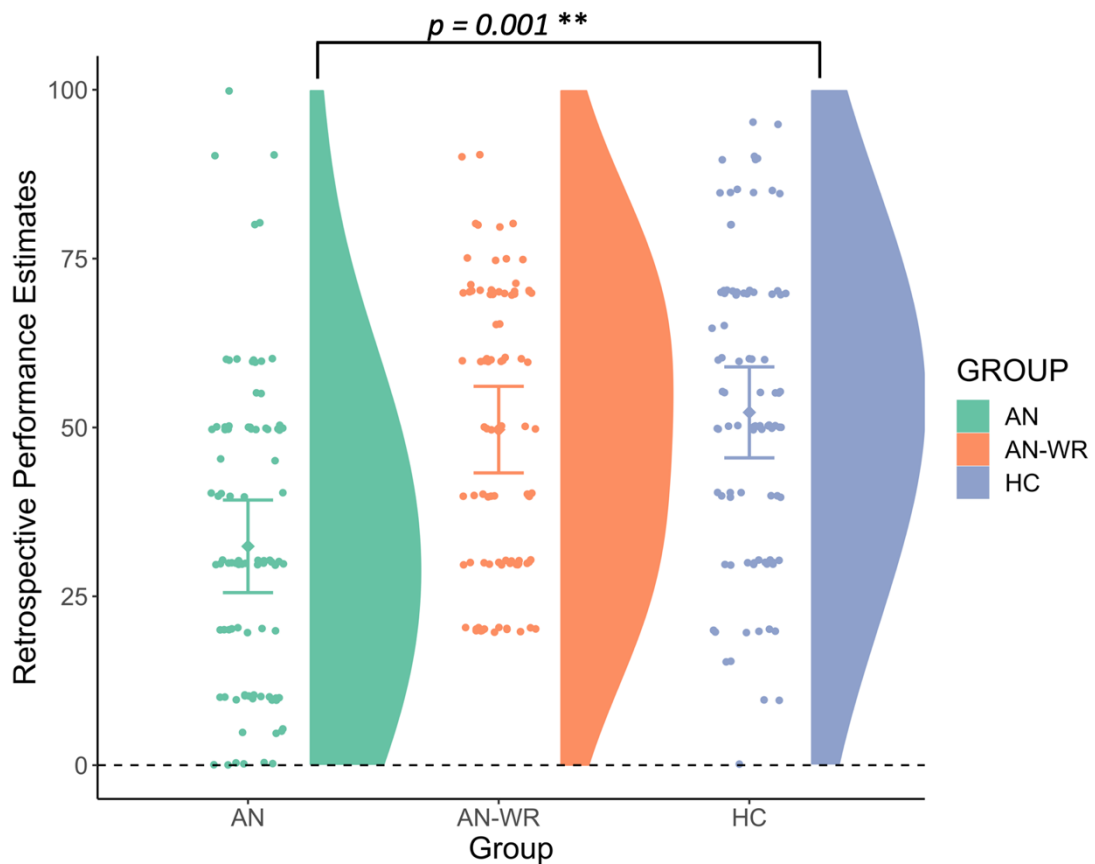


Figure 2.3. Retrospective Performance Estimates between the AN, AN-WR and HC Groups.

Women with AN, but not AN-WR women, were more influenced by their prospective performance beliefs than HCs.

The effect of Prediction Error per group on Retrospective Performance Estimates was examined using a linear regression. The AN group gave significantly lower Retrospective Performance Estimates than the HC group, and a small, non-significant difference was found between the Retrospective Performance Estimates of the AN-WR group and the HCs. A significant Group x Prediction Error interaction was also observed, which indicated that Retrospective Performance Estimates were also dependent on different levels of Prediction Error between groups (Table 2.4). Specifically, the interaction was influenced by the AN group's higher prediction error (they predicted to perform worse than they actually did). However, the relative difference between the AN-WR and HC groups was not significant.

Table 2.4. Hypothesis 3 and 4 Results

Predictors	β (SE)	<i>t</i>	95% CI	<i>p</i>
Retrospective Performance Estimates between Groups				
Intercept	3.41(5.97)	0.57	-8.40 – 15.21	0.569
Prospective Performance Estimate	0.53(0.07)	7.36	0.39 – 0.68	<0.001
Performance	0.33(0.07)	4.51	0.19 – 0.48	<0.001
AN	-13.36(3.80)	-3.52	-20.87 – -5.84	0.001
AN-WR	-2.44(3.80)	-0.64	-9.97 – 5.08	0.522
R^2 / Adj R^2 : 0.48 / 0.46				
Effect of Prediction Error on Retrospective Performance Estimates (between Groups)				
Intercept	51.51(3.34)	15.41	44.90 -58.12	<0.001
AN	-10.56 (5.30)	-1.99	-21.04 - -0.07	0.048
AN-WR	-2.90(5.21)	-0.56	-13.19 – 7.40	0.579
PE	0.055(0.12)	0.46	-0.18 - -0.09	0.645
AN x PE	-0.419(0.17)	-2.54	-0.74 - -0.09	0.012
AN-WR x PE	-0.03(0.18)	-0.18	-0.38 – 0.32	0.855
R^2 / Adj R^2 : 0.019 / 0.159				
N.B. For Hypothesis 3 (Group differences on Retrospective Performance Estimates) Prospective Performance Estimates and Performance were added as covariates. Abbreviations: AN (Acute Anorexia Nervosa Group); AN-WR (Group Weight-Restored from Anorexia Nervosa); HC				

(Healthy Controls Group); PE (Prediction Error). Bolded values denote statistical significance ($p < 0.05$).

False feedback resulted in a change in performance evaluation estimates in all groups. Specifically, positive false feedback resulted in higher performance evaluation estimates and Negative false feedback resulted in lower performance evaluation estimates.

First, three linear regressions were run (one per group), to control for potential effects of false feedback on performance. As expected, no statistically significant effect was found, i.e., the false feedback that participants received was not based on their performance in the HCT (Table 2.5). In the main analyses that were conducted to examine whether false feedback affected participants' post-false feedback performance estimates, the same pattern of results was observed across the three groups. That is, participants within each group gave significantly higher Post False-feedback Retrospective Performance Estimates if they received positive false feedback, in comparison to those who received negative false feedback (Figure 2.4; Table 2.5).

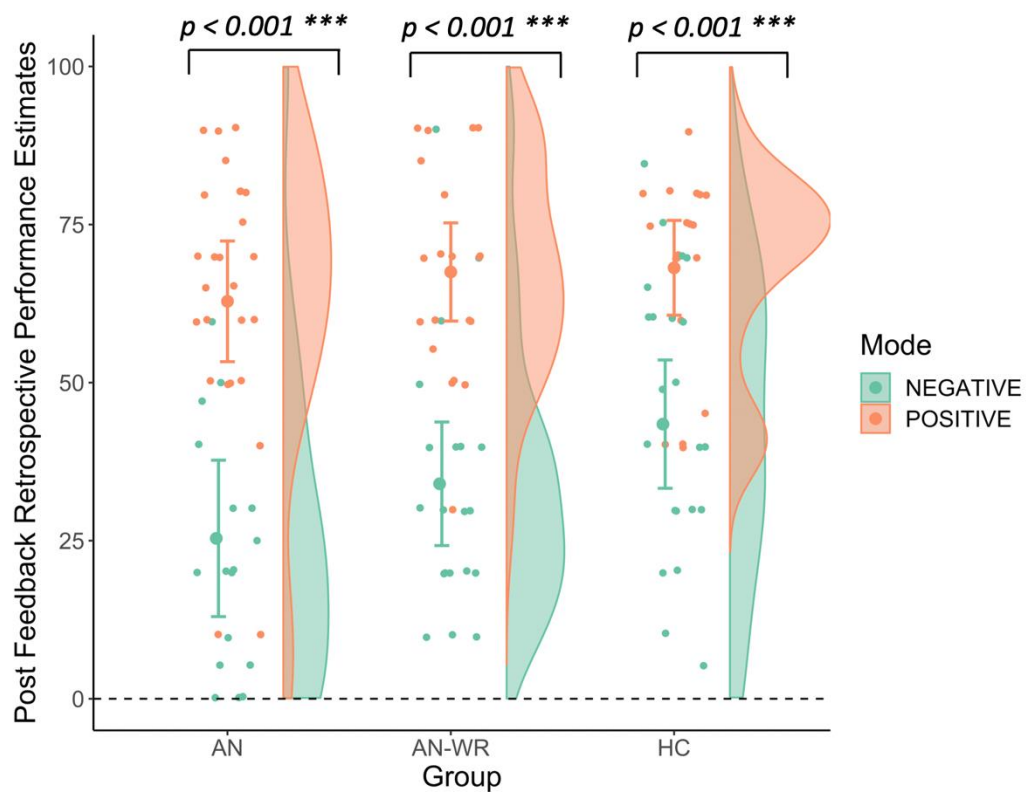


Figure 2.4. Post Feedback Retrospective Performance Estimates Following False Feedback

Table 2.5. Hypothesis 5 Results

Predictors	β (SE)	<i>t</i>	95% CI	<i>p</i>
Control Analysis for Effect of False Feedback on Performance				

Intercept ¹	61.32(3.95)	15.52	53.36 – 69.28	<0.001
Positive False Feedback	1.23(5.16)	0.24	-9.17 – 11.62	0.813
R ² / Adj R ² : 0.001 / -0.021; ¹ in the AN Group				
Intercept ²	61.14(4.80)	12.73	51.41 – 70.87	<0.001
Positive False Feedback	4.94(6.88)	0.72	-9.00 – 18.88	0.477
R ² / Adj R ² : 0.014 / -0.013; ² in the AN-WR Group				
Intercept ³	61.45(5.03)	12.21	51.28 – 71.62	<0.001
Positive False Feedback	-4.46(7.48)	-0.60	-61.45 – 34.62	0.554
R ² / Adj R ² : 0.009 / -0.016; ³ in the HC Group				
Effect of False Feedback on Post False-Feedback Retrospective Performance Estimate				
Intercept ¹	25.37(5.75)	4.42	13.80 – 36.94	<0.001
Positive False-Feedback	37.49(7.44)	5.04	22.50 – 52.48	<0.001
R ² / Adj R ² : 0.36 / 0.35; ¹ in the AN Group				
Intercept ²	34(4.22)	8.01	25.47 – 42.53	<0.001
Positive False-Feedback	33.50(5.96)	5.62	21.43 – 45.57	<0.001
R ² / Adj R ² : 0.45 / 0.44; ² in the AN-WR Group				
Intercept ³	43.44(4.23)	10.27	34.88 – 51.99	<0.001
Positive False-Feedback	24.72(6.29)	3.93	12.01 – 37.44	<0.001
R ² / Adj R ² : 0.28 / 0.26; ³ in the HC Group				
Abbreviations: AN (Acute Anorexia Nervosa Group); AN-WR (Group Weight-Restored from Anorexia Nervosa); HC (Healthy Controls Group); PE (Prediction Error). Bolded values denote statistical significance ($p < 0.05$).				

2.4 Discussion

The present experiment investigated the updating of beliefs in women with, and weight-restored from AN; specifically, it looked at the prospective and retrospective interoceptive self-efficacy beliefs individuals hold before and after completing a heartbeat counting (HCT) task (Schandry, 1981), and after receiving different types of information about their heartrate (i.e., embodied signals during HCT performance, and explicit, verbal feedback from the experimenter. Overall, the results showed that the AN group held more pessimistic self-efficacy beliefs (or, Prospective and Retrospective Performance Estimates) in comparison to the HCs, both before and after the HCT. The results and their implications are discussed in more detail below.

In line with the first hypothesis, women with AN, but not AN-WR women, provided lower prospective performance estimates (i.e., pessimistic prospective self-efficacy beliefs about cardiac awareness) in the HCT in comparison to HCs. Specifically, the AN group gave more pessimistic, predictive estimates about their future performance on a HCT that asked them to focus on and report felt heartbeats during certain, different time intervals. These results may

relate to previous findings in eating disorders suggesting differences between patients and control groups in the prediction and anticipation of various interoceptive signals (Crucianelli et al., 2021; Khalsa et al., 2015; see Section 1.3 (Chapter 1) and Section 2.1), although these studies tested only prospective beliefs at the perceptual level (e.g., the equivalent of “*How many heartbeats do you think you would perceive*”) and not at the metacognitive level tested here (i.e., “*How well do you think you will do in this task*”). Interestingly, in clinical studies, prospective self-efficacy at treatment onset has been found to be predictive of eating disorder treatment dropout (Keshen et al., 2017). To my knowledge, however, no studies had previously examined interoceptive self-efficacy prospectively and the present results warrant further investigation of the role of such beliefs in treatment response.

In the present sample, only those with acute AN and not the AN-WR group presented more pessimistic beliefs than HCs. This result may be explained by the relatively small sample, or factors relating to recovery (e.g., some weight-restored patients may have improved their self-efficacy more generally), or illness duration. Indeed, acute AN patients typically have a longer illness duration than those in the AN-WR group (e.g., Kinnaird et al., 2020; Konstantakopoulos, Georgantopoulos, et al., 2020; Konstantakopoulos, Ioannidi, et al., 2020; Miles et al., 2020). This pattern was also observed in the present sample. However, it is unlikely that the pessimistic prospective self-efficacy beliefs of the AN group were explained by illness duration given that the within-group analyses with illness duration as an IV did not yield significant effects (see Appendix A, Table A.3). Finally, this result may suggest a state, rather than trait effect. In fact, prolonged starvation can lead to severe modifications of brain and peripheral-organ function which in turn impact aspects of cognitive abilities (e.g., Abbate-Daga et al., 2011; Kaye et al., 2009; Miles et al., 2020). Thus, the lack of a significant difference between the AN-WR and HC group suggests that the significantly more pessimistic prospective self-efficacy estimates of the AN group as opposed to those of the HC group may have been a secondary effect of starvation, and not a pre-existing trait which persisted after weight-restoration. Future studies should explore these different possibilities further.

As predicted, HCT Performance did not differ between the three groups. These findings are consistent with previous studies reporting no differences in interoceptive accuracy between clinical and HC groups (e.g., Kinnaird et al., 2020; Lutz et al., 2019; Richard et al., 2019), although contrary findings also exist (e.g., Pollatos et al., 2008). These differences between the studies may be at least partly explained by sampling differences (e.g., comorbidities, inpatient

or outpatient treatment, illness stage; Fischer et al., 2016; Pollatos et al., 2008; Richard et al., 2019). In the present AN group I recruited individuals who met DSM-5 restrictive AN criteria (American Psychiatric Association, 2013) and were at different treatment stages, receiving either inpatient or outpatient care, and had various comorbidities unlike Pollatos and colleagues (2008) who recruited outpatients with either a restrictive or binge-purge AN diagnosis and with fewer comorbidities. Notwithstanding, heterogeneity in treatment stage and type, BMI, and clinical profiles did not affect the analysis outcomes, suggesting that in the present cohort these factors were unlikely to have influenced participant HCT performance. All of the groups performed above chance – something observed in most studies with similar methodologies (e.g., Kinnaird et al., 2020; Lutz et al., 2019; Pollatos et al., 2008, 2016). It is worth noting that studies in the general population have sometimes demonstrated poorer performance in HCs (even below chance, e.g., Desmedt et al., 2020), and this variability could be attributed to the instructions, duration of the counting trials, mixed gender samples, or other factors which could not be accounted for in the present experiment. Despite the three groups' scores being close to what has been observed in other studies, the validity and reliability of the HCT task itself has been criticised (Brener & Ring, 2016; Desmedt et al., 2018, 2020, 2022, 2023; Murphy, Millgate, et al., 2018) and thus the present experiment's methods regarding cardiac, perceptual accuracy are less conclusive. Instead, the emphasis is placed on the prospective self-efficacy beliefs patients generate before the HCT, and their retrospective beliefs after the HCT, and then after explicit, verbal feedback.

Indeed, participants' retrospective self-efficacy beliefs regarding cardiac interoceptive abilities were assessed. It was predicted that the retrospective self-efficacy estimates of the AN group would be lower than those of the HC group despite the fact that the two groups did not significantly differ in their performance *per se*; *Hypothesis 3*) and it was further hypothesised that pessimistic retrospective beliefs would be better explained by their similarity to prior prospective self-efficacy beliefs than performance levels (*Hypothesis 4*). In line with the hypotheses, the results showed that (i) the AN but not the AN-WR group were significantly more pessimistic than the HCs in their retrospective self-efficacy estimates, and (ii) as the discrepancy between performance and prior prospective self-efficacy beliefs (Prediction Error) of the AN group (but not of the AN-WR group) increased, the AN group's retrospective self-efficacy estimates were reduced. These results suggest that even though Performance on the task was better than these patients expected, their retrospective beliefs about performance were not updated and instead continued to reflect the initial, pessimistic priors. Although these

results are based on a task with few trials and known limitations, they are consistent with previous findings on impaired retrospective metacognition about interoception in AN, based on paradigms using confidence ratings on interoceptive performance from multiple trials (e.g., Kinnaird et al., 2020; Smith et al., 2020). Interestingly, by sampling also explicit beliefs prior to the task, it was possible to observe these metacognitive difficulties at the level of prospective, self-efficacy beliefs and not only retrospectively. Future studies could explore whether greater reliance on pessimistic, prior beliefs is present in the AN-WR individuals when employing a larger AN-WR sample, or whether the pessimism observed in the current study is explained by other factors relating to the acute stage of illness as outlined above.

Interestingly and as predicted, all groups provided lower estimates following false-negative feedback, and higher estimates following false-positive feedback, a manipulation not previously applied in experimental studies using the HCT. This finding demonstrates that the ability to use feedback to update beliefs about task performance in AN may not be compromised across all levels and self-efficacy beliefs may be modified by social feedback, even if false as in this case. The pessimistic beliefs about one's abilities observed in those with AN are also present in depression, wherein individuals have difficulty adapting their beliefs to novel situations (Stange et al., 2017) despite being able to use arbitrary positive feedback to update task-specific expectations (Kube et al., 2022; Kube & Rozenkrantz, 2021). Yet, the study's sample experienced difficulty generalising such expectations to other, future tasks (global prospective metacognition; Kube et al., 2022). Given clinical time constraints not all participants in the present experiment were assessed on depression (using the DASS-21 questionnaire; Lovibond & Lovibond, 1995) and I thus only directly examined the effects of depression on prospective self-efficacy beliefs and interoceptive accuracy. The results were non-significant (see Appendix A, Table A.2) and although less conclusive given the small sample size, it is suggested that the pessimistic beliefs observed throughout the study were not directly influenced by depressive traits or states (the AN group scored higher than the AN-WR and HC groups on the DASS-21) but may instead be attributed to a bias in metacognitive beliefs or other factors as described above. Future studies will need to test this speculative interpretation. Interestingly, although the AN group showed more pessimistic, retrospective beliefs in comparison to HCs, despite similar performance in the HCT, following explicit, albeit false, performance feedback, they were able to adjust their estimates towards the direction of the feedback received. Thus, in line with Kube's et al. (2022) findings, the present results showed that in certain circumstances people with AN are able to use external, (dis)confirming

information about their performance and update an isolated estimate. It is also likely that participants from all groups, especially those in the AN group were able to use false feedback to update their beliefs because of low confidence in their own interoceptive abilities. Future studies should further investigate the role of explicit, external feedback in individual with both high and low interoceptive self-efficacy beliefs.

Despite the insights this experiment provided on the evaluation of abilities regarding perception of interoceptive signals and related metacognitive beliefs across groups, there were several limitations which will be addressed in the follow-up, preregistered study (see Chapter 3). The HCT (Schandry, 1981) has received several criticisms for being a poor measure of interoceptive accuracy as approximately 40% of typically developing adults are not consciously aware of their heartbeat (Khalsa et al., 2009), making the task less sensitive when examining interoception in individuals with lower ranges of ability (Murphy, Millgate, et al., 2018). Additionally, heartbeats may be perceived via vibrations against the chest wall (Khalsa et al., 2009) and depend on factors such as body-fat percentage (Rouse et al., 1988), and resting heart-rate variability (HRV; Knapp-Kline & Kline, 2005). Some questioning the validity and reliability of the HCT (Brener & Ring, 2016; Desmedt et al., 2018) also suggested that participants may report their felt heartbeats based on an approximation of the time they think has passed, general knowledge about average heartbeats in the general population and a calculation of the number of heartbeats they should have felt based on time awareness and heartbeat knowledge (e.g., Ainley et al., 2014; Murphy, Geary, et al., 2018; Shah et al., 2016). Though, Murphy and colleagues suggested that the criticisms may be less valid when controlling for these aforementioned measures (Murphy, Millgate, et al., 2018)— measures which will be controlled for in the study presented in Chapter 3. Finally, interoception has been associated with symptoms of difficulty in emotional awareness and emotion identification and description (often referred to as alexithymia), given findings on poor interoception in individuals with higher alexithymia scores (Herbert et al., 2011; Murphy et al., 2017). However, this was not the case in the present sample as seen in the analyses available in Appendix A, A.2, suggesting that in this sample the differences in alexithymia scores did not explain the significant differences in prospective self-efficacy beliefs or interoceptive accuracy.

2.4.1 Conclusion

To conclude, in this first exploratory study, simple instances of explicit belief updating in the cardiac awareness domain in AN were investigated. The results showed that AN patients hold pessimistic interoceptive, metacognitive self-efficacy beliefs even after a relatively successful, heartbeat counting performance, and mainly rely on such pessimistic, prospective self-efficacy beliefs, even when rating performance retrospectively. Although the results suggest that these findings may be linked to the acute illness stage of AN and hence may be secondary to starvation and malnutrition rather than enduring traits, the results offer preliminary insights regarding belief updating difficulties in AN that warrant further study, in the interoceptive domains and in relation to insight deficits that are also relevant only in the acute state of illness. The findings also highlight the importance of studying belief updating processes across different domains in AN to better understand how deficits in one or more domains may have manifested during the anorectic state or have been traits that lead to the development of symptoms and/or endured post-weight restoration and subsequently influenced engagement in and outcomes of treatment interventions. Future studies should further disentangle the roles of belief updating across and within different domains of relevance to AN.

Chapter 3: Updating Prospective Self-Efficacy Beliefs About Cardiac Interoception in Anorexia Nervosa: An Experimental and Computational Study

3.1 Introduction

Given the findings of the experiment presented in Chapter 2, which suggest that patients with AN have low self-efficacy beliefs about their interoceptive abilities before they even engage with an interoceptive task (prospectively), and they do not seem to update their self-efficacy beliefs after trying out the task (retrospectively), but instead somehow adhere to their prospective, self-efficacy beliefs, a second experiment, presented in this Chapter was conducted. Specifically, the cardiac interoception task that was used in Chapter 2 was developed further to allow the experimental and computational investigation of how patients with and weight-restored from AN (AN and AN-WR, respectively), use nested local, and global retrospective beliefs about performance on a cardiac interoception task to update their explicit, prospective beliefs about their related abilities.

As previously discussed (see Section 1.2, Chapter 1), Anorexia Nervosa (AN) is an eating disorder characterised by severe restriction of caloric intake, intense fears, and preoccupations regarding gaining weight and a disturbed appreciation of current low body weight, or its health consequences (American Psychiatric Association, 2013). The aetiology of AN is multifaceted and poorly understood. Neurobiological models of eating disorders commonly emphasize perturbed appetitive, emotional or control neural circuits (Fassino et al., 2004; Friederich & Herzog, 2010; Kaye et al., 2009, 2013; Steinglass et al., 2019). More recently however accumulating behavioural and neurobiological evidence points to the central role of interoceptive abnormalities in the onset and maintenance of eating disorder symptomatology (Jenkinson et al., 2018; Khalsa et al., 2022; Martin et al., 2019). The term interoception refers to the process of the sensing, integrating and interpreting internal, physiological bodily signals arising from major visceral systems such as the gastrointestinal, respiratory, and cardiovascular systems (Khalsa, Adolphs, et al., 2018; see more in Section 1.3, Chapter 1). In some definitions interoception is also used to describe certain cutaneous sensations such as affective touch, pain and temperature (Craig, 2003; Crucianelli & Ehrsson, 2023; von Mohr & Fotopoulou, 2018). Interoception has not only been associated with physiological states but also with the awareness and regulation of cognitive and emotional states at a behavioural (Füstös et al., 2013; Schandry, 1981) and neural level (Craig, 2003; Pollatos et al., 2007). Disruptions of interoceptive

processing and damage, or functional alterations to a related neural network centred around the mid-dorsal insula have been proposed as a ‘transdiagnostic’ mechanism conferring a common vulnerability across multiple mental health disorders (Nord et al., 2021) and particularly those at the interface between physical and mental health, such as eating disorders (Berner et al., 2019; Brewer et al., 2021; Frank et al., 2013; Gordon et al., 2001; Kerr et al., 2016; Khalsa et al., 2018; Kojima et al., 2005; Nunn et al., 2008, 2011; Uher et al., 2003, 2004; Wagner et al., 2008; Zucker et al., 2017). Indeed, interoceptive deficits have been implicated in the aetiology and maintenance of eating disorders since their initial conceptualisation (Bruch, 1962), with patients presenting with difficulties to perceive and interpret signals from their body. Indeed, patients may present with difficulties in their ability to sense a depleted nutritional body state (for recent reviews see: Khalsa et al., 2022; Maniscalco & Rinaman, 2018; Martin et al., 2019) but also extending to non-gastrointestinal interoceptive domains, such as in cutaneous, cardiac or respiratory modalities (e.g., Berner et al., 2018; Crucianelli et al., 2021; Kerr et al., 2016; Lapidus et al., 2020; Pollatos et al., 2008, 2016). Moreover, contemporary network analyses, which are increasingly utilised to conceptualise and assess the importance of symptoms and interactions in the maintenance of psychological disorders (Borsboom, 2017; McNally, 2016), consistently identify dysfunctional interoception as a core component of eating disorder symptomatology (Brown et al., 2020; Cascino et al., 2019; Olatunji et al., 2018; Solmi et al., 2018; Vervaet et al., 2021). Crucially, the larger such interoceptive deficits are, the worse the treatment outcome in eating disorders (Bizeul et al., 2001; Carter et al., 2004). Finally, a small but fast-growing number of interoceptive-based interventions, including behavioural (for systematic review, see Heim et al., 2022) and neurostimulation interventions (Baker, 2020; Dalton et al., 2018), appear to be effective at treating key symptoms in eating disorders.

Studies in eating disorders often use self-report measures to target interoception, but recently some have also investigated patient accuracy in perceiving interoceptive signals during behavioural tasks (e.g., Crucianelli et al., 2016, 2021; Khalsa et al., 2015; Lutz et al., 2019; the latter also being referred to as ‘interoceptive accuracy, or sensitivity’; e.g., Garfinkel et al., 2015; Khalsa et al., 2018). For example, studies have found reduced cardiac (e.g., Pollatos et al., 2008, 2016) or affective touch signal sensitivity (Bischoff-Grethe et al., 2018; Crucianelli et al., 2016, 2021) in clinical groups in comparison to healthy samples and related differences in neural networks centred on the insula (e.g., Hassanpour et al., 2018; Kerr et al., 2016; Nunn et al., 2011). Interestingly, a few studies have also measured subjective beliefs on one’s performance and shown group differences in the evaluation and appraisal of interoceptive

signals metacognitively. For instance, patients may show lower confidence than controls when rating their ability to count their own heartbeats, despite not performing at lower levels (Khalsa et al., 2015; Kinnaird et al., 2020; see also Hoogervorst et al., 2023; Smith et al., 2020). These higher order levels of interoceptive processing, referred to as interoceptive metacognition or insight (see Critchley & Garfinkel, 2017; Garfinkel et al., 2015; Khalsa et al., 2018) may be of particular relevance to AN, given the well-recognised role of abnormal metacognitive beliefs in the onset, maintenance and treatment of AN (for systematic reviews see: Palmieri et al., 2021; Sun et al., 2017).

Unfortunately, these metacognitive deficits have been studied by different research traditions, focusing on different dimensions and methods (for more details see Section 1.4, Chapter 1). Specifically, while experimental research on metacognition typically focuses on ‘local’ (e.g., trial-by-trial confidence ratings) and ‘retrospective’ (e.g., confidence-accuracy correspondence on a preceding trial) measures and levels of metacognition (e.g., Fleming & Lau, 2014; Garfinkel et al., 2015; Rouault, McWilliams, et al., 2018), clinical traditions typically focus on global, retrospective and prospective metacognitive beliefs. These beliefs include explicit beliefs about one’s cognition and related (dis)abilities, such as beliefs about one’s cognition, behaviour or about one’s illness, or its consequences (termed clinical insight; David, 2004), and beliefs about one’s capacity to succeed in, or cope with different tasks, situations and contexts (termed self-efficacy; Bandura, 1977). Such metacognitive beliefs of patients with eating disorders and particularly patients with AN have been found to differ from those of the general population. For example, clinical psychology studies, using mostly questionnaires on explicit beliefs, have found consistent differences between AN samples and the general population in metacognitive beliefs, such as higher levels of negative beliefs about worry (i.e., beliefs that thoughts are uncontrollable and dangerous; Cooper et al., 2007; Davenport et al., 2015; McDermott & Rushford, 2011; Palomba et al., 2017) and higher positive beliefs about worry (e.g., beliefs that perseverative thinking is useful; Davenport et al., 2015; McDermott & Rushford, 2011). The latter findings are consistent with the egosyntonic nature of AN, i.e., the fact that patients with AN may regard their restrictive eating as adaptive to their psychological needs (Gregertsen et al., 2017; Serpell et al., 1999). Self-efficacy has been examined in relation to beliefs about one’s ability to regulate eating, other bodily domains and cognition, with questionnaire studies finding a relation between ‘eating self-efficacy’ and subclinical (e.g. Berman, 2006; Linardon et al., 2017) as well as clinical eating disorders (Goodrick et al., 1999; O’Leary, 1985). Finally, insight deficits (lacking awareness into one’s symptoms) have been

found to be a maintaining factor in AN, hindering treatment adherence and recovery (Abbate-Daga et al., 2013; Konstantakopoulos et al., 2012). Insight deficits are more frequently observed in restrictive AN than in other eating disorders (Greenfield et al., 1991; Konstantakopoulos et al., 2011) and are irreducible to other clinical, cognitive or emotional symptoms in AN (Konstantakopoulos et al., 2012). Instead, such deficits appear to have a specific, metacognitive basis (Arbel et al., 2013).

More recently, correlational, as well as network analysis studies have shown that metacognitive beliefs may also determine the association between interoceptive awareness and eating symptomatology. For example, ‘body mistrust’ and ‘fear about the safety of bodily signals’ has been linked to the association between interoceptive awareness and eating disorders (Olatunji et al., 2018). Thus, although it has become clear that deficits in interoception, self-awareness and metacognition, and their complex interrelations, are frequently associated with eating disorders and particularly AN (Cooper et al., 2021; Jenkinson et al., 2018; Khalsa et al., 2022), the separate and narrow study of these dimensions by different fields has not allowed the integration of theoretical insights generated by studies using neurophysiological and psychophysical paradigms with the theoretical insights generated by studies focusing on explicit, clinically-relevant beliefs. Moreover, while the importance of global (mostly retrospective) metacognition in mental health is getting some recognition among experimental traditions (Seow et al., 2021), such insights have not been applied to eating disorders research, or interoception research (see Stephan et al., 2016 for a first theoretical proposal and model). Finally, explicit, prospective beliefs about one’s interoceptive capacity are rarely sampled by current paradigms, despite the fact that exteroceptive, prospective metacognition has been shown to dissociate from retrospective metacognition (Fleming et al., 2016). Addressing these important gaps in the literature is the central aim of the interdisciplinary study presented in this Chapter.

Specifically, in the present study, the aim was to use a unifying Bayesian, computational approach (Friston, 2010; Petzschner et al., 2017; Smith et al., 2020; Smith, Mayeli, et al., 2021; Stephan et al., 2016) that allows to study under one continuous framework how interoceptive perception and local metacognition influence the updating of explicit global, prospective beliefs about interoception. This unified computational approach was recently used in a different neuropsychiatric context to characterise the continuity between perceptual and metacognitive beliefs (Kirsch et al., 2021), and to highlight the importance of ‘counterfactual’

beliefs (inferences about the causes of sensory inputs of predicted but not as-yet-experienced states; see Fotopoulou, 2015; Kirsch et al., 2021; Palmer et al., 2015) in self-awareness. Specifically, the researchers found that disruptions in the updating of explicit, prospective metacognitive beliefs, such as sensorimotor self-efficacy beliefs after stroke (e.g., “*Could I do this action tomorrow, or at home?*”), can be attributed to differences in the motivational salience of retrospective, perceptual beliefs (“*How much does it matter to me whether I did this exercise correctly?*”), rather than the accuracy of the perceptual beliefs themselves (“*Did I do this exercise correctly?*”). In the present experiment, this unifying Bayesian approach was applied to explicit, prospective, and retrospective capability beliefs in the domain of interoceptive inference. Specifically, this experiment assessed how information from one’s actual performance on a cardiac signal perception task (heartbeat counting) and related, local and global retrospective beliefs about this performance (the level of accuracy-confidence correspondence typically studied in experimental, perceptual metacognition studies) are combined and used to inform the updating of explicit, prospective metacognitive beliefs (the level of beliefs typically studied in clinical, self-efficacy or ‘worry’ studies) about one’s ability to monitor one’s own cardiac signals in AN. These investigations stand to inform our understanding of AN not only as regards the perception and evaluation of signals from the body in the here-and-now experience of the patient (“*How well did I feel my internal sensations?*”), but also as regards the processes that allow patients to use such ‘local’ perception and evaluation to update their ‘global’ beliefs about their interoceptive ability (“*How well do I perceive my internal sensations in general?*”).

The wider Bayesian framework used here has already been extensively applied in computational psychiatry studies to characterise neurocomputational abnormalities in exteroceptive perceptual and metacognitive Bayesian inference (a method of statistical inference in which new observations are used to update or infer the probability that a hypothesis was true) in many other psychopathologies (for interoceptive examples see: Fogelson et al., 2014; Gu, Eilam-Stock, et al., 2015; Gu, Lohrenz, et al., 2015; Owens, Allen, et al., 2018; Owens, Friston, et al., 2018; Petzschnner et al., 2021; Stephan et al., 2016; for non-interoceptive examples see: Adams, 2018; Adams et al., 2016; Fletcher & Frith, 2009; Huys et al., 2016, 2021; Stephan et al., 2017; Sterzer et al., 2018). The framework relies on the premise that afferent sensory input to the brain is shaped by previously learned expectations, and vice versa, at multiple levels of the neurocognitive hierarchy, from unconscious processing levels to explicit, conscious and metacognitive beliefs. Feedback and feedforward loops are thought to

constantly update and refine one's expectations of sensory or cognitive input in order to predict the causes and contexts of future states. Thus, the symptoms of multiple psychiatric conditions could be explained by faulty perceptual or metacognitive inference, due to an inappropriate weighting of prior beliefs and new sensory, or cognitive evidence, particularly regarding interoception (Barrett & Simmons, 2015; Owens, Allen, et al., 2018; Paulus et al., 2019; Petzschner et al., 2017; Smith et al., 2017; Stephan et al., 2016). Indeed, under this framework, interoception itself has been understood as a bidirectional process between the brain and the body, allowing a generative model aimed at predicting and regulating future states of the physiological condition of body (Allen & Friston, 2018; Barrett & Simmons, 2015; Owens, Allen, et al., 2018; Petzschner et al., 2017, 2021; Pezzulo et al., 2015; Seth, 2013). Thus, several theories have proposed that a key transdiagnostic factor in psychopathology may be an inability of the brain to update its body model based on interoceptive prediction errors (i.e., mismatches between expected and received afferent interoceptive signals from the body). Under this framework, aberrant belief updating is thought to originate through dysfunctions in various parameters such as either a dysfunctional "precision" mechanism (the mathematical equivalent of weighting the relevance or reliability of sensory signals), which governs how afferent bodily signals influence beliefs, or individual differences in how quickly one updates their prior beliefs over repeated observations based on how individuals optimise the relative precisions of sensory signals and prior beliefs across observations. Thus, the application of formal computational models in psychiatry can be used to distinguish which parameters show differences between individuals and best explain differences in belief updating, including explicit beliefs about one's body and self.

Interestingly, a handful of theoretical (Barca & Pezzulo, 2020) and formal (Smith, Mayeli, et al., 2021) approaches in eating disorders have used a similar Bayesian framework as the background for characterising disruptions in interoceptive Bayesian inference at the perceptual domain. A recent study fitted a formal computational model of interoception to behavioural responses during experimental modulations of interoception in individuals with various psychiatric disorders, including a small ($N = 14$) eating disorders sample and found precision weighting differences between clinical groups and healthy controls (Smith, Mayeli, et al., 2021; see Lavalley et al., 2023 for a recent replication study). Notwithstanding the value of this approach, most existing studies in computational psychiatry use binary perceptual tasks, requiring subjects to detect the absence or presence of a signal or stimulus, or to make a binary decision between two options. Some tasks also use explicit, retrospective confidence ratings

after such perceptual tasks. However, to my knowledge this approach has not been applied to explicit and prospective, or counterfactual metacognitive beliefs, such as those typically identified as aberrant in clinical psychology studies of eating disorders (see above). Thus, a simplified interoceptive belief updating task and a corresponding Bayesian modelling approach that can examine the key parameters involved in the updating of such explicit, prospective beliefs about one's interoceptive capacity in the cardiac domain were developed. When such metacognitive beliefs need to be updated, various sources of evidence, and corresponding precision and learning rate parameters are at stake, involving not only sensory signals and related beliefs but also cognitive beliefs about the underlying sensory beliefs and their precision.

To this end, the experiment in Chapter 3 investigated how women with AN at the acute stage of illness, women weight-restored from AN (AN-WR) and a group of healthy controls (HC) update their prospective (self-efficacy) beliefs about their ability to detect their heartbeats after engaging in a modified heartbeat counting task (HCT; Schandry, 1981). Comparisons across these three groups allowed to disentangle state (i.e., present only during the acute AN phase as secondary neurocognitive consequences to prolonged malnutrition) from trait mechanisms (i.e., pre-morbid deficits present in at-risk individuals, or deficits that endure beyond the acute phase and are present during remission; Jenkinson et al., 2018; Kaye et al., 2009; Saramandi et al., 2022). The heartbeat counting task (Schandry, 1981) has known methodological limitations (e.g., Brener & Ring, 2016; Desmedt et al., 2018; Murphy et al., 2018; Richard et al., 2019) but it was chosen at the time (before many of its limitations came to light) as it is feasible to do in an acute hospital setting and the requirement to form beliefs about one's performance can be intuitively and meaningfully understood by patients (e.g., "*How accurately will you be able to count your own heartbeats?*"). These 'counterfactual' prospective beliefs and their updating (*how well you think you will do in the future?*) relatively to one's 'perceived' interoceptive accuracy (*how well you think you did?*), rather than the precise measurement of interoceptive accuracy *per se*, are the focus of the present Chapter. Nevertheless, given these task limitations, after the first, exploratory experiment (see Chapter 2), the preregistered experiment of Chapter 3 included additional control measures, including time-estimation abilities and general knowledge about heart rate, to account for their potential role in participants' interoceptive accuracy performance. Hence these factors were taken into account in the preregistered statistical analyses of performance and key explicit beliefs between the three groups.

Furthermore, to understand how individuals with, or weight-restored from AN in comparison to HCs update their prospective beliefs about their ability to detect their heartbeats, a simplified version of the aforementioned Bayesian Learning Framework was adopted (Friston, 2003; Mathys et al., 2011), adjusted to take into account the nested nature of prospective and retrospective metacognitive beliefs about sensory performance (Kirsch et al., 2021). Specifically, the updating of explicit metacognitive beliefs can be influenced by at least two key sources of ‘evidence’: first, the perceptual performance itself (e.g., how accurate one was) and second, one’s global retrospective beliefs about such performance (e.g., how accurate one thought they were after the task ends). Similarly, there can be different sources of ‘evidence precision’ and here two experimental measures were considered as ‘proxies’ for such precision—namely, how confident individuals were about their performance during the task and how individuals rated their interoceptive abilities in everyday life as sampled by a standardised questionnaire. First, four models were built under the Bayesian Learning Framework in order to identify which combination of proxy measures of ‘evidence’ and of ‘precision of evidence’ would lead to computations of posterior prospective beliefs (*how well you would do in the future?*) that best approximate the actual posterior beliefs of the participants. The predictive power of these models (i.e., how close were the ‘posterior prospective beliefs’ predicted by each model comparing to the actual participant beliefs) was assessed using the Bayesian Information Criterion (BIC) and Mean Absolute Error (MAE) as measures of model fit. For validation purposes, two simpler baseline models were also created (see Section 3.2.3 below). While there were no fit differences between the ‘precision of evidence’ proxies that were used, retrospective metacognitive evidence was found to predict prospective posterior beliefs better than performance itself across and in each group. Using this winning model, i.e., the model that corresponded to these (optimal) choices of ‘evidence’ and ‘precision of evidence’, I then also compared this model’s predictive fit between our groups using the same criteria. Thus, this analysis allowed to identify which choice of perceptual performance versus metacognitive ‘evidence’ and ‘precision of evidence’ gave the most accurate ‘posterior prospective beliefs’ predictions across and between our groups.

Crucially, the winning model’s measures of ‘precision of prior beliefs’ and ‘precision of evidence’ were then used to compute Bayesian optimal, precision-weighted, prospective belief ‘learning rates’ per participant (Kirsch et al., 2021) and these ‘learning rates’ were compared between groups. Within this version of the Bayesian Learning Framework, these ‘learning rates’ are a measure of how much ‘evidence’ (here retrospective metacognitive beliefs about

performance rather than performance itself) is taken into account while updating prospective beliefs. Therefore, the importance of this comparison is that it helped explore if the AN group differs from the other groups in how much they are expected to take this kind of ‘evidence’ into account when updating their ‘prospective beliefs’ regarding their interoceptive abilities in the cardiac domain.

In brief, the experiment in Chapter 3 assessed how patients with AN, or AN-WR in comparison to HCs use local and global retrospective beliefs about performance on a cardiac signal perception task to update their explicit, prospective beliefs about their related abilities.

3.2 Methods

3.2.1 Participants

The participants recruited in the present experiment were different to those recruited for the experiment presented in Chapter 2. A total of $N_{AN}=40$, $N_{AN-WR}=40$, and $N_{HC}=121$ participants were screened, $N_{AN}=4$ were excluded due to other comorbid Eating Disorders; $N_{HC}=1$ was excluded as upon screening they said they were on psychiatric medication; and, $N_{HC}=1$ was excluded due to having a BMI >35 (significantly higher than our BMI upper limit of 25). The rest of the screened participants were recruited and tested, however $N_{AN}=1$ wished to withdraw halfway during the session; and $N_{HC}=2$ were excluded from all analyses due to a technical failure during the HCT and failure to comply with task instructions, respectively. The final sample consisted of $N_{AN}=35$, $N_{AN-WR}=40$, and $N_{HC}=117$ participants (see Table 3.1 for details on participant demographics and clinical characteristics).

The recruitment criteria were identical to those of the experiment presented in Chapter 2. AN patients met the restrictive subtype AN DSM-5 criteria (American Psychiatric Association, 2013) as diagnosed by an experienced clinician using standard clinical measures (patient history, psychometric questionnaires and a physical assessment), and had a BMI <18.5 . As discussed in Chapter 2 (see Section 2.2), given the growing concerns around weight-restoration criteria, a combination of objective and clinical criteria were chosen to best represent the patients’ clinical reality instead of relying only on BMI criteria, as it inadequately reflects the clinical complexity of the AN recovery stages and symptom evolution. The AN-WR group no longer met DSM-5 restrictive subtype AN criteria according to their psychiatrist and met at least two of the following: BMI >16.5 , clinical and behavioural signs of AN recovery (e.g., no restrictive eating patterns) for at least 6 months, and/or a global Eating Disorders Examination

Questionnaire (EDE-Q; Fairburn & Beglin, 1994) score <4 . The EDE-Q has been validated and is available in both Greek and Italian, and the participants from the Greek and Italian sites were administered it in the official language spoken at the respective local site. There were no menstrual cycle criteria, consistent with the DSM-5 (American Psychiatric Association, 2013). HCs had a BMI between 18.5 and 25 and were excluded if they or a first-degree relative had an ED history. All participants, aged 18 and 45, had no history of brain damage or disease, drug dependence and/or severe psychiatric disease (e.g., psychosis). Other exclusion criteria included pregnancy, being male, any report or indication of current substance and/or alcohol abuse, and having previously participated in the experiment presented in Chapter 2.

The AN and AN-WR participants tested in Italy were recruited from the Psychiatric Daycare Unit of Sao Paolo General Hospital (ASST Santi Paolo e Carlo) in Milan, from Comunità Villa Miralago in Cuasso al Monte (VI) and from Casa di Cura Villa Margherita in Arcugnano, Italy, while those tested in Greece were recruited from the collaborating 1st and 2nd psychiatric clinics of the National Kapodistrian University of Athens (NKUA). Healthy controls were recruited from the University of Milan (Italy), through SONA, the University College London (UCL) psychology subject pool, and the University of Athens (Greece).

Institutional ethics approval was granted and written informed consent was obtained from all participants prior to their participation. HCs from the UK site received either monetary compensation at a fixed rate of £7.5/hour, or university course credits. The participants from the Italian site received £10 Amazon vouchers. The participants from the Greek site were not allowed to receive compensation in line with the Ethics regulations from the local institutions.

Table 3.1. Participant Demographics and Characteristics

	AN	AN-WR	HC	Test Statistic [β , (SE)], p^1	
	Mean (SD)	Mean (SD)	Mean (SD)	AN vs HC	AN-WR vs HC
N	35	40	117	-	-
Age	22(6.19)	26.51(7.24)	24.67(5.76)	-2.61(1.19), 0.029	1.84(1.13), 0.105
BMI	16.13(1.61)	19.93(1.82)	20.48(4.77)	-4.35(0.93), <0.001	-0.54(1.05), 0.604
Illness Duration (years)	5.85(6.39)	7.29(6.57)	-	-	-
Heartbeats – Self*	65.94(18.72)	67.46(18.28)	68.94(31.89)	-3.00(-1.48), 0.684	-1.48(6.52), 0.820
Heartbeats at Rest*	69.29(11.91)	71.26(9.44)	73.03(9.32)	-3.74(3.08), 0.227	-1.77(2.34),
Heartbeats – Other*	60.11(17.8)	65.68(16.76)	68.72(30.08)	-8.60(5.00), 0.087	-3.03(4.85), 0.533
WCST – Percentage Correct	69.18(14.66)	74.45 (14.85)	72.76(15.16)	-3.58(4.89), 0.465	1.69(3.82), 0.658
WCST – Percentage Preservative	33.95(8.84)	30.76(10.1)	29.59(10.69)	4.35(3.38), 0.201	1.16(2.70), 0.668
EDE-Q Total	1.62(1.42)	1.44(1.58)	1.08(1.07)	0.55(0.24), 0.026	0.37(0.23), 0.116
EDI-3 Interoceptive Deficits	14(9.19)	13.16(10.15)	9.5(9.35)	4.50(1.85), 0.016	3.66(1.77), 0.040
OCI-R Total	1.32(0.79)	1.13(0.75)	0.81(0.62)	0.051(0.13), <0.001	0.32(0.13), 0.013
IUS-12 Total	39.53(12.23)	34.82(12.08)	27.04(9.36)	12.49(2.05), <0.001	7.77(1.96), <0.001
DASS-21 Depression	20.26(12.92)	14.00(13.4)	6.94(7.01)	13.32(2.19), <0.001	7.06(2.61), 0.008
DASS-21 Anxiety	13.13(9.96)	11.87(9.43)	6.92(6.55)	6.21(1.81), 0.001	4.95(2.15), 0.023

DASS-21 Stress	22.87(11.49)	21.20(10.92)	11.13(8.67)	11.74(2.25), <0.001	10.07(2.68), <0.001
TAS-20 Total	56.03(14.25)	47.79(13.87)	42.24(10.79)	13.79(2.36), <0.001	5.55(2.26), 0.015
	2 Depression	2 Depression			
	1 GAD	2 GAD			
Psychiatric	3 OCD	1 OCD			
Comorbidities	4 Personality Disorder	6 Personality Disorder	n/a	-	-
	1 Learning Disability				
	4 Antidepressants	6 Antidepressants			
	5 Sedatives	2 Sedatives	n/a	-	-
	4 Antipsychotics	7 Antipsychotics			
Current Medication	5 SSRIs	7 SSRIs			
	1 Anxiolytics	3 Anxiolytics			

N.B. Dashes indicate measure was not taken. Heartbeats – Self*, and -Other* refer to participants’ estimation of their own resting heart rate, and their estimation of the average resting heart rate in the general population (beats per minute (bpm), respectively). Heartbeats at Rest refer to participants’ actual average bpm when at rest. Abbreviations: AN (Anorexia Nervosa); AN-WR (Anorexia Nervosa Weight-Restored); HC (Healthy Control); WCST (Wisconsin Card Sorting Task); EDE-Q (Eating Disorder Examination Questionnaire); EDI-3 (Eating Disorders Inventory 3); OCI-R (Obsessive Compulsive Inventory Revised); IUS (Intolerance of Uncertainty Scale); DASS-21 (21-Item Depression, Anxiety and Stress Scale); TAS-20 (20-Item Toronto Alexithymia Scale); GAD (Generalised Anxiety Disorder); OCD (Obsessive Compulsive Disorder); SSRIs (Selective Serotonin Reuptake Inhibitors). ¹Linear regressions were run to examine group differences, with HC as the intercept. As expected, there were significant between group differences on BMI and EDE-Q – such differences are axiomatic to our groups and consistent with our inclusion criteria. Given also some of the expected group differences in psychometric traits, these were taken into account in the main analyses (see Section 3.2 and Appendix B). An unexpected group difference on age was found, and it was thus taken into account in the behavioural analyses on self-efficacy beliefs and Performance. Bolded values denote statistical significance ($p < 0.05$).

3.2.2 Belief Updating Task Design and Measures

Here, I further explored why AN patients who do not seem to perform worse than HCs struggle to update their (metacognitive, prospective) beliefs about cardiac interoceptive abilities even after trying a HCT (Schandry, 1981). The role of a general cognitive inflexibility (previously noted in AN; Arbel et al., 2013; Konstantakopoulos et al., 2011; Miles et al., 2020) on belief updating was also examined. Thus, the present experiment allowed to study how AN and AN-WR, compared to HCs women, form and update Prospective Self-Efficacy Beliefs (the main focus in this experiment; see Figure 3.1 and below) about interoception before and after completing a modified version of the HCT (Schandry, 1981). The widely used measure of Performance (or, Interoceptive Accuracy) as in Chapter 2 was used, as well as three additional self-efficacy measures, each with corresponding subjective confidence ratings. These measures included participants' (1) Prior Prospective Self-Efficacy Beliefs (i.e., how well they think they will do on the HCT, as in Chapter 2), (2) Posterior Retrospective Self-Efficacy Beliefs (i.e., how well they think they performed on the task, as in Chapter 2), and (3) Posterior Prospective Self-Efficacy Beliefs (i.e., how well they think they would do in the future; these self-efficacy beliefs about Performance, explicitly sampled here for the first time, are identical to the sampled Prior Prospective Self-Efficacy Beliefs and hence allowed to examine how such prospective, global beliefs are updated after Performance and how local and global retrospective beliefs are generated; Figure 3.1). Various control measures were also introduced (see below).

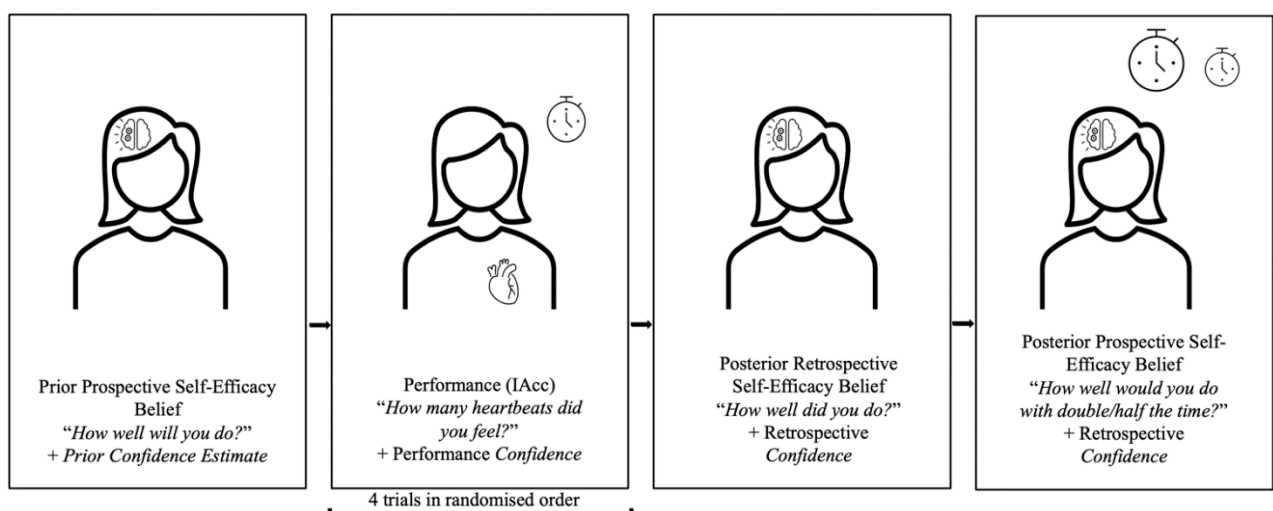


Figure 3.1. Trial Timeline for the Heartbeat Counting Task (HCT).

The figure shows the trial timeline of the HCT and the additional measures of the present experiment, including the main experimental target: the updating of prospective self-efficacy (metacognitive) beliefs about heart beat counting after trying out a heartbeat counting task and forming explicit retrospective beliefs about one's performance. Specifically, participants were asked to tell the experimenter how well they think they would do if they gave them half the time, and then separately if they gave them double the time to complete the task. These questions allowed to see how the participants used what they had observed and explicitly reported during the task (retrospective, metacognitive beliefs about their performance) to form self-efficacy beliefs that they could apply prospectively in a varied context. N.B. A similar procedure was used for Chapter 2, however for the experiment in Chapter 2 confidence estimates were not collected, participants completed 3 instead of 4 trials of heartbeat counting, and critically the participants' Posterior Prospective Self-Efficacy Beliefs were not sampled and hence it was not possible to study how self-efficacy beliefs are updated prospectively.

3.2.3 Primary Hypotheses, Measures and Analyses

3.2.3.1 Behavioural Analyses

After the results presented in Chapter 2, I first examined why AN patients struggle to update their 'pessimistic' prospective self-efficacy beliefs despite a comparable performance to HCs on formal interoceptive tasks. Thus, the updating of these prospective, self-efficacy beliefs was the main measure of interest here and the primary focus of the computational modelling analyses (see below). As preregistered, it was predicted that AN and AN-WR participants would have lower Posterior Prospective Self-Efficacy Beliefs than HCs, despite predicting to find no evidence of group differences on Performance (as in Chapter 2).

To examine Group differences in Performance, a multilevel model analysis (MLM) was run. Performance (Interoceptive Accuracy) was calculated using the following Schandry (1981) calculation, like in Chapter 2.

$$Performance = \left[1 - \frac{1}{4} \sum \frac{|Recorded\ Heartbeats - Counted\ Heartbeats|}{Recorded\ Heartbeats} \right] * 100 \quad 3.1$$

Scores ranged from 0 (*worst Performance*) and 100 (*best Performance*). The following measures were also controlled for: knowledge about heartbeats, time estimation abilities, HRV

and BMI, given criticisms on the effects of these aforementioned factors on Performance (Brenner & Ring, 2016; Knapp-Kline & Kline, 2005; Richard et al., 2019; details in Appendix B, Section 1 and Table B.1).

In addition to a frequentist approach, the analysis was supplemented with a Bayesian analysis, which presents the ratio of the likelihood of the alternative hypothesis relative to the likelihood of the null hypothesis. A Bayes Factor (BF_{10}) >3 indicates evidence for the alternative hypothesis, whereas a $BF_{10} <3$ indicates evidence for the null hypothesis. A BF_{10} between 0.3 and 3 indicates an inconclusive result which is not in favour of either hypothesis (Carey et al., 2021; Kass & Raftery, 1995).

Next, between-group differences on self-efficacy beliefs after completing the HCT were tested. To do this, the effect of Group on Posterior Prospective Self-Efficacy Beliefs was assessed, using Age as a control variable, and Study Site as a random effect (given the multi-site testing). In preregistered, exploratory analyses I tested whether differences in traits and behaviours often seen in the AN population and also seen in the present cohort (e.g., depression, and anxiety; Table 3.1), explained the group differences in self-efficacy beliefs (details in Appendix B, Table B.2). In non-preregistered analyses the role of set-shifting difficulties (as measured via the Wisconsin Card Sorting Test, WCST; Grant & Berg, 1948) on Posterior Prospective Self-Efficacy Beliefs was explored. The WCST is used to assess cognitive flexibility and set-shifting: individuals need to categorise response cards based on different, shifting criteria, e.g., colour and shape (Kopp et al., 2021; Westwood et al., 2016).

3.2.3.2 Computational Modelling

The behavioural analyses were complemented with modelling analyses to account for the role of the nested nature of prior prospective and retrospective beliefs in the updating of such posterior prospective beliefs (see Section 3.1) and other parameters such as precision and learning rate. I first examined which model best predicted the key measure (Posterior Prospective Self-Efficacy Beliefs) by constructing and comparing between models that included the scores of different proxy-measures for evidence and for the precision of this evidence. Then, I compared how the winning model predicted the groups' actual posterior prospective beliefs. Furthermore, the clinical groups' learning rate (using the winning model's measures of 'precision of prior beliefs' and 'precision of evidence'; see below) was examined,

expecting it to be lower than the HCs' when controlling for more general mental flexibility deficits.

Specifically, a posterior self-efficacy belief ($\mu_{\theta|y}$) was computed using a generic Bayesian equation for belief updating when receiving new information (or, evidence) under a Gaussian model with conjugate prior (Friston, 2017; Mathys et al., 2014): $\mu_{\theta|y} = \mu_{\theta} + \frac{\pi_{\epsilon}}{\pi_{\theta} + \pi_{\epsilon}}(y - \mu_{\theta})$, (3.2; Kirsch et al., 2021). This equation allowed to create a set of four target models to examine the hypotheses (see Table 3.2 for full model description). For all the models participants' Prior Prospective Self-Efficacy Beliefs were used as the prior (μ_{θ}), and the respective Prior Prospective Confidence estimates were used as a precision proxy of the prior (π_{θ}). These models differed in the measures that were used as evidence (y ; Performance versus Posterior Retrospective Self-Efficacy Beliefs) and as evidence precision proxies (π_{ϵ} ; Performance Confidence versus EDI-3-ID; Garner, 2004). For each model, the Learning Rate (λ ; also known as the Bayesian precision ratio) was computed per participant: $\lambda = \frac{\pi_{\epsilon}}{\pi_{\theta} + \pi_{\epsilon}}$ (3.3).

Table 3.2 Main Target Models for Computational Modelling Analyses

Model	Prior (μ_{θ})	Evidence (y)	Precision of the prior (π_{θ})	Precision of evidence (π_{ϵ})
Model 1	Prior Prospective Self Efficacy Belief	Performance	Prior Prospective Confidence	Performance Confidence
Model 2	Prior Prospective Self Efficacy Belief	Posterior Retrospective Self-Efficacy	Prior Prospective Confidence	Performance Confidence
Model 3	Prior Prospective Self Efficacy Belief	Performance	Prior Prospective Confidence	EDI-3-ID
Model 4	Prior Prospective Self Efficacy Belief	Posterior Retrospective Self-Efficacy	Prior Prospective Confidence	EDI-3-ID

N.B. Model 1 is the same as Models 1.1., 2.1, and 3.1, and 4.1; Model 2 is the same as Models 1.2, 2.2., 3.2, and 4.2; Model 3 is the same as Models 1.3, 2.3, 3.3, and 4.3; Model 4 is the same as 1.4, 2.4, 3.4, 4.4 all presented in Table 3.6a. The models presented in this Table show the Prior, Evidence and Prior and Evidence Precision Proxies that were used, while the models presented in Table 3.6a below show the model comparison for each of these models across all groups and within each group separately. Abbreviations: EDI-3-ID (Interceptive Deficits Subscale from the Eating Disorders Inventory 3; Garner, 2004).

For validation and control purposes of the main models of interest, two sets of baseline models were created and then compared. The first set of models (one model per evidence measure)

assumed a perfect learning rate ($\lambda = 1$, the participant uses the evidence as their posterior beliefs). The second set of baseline models (actually, one model given its baseline nature) assumed no learning ($\lambda = 0$, the participant uses the prior as their posterior beliefs). If $\lambda = 0$, the models would yield the same score, assuming the same values are used as priors, posteriors, and respective proxies for precision.

For all the models the Bayesian Information Criterion (BIC) and Mean Absolute Error (MAE) were computed in order to measure model fit. As preregistered, I initially examined which variable was the best measure of evidence, in the HCs only, by comparing the fit of the model that used Performance Confidence as an evidence precision proxy and ‘Posterior Retrospective Self-Efficacy Beliefs’ as evidence with the alternative learning model that used the same precision proxy but Performance as evidence, predicting that the former model would show a better fit than the alternative learning model, particularly in the HCs (Prediction A). This was repeated across groups, and within each clinical group separately. To further explore the winning model’s validity (i.e., whether the winning model was significantly better than the other competing models) two non-preregistered analyses on the winning model (one across groups and one within the HCs only) were run using as outcome variable the absolute error of the Predicted Posterior Prospective Self-Efficacy Belief (see Appendix B, Table B.8a). In an additional (exploratory and non-preregistered) step, the above model comparison was repeated (i.e., between the two different evidences but the same precision proxy) by using EDI-3-ID (an eating disorders-specific measure of interoceptive sensitivity) as an evidence precision proxy instead of Performance Confidence. This exploratory comparison was run across all groups and within each group separately (see Appendix B, Table B.8a).

For the measure of evidence which was associated with the best model fit from the previous analyses, a preregistered further modelling step was performed, wherein I examined whether subjective confidence ratings (Performance Confidence) or trait measures of interoceptive sensibility (EDI-3-ID; Garner, 2004) when used as precision proxies, best captured the groups’ belief updating, expecting the clinical groups to be more influenced by the trait measures than HCs (Prediction B). This precision-proxy comparison was complemented with two non-preregistered analyses to explore the winning model’s validity: I tested whether the precision-proxy of the winning model was significantly better than the other proxy used in the competing model (i.e., same evidence measure, but different precision proxy) – this was done both across groups and within the HCs only (see Appendix B, Table B.8b).

Next, between-group differences on precision-weighted Learning Rates were examined. Assuming that the Bayesian Model of choice is a representation of actual learning, then a higher Learning Rate would suggest that participants consider the Prediction Error to a greater degree when updating their beliefs. One precision-weighted Learning Rate per precision proxy was calculated to then look at group differences in two separate analyses, expecting the clinical groups' Learning Rates to be lower and less Bayesian optimal (i.e., greater absolute difference between participants' actual and precision-weighted Learning Rates) than the HCs' (Prediction C). Given the effect of Depression and Stress on self-efficacy beliefs (see behavioural results), in non-preregistered analyses I then explored whether Depression and Stress scores influenced participants' precision-weighted Learning Rates.

Finally, as preregistered, group differences on actual Learning Rates were examined, also expecting the Learning Rates of the clinical groups to be lower than those of the HCs (Prediction D; see Appendix B, Table B.10). Actual Learning Rates (λ_{Actual}) were calculated using the following equation: $\lambda_{Actual} = (\mu_{\theta|y} - \mu_{\theta}) / (y - \mu_{\theta})$, (3.4) where $\mu_{\theta|y}$ represents participants' Posterior Prospective Self-Efficacy Beliefs, μ_{θ} represents participants' Prior Prospective Self-Efficacy Beliefs, and y represents participants' Posterior Retrospective Self-Efficacy Beliefs.

3.2.3.3 Secondary Hypotheses, Measures and Analysis

Additionally, the following and mostly preregistered (as designated in each case) series of secondary questions were also examined. The hypotheses, main methods and results regarding these analyses are briefly presented below, and a thorough presentation is available in Appendix B (see Appendix B, Sections 1 and 2).

First, group-differences on participants' evaluation of their abilities before and after completing the HCT (i.e., Prior Prospective and Posterior Retrospective Self-Efficacy Beliefs) were examined. Group differences in the correspondence between Performance and respective ratings of subjective confidence (interoceptive awareness), and group differences in Confidence throughout the task were also examined. It was predicted that the two clinical groups would be more "pessimistic" about their abilities and show poorer interoceptive awareness compared to HCs. Finally, it was predicted that the two clinical groups would be

less confident in the accuracy of their answers throughout the experiment and especially during the HCT task and in their overall retrospective Performance evaluation.

In the AN and AN-WR groups only, I explored whether aspects of everyday symptoms and beliefs (e.g., illness severity and weight-gain fears; details in Table 3.3 and Appendix B, Table B.11) correlated with their beliefs regarding their task-performance abilities.

The experimental design and analyses were preregistered in OSF and the preregistration can be found here: <https://osf.io/x4ysv>. The data were analysed using R and the figures were created using ggplot2 (Wickham, 2016).

3.2.4 Psychometric Measures

To examine whether traits and symptoms associated with AN had any effects on the main, behavioural results, exploratory analyses were run questionnaires assessing eating behaviours and responses to emotional states (EDI-3; Garner, 2004), obsessive and compulsive symptoms (Foa et al., 2002), reactions to uncertainty, ambiguous situations and the future (IUS-12; Carleton et al., 2007), depression, anxiety, stress (DASS-21; Lovibond & Lovibond, 1995), alexithymia (TAS-20; Bagby et al., 1994), and insight into dimensions of restrictive, anorexic symptoms (see Table 3.3).

Eating Disorder Inventory (EDI-3; Garner, 2004): The EDI-3 is a self-report measure used to assess disordered eating behaviour (e.g., bulimia and the tendency to think about and engage in binge-eating), as well as other psychological behaviours that are highly relevant, but not necessarily specific, to eating disorders. The EDI-3 is widely used in interoception research given that certain of its items assess response to emotional states, confusion related to affective and bodily functioning and is also used to assess the presence of an eating disorder. The 12-subscale questionnaire consists of 91 items rated on a 6-point Likert scale, ranging from 0 (*Always*) to 4 (*Rarely*; *Always* and *Usually* both receive a score of 0). Here, the Interoceptive Deficits subscale (EDI-3-ID) which consists of 9 items was used as a precision proxy. The internal consistency of the total EDI-3 score and the EDI-3-ID subscale in the present sample was excellent with Cronbach's α scores of 0.96 and 0.95, respectively.

Obsessive-Compulsive Inventory (OCI-R; Foa et al., 2002): The OCI-R is an 18-item self-rating scale typically used to assess the severity and type of obsessive-compulsive symptoms.

The OCI-R evaluates six groups of OCD symptoms on a 5-point Likert scale ranging from 0 (*not at all*) to 4 (*extremely*), with three items in each group and a total score obtained by summing all items and a total score range of 0-72. Higher scores indicate greater OCD symptoms. A cut-off score of 21 has been introduced, with scores equal to or higher than 21 indicating the possible presence of OCD symptoms. Overall, the OCI-R has shown great overall internal consistency and good test-retest reliability (OCI-R; Foa et al., 1998; Wootton et al., 2015). The discriminant validity of the OCI-R remains unclear; namely, the ability of the OCI-R to distinguish OCD from other types of anxiety (e.g., worry and general anxiety) is not fully examined. The internal consistency in the present sample was excellent, with Cronbach's $\alpha=0.91$.

Intolerance of Uncertainty (IUS-12; Carleton et al., 2007): The IUS-12 questionnaire was used to assess reactions to uncertainty, ambiguous situations and the future. Items are scored on a 5-point Likert scale ranging from 1 (*Not at all characteristic of me*) to 5 (*Entirely characteristic of me*), with possible total scores ranging from 27 to 135, and higher scores suggest greater intolerance of uncertainty. The IUS-12 consists of two subscales, the prospective anxiety subscale (PAS) and the inhibitory anxiety subscale (IAS) and has good internal consistency. Finally, the IUS-12 is significantly correlated with state measures of worry (Penn State Worry Questionnaire; Meyer et al., 1990; Van Rijsoort et al., 1999) and has a lower correlation with measures of trait anxiety (trait version of the State-Trait Anxiety Inventory; Spielberger et al., 1971; Van der Ploeg, 1980). Internal consistency in the present sample was excellent, with Cronbach's $\alpha = 0.94$.

Depression, Anxiety, Stress Scale (DASS-21; Lovibond & Lovibond, 1995): The DASS-21 was used as a scale to assess symptoms relating to negative mood and emotions on a 4-point likert scale ranging from 0 (*did not apply to me at all*) to 3 (*applied to me very much*; see Section 2.2, Chapter 2 for more details). The internal consistency in the current sample was excellent with Cronbach's $\alpha = 0.95$.

Toronto Alexithymia Scale (TAS-20; Bagby et al., 1994): The TAS-20 is the most used self-report measure of alexithymia with good reliability and factorial validity in both clinical and non-clinical populations (González-Arias et al., 2018). All participants, except for two AN and one AN-WR, completed the TAS-20, which consists of 20 items rated on a 5-point Likert scale

ranging from 1 (*strongly disagree*) to 5 (*strongly agree*; see Section 2.2, Chapter 2 for more details). The internal consistency in the current sample was excellent with Cronbach's $\alpha = 0.87$.

Insight into Illness: The AN and AN-WR groups completed a 12-item questionnaire on explicit beliefs about illness and symptomatology developed by the research team (Table 3.3). The questionnaire measures four dimensions of restrictive, anorexic symptom awareness, namely the (i) General, Allocentric Perspective of Self, (ii) Health Consequences, (iii) Perspective Taking, and (iv) Future Perspectives: Hopes and Fears (hereafter referred to as Counterfactual beliefs). The self-report items measure patients' awareness of their emaciated appearance and illness severity, the negative consequences of their disordered eating, their beliefs regarding others' perspectives on their conditions, and finally, their fears and beliefs regarding behaviours which may hinder prognosis. Each question is rated on a 0 (*not at all*) to 10 (*very much*) scale, and each participant's clinician (both for AN and AN-WR groups) also rates symptom and illness severity on these questions based on their professional evaluation of the individual's symptoms and the individual's medical files. To obtain insight scores an error score was calculated using the clinician-score weighted difference between patient and clinician scores, with a negative score suggesting deficits into insight. An error score was obtained for each of the 8 insight questionnaire items, and as well as a score per dimension: (i) General, Allocentric Perspective of Self, (ii), Health Consequences, (iii) Perspective Taking).

Table 3.3. Insight into Illness Questionnaire

Item	Subscale
1 How severe is your condition? (<i>Eating disorder</i>)	General, Allocentric Perspective of Self
2 How pathological do you think your approach to food is?	
3 Do you think that you are too thin in comparison to the average woman in the country?	
4 How pathological do you think your eating behaviour is?	
5 Do you think that your eating behaviour could have negative consequences for your health?	Health Consequences
6 Do you think that your body weight and shape could have negative consequences for your health?	

7	Do you think that people are correct when they say that you are too thin?	<i>Perspective Taking</i>
8	Do you think that people are correct when they say that you do not eat enough?	
9	Are you afraid of losing control of your eating behaviour? *	
10	Do you believe that the severity of your condition will improve in the future?	<i>Future Perspectives: Hopes and Fears</i>
11	Are you afraid of gaining weight? *	
12	Do you believe that your approach towards food could improve in the future?	

N.B. All items were scored on a scale of 0 (*not at all*) to 10 (*extremely*). * Denotes that items were reverse scored. Each patient’s clinician was given items 1 to 8 and asked to answer based on their clinical evaluation of the patient’s state and the respective clinical file. Patient scores on items 9-12 were used in separate analyses on everyday, counterfactual beliefs and their relation to prospective beliefs about interoception (see Appendix B, Tables B.11 – B.13).

As in Chapter 2, where available, the Greek- and Italian-validated questionnaires were administered to the participants at the Greek and Italian sites, respectively. This was the case for the OCI-R, TAS-20, and DASS-21 questionnaires. The validated version of the Italian IUS-12 was used, and although a validated version of the IUS-12 now exists in Greek, it was not available at the time of design and testing of the present experiment. Validated versions of the EDI-3 and insight into illness questionnaire also do not exist in Greek and Italian. Therefore, two members of the research team (including the PhD student) who are fluent in Greek and English translated the IUS-12, EDI-3, and insight into illness in Greek, and two other members backtranslated it into English. The same was repeated for the translation and backtranslation of the EDI-3 and insight into illness questionnaires in Italian.

3.2.5 Experimental Materials and Procedure

Following baseline, demographic and psychometric assessments (see below), participants were given an Empatica E4 watch (a medical-grade wearable device that records real-time physiological data; Empatica Srl, Italy; see <https://www.empatica.com/research/e4/>) to wear on their left wrist and the heart rate monitoring throughout the experiment was explained to them. The same principles, as presented in Chapter 2, were applied here to ensure no movement

activity or noise was generated during the heart rate activity recording, with some differences on the following: unlike the Polar watch that was used in Chapter 2, here, the beginning and end of a recording phase was marked by the experimenter (sat across the participant) who would manually press the button on the Empatica E4 watch worn by the participant. A phone timer (not visible to the participant) was used to inform the experimenter when it was time to mark the end of a trial. As in Chapter 2, two members of the team checked the recording for artifacts and noise, and counted the number of recorded heartbeats. Next, participants silently sat on a chair with their legs uncrossed and their wrist gently resting on the table in front of them to obtain a 5-min baseline recording of their heart rate and heart rate variability (HRV; used as control measures; see Table 3.1 and Appendix B, Section 1 and Table B.1).

Participants then completed the HCT task in the same, relaxed position, with their eyes open or closed, and without attempting any physical manipulation to facilitate heartbeat detection (with the same instructions as presented in Chapter 2; Figure 3.1; see also Appendix A for detailed instructions script). In the present experiment however, there were four instead of 3 counting phases: the duration of the heartbeat counting phases matched that of the time-estimation (see below) trials (25s, 35s, 45s, 65s, with 30s rest breaks in between). After each trial, participants verbally reported how many heartbeats they felt, followed by a confidence estimate (ranging from 0, *not at all confident* to 100, *extremely confident*) on the accuracy of their response. Before beginning the HCT, participants gave a prospective self-efficacy belief and a confidence estimate. After completing all heartbeat counting phases participants provided a Posterior Retrospective and two Posterior Prospective Self-Efficacy beliefs, and the respective confidence estimates.

At the end of the HCT participants reported how many heartbeats they think they typically have when at rest, and the general population average (per minute). The answers were used in control analyses (see Appendix B, Table B.1). Participants also completed a time-estimation task (same time intervals as the heartbeat counting trials). Participants silently counted seconds until prompted to stop and then verbally reported how many seconds they counted. The order of both the time-estimation and heartbeat counting trials was randomised between participants. Participants were not given any feedback on trial length or performance at any point.

Finally, participants completed a series of psychometric questionnaires and the WCST (Grant & Berg, 1948), and were fully debriefed at the end.

3.3 Results

3.3.1 Primary Hypotheses Results

HCT Performance did not differ significantly between the three groups.

As predicted, the frequentist analysis did not yield a significant result (Figure 3.2; Table 3.4) on group differences in Performance. Moreover, the Bayes Factor analysis suggested that there is moderate evidence for equivalence regarding our groups' Performance ($BF_{10}=0.28$), indicating that as tested here, the AN and AN-WR groups did not perform differently than HCs. Then, preregistered control analyses to account for potential confounding variables were run and the pattern of results remained the same (see Appendix B, Table B.1).

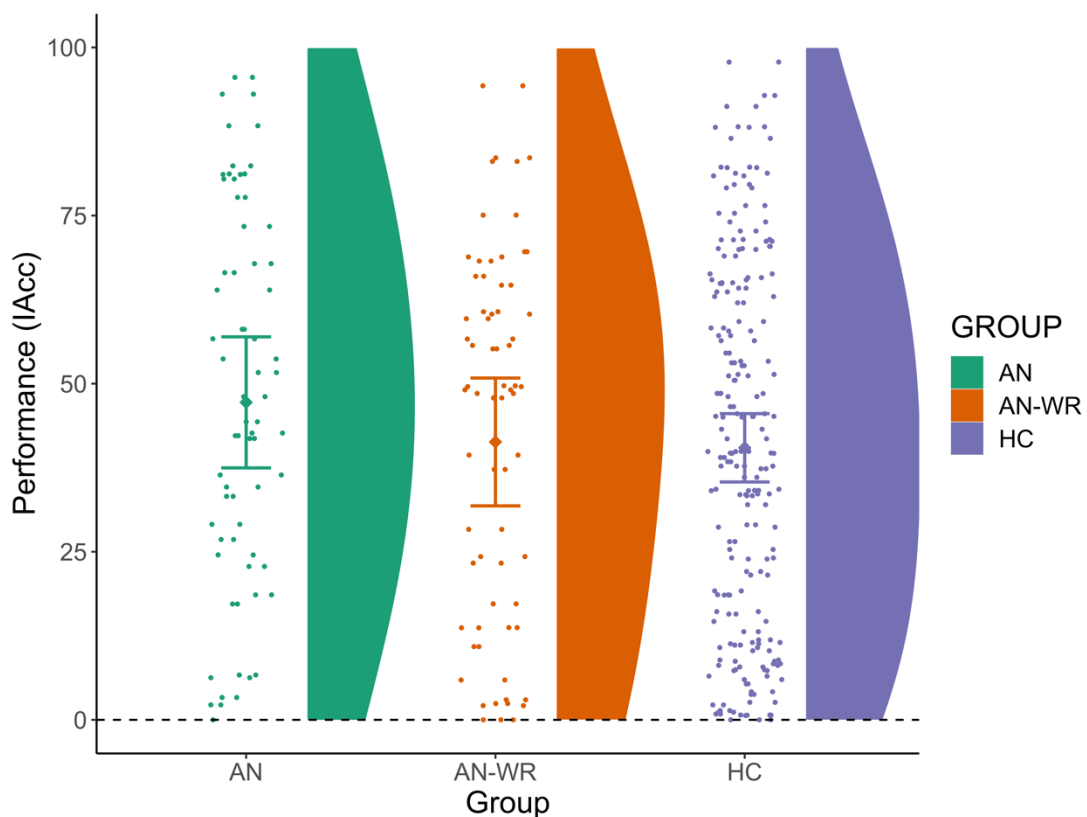


Figure 3.2. Average Performance (IAcc) Scores at the heartbeat counting task per group.

N.B. The overall average Performance scores of the present sample are lower than the average scores typically observed in other studies with both healthy and clinical populations (e.g., Desmedt et al., 2018; Kinnaird et al., 2020; Koreki et al., 2021), though it should be noted that they still fall within the spectrum of average Performance scores observed when participants are encouraged to report only number of actually felt heartbeats and not estimates (e.g., Desmedt et al., 2018).

Table 3.4. Results for Analysis on Performance Group Differences

Predictors	β(SE)	t	95% CI	p
(Intercept)	50.05(9.00)	5.56	32.41 – 67.70	<0.001
AN	2.12(5.72)	0.37	-9.08 – 13.32	0.711
AN-WR	0.36(5.51)	0.07	-10.43 – 11.14	0.947
AGE	-0.38(0.33)	-1.16	-1.04 – 0.27	0.248
Random Effects				
σ^2	227.25			
τ_{00} ID	678.62			
τ_{00} STUDY_SITE	30.48			
ICC	0.76			
N ID	185			
N STUDY_SITE	4			
Observations	672			
Marginal R ² / Conditional R ²	0.008 / 0.759			
Abbreviations: AN (Acute Anorexia Nervosa Group); AN-WR (Weight-Restored Anorexia Nervosa Group); HC (Healthy Controls Group). N.B. Age was used as the Control Variable; Participant ID and Study Site were added as Random Effects. Bolded values denote statistical significance ($p < 0.05$).				

Updating Prospective Beliefs: Behavioural and Computational Analyses.

AN participants gave significantly lower Posterior Prospective Self-Efficacy Beliefs compared to HCs, and the same effect was present as a statistical trend in the AN-WR group (Figure 3.3; Table 3.5). That is, both clinical groups expected, on average, to perform worse in a future HCT with half or double the available time, compared to HCs. To further examine what explained this observed ‘pessimism’ in the two clinical groups, the effect of comorbid traits and behaviours was examined (I used all the variables in which significant group differences were found, see Table 3.1). The results showed that only depression and stress, as measured via the Depression, Anxiety and Stress Scale (DASS-21; Lovibond & Lovibond, 1995) explained the clinical populations’ pessimistic Posterior Prospective Self-Efficacy Beliefs (see Appendix B and Table B.4a for details and see follow-up, exploratory analyses below). I also explored whether a more general set-shifting difficulty could have explained the pessimistic beliefs of the clinical groups (in comparison to the HCs’) but found no significant effect of the WCST performance on Posterior Prospective Self-Efficacy Beliefs (see Appendix B, Table B.4b). Next, using computational modelling, the parameters which could affect the formation of these posterior prospective beliefs were examined.

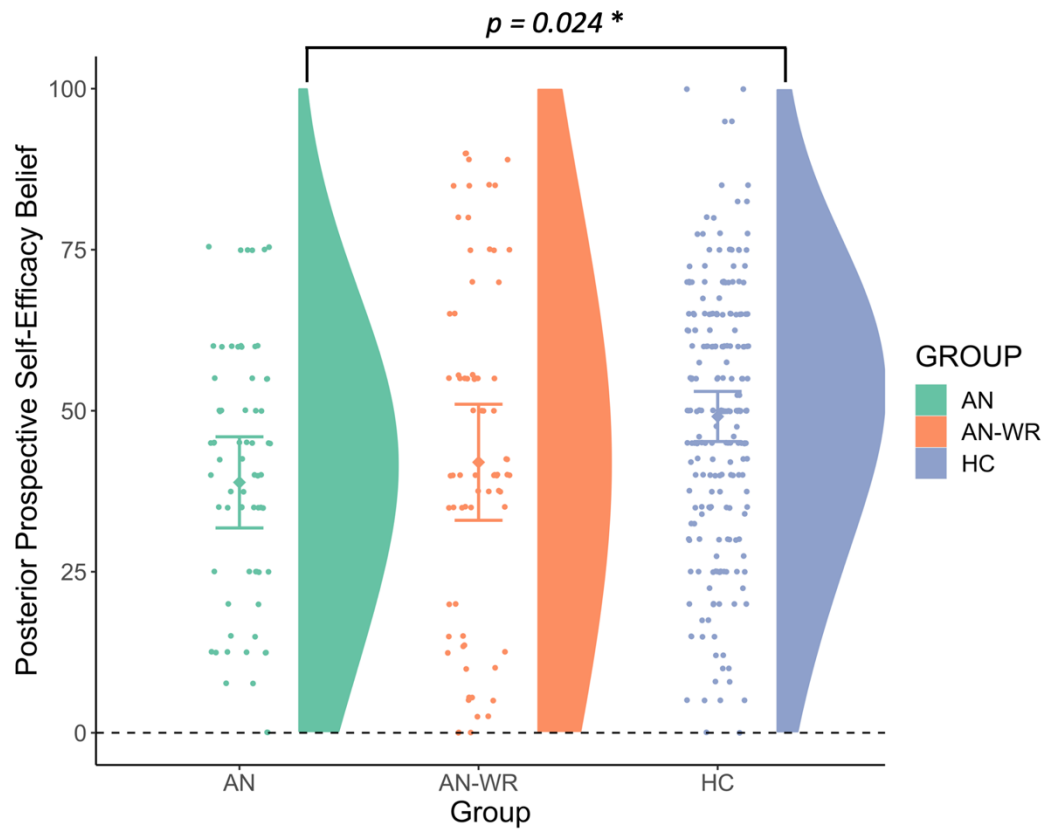


Figure 3.3. Average of Posterior Prospective Self-Efficacy Beliefs per group.

Table 3.5. Analysis for Group Differences on Posterior Prospective Self-Efficacy Beliefs.

<i>Predictors</i>	β (SE)	<i>t</i>	95% CI	<i>p</i>
(Intercept)	50.25(7.13)	7.05	36.27 – 64.23	<0.001
AN	-10.31(4.56)	-2.26	-19.26 – -1.37	0.024
AN-WR	-7.96(4.34)	-1.83	-16.47 – 0.55	0.067
AGE	-0.05(0.27)	-0.20	-0.59 – 0.48	0.841
Random Effects				
σ^2	500.37			
τ_{00} STUDY_SITE	8.85			
ICC	0.02			
N STUDY_SITE	4			
Observations	189			
Marginal R^2 / Conditional R^2	0.038 / 0.053			

Abbreviations: AN (Acute Anorexia Nervosa Group); AN-WR (Weight-Restored Anorexia Nervosa Group); HC (Healthy Controls Group). N.B. Age was used as a Control Variable and Study Site as a Random Effect Bolded values denote statistical significance ($p < 0.05$).

3.3.2 Computational Modelling

Preregistered Prediction A and further validation analyses

In the first step, I examined which model showed a better fit in the HC group by comparing between two models that both used Performance Confidence as a precision proxy but differed in that Performance versus Posterior Retrospective Self-Efficacy Beliefs were used as alternative measures of evidence (Table 3.6a; Models 2.1 and 2.2; see Appendix B Table B.7 for comparison of the Baseline models). Model 2.2 was the winning model in the HCs, but also across groups, and in each clinical group (Table 3.6a; Models 1.2, 3.2 and 4.2). This model comparison was complemented by exploring the difference between the predicted Posterior Prospective Self-Efficacy beliefs, using the two alternative measures of evidence, and Performance Confidence as the precision proxy. Indeed, the model that used Posterior Retrospective Self-Efficacy Beliefs as evidence was significantly better than that using Performance as evidence (both across all participants, and within the HCs; Appendix B, Table B.8a).

Table 3.6. Main output of Bayesian Belief Updating Analyses.

Table 3.6a: Mean Absolute Error (MAE) and Bayesian Information Criterion (BIC) model comparisons			
Model Comparison across all Participants	N	MAE	BIC
1.1 Evidence: Performance; Proxy for Precision: Performance Confidence	183	14.85	1591.5
1.2 Evidence: Posterior Retrospective Self-Efficacy; Proxy for Precision: Performance Confidence	183	9.46	1470.9
1.3 Evidence: Performance; Proxy for Precision: EDI-3-ID (Interceptive Deficits Subscale)	183	15.49	1602.6
1.4 Evidence: Posterior Retrospective Self-Efficacy; Proxy for Precision: EDI-3-ID	183	8.97	1459.7
Model Comparison in the Healthy Control Group (HC) only			
2.1 Evidence: Performance; Proxy for Precision: Performance Confidence	114	14.34	981.6
2.2 Evidence: Posterior Retrospective Self-Efficacy; Proxy for Precision: Performance Confidence	114	8.7	893.7
2.3 Evidence: Performance; Proxy for Precision: EDI-3-ID	114	15.03	985.4
2.4 Evidence: Posterior Retrospective Self-Efficacy; Proxy for Precision: EDI-3-ID	114	8.45	881.1
Model Comparison in the Anorexia Nervosa (AN) group only			
3.1 Evidence: Performance; Proxy for Precision: Performance Confidence	34	13.98	292.6
3.2 Evidence: Posterior Retrospective Self-Efficacy; Proxy for Precision: Performance Confidence	34	11.17	285.3
3.3 Evidence: Performance; Proxy for Precision: EDI-3-ID	34	15.84	296.2
3.4 Evidence: Posterior Retrospective Self-Efficacy; Proxy for Precision: EDI-3-ID	34	10.43	282.9
Model Comparison in the AN weight-restored (AN-WR) group only			
4.1 Evidence: Performance; Proxy for Precision: Performance Confidence	35	17.36	320.1
4.2 Evidence: Posterior Retrospective Self-Efficacy; Proxy for Precision: Performance Confidence	35	10.25	299.2
4.3 Evidence: Performance; Proxy for Precision: EDI-3-ID	35	16.64	323.7

4.4 Evidence: Posterior Retrospective Self-Efficacy; Proxy for Precision: EDI-3-ID

35

9.23

302.5

Table 3.6b: Precision-Weighted Learning Rate Comparisons

Learning Rate with Performance Confidence as a Precision Proxy

<i>Predictors</i>	β(SE)	t	95% CI	p
Intercept	0.47(0.01)	31.53	0.44 - 0.50	<0.001
AN	-0.09(0.03)	-2.78	-0.15 - -0.02	0.006
AN-WR	-0.02(0.03)	-0.76	-0.08 - -0.04	0.446

Learning Rate with EDI-3-ID as a Precision Proxy

Intercept	0.52(0.02)	33.53	0.49 - 0.55	<0.001
AN	-0.03(0.03)	-0.93	-0.09 - 0.03	0.352
AN-WR	-0.01(0.03)	-0.41	-0.07 - 0.05	0.686

Table 3.6c: Comparisons to Examine the Bayesian Optimality of the Precision-Weighted Learning rates

Absolute Difference Between Actual Learning Rate and Learning Rate 1*

(Intercept)	0.59(0.05)	11.07	0.49 – 0.70	<0.001
AN	0.16(0.12)	1.42	-0.06 – 0.39	0.159
AN-WR	-0.01(0.11)	-0.07	-0.22 – 0.21	0.946

Absolute Difference Between Actual Learning Rate and Learning Rate 2*

(Intercept)	0.57(0.05)	10.80	0.47 – 0.68	<0.001
AN	0.12(0.12)	1.03	-0.11 – 0.35	0.306
AN-WR	-0.03(0.11)	-0.23	-0.24 – 0.19	0.817

N.B. MAE and BIC are two measures are used to examine model fit, with smaller values suggesting better model fit. Bolded values denote statistical significance ($p < 0.05$). * Learning Rate 1 refers to the precision-weighted learning rate with Performance Confidence as a proxy for precision of the evidence; Learning Rate 2 refers to the precision-weighted learning rate with EDI-3-ID as a proxy for precision of the evidence.

Preregistered Prediction B

In a further modelling step, I examined which precision proxies best captured the groups' belief updating when Posterior Retrospective Self-Efficacy Beliefs were used as the evidence (winning model from above) and found that the model which used EDI-3-ID as a precision proxy (vs. the other competing precision proxy; Performance Confidence) was the winning model (both across groups and within each group; Table 3.6a models 1.4, 2.4, 3.4, and 4.4), as predicted. Then, I explored (both across groups and within the HCs only) whether the difference between the two precision proxies (when using Posterior Retrospective Self-Efficacy Beliefs as evidence) was significant. However, the model comparisons showed that although EDI-3-ID was a better precision proxy of evidence, it was not significantly better than its competing precision proxy (Performance Confidence; Appendix B, Table B.8b). This suggests that irrespective of which precision proxy for the evidence is examined, Posterior Retrospective Self-Efficacy Beliefs as evidence best predict the groups' Posterior Prospective Self-Efficacy Beliefs.

Preregistered Prediction C

Finally, for prediction C, two further steps were performed. Firstly, I looked at between-group differences in the precision-weighted Learning Rates (for both precision proxies of the evidence, given the inconclusive results of prediction B). Secondly, I looked at whether the Bayesian optimality of this rate differed between groups, by assessing the absolute difference between the groups' actual and precision-weighted Learning Rates ($|\lambda_{Actual} - \lambda|$). Specifically, two separate linear regressions were run to explore the Group effect on the precision-weighted Learning Rate. When using Performance Confidence as a precision proxy of evidence to compute participants' Learning Rates, AN participants but not AN-WR had a significantly lower Learning Rate than HCs (Table 3.6b; Prediction C). No significant between-group differences on Learning Rates were found when using EDI-3-ID as a precision proxy (Table 3.6b). Then, two separate linear regressions were run: when examining the optimality of the groups' Learning Rates (calculated as the absolute difference between their actual and precision-weighted Learning Rates) no statistically significant between-group difference was observed - with either proxy for precision (Table 3.6c). These results suggest that when considering AN patients' reported subjective uncertainty (confidence) regarding Performance they have a lower Learning Rate than HCs, but contrary to the prediction the AN group's learning mechanism is not significantly less Bayesian optimal than the HCs'.

Exploring the Role of Depression and Stress

Given the effects of Depression and Stress on the clinical groups' Posterior Prospective Self-Efficacy Beliefs (see Appendix B, Table B.4), in non-preregistered analyses I explored whether Depression and Stress also explained the difference between the AN and HC groups' precision-weighted Learning Rates (when using Performance Confidence as an evidence precision proxy). First, a linear regression using Depression and Stress as predictor variables was run: Learning Rate lowered as Depression and Stress scores increased (Table 3.7). I then explored whether Depression and Stress explained the group effect: we found a non-significant trend of Depression and a significant effect of Stress on Learning Rates, which would explain the difference between the AN and HC groups' Learning Rates. These were complemented with Holm-corrected linear regressions on the AN and AN-WR groups separately, using Depression and Stress scores as predictors to examine trait and state effects. Higher Depression scores significantly influenced the AN-WR group's Learning Rates. Although a non-significant effect was observed in the AN group, it is suggested that this is due to the already high Depression scores within the entire sample (same as in the self-efficacy beliefs) and discuss this finding further in the Discussion (Section 3.4). Next, given that the precision-weighted Learning Rates were computed using Performance Confidence as a precision proxy of evidence and that the AN group had a lower Performance Confidence than HCs (Table S8b), I examined whether Depression and Stress also explained the group differences on Performance Confidence. Higher Depression and Stress scores explained lower Performance Confidence Estimates. This suggests that as the uncertainty on the evidence increases, there is less evidence- and more prior-weighting, explaining the observed lower Learning Rate of the AN group in comparison to that of HCs. For control purposes I examined whether these findings also applied when looking at the effects of Depression and Stress on the actual Learning Rates and not only on the modelled, precision-weighted Learning Rates, but found a non-significant trend of Depression scores on actual Learning Rates, and no significant effect of Stress. Although Depression and Stress influenced the precision-weighted Learning Rates, they did not significantly influence the Learning Rates' Bayesian optimality (Table 3.7).

Table 3.7. Analyses to Explore the Role of Depression and Stress on Learning Rates*

Predictors	β(SE)	t	95% CI	p
Analysis on (precision-weighted) Learning Rates using Depression and Stress as the Predictors				
(Intercept)	0.45(0.01)	33.32	0.42 – 0.47	<0.001
Depression	-0.03(0.01)	-2.62	-0.06 – -0.01	0.010
Stress	-0.05(0.01)	-3.85	-0.08 – -0.03	<0.001
Analysis on (precision-weighted) Learning Rates using Group, Depression and Stress as the Predictors				
(Intercept)	0.46(0.02)	27.28	0.43 – 0.50	<0.001
AN	-0.07(0.04)	-1.88	-0.14 – 0.00	0.063
AN-WR	-0.05(0.04)	1.15	-0.13 – 0.04	0.252
Depression	0.03(0.01)	-1.81	-0.05 – 0.00	0.073
Stress	-0.04(0.01)	-3.08	-0.07 – -0.02	0.003
Holm-corrected analysis on (precision-weighted) Learning Rates of the AN group using Depression and Stress as the Predictors				
(Intercept)	0.38(0.05)	7.02	0.27 – 0.50	<0.001
Depression	-0.00(0.04)	-0.09	-0.08 – 0.08	0.933
Stress	-0.06(0.04)	-1.25	-0.15 – 0.04	0.675
Holm-corrected analysis on (precision-weighted) Learning Rates of the AN-WR group using Depression and Stress as the Predictors				
(Intercept)	0.38(0.03)	11.04	0.30 – 0.45	<0.001
Depression	-0.09(0.03)	-3.32	-0.14 – -0.03	0.024
Stress	0.04(0.04)	0.95	-0.05 – 0.12	0.362
Analysis on Performance Confidence using Group, Depression and Stress as the Predictors				
(Intercept)	56.45(2.60)	21.68	51.30 – 61.51	<0.001
AN	-8.72(5.94)	-1.47	-20.49 – 3.05	0.145
AN-WR	-11.39(6.72)	-1.69	-24.70 – 1.93	0.093

Depression	-4.81(2.20)	-2.19	-9.16 – -0.45	0.031
Stress	-5.35(2.23)	-2.40	-9.76 – -0.94	0.018

Analysis on actual Learning Rates using Depression and Stress as the Predictors				

(Intercept)	0.71(0.07)	10.76	0.58 – 0.84	<0.001
Depression	-0.12(0.06)	-1.83	-0.25 – 0.01	0.070
Stress	0.06(0.07)	0.88	-0.08 – 0.20	0.379

Analysis on the absolute difference between actual and precision-weighted Learning Rates using Depression and Stress as Predictors				

(Intercept)	0.60(0.05)	11.05	0.49 – 0.71	<0.001
Depression	0.06(0.05)	1.19	-0.04 – 0.17	0.238
Stress	-0.05(0.06)	-0.95	-0.17 – 0.06	0.347

Abbreviations: AN (Acute Anorexia Nervosa Group); AN-WR (Weight-Restored Anorexia Nervosa Group); HC (Healthy Controls Group).
 N.B. *Here only the role of Depression and Stress (as measured using the DASS-21) on the precision-weighted Learning Rates which used Performance Confidence as a precision proxy of evidence was examined, given that I only found significant group differences on these learning rates and not the ones computed using EDI-3-ID as a precision proxy of evidence. Bolded values denote statistical significance ($p < 0.05$).

Control Analyses

To further understand if other cognitive factors may have also explained the lower precision-weighted Learning Rate in AN (when using Performance Confidence as a precision proxy) the role of set-shifting abilities was explored. However, the difference between the three groups' percentage of correct WCST answers, or their WCST preservative score was non-significant (see Appendix B, Table B.9). Moreover, neither of the WCST scores explained the group differences in Learning Rates (see Appendix B, Table B.9). Therefore, general cognitive flexibility and set-shifting abilities did not influence the patients' learning rates, or their optimality.

3.3.3 Further control and exploratory analyses

When examining subjective confidence *per se*, as predicted, AN patients were significantly less confident after the HCT than HCs (see Appendix B, Table B.6b), and there was a non-significant trend in the Posterior Prospective Confidence estimates of the AN group vs HCs. However, the other between-group confidence estimate comparisons were non-significant (see Appendix B, Tables B.6a and B.6c). Similarly, the rest of the control and exploratory analyses were also non-significant (see Appendix B, Tables B.11-B.13; full results are presented in Appendix B and key findings are further elaborated in the Discussion, Section 3.4), with one exception: when considering the relation of the experimental Posterior Prospective Self-Efficacy beliefs, a relationship between higher beliefs and higher ratings of everyday 'optimism' about eating difficulties in the future in both clinical groups and less fear about eating difficulties in the future in the AN group was found (see Appendix B, Tables B.11d and B.11f).

3.4 Discussion

The experiment presented in this Chapter examined how patients with and weight-restored from anorexia nervosa (AN and AN-WR, respectively), in comparison to HCs update their self-efficacy beliefs about their ability to count their own heartbeats. The preregistered experiment aimed to examine specifically how pessimistic prospective self-efficacy beliefs about cardiac signal perception abilities are formed and maintained, despite patients with AN performing comparably to HCs. As predicted, results confirmed that although AN patients do not perform worse than controls, when asked how they think they would do in a similar task in the future, they tend to be more pessimistic than HCs. Furthermore, computational analyses, taking into account the nested nature of prospective and retrospective metacognitive beliefs about sensory

performance (Kirsch et al., 2021) and aiming to determine some of the parameters that may lead to such belief updating, revealed that AN patients seem to rely more on their pessimistic retrospective beliefs about interoceptive performance rather than their actual performance. AN patients also show low confidence in the accuracy of their performance, which when accounted for reveals a smaller learning rate in AN patients than controls, indicating that they make less use of prediction errors. Interestingly, the critical parameters revealed by the computational analyses were associated more with mood than with cognitive traits, supporting a previously hypothesised relation between negative mood, interoception, and self-efficacy in other psychopathologies such as depression (Stephan et al., 2016).

As expected, the three groups did not perform differently to each other in the actual HCT (also supported by Bayesian equivalence testing). This finding is observed in most studies of similar populations using the HCT (e.g., Eshkevari et al., 2014; Kinnaird et al., 2020; Lutz et al., 2019; see Section 3.1, Chapter 2, and Section 1.3 from Chapter 1), although contrary findings also exist (e.g., Pollatos et al., 2008, 2016). Differences between the studies may be explained by the noted, low validity and reliability of the HCT and its confounds (Brener & Ring, 2016; Desmedt et al., 2018, 2020, 2022; Legrand et al., 2022), as interoceptive accuracy scores may differ depending on task demands, time estimation abilities and heartbeat knowledge. Here, control analyses (e.g., on time estimation abilities, heartbeat knowledge, and HRV) suggest that in the present sample, these factors are unlikely to have influenced Performance (in line with Ferentzi et al., 2022). Sampling differences (e.g., comorbidities, treatment type, illness stage; Fischer et al., 2016; Pollatos et al., 2008; Richard et al., 2019) could also account for the between-study differences on interoceptive accuracy. For example, similar to the sample of the experiment presented in Chapter 2, the AN sample only included individuals who met restrictive AN criteria (American Psychiatric Association, 2013) and no other ED diagnoses, while others may recruit patients with both a restrictive or binge-purge AN diagnosis. Notwithstanding heterogeneity in treatment stage and type, BMI and clinical profiles did not affect the findings (see Appendix B), suggesting that these factors were unlikely to have influenced the current sample's Performance/interoceptive accuracy. Despite controlling for some of the possible task confounds, future studies on interoceptive metacognition in AN could use the more recently developed tasks to capture interoceptive accuracy (see Desmedt et al., 2023; Garfinkel et al., 2022; Harrison et al., 2021; Legrand et al., 2022; Palmer et al., 2019; Plans et al., 2021), or use pharmacological (Khalsa et al., 2009, 2015) or behavioural (Fitz-Clarke, 2007; Smith, Feinstein, et al., 2021) heart rate manipulations to increase the signal

strength. It should be noted however, that as the emphasis here was on testing explicit belief updating and not interoceptive accuracy *per se*, the HCT has the advantage of good patient acceptability and was easily understood as a brief task that one can have meaningful self-efficacy beliefs about.

A key finding of the present study is that AN patients (and AN-WR patients at trend levels) do not seem to place sufficient emphasis on their performance when forming prospective beliefs about their abilities (self-efficacy beliefs). Hitherto, questionnaire studies have consistently shown that AN patients are characterised by higher levels of worry, rumination, and other maladaptive, metacognitive beliefs, such as low self-efficacy about eating regulation abilities, body mistrust, and fear about the safety of bodily signals than the general population (e.g., Berman, 2006; Davenport et al., 2015; Palmieri et al., 2021). However, these studies have not examined beliefs about interoceptive accuracy, nor have they investigated the various, interwoven parameters that may underlie these beliefs, as was done with computational modelling here. Moreover, the findings also demonstrated that the experimental measure of ‘Posterior Prospective Self-Efficacy Beliefs’ was associated with beliefs about everyday life, such as fear of gaining weight, and less hope for improvement of clinical symptoms and behaviours. These findings suggest that the task had good ecological validity and examining computationally the various parameters that influence such beliefs in the present task could provide insights regarding patients’ everyday negative metacognitive beliefs about interoception.

Specifically, the present study leads to three main insights regarding such beliefs. First, AN patients’ updating of explicit, prospective beliefs was influenced more by “pessimistic” global retrospective beliefs about such performance (e.g., how accurate one thought they were after the task ends) rather than the perceptual performance itself (e.g., how accurate one was). Specifically, model comparisons revealed that retrospective beliefs were better at predicting posterior prospective beliefs than performance, with the winning model (i.e., the one using retrospective beliefs as evidence) best predicting the posterior prospective beliefs of the HCs than those of the clinical populations. This observation was also supported by poorer retrospective, interoceptive awareness (as indexed by the Interoceptive Trait Prediction Error (ITPE) z-score analyses; see Appendix B, Section 1 and Table B.5) in the clinical groups, in comparison to HCs who evaluated their performance as better than it was. Moreover, the AN had lower Performance Confidence than HCs, and this lower rated confidence, combined with

lower, explicit global retrospective beliefs about performance suggests that AN patients struggle more than HCs in incorporating new evidence (from the HCT) to update beliefs retrospectively. This pattern of findings suggests that irrespective of any individual differences in interoceptive accuracy *per se*, AN patients may face difficulties in how they draw metacognitive conclusions about their cardiac interoceptive performance retrospectively.

Moreover, these retrospective metacognition difficulties seem to also extend to prospective metacognition. Specifically, when participants needed to then use such retrospective beliefs to update their self-efficacy prospectively, AN patients showed lower precision-weighted Learning Rates (with Performance Confidence as a precision proxy of evidence) than HCs, suggesting that patients consider the ‘evidence’ (here, Posterior Retrospective Self-Efficacy Beliefs) less than HCs when updating such beliefs. As a result, it is plausible that AN patients rely more on prior prospective beliefs regarding their cardiac interoceptive abilities, and this ultimately influences their explicit, global (posterior) prospective beliefs. Therefore, neither local evidence from performance, nor retrospective beliefs following that performance (that themselves were poorly updated by prediction errors, and hence may be hard to target therapeutically) seem to be sufficient to counter and update pessimistic, self-efficacy beliefs in AN. Instead, it may be the ‘precision-based’, mechanism of belief updating itself that requires therapeutic targeting. Indeed, it has been previously hypothesised in two, transdiagnostic studies (Lavalley et al., 2023; Smith et al., 2020; see Section 3.1) that an inappropriate weighting of prior beliefs versus new evidence, especially regarding interoception, may lead to the manifestation of several psychiatric symptoms resulting from faulty inference (Barca & Pezzulo, 2020; Barrett & Simmons, 2015; Owens, Allen, et al., 2018; Paulus et al., 2019; Petzschner et al., 2017). Indeed, among existing interoception studies in AN, pessimistic estimates about one’s abilities have been documented on an anticipatory and perceptual level (e.g., Crucianelli et al., 2016, 2021; Khalsa et al., 2015; Lutz et al., 2019), although only a handful of proposals had envisioned such faulty updating at metacognitive, and prospective cognition levels (e.g., Stephan et al., 2016).

Second, the potential role of other higher-order cognitive difficulties previously noted in AN, such as disruptions in mental flexibility, abstraction or set-shifting (Lang et al., 2014; Miles et al., 2020, 2022; Tchanturia et al., 2012), which could have explained the reduced belief updating in AN was also explored. However, it was found that cognitive, set-shifting abilities as tested here by the Wisconsin Card Sorting Task (WCST) in a sample subset did not explain

group differences in posterior beliefs or learning rates. Indeed, prior studies have shown that such domain general, ‘frontal’ functions cannot explain some of the delusionality in AN (Konstantakopoulos et al., 2011, 2012; see also Section 3.1), and as seen also in Chapter 2 AN patients seem to be able to update their beliefs based on external (albeit random) feedback (see also Kube et al., 2022). Notwithstanding, the present study did not systematically test belief updating across different modalities. It is worth pointing out that patients with AN had more difficulty, in comparison to controls, in updating metacognitive beliefs (where metacognitive updating is one’s own evaluation of abilities, and not dependent on external or other feedback modes). Yet, as seen in Chapter 2, AN patients, in light of feedback, could update self-efficacy beliefs about interoceptive abilities. Therefore, this could suggest that while patients may be pessimistic about their interoceptive self-efficacy beliefs and abilities when metacognitively evaluating and updating their beliefs, their evaluation could be facilitated through feedback. As a result, no conclusions about the interoceptive or more general domain-specificity of the present findings are warranted. This point is also further elaborated in the General Discussion (Chapter 6), also following the observations from the clinical audit presented in Chapter 5.

Third, it was found that depression and stress scores predicted the degree of pessimism in self-efficacy beliefs and explained why the clinical groups who notably had greater depression and stress levels also had more pessimistic Posterior Prospective Self-Efficacy Beliefs compared to HCs. Crucially, in within-group analyses, depression (but not stress) significantly influenced the AN-WR, but not the AN group’s posterior self-efficacy beliefs (i.e., the more severe the depression, the more pessimistic the beliefs). The AN group was overall more pessimistic and had more severe depression and stress than the HCs, making it difficult to disentangle which of these factors is more predictive of their lower self-efficacy beliefs, but the results showed the overall effect of depression and stress across all groups. Depression, anxiety and stress are key comorbidities with eating disorders, including AN (Andrés-Pepiñá et al., 2020; Eskild-Jensen et al., 2020; Himmerich et al., 2019) and the present findings also support previously noted associations between depression, negative metacognitive thoughts, and ED symptoms (e.g., Cooper et al., 2007; Palomba et al., 2017; see Section 3.1). Notably, although in some research traditions, these findings could be interpreted as though negative mood could ‘explain away’ the pessimistic beliefs about interoception, from a more transdiagnostic, computational psychiatry perspective, the present study can elucidate the mechanisms by which depression and stress symptoms could influence belief updating (e.g., Aylward et al., 2020; Katyal et al., 2023; Kim et al., 2020; Ruppel et al., 2018; Stankevicius et al., 2014). Specifically, here

it was explored if and how these two traits explained the observed group effect on the precision-weighted learning rates (when using Performance Confidence as a precision proxy of evidence). Interestingly, higher uncertainty on the precision of the evidence (as indexed by lower Performance Confidence estimates) was influenced by higher depression and stress scores, which in turn also explained the lower precision-weighted learning rates of the AN group compared to HCs. Other recent studies have also found associations between local confidence and mood disorder symptom severity (Benwell et al., 2022; Rouault, Seow, et al., 2018; Seow et al., 2021). The present findings, however, further suggest that mood may relate to a dysfunctional precision-weighting mechanism extending beyond local metacognition to global, prospective beliefs about one's performance. Moreover, interestingly, a recent theoretical framework has posited depression as the eventual consequence of chronic dysregulation in the domain of interoception and thus also a chronic low self-efficacy regarding homeostatic/allostatic control (Stephan et al., 2016). The present study was cross-sectional and did not target interoceptive control specifically, so the hypothesised chronic mechanisms cannot be characterised here. Nevertheless, the fact that AN-WR participants also had more pessimistic posterior prospective beliefs than the HCs at a trend-level and these were associated with their (less severe) depressive symptoms suggests that this dimension of illness may not be restricted only to the acute state and may be a more enduring trait of these individuals.

In addition to some of the aforementioned limitations of the experimental task and methods, the present study also had several limitations related to clinical research. A multi-site clinical study presents challenges, as participants are tested by different experimenters and under different clinical, cultural and practical restrictions, potentially contributing to between-site variability and some missing values. However, meticulous protocol translations, and experimenter training were ensured, and study site as a random effect in the analyses did not significantly affect the main results. Moreover, a multi-site approach contributes to current efforts in neuroscience, psychology and psychiatry research practices for more representative and diverse samples (Naddaf, 2023). Patient populations also had varying pharmacotherapy based on national guidelines, making standardised control for its effects beyond assessment via self-report questionnaires and medical records impossible. This study also focussed on women with AN, due to the availability of these patients in the collaborating clinics and given hypothesised sex-based neurobiological differences (see more in Chapter 6), but future research should consider more diverse populations when studying interoception in AN.

3.4.1 Conclusion

To conclude, the study in this Chapter investigated explicit, interoceptive belief updating in AN, with a focus on cardiac awareness and related, prospective self-efficacy beliefs using a computational Bayesian Learning Framework, adjusted to account for the nested nature of prospective and retrospective metacognitive beliefs about sensory performance (as in Kirsch et al., 2021). While the AN participants did not differ from controls in interoceptive accuracy *per se*, they failed to update their interoceptive, metacognitive self-efficacy beliefs after performance. Computational modelling showed that neither local evidence from performance, nor retrospective beliefs following that performance (that themselves were suboptimally updated) seemed to be sufficient to counter and update pessimistic, self-efficacy beliefs in AN. AN patients showed lower learning rates than controls, revealing a tendency to base their posterior beliefs more on prior beliefs rather than prediction errors in both retrospective and prospective belief updating. These differences in both explicit beliefs, and the latent mechanisms of belief updating, were not explained by cognitive inflexibility, but by depression and stress measures, even after the acute AN phase. The findings offer novel insights regarding clinically relevant, metacognitive belief updating difficulties about interoception in acute AN and their relation to depression and stress. However, future studies should aim to further explore belief updating across different modalities and domains, and in longitudinal designs, to draw conclusions regarding the specificity of these belief update difficulties and also to examine whether belief updating could be facilitated by the presence of feedback (e.g., as seen in Chapter 2; see also Chapter 6).

Chapter 4: Am I as full as I predicted I would be? An exploration of changes in fullness estimates before, during and after a Water-Sipping Task in women with and weight restored from Anorexia Nervosa

4.1 Introduction

Patients with Anorexia Nervosa (AN) present with various primary symptoms and behaviours that are used to make the diagnosis (e.g., restricted caloric intake, a disturbed view of one's own weight or body shape, undue influence of weight and shape on self-evaluation, or denial of the seriousness associated with current low body weight; Diagnostic and Statistical Manual 5, DSM-5; American Psychiatric Association, 2013; see more details in Section 1.1., Chapter 1). Patients with AN also frequently present with gastrointestinal complaints (Westmoreland et al., 2016), including early and exaggerated fullness, symptoms of gastroesophageal reflux, abdominal pain, bloating and fullness outside of mealtimes, delayed gastric emptying, regurgitation symptoms, as well as with gastrointestinal disorders, including functional symptoms and conditions (Dhopatkar et al., 2023; Halmi & Sunday, 1991; Lin et al., 2021; Norris et al., 2016; Robinson, 1989; Sato & Fukudo, 2015; Schalla & Stengel, 2019; Wiklund et al., 2021). Although the gastrointestinal system and related deficits are key to the AN symptomatology and prolonged restriction in caloric intake has been found to impact the state of the gastrointestinal tract (Westmoreland et al., 2016), research to date has not disentangled which of these symptoms and behaviours are secondary consequences of the acute AN phase (i.e., a state), or enduring traits which are risk factors for AN and/or may persist beyond weight-restoration and recovery. For instance, some suggest that gastrointestinal issues are secondary consequences to the restricted caloric intake and improve after weight restoration (e.g., Benini et al., 2010; Rigaud et al., 1988; Szmukler et al., 1990), i.e., gastrointestinal issues are a state resulting from the prolonged starvation and malnutrition. Others have shown that symptoms such as dysphagia, heartburn and regurgitation may persist (e.g., Riedlinger et al., 2020; Santonicola et al., 2019; Zipfel et al., 2006), or even complicate overall therapy outcomes (e.g., Mascolo et al., 2017; for review see West et al., 2021). Accounts also suggest that the reduced sensitivity to gastric signals and gastrointestinal complaints could be a premorbid risk factor that results in disordered eating behaviours (e.g., van Dyck et al., 2016). For example, childhood onset of gastrointestinal complaints has been linked to an earlier onset of bulimic symptoms (Gendall et al., 2005), and adverse gastrointestinal symptoms or gastrointestinal-related autoimmune diseases (e.g., celiac disease or Crohn's) in childhood have also been associated with a higher risk for developing an eating disorder in adolescence (e.g., Bern &

O'Brien, 2013; Hedman et al., 2019; Jacobi et al., 2004; Mårild et al., 2017; Wiklund et al., 2021). Different approaches have thus been employed to understand appetite and food intake regulation in health and disease, and particularly the noted differences in the hunger, fullness, satiety, and thirst levels of AN patients in comparison to those of controls.

Neuroendocrine and neurobiological approaches, as well as biopsychosocial evidence suggest differences between AN and healthy individuals may be crucial to the understanding of the distorted perception of gastrointestinal signals (e.g., hunger, fullness, satiety, thirst). As briefly reviewed below, the different perspectives have provided useful insights in the understanding of eating motivation and appetite processes and how disruptions across the mechanisms and processes underlying eating motivation and appetite processes may explain AN symptoms relating to eating restriction and distortions in the perception of hunger, fullness, thirst, and satiety. Yet, the relation between the different approaches (e.g., neurobiological, and psychosocial factors) has not always been mapped precisely, and although studies have shown that psychological behaviours may be explained by disruptions in brain function, accounts on how the interpretation of hunger, fullness, thirst, and satiety signals is also contributing to eating motivation and appetite processes are scarce. From a neuroendocrine perspective, hormones and peptides (e.g., leptin, motilin, ghrelin – a marker for body fat, an inducer for stomach contractions, and an appetite stimulating hormone, respectively) are important in aiding the body to adapt its nutritional state and have important roles in the modulation of brain reward circuitry (Monteleone et al., 2018). Aberrant neuroendocrine processes and levels, as seen in AN (e.g., Monteleone et al., 2008; Viltart et al., 2018; Westwater et al., 2021; Xiao et al., 2020) in turn influence one's ability to perceive and respond to hunger and appetite states. Others have looked at the brain reward system, such as the taste pathway which projects inputs from one's tongue taste receptors to the insula (Kurth et al., 2010), and the reward system which receives its input regarding desires to consume foods from frontal cortical regions (Kelley et al., 2005). The frontal cortical regions are also connected with the hypothalamus which incorporates signals from the body's periphery (e.g., blood sugar levels) to maintain the body's energy homeostasis by regulating food intake (Kelley et al., 2005). Next, the learned associations that are formed between taste and subjective hedonic experiences create an internal cognitive and emotional representation of food stimuli which activates at the sight, olfaction, or taste of food (Rolls, 2011). These associations are used by the dopamine reward systems which mediate the motivational and emotional drive for food, also influencing the kind and amount of future food intake (Kennerley et al., 2006; Kringelbach & Rolls, 2004). Both

evidence from animal models and from functional brain imaging studies in AN have shown altered activation in reward-processing regions and brain dopamine circuits are linked to food restriction (Avena et al., 2008; Frank, 2016; Johnson & Kenny, 2010; Kelley, 2004; Monteleone et al., 2018). While a large body of research has focussed on the neuroendocrine and neurobiological factors that drive eating and appetite as briefly reviewed here, there are also biopsychosocial perspectives. For instance, one model, the habit-based model, suggests that food restriction is related to habit strength and not cognitive restraint and the learning occurs based on past interactions with the environment (Coniglio et al., 2017; Dayan & Berridge, 2014). At the same time, a goal-directed approach puts forward the idea that eating restriction is perpetuated by an anxiety-driven motivation to restrict food, leading to alterations in brain and bodily functions - including the aforementioned dopamine reward system - and creating a vicious circle of fear/anxiety, food restriction and weight loss (Frank, 2016, 2021). Further, individuals with AN show enhanced medial prefrontal cortex activity and amygdala activation in response to food image cues they found threatening and disgusting (Uher et al., 2004) and this kind of fear is believed to be a contributor to the eating restriction. A recent framework also put forward the idea that starvation in AN may be a coping mechanism to the discomfort adolescent girls experience during menstruation (Zucker & Bulik, 2020). Specifically, they propose that since starvation may eliminate or minimise menstruation, one may thus ‘prefer’ the sensations that can be controlled (in this case hunger and starvation) over those that cannot (e.g., menstruation-related abdominal discomfort), in turn ‘replacing’ uncontrollable somatic events with somatic sensations that can be attributed as the direct consequence to one’s behaviour (Zucker & Bulik, 2020). This theory however remains speculative and warrants further investigations. A common denominator across these different approaches and models remains the notion of a disruption or disconnection between the gut-brain signalling.

Being able to sense a depleted nutritional body state and subsequently engage in food-seeking behaviours is also believed to rely on gastrointestinal interoception (for recent reviews see Jacquemot & Park, 2020; Khalsa et al., 2022). More recent approaches have thus added onto existing neuroendocrine, neurobiological and biopsychosocial models to further study the bidirectional interactions between the brain and the gastrointestinal tract and understand how a disturbance in interoceptive, gastrointestinal signals may explain why AN patients show difficulties in perceiving gastrointestinal signals and related symptoms. Interoception (which is broadly characterised as the process of sensing, integrating, and interpreting internal, physiological systems (Critchley & Garfinkel, 2017; Garfinkel et al., 2015; Heim et al., 2022;

Khalsa, Adolphs, et al., 2018; for a detailed overview of interoception see Section 1.3, Chapter 1) has been found to be affected in AN. For instance, studies have shown that AN patients' ability to feel their heartbeat and their beliefs regarding such abilities, as well as their sensations and feelings of hunger and satiety seem to be compromised (e.g., Brown et al., 2022; Fassino et al., 2004; Kerr et al., 2016; Saramandi et al., 2022, 2023). The first two empirical chapters of the present PhD thesis (Chapters 2 and 3) examined how explicit metacognitive beliefs about interoception are updated in AN by looking at the updating of self-efficacy beliefs regarding one's ability to count their own heartbeats. The results showed that patients with and weight-restored from AN (AN and AN-WR, respectively) are more pessimistic than healthy controls (HCs) in their abilities, despite not showing differences in their actual performance abilities during the task (i.e., in cardiac interoceptive accuracy), and this pessimism in their beliefs also influenced their belief updating as indexed by patients' precision-optimisation difficulties and their lower learning rates than HCs. A relation between these cardiac, self-efficacy beliefs and clinical traits associated with AN, such as depression and stress, was also found. Specifically, patients' overall pessimistic self-efficacy beliefs in a cardiac interoception task were in part explained by the more severe depression and stress levels they have in comparison to HCs. Moreover, there were indications that the role of depression in the formation and updating of such self-efficacy beliefs may not be restricted only to the acute state of AN given the trend-level differences between the AN-WR and HC groups (but see Section 3.4 (Discussion) of Chapter 3 for a more detailed elaboration of these findings and their implications). However, the results from Chapters 2 and 3 are only focussed on the cardiac modality of interoception which is not an interoceptive system directly linked to the key AN symptomatology like other systems, such as the gastrointestinal system. Given the findings of Chapters 2 and 3, the experiment presented in Chapter 4 looked at whether the results observed in the cardiac system can generalise across other interoceptive systems, namely the gastrointestinal system.

Additionally, the gastrointestinal system is also closely related to body image (Todd, 2020; Todd et al., 2019). AN patients often starve themselves to achieve their ideal body image, with empirical findings supporting that patients will report prolonged fullness, and clinical observations showing that one in three AN patients will also report satiety and lack of hunger cues even when on an empty stomach (Bluemel et al., 2017; Robinson, 1989), all of which may contribute to the observed gastrointestinal interoception deficits. There are several explanations for altered gastric interoception in AN; distorted sensitivity to interoceptive signals has been attributed to gastric distention differences in AN patients in comparison to controls

(Coddington & Bruch, 1970), to predisposing risk factors related to metabolic processes (Couturier & Lock, 2006; Khalsa et al., 2017; Pike, 1998), and more recently to differences in neural activation patterns observed in AN and not in HCs (Kerr et al., 2016, 2017, 2022; Wang et al., 2008). The latter is a key interoceptive region in the brain, with one study showing that behavioural markers of illness severity were related to atypical insular activity during gastric interoception (Kerr et al., 2016), and highlighting that responses and sensations related to eating may be linked to brain-body signalling. The eating restrictions are also theorised to be partly ascribed to interoceptive uncertainty regarding gastrointestinal states (Barca & Pezzulo, 2020). In particular, the noted interoceptive deficits in AN may be influencing the way in which patients with AN experience and handle hunger and satiety cues. For instance, patients with AN may have a difficulty in using the interoceptive signals and related beliefs to form an accurate representation of their inner, physiological self. Thus, patients may engage in (intentional yet unconscious) strategies of eating restriction to suppress or minimise the interoceptive uncertainty that may arise from the noisy and ambiguous interoceptive signals they receive from their body. Despite these initial indications of an association between gastrointestinal interoception and AN, most studies focus on the cardiac modality, while the methods to study gastrointestinal sensations and interoception are not yet standardised.

Most attempts to investigate the perception of gastrointestinal sensations typically employ invasive techniques, and have been tested in healthy populations, whilst examinations in clinical populations remain predominantly tested at a pilot or subclinical population level, and they have not been extensively used to investigate gastrointestinal interoception in AN in depth. For instance, certain approaches involve inserting inflatable balloons into the oesophagus (Aziz et al., 1997; Hobson et al., 2000; Remes-Troche et al., 2012), stomach (Geeraerts et al., 2011), colon (Hölzl et al., 1996), or rectum (Haas et al., 2014; Sun et al., 1990; Toma et al., 1996), or using chemical irritants to trigger direct gastric perfusion (Yang et al., 2006). Though these methodologies are adequate for the examination of presumed interoceptive cortical neural circuits (Aziz, Schnitzler, et al., 2000; Aziz, Thompson, et al., 2000; Naliboff et al., 2006), they come with notable limitations, such as the risk of provoking adverse effects in a population with recognised increased body image concerns (Khalsa et al., 2022), and they can also not be easily implemented outside medical settings. Another recently developed protocol uses a minimally invasive approach, in which participants ingest a vibrating capsule which then allows the estimation of aspects involved in this kind of gut stimulation (e.g., perceptual

accuracy of vibration detection in the gut; (Mayeli et al., 2023; Smith et al., 2021)), but this remains untested in AN. Other, less invasive protocols - also tested in a clinical sample (in patients with Bulimia Nervosa and Binge-Eating Disorder) - use a water ingestion task, during which participants ingest water until reaching a feeling of fullness (also referred to as the Water-Load Test; van Dyck et al., 2016, 2021). Notwithstanding, these approaches have not investigated the perceptual gut processes that occur post-swallowing (Bohon et al., 2009; Frank et al., 2008; Khalsa et al., 2022). Moreover, these approaches have not been examined in an AN population; to date there exists only one pilot study of the Water Load Test as a measure of gastric interoception in AN and affect (mood) changes (Brown et al., 2022). Women with AN and gender-matched controls were asked to ingest as much water as they pleased until feeling ‘completely full’. The quantity consumed by the AN patients to reach complete fullness was found to be less than that consumed by HCs, despite all participants having complied with a 4-hour fast before the experiment (Brown et al., 2022). Interestingly, AN patients also exhibited greater increases in negative affect during the task, suggesting that the water consumption may have been perceived as an aversive experience (Brown et al., 2022). Although these results point towards differences between individuals with AN and healthy populations in the perception of satiety and related affects, hitherto no study has looked at what may drive the perception of fullness states in AN, such as the formation and updating of fullness state beliefs before and after completing a water-sipping task, respectively.

Several aspects thus need to be considered more carefully when studying gastrointestinal interoception in AN. Firstly, the experience of satiety and feelings of fullness are a subjective and different experience both between individuals but also during distinct individual states. Results in other domains (e.g., external body image and body size estimation; see Saramandi et al., 2022) suggest that the experience of estimating one’s body size in AN is different when asked to report one’s size based on ‘objective’ parameters (e.g., weight and body shape) and ‘subjective’ parameters (e.g., based on feelings of how one may look or feel in their body; see Saramandi et al., 2022 for more). These results indicate that aspects of body image can be elicited and tested by different instructions. As mentioned above, the gastric system is also closely related to one’s body image; thus, how one may experience their inner physiological state can influence how they experience their body ‘externally’ (e.g., feeling bloated may elicit a different body image compared to when not feeling bloated) there needs to be an exploration of the different types of fullness/satiety one may report, e.g., how full/satisfied one may *be* vs. how full/satisfied one may *feel*. Secondly, a better exploration of how aspects of eating and

appetite are formed, predicted, and updated over time is needed to acquire a more in-depth understanding of how a disconnection between bottom-up gastrointestinal signals and top-down beliefs about fullness may manifest through AN-relevant symptoms and behaviours, such as overestimation of and prolonged fullness and reduced hunger.

Specifically, recent theories have proposed the need to understand disturbances in gastrointestinal interoception in AN as a dysregulation between bottom-up and top-down interactions, leading to erroneous perceptual inferences (e.g., about interoceptive and appetitive phenomena), and a mismatch between a predicted and an experienced physiological state (also known as prediction error; Khalsa et al., 2022). As described in Chapter 1 (Section 1.5) and Chapter 3 (Section 3.1), recent approaches have attempted to study under one continuous framework – largely influenced by the wider Bayesian framework – how interoceptive perception and local metacognition influence the updating of explicit, global beliefs about interoception. Specifically, the wider Bayesian framework can be used to study perceptual and metacognitive inference (a method of statistical inference in which new observations are used to infer the probability that a hypothesis was true). For this to happen, afferent sensory input to the brain is influenced by previously learned expectations (or, priors), and vice versa, at multiple levels of the neurocognitive hierarchy, from unconscious processing levels to explicit, conscious, and metacognitive beliefs. Specifically, under a Bayesian inference framework, future body states are predicted and regulated with feedback and feedforward loops that constantly update an individual's internal model and expectations (Barrett & Simmons, 2015; Khalsa et al., 2022; Petzschner et al., 2021; Pezzulo et al., 2015; Seth et al., 2012; Stephan et al., 2016). This process contributes to the minimisation of uncertainty about future states, but also the minimisation of prediction errors. However, the extent to which beliefs are updated under this framework relies on parameters such as precision-weighting mechanisms (the mathematical equivalent of weighting the relevance or reliability of sensory signals). Dysfunctions in the way precision is weighted when updating beliefs and making inferences about metacognitive, interoceptive states is thus believed to be contributing to the observed symptomatology and behaviours in AN (e.g., restrictive eating as a result of a distorted perception of satiety, binge eating, or purging; e.g., Culbert et al., 2015; Herbert, 2021; Quadt et al., 2018; Smith et al., 2020). In Chapter 3, for instance, the weighting of prior beliefs about one's abilities to accurately perceive heartbeats (self-efficacy beliefs) and new sensory evidence was found to be maladaptive in AN. Specifically, AN patients did not weigh prediction errors in similar ways to control when evaluating their self-efficacy beliefs about

cardiac abilities and in turn such differences influenced how they updated their beliefs about their abilities to accurately feel heartbeats over time.

A recent perspective on AN theorises that eating restrictions may be used as a means to minimise interoceptive uncertainty (or, prediction errors; Barca & Pezzulo, 2020; Friston, 2010). Broadly speaking, this would mean that less precision (formally, an inverse variance) would be given to the new observations (here, evidence), and more precision-weighting would be given to prior beliefs (here, an individual's implicit and explicit knowledge regarding a bodily state) when predicting a future interoceptive, physiological state (or, posterior belief). As aforementioned, the quantity of water AN patients will consume until feeling completely full is less than that of HCs (Brown et al., 2022), however, it remains unclear how beliefs about a fullness state they may experience in the future and the quantity of water they believe is needed to reach complete fullness may change over time in combination of this 'one-off' experience and prior beliefs regarding quantities and fullness states. In Chapter 3, the AN group showed difficulty (in comparison to HCs') in incorporating new evidence from the cardiac interoception task they had just completed to update their self-efficacy beliefs retrospectively and prospectively (i.e., '*How well did you do just now*' and '*How well will you do in a similar task in the future*', respectively), and they also showed dysfunctional precision-weighting of the new evidence. These difficulties manifested as poor belief updating of cardiac self-efficacy beliefs and a lower rate at which new, precision-weighted information was used to update their prior beliefs. In the gastrointestinal modality, this potential difficulty in assessing new evidence such as hunger and fullness could lead to failed attempts at reducing interoceptive uncertainty and in turn explain part of the resulting maladaptive behaviours and interoceptive deficits (e.g., food restriction, impaired hunger and satiety signalling; Barca & Pezzulo, 2020). However, despite research looking at perceptual and affective state differences in fullness in AN compared to controls, no study has looked at how explicit beliefs about one's abilities to predict future states of fullness may change and update over time during a task in which individuals are asked to look at a range of water quantities (one at a time), predict how full they will be post-consumption and update such beliefs post-consumption. This examination of fullness beliefs and updating, as well as the examination of parameters which may influence the prediction of a fullness state (e.g., seeing the quantity and using this visual cue to make an interoceptive inference) was the key aim of the present experiment.

To this end, the present experiment investigated how women with AN at the acute stage of illness, AN-WR women and HCs generate and update perceptual beliefs about fullness before and after water consumption. Specifically, the aim was to examine if patients with AN demonstrate perceptual differences in their ability to predict how full drinking certain amounts of water will make them, how full they actually are after consuming the different amounts of water, and how they update predictions about future fullness states based on the experience they just had about drinking a certain water quantity, all compared to HCs. Moreover, the role of parameters which could explain group differences such as comorbid traits which could influence one's ability to accurately perceive fullness states and the role of subjective uncertainty on the accuracy of one's predicted and updated estimates of fullness (hereafter referred to as confidence), were also examined. To also delineate whether the predicted differences between the AN and HCs are state or trait effects, AN weight-restored individuals were also recruited. To address these aims, a water-sipping task was developed – specifically, participants were asked to report how full they were before and after drinking fixed amounts of water. Before drinking the water, participants were shown the amount of water (which they did not know yet they would then have to drink) and were asked to predict how full they would be if they were to drink that amount. Participants were asked to make another fullness prediction after consuming the water to examine the way in which predictions of perceived fullness states are updated. To account for perceptual differences in one's ability to estimate and predict the effect of different water quantities on fullness, the quantities of water they were shown and consumed corresponded to a lower or higher quantity with respect to the total volume of the pre-marked, unopened bottle (hereafter referred to as Low and High Quantity condition). The measures of interest were the following: (i) the difference between participants' reported state of fullness at the beginning of the trial and predicted fullness estimate (hereafter referred to as Prior Prospective Fullness Belief), and (ii) subsequent estimates (hereafter referred to as Experienced Retrospective Fullness Belief and Posterior Prospective Fullness Belief) to then estimate (iii) the prediction error (difference between evidence, i.e., Experienced Retrospective Fullness Belief (where it is calculated as the incremental change from the pre-sipping to post-sipping fullness state) and Prior Prospective Fullness Belief), and (iv) updating rate (here, updating rate refers to how participants adjust their Posterior Prospective Fullness Belief based on the evidence they have from the task).

In line with evidence in another domain (perceived and emotional body size estimation; see Saramandi et al., 2022), the noted difficulties in updating beliefs (see Chapters 2 and 3 for

belief updating difficulties in the cardiac interoception modality) and the difficulty of AN patients in perceiving gastrointestinal cues (see above), it was predicted that the difference in Prior Prospective Fullness Beliefs would be greater in the two clinical groups as compared to the HCs, and this would also be the case in the difference between the Experienced Retrospective and Posterior Prospective Fullness Beliefs, especially in the High vs Low Quantity condition. It was also predicted that the Prediction Error would be greater in the clinical groups as compared to the HCs, especially in the High Quantity condition in comparison to the Low Quantity condition, and that the two clinical groups would update their fullness estimates at a lower rate than the HCs. Specifically, the clinical groups would rely less on the fullness experience they had post-water consumption and more on prior beliefs or even their fullness state before engaging in the water sipping task when asked to predict how full consuming certain amounts of water (indicated to them from the experimenter) would make them. Given previous findings on the role of comorbid traits, namely depression and stress, on cardiac belief updating (see Chapter 3), this present experiment also looked at whether these two traits as well as anxiety, alexithymia and intolerance of uncertainty would be mediating the predicted group differences in fullness estimates (using the Baron and Kenny (1986) for mediation analysis; see below). In secondary analyses differences in confidence estimates (which can be used as precision-proxies) regarding the groups' fullness estimates were also explored, expecting the AN group to have greater subjective uncertainty regarding fullness state beliefs (i.e., poorer confidence). This prediction is in line with the findings of Chapter 3 (see also Saramandi et al., 2023). To further explore the role of state-trait effects in the ability to generate and update fullness estimates, analyses were conducted to explore if illness duration (in years) had an effect on patients' experience of fullness (in each clinical group separately). Finally, given the aforementioned links between the gastric system and body image, this experiment also explored whether overestimation of perceived and emotional states also occurs when one is asked to report how full they *are* vs. how full they *feel*. Specifically, all participants were asked to estimate how full they were and how full they felt at the beginning of the experiment (before receiving task instructions). It was predicted that, overall, the AN and AN-WR groups would overestimate their fullness in comparison to HCs, and especially when reporting feelings of fullness (i.e., "*How full do you feel?*") vs. when estimating perceived fullness (i.e., "*How full are you?*"). Moreover, the difference between the emotional and perceived estimates of fullness were predicted to be greater within the AN and AN-WR groups but not in the HC group.

4.2 Methods

4.2.1 Participants

Twenty-seven women with acute AN and 22 women weight-restored from AN (AN, and AN-WR, respectively) were recruited from the Psychiatric Daycare Unit of Sao Paulo General Hospital (ASST Santi Paolo e Carlo) in Milan, from Comunità Villa Miralago in Cuasso al Monte (VI), and from Casa di Cura Villa Margherita in Arcugnano, Italy. Age- and gender-matched healthy controls (HCs, N=48) were also recruited from the University of Milan (Italy) and University College London (UK). Due to the COVID-19 lockdown restrictions, data collection was paused in February 2020, and resumed in September 2021. Two AN-WR and one HC participants failed to comply with task instructions as confirmed by control checks during the session and had to be excluded from the final sample. As a result, the final sample consisted of $N_{AN}=27$, $N_{AN-WR}=20$ and $N_{HC}=47$ participants (Table 4.1). Patients with AN met DSM-5 criteria for restrictive subtype AN. For the same reasons as outlined in Chapter 2 (see Section 2.2) and also applied in Chapter 3 (see Section 3.2), a combination of indicators (e.g., a BMI >16.5, and clinical or behavioural signs of AN recovery for at least 6 months, and/or a global Eating Disorders Examination Questionnaire (EDE-Q; Fairburn & Beglin, 1994) score <4 (similar to the inclusion criteria of Jenkinson et al., 2023)). All HCs had a healthy BMI between 18.5 and 25 and were not included if they had a history of an eating disorder, or a first-degree relative with a history of an eating disorder. All participants were aged between 18 and 45 years, and had no history of neurological disease, brain damage, drug dependence or severe psychiatric disease (e.g., psychosis). Other exclusion criteria included being male, being pregnant, substance and/or alcohol abuse.

The participants from the Italian site received £10 Amazon vouchers, in line with local site regulations on participation compensation. The HCs from the British site were compensated at a fixed rate of £7.5/hour, in line with the participant payment regulations at the time of recruitment and testing. Institutional ethics approval was granted for the study which was conducted in accordance with the World Medical Association Declaration of Helsinki. Written informed consent was obtained from all participants prior to their participation.

Table 4.1. Participant Characteristics and Clinical Profiles.

	AN	AN-WR	HC	Test Statistic [β , (SE)], p^1	
	Mean (SD)	Mean (SD)	Mean (SD)	AN vs HC	AN-WR vs HC
N	27	20	47	-	-
Age	21.52(6.46)	23.2(5.47)	23.04(4.36)	-1.52(1.27), 0.234	0.16(1.41), 0.911
BMI	16.47(1.47)	19.92(1.74)	21.45(2.21)	-4.99(0.47), <0.001	-1.53(0.51), 0.004
Illness Duration (years)	4.15(6.33)	6.52(7.11)	-	-	-
EDE-Q Total	1.59(1.52)	0.92(1.00)	1.02(1.13)	0.56(0.30), 0.061	-0.10(0.33), 0.752
DASS-21 Depression	20.81(12.28)	12.2(10.03)	7.62(7.48)	13.20(2.32), 0.001	4.58(2.57), 0.080
DASS-21 Anxiety	13.63(10.08)	10.7(8.29)	6.64(5.80)	6.99(1.88), <0.001	4.06(2.07), 0.053
DASS-21 Stress	23.56(11.02)	20.7(10.39)	13.23(8.60)	10.32(2.35), <0.001	7.46(2.60), 0.005
TAS-20 Total	56.37(15.05)	47.1(9.15)	43.74(11.3)	12.63(2.93), <0.001	3.36(3.23), 0.302
EDI-3-ID	12.11(8.44)	15.65(6.73)	15.30(3.72)	-3.19(2.20), 0.151	0.35(2.43), 0.885
IUS-12 Total	36.81(13.00)	34.60(12.73)	27.62(8.29)	9.19(2.64), 0.001	6.98(2.88), 0.017

Psychiatric Comorbidities	5 Personality Disorder 2 Depression 1 Learning Disability 0 GAD	4 Personality Disorder 1 Depression 0 Learning Disability 1 GAD	N/A	-	-
Current Medication	8 Antidepressant 3 Anxiolytics 1 Antipsychotics 1 Sedatives 2 SSRIs	5 Antidepressants 0 Anxiolytics 0 Antipsychotics 1 Sedatives 1 SSRIs	N/A	-	-

N.B. Dashes indicate measure not taken. The HC group had no psychiatric comorbidities and was on no medication (N/A). Abbreviations: AN (Anorexia Nervosa); AN-WR (Anorexia Nervosa Weight-Restored); HC (Healthy Control); GAD (Generalised Anxiety Disorder); SSRIs (Selective Serotonin Reuptake Inhibitors); EDE-Q (Eating Disorder Examination Questionnaire; Cronbach's $\alpha=0.92$); DASS-21 (21-Item Depression, Anxiety and Stress Scale; Cronbach's $\alpha=0.95$); TAS-20 (20-Item Toronto Alexithymia Scale; Cronbach's $\alpha=0.88$); EDI-3-ID (Interoceptive Deficits subscale from the Eating Disorders Inventory 3; Cronbach's $\alpha=0.85$); IUS-12 Total (12-Item Intolerance of Uncertainty Scale; Cronbach's $\alpha=0.94$). ¹Linear regressions were run to examine group differences, with HC as the intercept. As expected, there were significant between group differences on BMI, consistent with the inclusion criteria. In this sample, no statistically significant differences were found in EDE-Q scores. Given some of the expected group differences in psychometric traits, these were taken into account in the main analyses (see Data Analysis and Results). Bolded values denote statistical significance ($p<0.05$).

4.2.2 Design and Data Analysis

The existing Water Load Test-II (van Dyck et al., 2016, 2021) was adapted to create a task which assesses perception of fullness. The aim was to explore how perceived gastric fullness and related estimates are formed and updated and how they differ between clinical groups (AN and AN-WR) in relation to HCs. The task is hereafter referred to as the Water-Sipping Task. Specifically, in the Water-Sipping Task the experimenter showed the participant a pre-marked amount of water on an unopened bottle (750ml in total volume) and asked them to provide an interoceptive, fullness estimate before and after consuming the indicated amount. The different types of estimates (estimating fullness at baseline, upon seeing the amount on the bottle, and after drinking the water) entail an interoceptive, and an external estimate and a combination of the two (see below for the specifics). Further, to examine how the varying estimates of fullness participants gave throughout the task differed, the amount of water participants were shown at each trial and subsequently asked to consume corresponded to a “Low Quantity” or “High Quantity” condition. Specifically, the “Low Quantity” and “High Quantity” condition had two levels each, as follows. In the “Low Quantity” condition, the pre-marked quantity that the experimenter showed the participants and asked to estimate how full they would be after consuming it corresponded to 2% and 5% of the entire bottle’s volume (in separate trials; equal to 15ml and 37.5ml respectively). In the “High Quantity” condition participants were shown - across two separate trials - amounts that corresponded to 12% and 15% of the entire bottle’s volume (equal to 90ml and 112.5 ml, respectively). However, at no point during the session did the experimenter reveal the total volume of the water bottle to the participants, or the actual amount they were asked to consume (see below for how this was achieved).

The interest was in examining if the groups’ additional fullness (in comparison to their initial fullness state at the beginning of each trial) would be significantly different, but also how they would differ in retrospective fullness estimates, and future fullness state predictions. To do this, fullness measures at different timepoints were obtained (Figure 4.1), and these measures included: (i) a Post-Instructions Pre-Sipping Baseline Fullness State Rating (i.e., “*How full are you now?*”), taken at the beginning of each trial, (ii) a First Fullness State Guess (i.e., “*If I gave you this amount of water to drink (experimenter showed the point on the bottle which corresponded to the respective water percentage, e.g., 2%), how full do you think you would be?*”), taken before participants were asked to drink the corresponding quantity from a water cup, (iii) a Post-Sipping Fullness State Rating (i.e., “*How full are you now?*”), taken after

participants consumed the requested water quantity and (iv) a Second Fullness State Guess (“*How full will you be if you were to drink the same amount again?*”), taken after the participants was shown the same quantity again and asked to make a future fullness state prediction; all fullness estimates were measured on a 0 (*not at all*) to 100 (*extremely full*) scale. Each rating was obtained four times in total – once per trial. The two conditions (Low and High Quantity) were counterbalanced between participants; however, the water percentages within each condition were not counterbalanced - all participants completed the lower percentage first and then the higher percentage (i.e., 2% then 5%, and 12% then 15%).

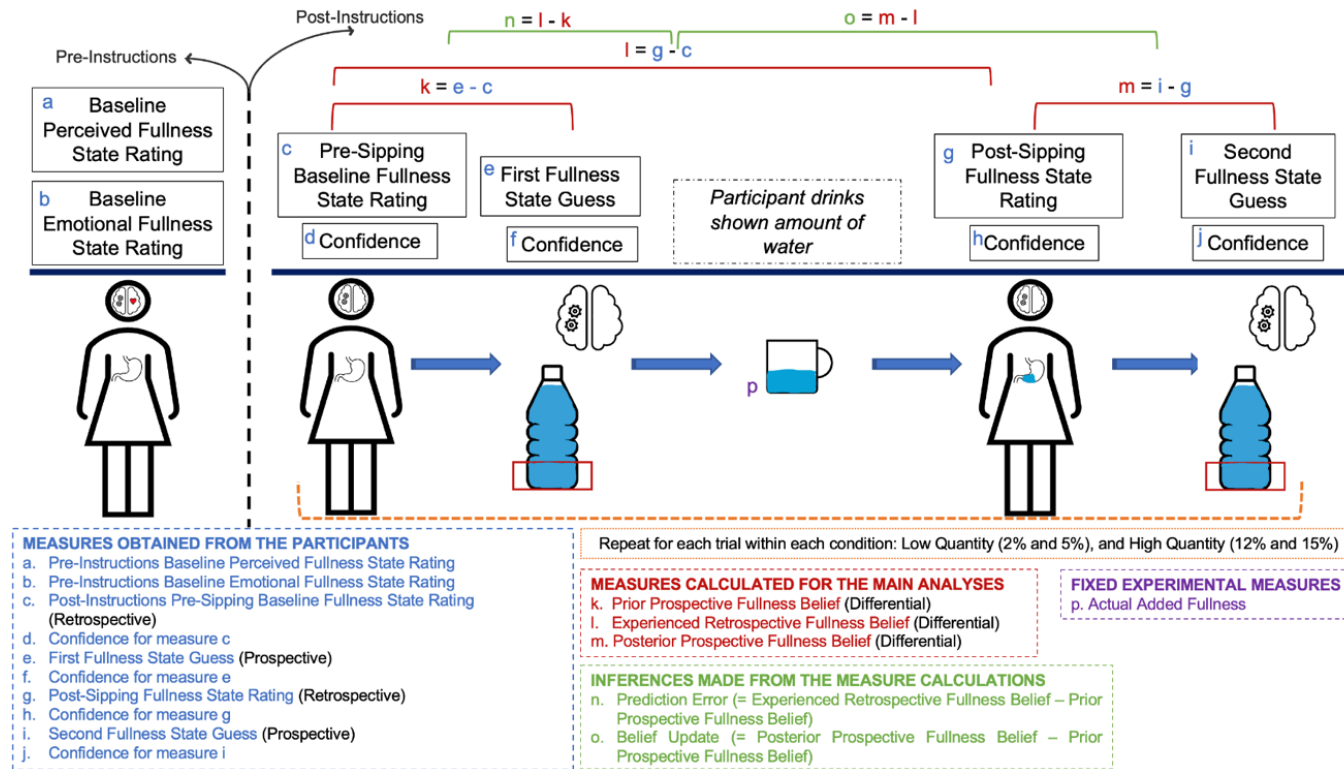


Figure 4.1. Trial Timeline of the Experiment and Experimental Measures.

All participants gave two baseline estimates (Measures a and b) before beginning the Water-Sipping Task. Participants then completed four trials (2 per Condition). They were first asked to report how full they were (Measure c) and were then shown the pre-marked quantity on the water bottle and asked to report how full they would be after consuming that amount (Measure e). Next, the experimenter poured the indicated amount into a glass of water and asked the participant to drink it. After drinking the water, participants were asked to report how full they were (Measure g). Then, the experimenter showed them the same pre-marked bottle, and asked them (again) how full they would be if they were to drink that much water again (Measure i). N.B. Measure p was fixed across all participants as it corresponds to the amount of additional fullness participants should have reported if their Experienced Retrospective Fullness Belief (Measure l) reflected the amount of water they consumed (with respect to the total volume of the water bottle). Thus, the Actual Added Fullness would equal to 2%, 5%, 12% and 15% (depending on the trial) and would be fixed across all participants.

4.2.2.1 *The Visual/Exteroceptive Aspects of the Water-Sipping Task*

Although participants were asked to provide interoceptive estimates (“*How full are you?*” and “*How full will you be?*”) the task design also consisted of an exteroceptive component – when guessing how full drinking the different water quantities would make them, participants had to use the external, visual information they had in combination with interoceptive information and inferences. Thus, this visual aspect of the water task was also taken into consideration when analysing the data and interpreting the results. Participants were shown the different water quantities on a separate bottle for each trial on which the experimenter had pre-marked the water quantity they would then be asked to consume. Any information regarding the bottle’s brand, or other identifiable characteristics and features (e.g., the label) were carefully removed so that participants would not make inferences based on their general knowledge regarding the water bottle’s brand and its typical, overall volume. The water bottle brand that was used was not a popular brand to reduce the likelihood of participants recognising the water bottle even without any labels or branding information, and the bottle shape was also atypical, i.e., although all water bottles contained 750ml, their height and/or circumference was slightly different which meant that in some bottles the water went all the way up to the middle of the bottleneck, while in other bottles the water reached just below the bottleneck. At the very end of the experiment, the experimenter asked the participants to estimate how many ml’s they thought were contained in the water. Their answer was used to control for potential effects on their fullness estimates (hereafter referred to as knowledge about the bottle quantity).

4.2.3 **Analyses on Primary Hypotheses**

4.2.3.1 *Main Behavioural Analyses*

To test the hypotheses on differences regarding how individuals with AN and weight-restored from AN think about their gastric fullness states, and how they update their estimates of fullness after completing a water-sipping task in comparison to HCs, measurements of perceived fullness were obtained at three stages during the experiment which were then used to calculate a (i) Prior Prospective Fullness Belief (measure k, Figure 4.1), an (ii) Experienced Retrospective Fullness Belief (measure l, Figure 4.1), and a (iii) Posterior Prospective Fullness Belief (measure m, Figure 4.1). The Prediction Error (measure n, Figure 4.1) and the Belief Update (measure o, Figure 4.1) scores of the participants were also calculated (see Figure 4.1 and below) to then also compare Updating Rates between the participants, i.e., examine the rate at which participants changed and updated their fullness state beliefs/ratings after completing the water-sipping trials. Secondary and exploratory analyses were also run to

examine the secondary aims. The process for each hypothesis and respective analysis is described below, separately.

To examine if the groups (independent variable; IV) differed in their Prior Prospective Fullness Belief, expecting the clinical groups to show greater changes in their fullness predictions, the difference between participants' First Fullness State Guess (measure e, Figure 4.1) and Post-Instructions Pre-Sipping Baseline Fullness State Rating (measure c, Figure 4.1) was first calculated and then a multilevel model analysis (MLM) was run using Prior Prospective Fullness Belief (measure k, Figure 4.1) as the dependent variable (DV).

Next, the groups' Experienced Retrospective Fullness Belief (measure l, DV) was calculated by subtracting their Post-Sipping Fullness State Rating (measure g, Figure 4.1) from the Post-Instructions Pre-Sipping Baseline Fullness State Rating (measure c, Figure 4.1). Group differences were then examined in an MLM, predicting that the difference between the two estimates would be greater in the AN and AN-WR groups in comparison to the HCs.

Similarly, participants' Posterior Prospective Fullness Belief (measure m, Figure 4.1; DV) was calculated by subtracting the groups' Post-Sipping Fullness Rating (measure g, Figure 4.1) from their Second Fullness State Guess (measure i, Figure 4.1). Group differences (IV) were examined in an MLM, expecting to see a greater difference in the two clinical groups than the HCs.

Moreover, a Prediction Error (measure n, Figure 4.1) score was calculated for each group [calculated as the difference between the Experienced Retrospective Fullness (measure l, Figure 4.1) and the Prior Prospective Fullness Belief (measure k, Figure 4.1)], expecting that the Prediction Error scores of the AN and AN-WR groups would be greater than those of the HCs. This was examined in an MLM using the Prediction Error as the DV.

For all these aforementioned analyses Group was first used as a main effect (with Condition as a random effect), and then the interaction between Group and Condition (Low and High Quantity) was also examined.

4.2.3.1.1 Random Effects, Control Variables and Data Preparation

Participant ID and trial order were added as random effects. If the variance of a random effect was equal to 0 (and Conditional R^2 output displayed as “N/A”) the random effect was removed from the model and the analysis was rerun. In a subset of participants ($N_{AN}=11$, $N_{AN-WR}=8$, $N_{HC}=37$), the effects of the following were also controlled for: feelings of fullness in the bladder and the stomach (on a scale of 0 - *not full at all* to 100 – *extremely full*), and knowledge about the bottle’s actual quantity (in ml) on participants’ estimates.

Participants completed four trials, each of which belonged to a different quantity. The measures obtained (e.g., Prior Prospective Fullness Belief) represent how much fuller participants thought they would be after consuming a certain water quantity that was indicated to them (and which represented – unbeknownst to them – 2%, 5%, 12%, or 15% of the entire bottle’s quantity). If participants’ answers were kept at face value for the analyses, the scores would not be able to accurately inform how fullness predictions and experiences changed across the different quantities. For instance, if a participant reported they are 50% full, and drinking 5% of the bottle’s quantity would make them 53% full (i.e., 3% fuller), but they also said that after drinking what equals to 12% of the bottle’s quantity would make them be 58% full (e.g., after reporting to be 55% full before drinking the water), then in both instances the participant predicts a 3% change in fullness. However, despite the additional predicted fullness in both instances being equal to 3%, the predicted fullness in proportion to the amount of water that the participant saw and used to make the fullness inference is higher in the 5% trial in comparison to the 12% trial. Thus, comparing the differences across trials by keeping the measures at face value would not accurately show how each different percentage is expected to influence participants’ fullness. To account for this, all measures obtained from the participants when asked to report how full they are and how full they will be (see Figure 4.1) were factored to one scale unit and represent the change in participants’ fullness state predictions and expectations if they were shown and subsequently drank what would be equal to 1% of the bottle’s entire quantity each time.

4.2.3.1.2 Mediation Analyses

As a second step to the main analyses, where significant differences were observed between one or both of the clinical groups in comparison to the HCs, the potentially mediating role of comorbid traits using the steps for mediation analysis was explored (Baron & Kenny, 1986). The following mediation analysis steps were applied (see Figure 4.2). First, I examined the

differences on the main DVs, as aforementioned (i.e., the effect of the IV (Group) on the separate DVs [namely, Prior Prospective Fullness Belief, Experienced Retrospective Fullness Belief, Posterior Prospective Fullness Belief, and Prediction Error; measures k, l, m, and n (Figure 4.1), respectively]. If significant, the role of traits was then explored. To select which traits would be used as mediator variables, I examined if the IV (Group) was a significant predictor of the traits of interest. Only the traits which differed significantly between the three groups were kept (see Table 4.1 for the traits, group scores (M and SD) and statistical differences), namely Depression, Stress, Anxiety, alexithymia and intolerance of uncertainty, as measured via the Depression, Anxiety and Stress scales of the DASS-21, TAS-20 and IUS-12, respectively. Then, the next step examined if the mediator variables also predicted the DVs of interest, in separate analyses (i.e., one trait at a time for each DV separately). However, none of the analyses were significant and thus the second part of the mediation analysis to determine if there was a partial or full mediation effect was not conducted.

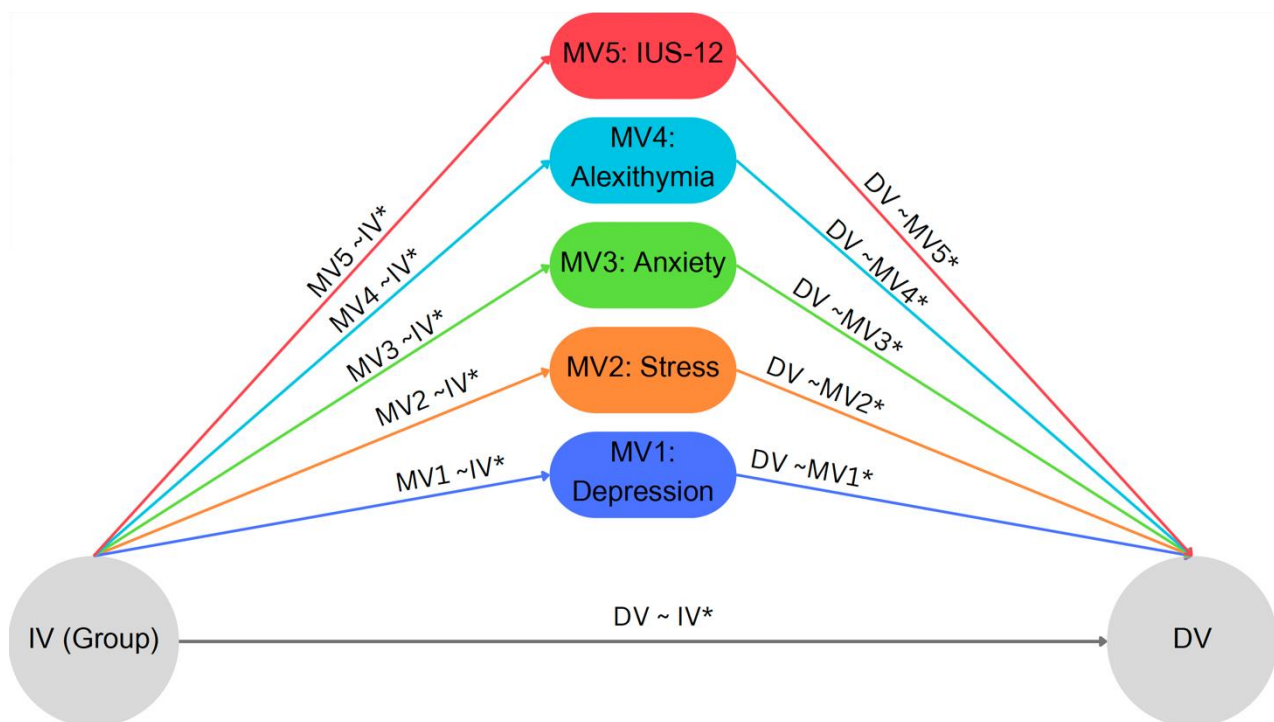


Figure 4.2. Schematic illustration of the Mediation Analysis steps (Baron & Kenny, 1986).

The Figure shows that initially one examines if the IV (Independent Variable; Group) significantly predicts the DV (Dependent Variable). If significant, one progresses to the steps of the mediation analysis. That is, first one examines if the IV is a predictor of the Mediator Variables (MVs) of interest (here, MV1 – MV5). If significant, one then examines if the MVs of interest (in separate analyses for each MV) significantly predict the DV. If significant, both IV

and $MV(s)$ are added onto a regression model to examine if there is a partial, full, or no mediation effect at all ($DV \sim IV + MV$). Abbreviations: IUS-12 (12-Item Intolerance of Uncertainty Scale). * represents the examination of whether the prediction is statistically significant (p values <0.05)

4.2.3.2 Updating Rate Analyses

Next, to examine if the groups differed in the rate at which they updated their fullness estimates, hereafter defined and referred to as “Perceived Fullness Updating Rate”, I calculated the difference between their Posterior Prospective Fullness Belief (measure m , Figure 4.1) and Prior Prospective Fullness Belief (measure k ; hereafter referred to as Belief Update; see Measure o in Figure 4.1) and investigated how the slopes of the two clinical groups compared to the slope of the HCs (main effect). Specifically, given the design constraints of the experiment, one cannot properly and accurately examine which parameters best predicted the participants’ Posterior Prospective Fullness Belief and relevant measures under the Bayesian Belief Updating Framework, such as the optimality of the participants’ learning rate [see Sections 3.2 and 3.3 (Chapter 3) for group comparisons of Bayesian optimality, model fit and Learning Rates in the cardiac interoception modality]. Instead, following the assumptions of a Bayesian Learning Framework, the data were analysed in the form of a linear regression based on the structural equivalence between the two (as seen in Saramandi et al., 2022). To examine the between-group differences a comparison was conducted between the interaction represented by the linear regression (Group $\times$ Prior Prospective Fullness Belief) and the equation representing the Bayesian Learning Framework. The Perceived Fullness Updating Rate of the HCs was the slope of the Prior Prospective Fullness Belief and the interaction between Group and the Prior Prospective Fullness Belief could be interpreted as the difference in Perceived Fullness Updating Rates between the AN and AN-WR groups in comparison to HCs (main effect). Therefore, this allowed to compare the rates at which the three groups updated their fullness estimates. It was expected that the two clinical groups would update their fullness estimate at a lower rate than HCs, as indexed by the less steep slope in comparison to the slope of HCs. To do this comparison I applied the following methodology:

$$\text{Belief Update} \sim \text{Group} \times \text{PPFB} + \text{ERFB} - \text{Group} + (1|\text{Participant\_ID}) \quad (4.1)$$

The abbreviation ‘PPFB’ refers to the Prior Prospective Fullness Belief and ‘ERFB’ refers to the Experienced Retrospective Fullness Belief. Equation 4.1 is the linear regression that was

ran in R to compare the slopes of the three groups: as explained below, the slope differences of *Group x PPFB* were compared; and Participant ID was added as a random effect to account for the variance between participants. This equation was generated using the structural equivalence of the two following models:

Linear Regression Predictive Equation:

$$\text{Belief Update} = \text{intercept} + \text{slope}_b * \text{ERFB} + \text{slope}_a * \text{PPFB} \quad (4.2)$$

Bayesian Learning Framework Predictive Equation:

$$\text{Belief Update} = \text{UpdatingRate} * \text{ERFB} - \text{UpdatingRate} * \text{PPFB} \quad (4.3)$$

The structure of Equations 4.2 and 4.3 is equivalent and both predict changes in Belief Update (DV), allowing to examine how the groups differ in their Perceived Fullness Updating Rates by doing a slope comparison of the Prior Prospective Fullness Beliefs. Specifically, the predictive linear regression equations for the groups were:

For HCs:

$$\text{Belief Update} = b0 + b1\_b * \text{ERFB} + b1\_a * \text{PPFB} \quad (4.4)$$

For ANs:

$$\text{Belief Update} = b0 + b1\_b * \text{ERFB} + (b1\_a + b2) * \text{PPFB} \quad (4.5)$$

where, $b0$ is the intercept, $b1\_a$ is the coefficient of the Prior Prospective Fullness Belief, $b1\_b$ is the coefficient of the Experienced Retrospective Fullness Belief, $b2$ is the coefficient of the interaction between Group and Prior Prospective Fullness Belief. The AN equation also applies to the AN-WR group (with the corresponding value for $b2$).

Next, under the Bayesian Learning Framework, the predictive equations for each group, respectively, would be:

$$\text{Belief Update} = \text{UR}_{hc} * \text{ERFB} - \text{UR}_{hc} * \text{PPFB} \quad (4.6)$$

$$\text{Belief Update} = \text{UR}_{an} * \text{ERFB} - \text{UR}_{an} * \text{PPFB} \quad (4.7)$$

where UR_{hc} , and UR_{an} refer to the Perceived Fullness Updating Rates of the HC and AN groups, respectively. Assuming the Bayesian Learning Framework, one can exploit the above equivalence between equations 4.4, 4.5, 4.6, and 4.7 to claim that the intercept (b_0) should be equal to 0, $b1_a$ should be equal to $(-UR_{hc})$ and $b1_a + b2$ should be equal to $(-UR_{an})$. Taken together, the difference between the Perceived Fullness Updating Rates of the two groups would be the following:

$$UR_{hc} - UR_{an} = -b1_a + b1_a + b2 = b2 \quad (4.8)$$

where, $b2$ is the coefficient of the interaction between Group and the Prior Prospective Fullness Belief. A statistically significant $b2$ would therefore imply that the difference in the Perceived Fullness Updating Rates between the groups was significant.

This linear regression was run twice. The first analysis looked at the group differences in Updating Rate when the Experienced Retrospective Fullness Belief (measure l, Figure 4.1) is calculated as described above and in Figure 4.1, i.e., the result of the difference between participants' Post-Sipping Fullness State Rating (measure g, Figure 4.1) and First Fullness State Guess (measure e, Figure 4.1). In the second analysis the concept regarding the structural equivalence between equations (4.2) and (4.3) was used but the Experienced Retrospective Fullness Belief was replaced with the Actual Added Fullness (measure p, Figure 4.1). The Actual Added Fullness was calculated as the amount of additional fullness (in comparison to the Post-Instructions Pre-Sipping Baseline Fullness State Rating participants provided) participants should have reported if their answer in the Post-Sipping Fullness State Rating was based on the quantity they consumed with respect to the bottle's volume, and not the subjective experience of increasing fullness which varies both between individuals but also on contextual and circumstantial individual factors. This means that participants' Actual Added Fullness would have been 2%, 5%, 12% and 15%, depending on the trial they were completing. After scaling the measure down to one unit for the aforementioned reasons, the Actual Added Fullness that was used in the analysis and was fixed across all participants had a fixed value of 1. The second analysis was run to explore if using the Actual Added Fullness instead of the Experienced Retrospective Fullness Belief would lead to the same predicted differences in the Updating Rates of the, also allowing to explore Updating Rate differences between the three groups when the 'embodied feedback' they received from the water-sipping experience is assumed to be the same across all subjects.

4.2.4 Secondary and Exploratory Analyses

4.2.4.1 Exploration of Confidence Differences

Secondary and exploratory analyses were run on additional measures. First, the effect of group and the interaction with Condition on Confidence estimates was examined. Participants were asked to report how confident they were on the accuracy of their fullness estimate (i.e., after their Post-Instructions Pre-Sipping Baseline Fullness State Rating, their First Fullness State Guess, their Post-Sipping Fullness State Rating, and their Second Fullness Guess; see measures d, f, h and j in Figure 4.1) on a scale of 0 (*not full at all*) to 100 (*extremely full*). If an analysis yielded a significant group effect, the mediating role of comorbid traits on the groups' confidence estimates was then examined (see above for mediation analysis steps) - also given previous findings and theoretical propositions on the role of psychological traits on overall confidence (e.g., Hoven et al., 2019, 2023; Katyal et al., 2023).

4.2.4.2 Effects of Illness Duration

Next, an exploration of illness duration (measured in years) on fullness was conducted (in each clinical group separately). Specifically, the analyses looked at whether illness duration (defined as the years during which the participant was in the acute state before achieving and maintaining weight-restoration and without any further relapses) influenced patients' perception of fullness (measures k, l, and m, Figure 4.1).

4.2.4.3 Feeling vs Being Full – Exploring Fullness State Differences

Finally, for exploratory purposes, participants were asked to provide two baseline ratings of fullness on a scale of 0 (*not full at all*) to 100 (*extremely full*) at the very beginning of the experiment before receiving task instructions (measures a and b, Figure 4.1). For the first rating, hereafter referred to as “Perceived Fullness”, participants were asked ‘*How full are you now?*’. For the second rating, hereafter referred to as “Emotional Fullness”, participants were asked “*How full do you feel now?*”. Both ratings assessed participants' fullness but in the former participants were asked to answer based on objective parameters, e.g., how full they were given their last meal and drink time, whereas in the latter participants were encouraged to focus on gastric feelings of fullness in that moment, irrespective of how full they should have been given their last meal and drink time. Between- and within-group differences in the rating mode (i.e., Perceived vs Emotional) were then examined, expecting the AN and AN-WR groups' Perceived and Emotional Fullness Ratings to be significantly higher than the respective ratings of the

HCs. It was also expected that the two clinical groups' Emotional Fullness Ratings would be significantly higher than their respective Perceived Fullness Ratings, in line with previous findings in another domain showing that feelings about one's body size and shape are associated with greater overestimation than perceptual ratings (see Introduction for more and (Saramandi et al., 2022)).

4.2.5 Transparency and Openness

All statistical analyses were performed using R and figures were generated using ggplot 2 (Wickham, 2016). The results section only present the main statistical results for brevity and coherence purposes and the control and supplementary analyses are presented in Appendix C. The experimental design but not analyses were preregistered as part of the main preregistration of the belief updating about cardiac interoception task aims, design and analysis (presented in Chapter 3; see preregistration here: <https://osf.io/x4ysv>). The Water-Sipping Task was added as an exploratory task to test whether the patterns observed in the cardiac modality of women with and weight-restored from AN, and matched HCs would be similar to those observed in the gastrointestinal modality of interoception. However, due to limitations and errors observed during the testing and data checks, in the present Chapter the design of and results from the Subjective block (i.e., wherein participants estimated how much water they need to consume in order to feel 2%, 5%, 12%, and 15% fuller) are not presented. Data and code used in the manuscript are available via GitHub: <http://tinyurl.com/ytp968c6>.

4.2.6 Materials and Measures

To conduct the experiment, four bottles of still water (750ml) were used. Each bottle was pre-marked at the height in which the water equalled to each corresponding percentage. The experimenter then used a fifth bottle to pour the pre-measured quantity in the cup the participant drank the water from.

4.2.6.1 Psychometric Measures

To examine the role of comorbid traits in fullness estimation, participants also filled out self-reported questionnaires used to assess depression, anxiety, stress, interoceptive deficits, alexithymia and intolerance of uncertainty (DASS-21, Lovibond & Lovibond, 1995; EDI-3-ID, Garner, 2004; TAS-20, Bagby et al., 1994; and, IUS-12, Carleton et al., 2007, respectively; see Chapters 2 and 3 for details on these measures' characteristics and descriptions). As detailed

in Chapters 2 and 3, the Italian-validated versions of the DASS-21, TAS-20, and IUS-12 were used, and the EDI-3 were translated and back-translated by members of the research team.

4.2.7 Procedure

First, control measurements for the task (i.e., time since last meal and drink) and the Pre-Instructions Baseline Perceived and Emotional Fullness Ratings were obtained. Before providing the two Pre-Instructions Baseline Fullness Ratings, participants were explained that fullness can depend on multiple factors (e.g., time elapsed since the last meal, the portion that was consumed, potential bloating experienced after consuming a certain type of food) and in the next two ratings they would have to provide one rating based on objective parameters (i.e., how full they should be in that moment given their last meal time and size; i.e., Pre-Instructions Baseline Perceived Fullness Rating), and one based on subjective parameters (i.e., how full they feel in that moment; i.e., Pre-Instructions Baseline Emotional Fullness Rating; instructions script in Appendix C). Then, after participants were given detailed instructions about the task, they completed the four trials across the two (Low and High Quantity) conditions (Figure 4.1). For each trial within the Low and High Quantity condition (2% and 5%; 12% and 15%, respectively), the experimenter first asked the participant to report how full they were and then showed them the first pre-marked bottle and asked them to estimate how full they would be after consuming the suggested quantity. Next, the experimenter poured the equivalent amount of water in a clean cup and asked the participant to drink it. After the participant consumed the water, they were asked to report how full they were. Participants were then shown the same bottle again and were asked to make a new fullness estimate prediction. However, they were not asked to drink that suggested quantity again. This procedure was completed for the four trials in total (i.e., two in the Low and two in the High Quantity condition).

At the end of the Water-Sipping Task, participants were asked to estimate the total volume of water (in ml's) they thought was contained in the bottle. Not all participants came to the session on an empty bladder and stomach and although Pre-Instructions Baseline Fullness Estimates (Perceived and Emotional) these estimates do not capture the extent to which feelings of fullness may have influenced participants' subsequent answers. To measure this, participants were asked to report the extent to which their fullness estimates during the session were influenced by feelings of fullness in their bladder and stomach – this was asked at the end of the session, after participants had completed all four water-sipping trials. These two measures (feelings of fullness in the bladder and stomach) were added as control variables in the analyses.

If a participant had to use the bathroom during the session, this was noted and reported below. Moreover, it was expected that some AN patients would potentially not wish to consume the water out of fear and other symptoms linked to the AN symptomatology and clinical behaviour. Where this was the case, it was noted and taken into consideration in the analyses.

4.3 Results

4.3.1 Primary Hypotheses Results

Women weight-restored from AN had a significantly higher Prior Prospective Fullness Belief than HCs.

It was predicted that both clinical groups would expect to become significantly fuller than HCs after drinking the indicated amount of water, in comparison to HCs and especially in the High Quantity Condition. That is, the predicted change in the AN and AN-WR groups' fullness in comparison to how full they report to be at the beginning of the trial will be greater than that of the HCs. Contrary to the prediction, there was only a significant main effect of Group in that the AN-WR group's Prior Prospective Fullness Belief was higher than that of HCs, but this was not the case for the AN group (Figure 4.3; Table 4.2). Although it was also predicted that participants, especially those in the clinical groups, would have a higher Prior Prospective Fullness Belief in the High vs Low Quantity condition, there was no significant interaction between Group and condition. However, there was a marginally significant main effect of condition. Specifically, the Prior Prospective Fullness Belief of all participants in the High Quantity condition was significantly lower than that of the Low Quantity condition (although this should be interpreted with caution as the result was marginally significant, $p=0.05$). This result suggests that the change in fullness participants predicted to experience was greater when they were shown the two lower percentages (2% and 5%) in comparison to the higher percentages (12% and 15%) – this finding is discussed further in the Discussion (Section 3.4). In a subset of participants, the effects of stomach and bladder fullness, as well as knowledge about the bottle's total volume on Prior Prospective Fullness Beliefs were also controlled for but non-significant effects were found (all $p>0.05$). The main effect of group remained significant, while the Condition main effect was no longer statistically significant ($p=0.796$; Appendix C, Table C.1a). To further explore what could drive the overestimation of the AN-WR group only and not the AN group, in comparison to HCs, a control analysis of group-differences on the Post-Instructions Pre-Sipping Baseline Fullness State Rating was conducted. One linear regression per quantity trial (i.e., 2%, 5%, 12%, and 15%) was conducted in order to investigate the baseline differences in fullness at the beginning of each trial, separately.

Interestingly, although the group differences were not statistically significant, the average Post-Instructions Pre-Sipping Baseline Fullness State Rating (Baseline) of the AN-WR group was lower than that of HCs and that of the AN group was higher (Appendix C, Table C.2a – C.2e). This was the case across all four trials and the role of this observation in relation to the results observed in this experiment is discussed in further detail in the discussion.

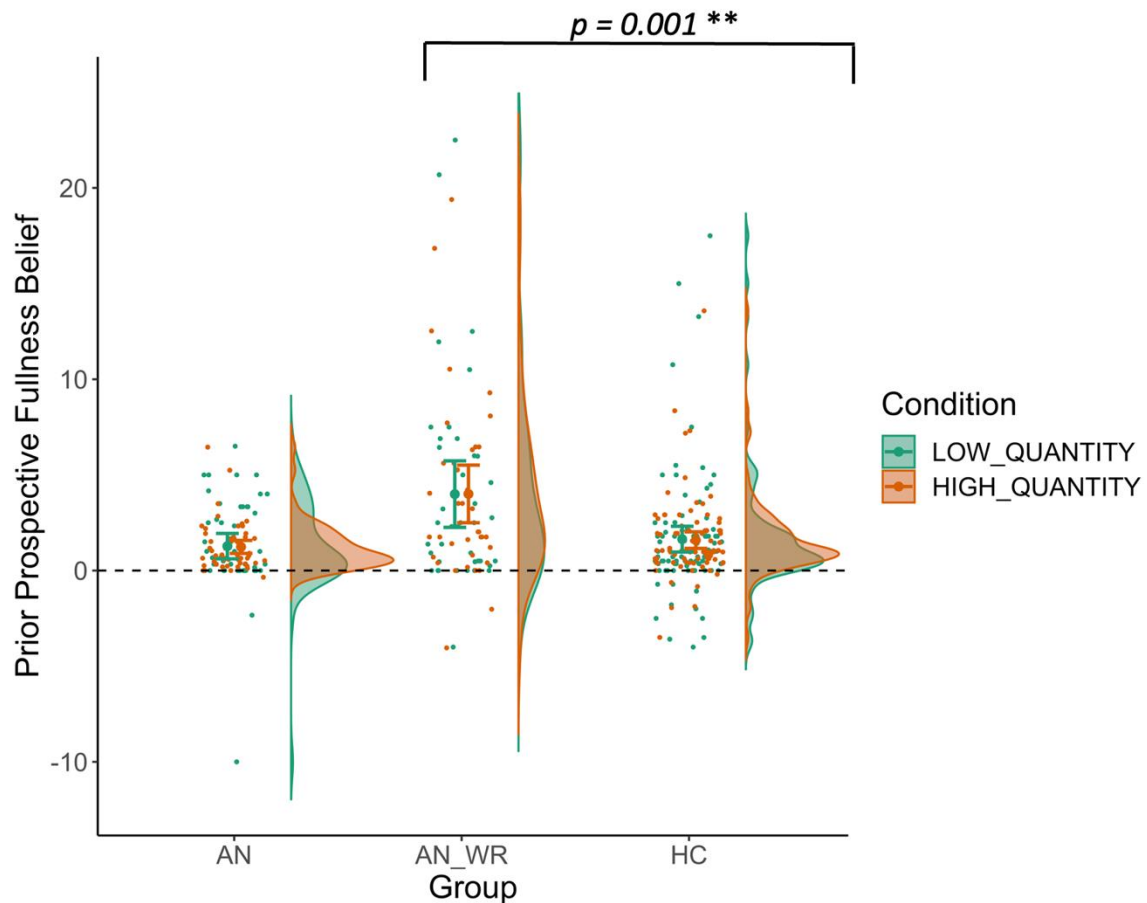


Figure 4.3. Between-group differences on Prior Prospective Fullness Beliefs separated by Condition (Low and High Quantity, i.e., 2%, 5% and 12%, 15%, respectively)

Table 4.2. Prior Prospective Fullness Belief Analyses

<i>Predictors</i>	β (SE)	t	95% CI	<i>p</i>
Group Main Effect				
(Intercept)	2.04(0.53)	3.88	1.01 - 3.08	<0.001
AN	-0.26(0.68)	-0.38	-1.59 - 1.08	0.706
AN-WR	2.47(0.74)	3.34	1.01 - 3.93	0.001
Marginal R ² / Conditional R ²	0.107 / 0.785			
Interaction between Group x Condition				
(Intercept)	2.29(0.48)	4.72	1.34 - 3.24	<0.001
AN	-0.02(0.73)	-0.30	-1.45 - 1.40	0.976

AN-WR	2.32(0.78)	2.97	0.79 - 3.86	0.003
Condition (High Quantity)	-0.48(2.25)	-1.95	-0.97 - 0.00	0.051
AN x Condition (High Quantity)	-0.38(0.42)	-0.89	-1.21 - 0.45	0.375
AN-WR x Condition (High Quantity)	0.27(0.46)	0.59	-0.63 - 1.17	0.556
Marginal R ² / Conditional R ²	0.116/ 0.785			

Abbreviations: AN (Acute Anorexia Nervosa Group); AN-WR (Weight-Restored from Anorexia Nervosa Group); HC (Healthy Controls Group). Bolded values denote statistical significance ($p < 0.05$).

Women weight-restored from AN reported to be significantly fuller than HCs after drinking the indicated quantity of water.

After ingesting the indicated water quantity, the Experienced Retrospective Fullness Belief of the AN-WR group was significantly higher than that of HCs, but this was not the case between the AN and HC groups (**Error! Reference source not found.**; Table 4.3). This finding suggests that even after drinking the same amount of water, the AN-WR group showed a greater change from their Pre-Instructions Pre-Sipping Baseline Fullness State Rating in comparison to HCs. However, when adding Condition as an interacting IV, the main effect of Group was no longer significant and the two-way interaction between Group and Condition was also not statistically significant. Thus, further control tests to examine the effects of feelings of fullness in the bladder and stomach and knowledge about the bottle's total volume on the DV were not performed.

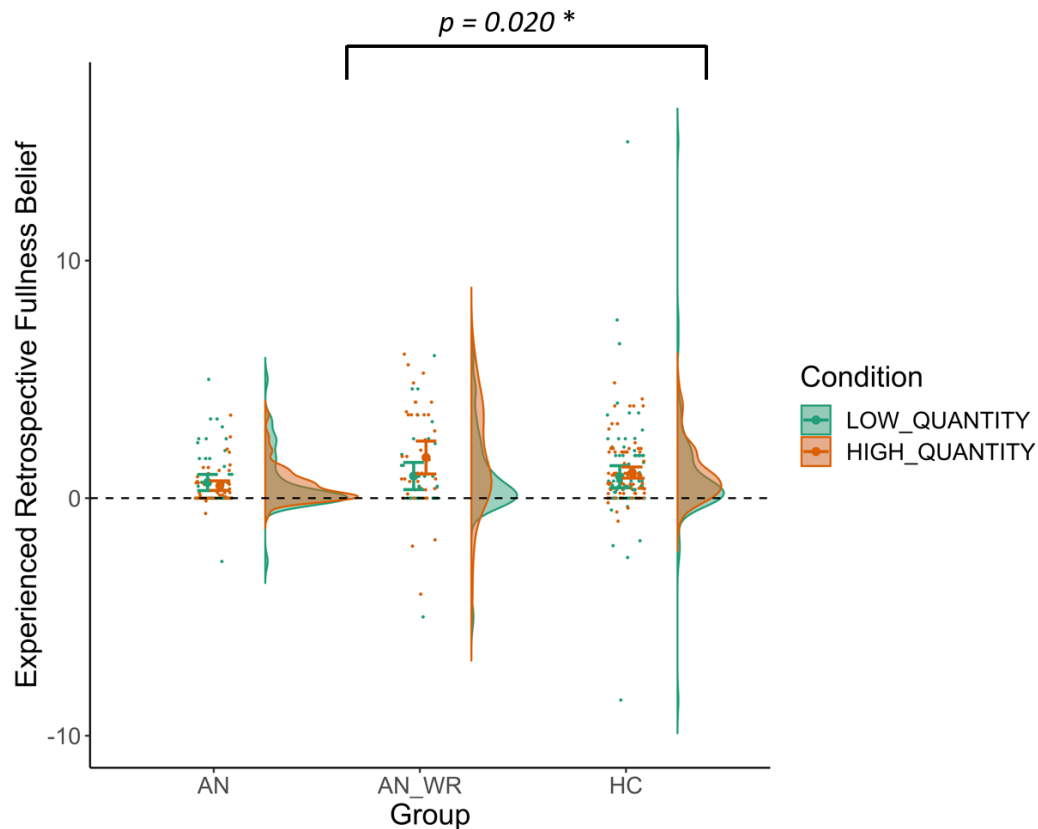


Figure 4.4. Between-group differences on Experience Retrospective Fullness Beliefs separated by Condition (Low and High Quantity, i.e., 2%, 5%, and 12%, 15%, respectively).

Table 4.3. Experienced Retrospective Fullness Estimate Analyses

<i>Predictors</i>	β (SE)	t	95% CI	<i>p</i>
Group Main Effect				
(Intercept)	1.50(0.18)	8.57	1.16 - 1.85	<0.001
AN	-0.33(0.29)	-1.13	-0.89 - 0.24	0.258
AN-WR	0.70(0.30)	2.33	0.11 - 1.29	0.020
Marginal R ² / Conditional R ²	0.049 / 0.226			
Interaction between Group and Condition				
(Intercept)	1.67(0.22)	7.65	1.24 - 2.10	<0.001
AN	0.01(0.41)	0.02	-0.79 - 0.81	0.982
AN-WR	0.20(0.42)	0.47	-0.62 - 1.01	0.637
Condition (High Quantity)	-0.30(0.25)	-1.18	-0.79 - 0.20	0.238
AN x Condition (High Quantity)	-0.52(0.47)	-1.10	-1.45 - 0.41	0.271
AN-WR x Condition (High Quantity)	0.83(0.48)	1.74	-0.10 - 1.76	0.082
Marginal R ² / Conditional R ²	0.071 / 0.242			
Abbreviations: AN (Acute Anorexia Nervosa Group); AN-WR (Weight-Restored from Anorexia Nervosa Group); HC (Healthy Controls Group). Bolded values denote statistical significance (<i>p</i> <0.05).				

Women weight-restored from AN had higher Posterior Prospective Fullness Beliefs in comparison to the HCs, especially in the High Quantity Condition.

It was predicted that both clinical groups' Posterior Prospective Fullness Beliefs would be significantly higher than those of HCs, and especially in the High Quantity condition. The difference between how full participants predicted to be in the future, if they were to consume the same water quantity again, and how full they reported to be after consuming that amount (i.e., Post-Sipping Fullness Rating) was only significantly larger in the AN-WR and not the AN group in comparison to HCs when looking at the main effect of Group (Figure 4.5.; Table 4.4a). That is, the AN-WR group predicted that drinking a certain amount of water would make them significantly fuller in the future, in comparison to the estimates HCs gave (Table 4.4b). There was also a significant main effect of condition, in that participants' Posterior Prospective Fullness Belief was significantly lower in the High Quantity condition in comparison to the Low Quantity condition, contrary to the initial prediction which was of the opposite direction. Finally, a significant interaction between group and condition was found. To further determine the nature of the significant interaction effects identified in the main analysis, two Holm-corrected planned contrasts were conducted. The first planned contrast was a pairwise comparison between the AN-WR and HC groups for the Low Quantity condition, and there was a non-significant effect of Group on the Posterior Prospective Fullness Beliefs (Table 4.4b). The second planned contrast was run on the same group pairwise comparison but in the High Quantity condition instead, and it revealed that the AN-WR group had significantly higher Posterior Prospective Fullness Beliefs than the HC group when predicting a future fullness state (Table 4.4c). The AN-WR group main effect remained significant when controlling for the effects of bladder and stomach fullness, and knowledge regarding the bottle's total volume in a subset of our entire sample but the condition effect became non-significant, and the effects of the control variables were also non-significant (Appendix C, Table C.1b).

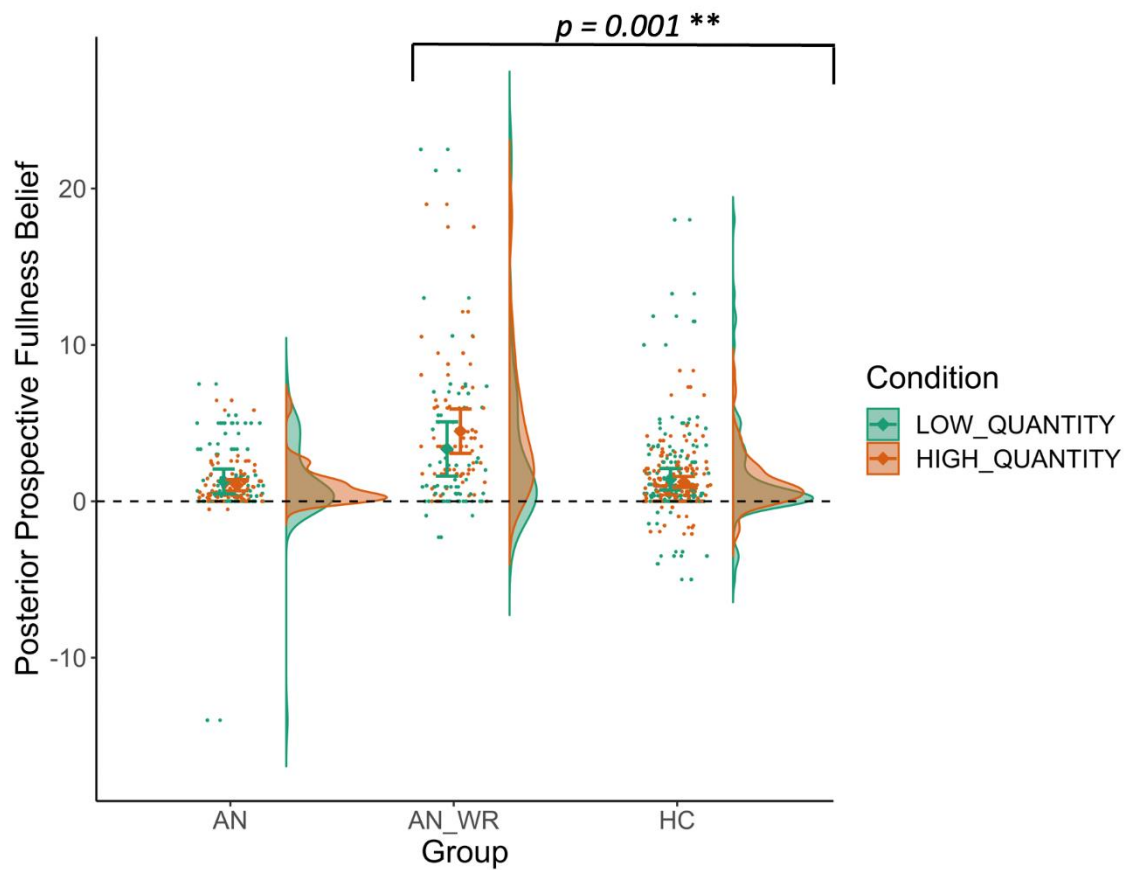


Figure 4.5. Between-group differences on Posterior Prospective Fullness Beliefs separated by Condition (Low and High Quantity, i.e., 2%, 5% and 12%, 15%, respectively).

Table 4.4. Group Differences on Posterior Prospective Fullness Beliefs.

Predictors	β (SE)	t	95% CI	p
Table 4.4a: Main Analyses				
Group Main Effect				
(Intercept)	2.10(0.77)	2.74	0.60 - 3.61	0.006
AN	-0.34(0.70)	-0.49	-1.72 - 1.03	0.627
AN-WR	2.55(0.77)	3.31	1.04 - 4.06	0.001
Marginal R^2 / Conditional R^2	0.116 / 0.787			
Interaction between Group and Condition				
(Intercept)	2.56(0.73)	3.4	1.12 - 3.99	<0.001
AN	-0.08(0.76)	-0.11	-1.56 - 1.40	0.914
AN-WR	1.63(0.83)	1.97	0.01 - 3.26	0.049
Condition (High Quantity)	-0.86(0.28)	3.10	-1.40 - -0.31	0.002
AN x Condition (High Quantity)	-0.45(0.46)	-0.99	-1.34 - 0.44	0.323
AN-WR x Condition (High Quantity)	1.51(0.48)	3.11	0.56 - 2.45	0.002
Marginal R^2 / Conditional R^2	0.131 / 0.802			

Table 4.4b: AN-WR vs. HC Group differences on Posterior Prospective Fullness Beliefs (in the Low Quantity Condition)

(Intercept)	2.63(1.24)	2.20	0.20 – 5.07	<i>0.034</i>
AN-WR	1.78(1.27)	1.39	-0.72– 4.27	<i>0.163</i>
Marginal R ² / Conditional R ²	0.034 / 0.888			

Table 4.4c: AN-WR vs. HC Group differences on Posterior Prospective Fullness Beliefs (in the High Quantity Condition)

(Intercept)	1.70(0.65)	2.64	0.44 – 2.97	<i>0.008</i>
Condition (High Quantity)	3.15(0.75)	4.20	1.68 – 4.62	0.002
Marginal R ² / Conditional R ²	0.213 / 0.879			

Abbreviations: AN (Acute Anorexia Nervosa Group); AN-WR (Weight-Restored from Anorexia Nervosa Group); HC (Healthy Controls Group). Bolded values denote statistical significance ($p < 0.05$).

Women weight-restored from AN had a greater Prediction Error than HCs.

It was predicted that the two clinical groups would have a greater Prediction Error, which would be suggestive of an overestimation of how much fuller they would be (Prior Prospective Fullness Belief) in comparison to how much fuller they were after consuming the required water quantity (Experienced Retrospective Fullness Belief). Interestingly, only the AN-WR group had a significantly greater Prediction Error score in comparison to HCs, and this was not the case when comparing the Prediction Error of the AN and HC groups (Figure 4.6.; Table 4.5). The effects remained the same when controlling for fullness estimates in the stomach and bladder and knowledge of the bottle's entire volume (Appendix C, Table C.1c). However, the interaction with Condition was not significant (Table 4.5). Interestingly, the interaction with Condition was significant when controlling for covariates (Appendix C, Table C.1c).

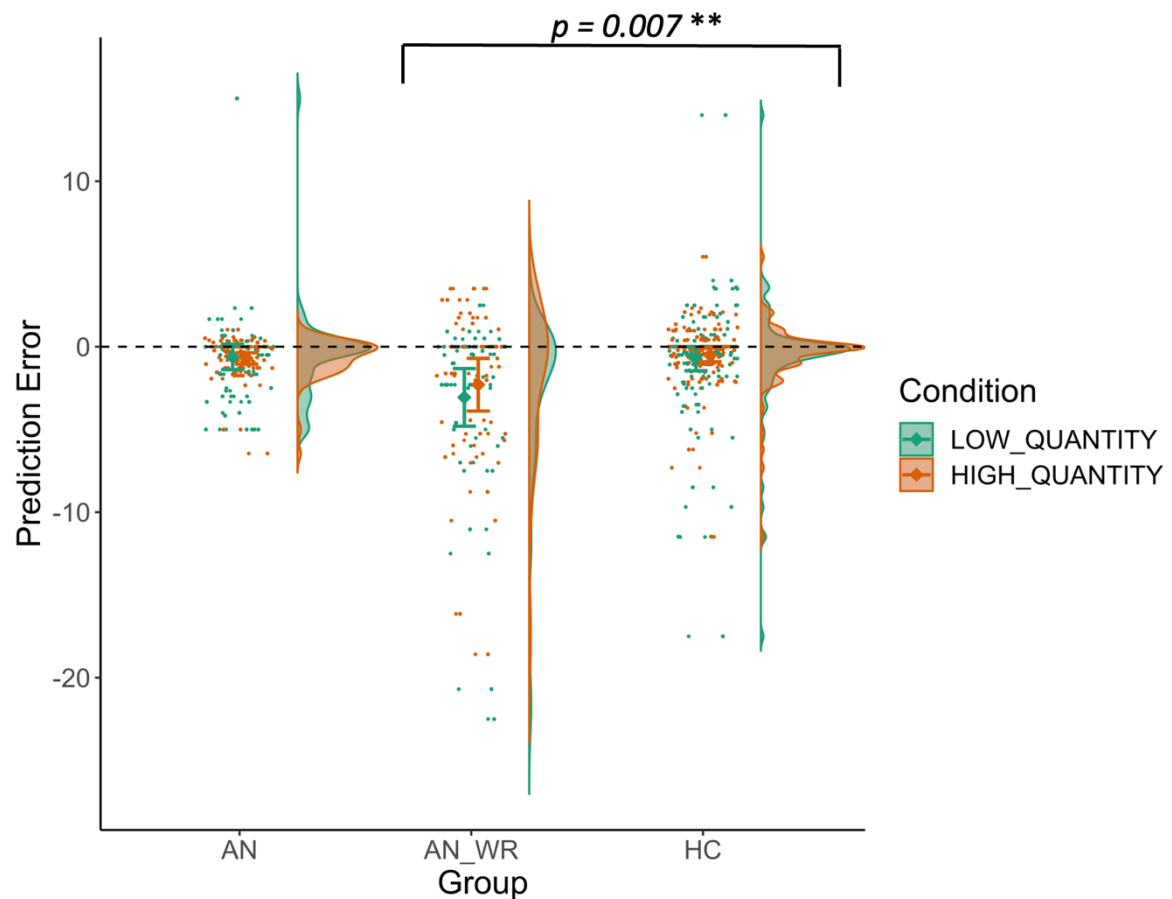


Figure 4.6. Between-group differences on Prediction Error scores separated by Condition (Low and High Quantity, i.e., 2%, 5% and 12%, 15%, respectively).

N.B. The prediction Error here was calculated as the difference between participants' Experienced Retrospective Fullness Belief and Prior Prospective Fullness Belief.

Table 4.5. Group Differences in Prediction Error Scores.

Predictors	β (SE)	t	95% CI	p
Group Main Effect				
(Intercept)	-0.60(0.43)	-1.40	-1.45 - 0.24	0.162
AN	-0.06(0.70)	-0.09	-1.43 - 1.31	0.931
AN-WR	-2.07(0.77)	-2.69	-3.59 - -0.56	0.007
Marginal R^2 / Conditional R^2	0.069 / 0.677			
Interaction between Group and Condition				
(Intercept)	-0.74(0.45)	-1.66	-1.61 - 0.13	0.097
AN	0.12(0.74)	0.16	-1.33 - 1.57	0.870
AN-WR	-2.32(0.82)	2.85	-3.92 - -0.72	0.004
Condition (High Quantity)	0.27(0.29)	0.94	-0.29 - 0.84	0.346
AN x Condition (High Quantity)	-0.36(0.48)	-0.76	-1.30 - 0.57	0.445

AN-WR x Condition (High Quantity)	0.50(0.53)	0.95	-0.53 - 1.53	0.344
Marginal R ² / Conditional R ²	0.062 / 0.678			

Abbreviations: AN (Acute Anorexia Nervosa Group); AN-WR (Weight-Restored from Anorexia Nervosa Group); HC (Healthy Controls Group). Bolded values denote statistical significance ($p < 0.05$).

Mediation Analyses

To examine the role of traits as mediator variables in the aforementioned analyses, I then explored if the aforementioned DVs of interest were significantly predicted by each trait (depression, anxiety, stress, alexithymia, intolerance of uncertainty), separately. Although these traits were significantly different between groups (see Table 4.1), they did not significantly predict the main DVs of interest (all $p > 0.05$; see statistical results in Appendix C, Table C.3). Thus, the next steps of the mediation analysis described in the Data Analysis section were not performed.

The rate at which women with and weight-restored from AN updated their fullness estimates was lower than that of HCs.

The analysis showed a significant interaction between Group and the Prior Prospective Fullness Belief variable. Based on the structural equivalence between the linear regression and belief update equations (see above), the findings indicate that the slope of the AN and AN-WR groups was steeper than that of the HC group, suggesting that the rate at which they updated their Posterior Prospective Fullness Beliefs was significantly lower than the HCs'. Interestingly, this was the case in both clinical groups (in comparison to HCs) when using the Actual Added Fullness (i.e., how much fuller all participants should have been following the water consumption; Figure 4.7A; Table 4.6a) and not when using participants' Experienced Retrospective Fullness Belief in the equation (i.e., how much fuller participants reported to be in comparison to their Post-Instructions Pre-Sipping Baseline Fullness State Rating; Figure 4.7B; Table 4.6b). When using the latter estimate, the difference in slopes was only significant between the AN-WR and HC group and not the AN and HC group.

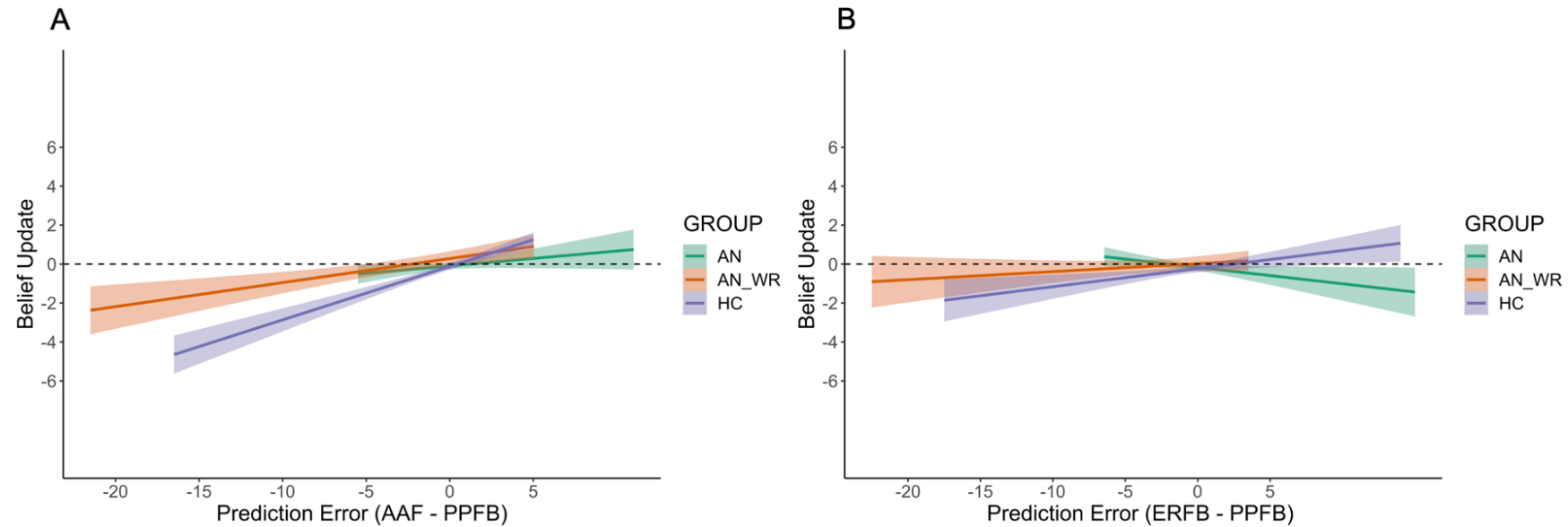


Figure 4.7. Group Differences in Updating Rates.

Panel (A) shows the slope differences when the Prediction Error is calculated using the difference between the Actual Added Fullness participants should have experienced (fixed to “1” across all participants after all ratings were scaled down to 1 unit to represent the changes participants would predict and report if they were to drink 1% of the bottle’s entire water quantity) and their Prior Prospective Fullness Belief. Panel (B) shows the slope differences when using participants’ ratings to calculate their Prediction Error scores, i.e., the difference between participants’ Experienced Retrospective Fullness Belief and Prior Prospective Fullness Belief. Abbreviations: AAF (Actual Added Fullness); PPFB (Prior Prospective Fullness Belief); ERFB (Experienced Retrospective Fullness Belief). N.B. In both graphs, Belief Update refers to the difference between participants’ Posterior Prospective Fullness Belief and Prior Prospective Fullness Belief.

Table 4.6. Group Differences in Updating Rate.

Predictors	β(SE)	t	95% CI	p
Table 4.6a: Updating Rate using Participants' Actual Added Fullness				
(Intercept)	0.26(0.14)	1.92	-0.01 - 0.53	0.055
Initial Predicted Fullness	-0.35(0.05)	-6.83	-0.45 - -0.25	<0.001
AN x Initial Predicted Fullness	0.22(0.09)	2.38	0.04 - 0.41	0.017
AN-WR x Initial Predicted Fullness	0.19(0.06)	3.03	0.07 - 0.31	0.020
Marginal R ² / Conditional R ²	0.166 / 0.357			
Table 4.6b: Updating Rate using Participants' Experienced Retrospective Fullness Belief				
(Intercept)	0.50(0.14)	3.54	0.22 – 0.78	<0.001
Initial Predicted Fullness	-0.26(0.05)	-5.42	-0.35 – -0.16	<0.001
Actual Fullness Estimate	-0.37(0.05)	-6.99	-0.48 – -0.27	<0.001
AN x Initial Predicted Fullness	0.12(0.00)	1.40	-0.05 – 0.29	0.162
AN-WR x Initial Predicted Fullness	0.16(0.06)	2.80	0.05 – 0.27	0.005
Marginal R ² / Conditional R ²	0.245 / 0.372			
Abbreviations: AN (Acute Anorexia Nervosa Group); AN-WR (Weight-Restored from Anorexia Nervosa Group); HC (Healthy Controls Group). Bolded values denote statistical significance ($p < 0.05$).				

4.3.2 Secondary Hypotheses and Exploratory Analyses

Women with AN reported significantly higher confidence than HCs following their Predicted Fullness Estimate.

Contrary to the prediction, the AN group reported significantly higher confidence estimates following their First Fullness State Guess in comparison to HCs (Table 4.7a; although note that this effect was a non-significant trend when looking at the main effect of Group only and not at the interaction between Group and Condition). Contrary to predicted between-group differences on confidence estimates following the other three fullness estimates (Pre-Sipping and Post-Sipping Fullness State Rating, and Second Fullness State Guess), no statistically significant differences were found (Tables C.4a and C.4b). Interestingly, there was a significant main effect of condition; specifically, the confidence estimates of participants following their Post-Sipping Fullness Rating were significantly lower in the High Quantity condition in comparison to the Low Quantity condition. It was further explored whether the difference in the confidence estimates after participants gave their First Fullness State Guess was mediated by any comorbid traits, namely depression, anxiety, stress, alexithymia, or intolerance of uncertainty. Using the mediation analysis steps a significant effect of stress on the DV of interest (Confidence for the First Fullness State Guess) was found, while the effect of the other traits was not significant. I then examined if stress explained the group effect and found that

stress had a partial mediation effect as the effect of group also remained significant (Table 4.7b-7c and C5.1-C5.2).

Table 4.7. Group Difference on Confidence Estimates.

<i>Predictors</i>	β (SE)	t	95% CI	p
Table 4.7a: Confidence Estimate for the First Fullness State Guess				
Main Effect of Group				
(Intercept)	69.90(2.72)	25.73	64.58 - 75.23	<0.001
AN	7.22(4.34)	1.66	-1.28 - 15.72	0.096
AN-WR	1.11(4.80)	0.23	-8.29 - 10.51	0.817
Marginal R ² / Conditional R ²	0.021 / 0.728			
Interaction between Group x Condition				
(Intercept)	70.74(2.88)	24.59	65.10 - 76.38	<0.001
AN	9.31(4.61)	2.02	0.27 - 18.35	0.044
AN-WR	2.79(5.10)	0.55	-7.21 - 12.78	0.585
Condition (High Quantity)	-1.67(1.89)	-0.88	-5.38 - 2.04	0.378
AN x Condition (High Quantity)	-4.18(3.13)	-1.33	-10.33 - 1.96	0.182
AN-WR x Condition (High Quantity)	-3.35(3.47)	-0.97	-10.15 - 3.44	0.333
Marginal R ² / Conditional R ²	0.030 / 0.637			
Table 4.7b: Post-Sipping Confidence Estimate				
Main Effect of Group				
(Intercept)	72.96(3.28)	22.22	66.52 - 79.39	<0.001
AN	6.21(4.51)	1.38	-2.62 - 15.04	0.168
AN-WR	-3.56(4.98)	-0.71	-13.32 - 6.21	0.475
Marginal R ² / Conditional R ²	0.028 / 0.737			
Interaction between Group x Condition				
(Intercept)	74.59(2.84)	26.28	69.02 - 80.15	<0.001
AN	6.60(4.70)	1.41	-2.61 - 15.81	0.160
AN-WR	-2.71(5.19)	-0.52	-12.89 - 7.47	0.602
Condition (High Quantity)	-3.26(1.61)	-2.03	-6.40 - -0.11	0.043
AN x Condition (High Quantity)	-0.78(2.66)	-0.29	-5.99 - 4.43	0.769
AN-WR x Condition (High Quantity)	-1.69(2.94)	0.58	-7.45 - 4.06	0.564
Marginal R ² / Conditional R ²	0.036 / 0.734			
Table 4.7c: Confidence Estimate for the First Fullness State Guess ~ Stress				
(Intercept)	78.11(3.58)	21.83	71.10 - 85.12	<0.001
Stress	-0.33(0.17)	-1.93	-0.67 - 0.01	0.054
Marginal R ² / Conditional R ²	0.027 / 0.624			
Confidence Estimate for the First Fullness State Guess ~ Group*Condition + Stress				
(Intercept)	78.00(3.65)	21.36	70.83 - 85.15	<0.001
AN	14.97(4.85)	3.09	5.46 - 24.47	0.002

AN-WR	6.87(5.11)	1.34	-3.16 - 16.89	0.18
Condition (High Quantity)	-1.67(1.89)	-0.88	-5.38 - 2.04	0.378
Stress	-0.55(0.19)	-2.95	-0.91 - -0.18	0.003
AN x Condition (High Quantity)	-4.18(3.13)	-1.33	-10.33 - 1.96	0.182
AN-WR x Condition (High Quantity)	-3.35(3.47)	-0.97	-10.15 - 3.44	0.333
Marginal R ² / Conditional R ²	0.090 / 0.638			

N.B. Stress was measured using the DASS-21. Abbreviations: AN (Acute Anorexia Nervosa Group); AN-WR (Weight-Restored from Anorexia Nervosa Group); HC (Healthy Controls Group). Bolded values denote statistical significance ($p < 0.05$).

Illness duration only influenced the AN group's Initial Predicted Estimates.

The AN group's Prior Prospective Fullness Beliefs were higher (i.e., expected to become a lot fuller after consuming the indicated amount of water) the longer they had been in the acute AN state (Table 4.8a). The same pattern was observed for their Posterior Prospective Fullness Beliefs (Table 4.8c). Interestingly, their Experienced Retrospective Fullness Beliefs were not influenced by illness duration (measured in years; Table 4.8b). Also, the effect of illness duration was non-significant when looking at its influence on the AN-WR group's fullness beliefs (Table 4.8d-f).

Table 4.8. Effects of Illness Duration on Fullness.

Predictors	β(SE)	t	95% CI	p
Table 4.8a: Effect of Illness Duration on the AN Group's Prior Prospective Fullness Beliefs				
(Intercept)	1.42(0.28)	5.00	0.86 - 1.97	<0.001
Illness Duration (in Years)	0.09(0.04)	2.02	0.00 - 0.17	0.043
Marginal R ² / Conditional R ²	0.098 / 0.498			
Table 4.8b: Effect of Illness Duration on the AN Group's Experienced Retrospective Fullness Beliefs				
(Intercept)	0.45(0.13)	3.37	0.19 - 0.71	0.001
Illness Duration (in Years)	0.02(0.02)	0.85	-0.02 - 0.05	0.395
Marginal R ² / Conditional R ²	0.013 / 0.128			
Table 4.8c: Effects of Illness Duration on the AN Group's Posterior Prospective Fullness Beliefs				
(Intercept)	1.43(0.37)	3.85	0.70 - 2.16	<0.001
Illness Duration (in Years)	0.09(0.05)	1.93	-0.00 - 0.19	0.054
Marginal R ² / Conditional R ²	0.091 / 0.466			
Table 4.8d: Effect of Illness Duration on the AN-WR Group's Prior Prospective Fullness Beliefs				
(Intercept)	4.62(1.38)	3.34	1.91 - 7.33	0.001

Illness Duration (in Years)	-0.03(0.17)	-0.17	-0.36 – 0.30	0.862
Marginal R ² / Conditional R ²	0.002 / 0.914			

Table 4.8e: Effect of Illness Duration on the AN Group's Experienced Retrospective Fullness Beliefs

(Intercept)	1.48(0.38)	3.87	0.73 – 2.24	<0.001
Illness Duration (in Years)	-0.00(0.04)	-0.08	-0.09 – 0.08	0.939
Marginal R ² / Conditional R ²	0.000 / 0.156			

Table 4.8f: Effect of Illness Duration on the AN Group's Posterior Prospective Fullness Beliefs

(Intercept)	4.61(1.44)	3.21	1.79 – 7.43	0.001
Illness Duration (in Years)	-0.01(0.17)	-0.03	-0.34 – 0.33	0.974
Marginal R ² / Conditional R ²	0.000 / 0.918			

Abbreviations: AN (Acute Anorexia Nervosa Group); AN-WR (Weight-Restored from Anorexia Nervosa Group). Bolded values denote statistical significance ($p < 0.05$).

The Baseline Emotional Fullness Ratings were greater than the Baseline Perceived Fullness Ratings.

In line with the prediction, all participants gave higher fullness estimates when asked how they 'felt' vs how full they 'were'. Specifically, Emotional Fullness Ratings at baseline, i.e., at the very beginning of the experiment and before participants were shown the different quantities of water which they also consumed, were significantly higher than the Perceived Fullness Ratings. Although the interaction between the rating mode (Perceived vs Emotional) and group was not significant, within-group analyses were also run to examine such differences within each group separately (Figure 4.8.; Table 4.9a). As predicted, both the AN and AN-WR groups provided significantly higher Baseline Emotional Fullness Ratings than Baseline Perceived Fullness Ratings (Table 4.9b-c). Contrary to the prediction, this was also the case in the HC group (Table 4.9d). Although the interaction (between Mode and Group) was not significant (Table 4.9a), a data visualisation indicates that the ratings of the AN group were overall higher than those of the HC group, as well as the difference between the two types of rating – namely Emotional and Perceived (Figure 4.8.). However, these findings warrant further investigation.

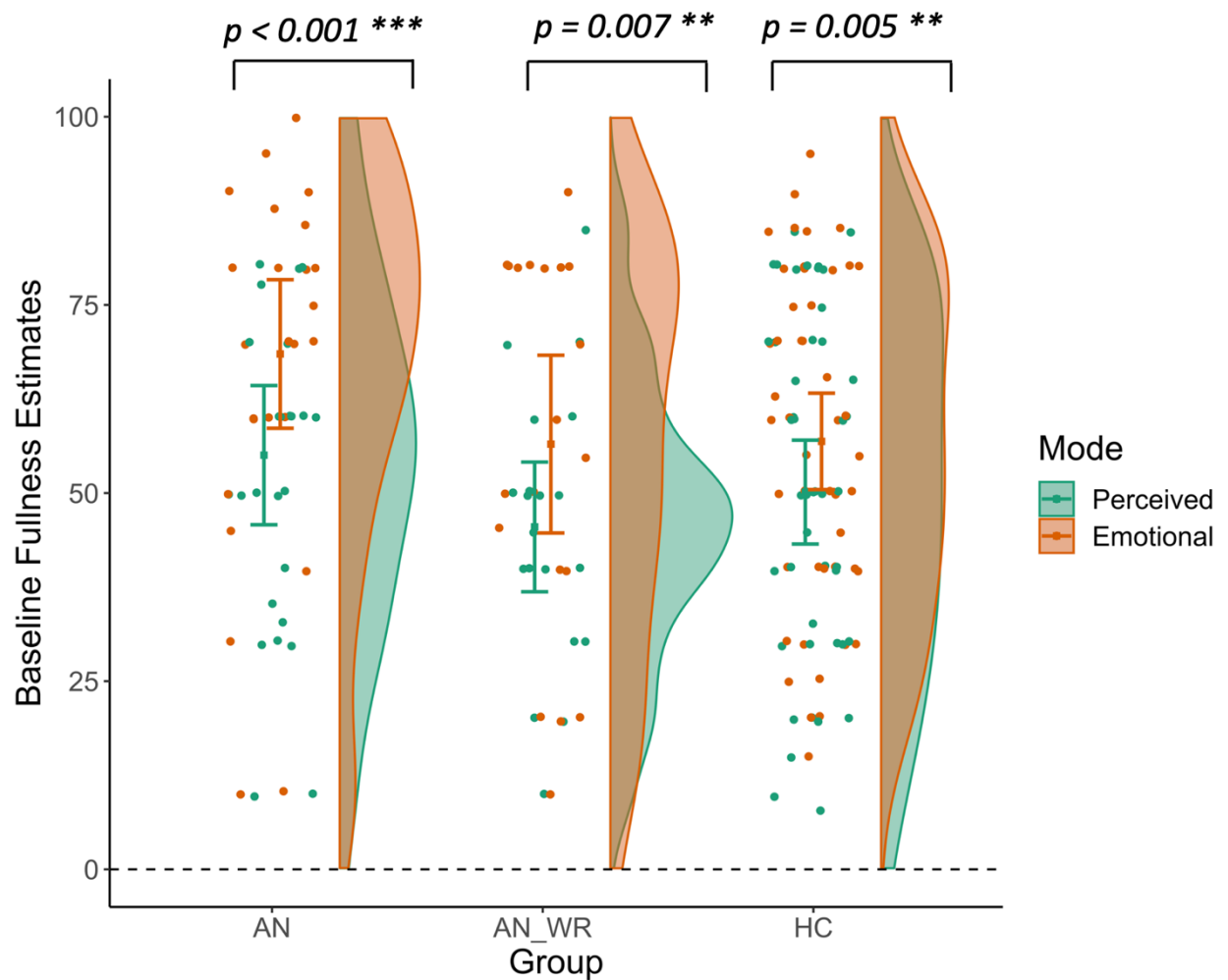


Figure 4.8. Between-group differences in Baseline Fullness Ratings, separated by mode of rating (Perceived, i.e., ‘How full are you’, vs Emotional, i.e., ‘How full do you feel’).

Table 4.9. Effect of Fullness Estimate Mode - Between and Within Groups.

<i>Predictors</i>	β (SE)	t	95% CI	<i>p</i>
Table 4.9a: Interaction between Group and Mode				
(Intercept)	50.13(3.35)	14.94	43.55 - 56.70	<0.001
AN	4.91(5.55)	0.88	-5.98 - 15.80	3.77
AN-WR	-4.63(6.14)	-0.75	-16.66 - 7.41	0.451
Mode (Emotional)	6.74(2.47)	2.74	1.91 - 11.58	0.006
AN x Mode (Emotional)	6.70(4.08)	1.64	-1.30 - 14.70	0.101
AN-WR x Mode (Emotional)	4.26(4.51)	0.94	-4.59 - 13.10	0.346
Marginal R ² / Conditional R ²	0.075 / 0.750			
Table 4.9b: Within-Group Effect of Mode (in the AN Group)				
(Intercept)	55.04(4.65)	11.83	45.92 - 64.16	<0.001
Mode (Emotional)	13.44(3.18)	4.22	7.20 - 19.68	<0.001
Marginal R ² / Conditional R ²	0.073 / 0.783			

Table 4.9c: Within-Group Effect of Mode (in the AN-WR Group)				
(Intercept)	45.50(4.94)	9.21	35.82 - 55.18	<i><0.001</i>
Mode (Emotional)	11.00(4.05)	2.72	3.06 - 18.94	0.007
Marginal R ² / Conditional R ²	0.060 / 0.684			
Table 4.9d: Within-Group Effect of Mode (in the AN-WR Group)				
(Intercept)	50.13(3.31)	15.15	43.64 - 56.61	<i><0.001</i>
Mode (Emotional)	6.74(2.42)	2.79	2.00 - 11.49	0.005
Marginal R ² / Conditional R ²	0.022 / 0.739			
N.B. Mode rating: Perceived ('How full are you') vs. Emotional ('How full do you <i>feel</i> '). Abbreviations: AN (Acute Anorexia Nervosa Group); AN-WR (Weight-Restored from Anorexia Nervosa Group); HC (Healthy Controls Group). Bolded values denote statistical significance (<i>p</i> <0.05).				

4.4 Discussion

The present study aimed to investigate if perception of changing gastric states (here, measured in terms of subjective fullness and manipulated by asking participants to drink different quantities of water at specific phases during the experiment) is different in Anorexia Nervosa (AN) in comparison to matched healthy controls (HCs), how beliefs regarding expected states of fullness can be updated retrospectively and prospectively using relevant (in the moment and prior) experiences, as well as the potential impact of weight restoration on dimensions of gastrointestinal interoception. Forming and updating beliefs about stomach fullness can guide decisions about whether and how much to eat or drink, and the perception of food size and estimation of meal portions are an acknowledged challenge in AN (Dörsam et al., 2020; Milos et al., 2013; Zitron-Emanuel et al., 2022). To address these aims, participants completed a Water-Sipping Task and were asked to give estimates of fullness before and after ingesting indicated amounts of water. The reported estimates of fullness of the groups were then compared, expecting to find greater overestimations of fullness in the two clinical groups in comparison to HCs. This hypothesis was only partly confirmed in that only AN-WR women, in comparison to the HCs, overestimated their fullness estimates, both when asked to form fullness beliefs and when updating them retrospectively and prospectively. By contrast, there were no significant differences between the acute AN and HC groups. In exploratory analyses, it was shown that a longer duration of illness (in the AN but not in the AN-WR group) was associated with an overestimation in both the Prior and Posterior Prospective Fullness Beliefs. These findings are discussed below in the context of state-trait effects and why significant overestimations of fullness may have been observed between the AN-WR and HC groups and not between the AN and HC groups. Other factors, such as clinical traits of depression, anxiety

and intolerance of uncertainty were not found to influence fullness beliefs, and nor did participants' reported fullness in the bladder and stomach. One of the central aspects of this experiment was however placed on how individuals prospectively update the beliefs they have formed about fullness. As predicted, both clinical groups had a lower belief updating rate than HCs, i.e., they updated their fullness beliefs less than HCs after drinking water, prospectively. In other terms, the AN and AN-WR groups did not update their prospective fullness beliefs at the rate at which HCs did, which could suggest that the retrospective (embodied) feedback they received after the water-sipping trials was considered less than prior beliefs they had regarding fullness. Moreover, although it was expected that both clinical groups would be less confident regarding the accuracy of their fullness state ratings and predictions, results showed that women with AN (but not AN-WR) had greater confidence than HCs regarding their prediction of how full drinking a certain amount of water will make them. However, the AN group's confidence levels seemed to be influenced by stress traits (as indexed using the DASS-21 stress subscale, a measure also used to assess trait anxiety and depression; Lovibond & Lovibond, 1995), with confidence scores in the AN group decreasing as the severity of premorbid stress traits increased. Notably, there were differences in the way all participants experienced different types of fullness – namely, there was a main effect of fullness estimate mode (perceived vs emotional; and no main effect of group or interaction with mode) and all participants across the three groups seemed to give higher emotional ratings of fullness vs objective, perceptual states of fullness. The results and insights from this experiment are discussed in detail below.

One of the primary aims was to examine the differences between the AN and AN-WR groups in comparison to HCs regarding how they predict future fullness states. Patients with and weight-restored from AN were expected to show aberrant perception of fullness in line with previous findings pointing towards a general difficulty in satiety and hunger processing - with patients often reporting not feeling hungry even when on an empty stomach and also reporting prolonged fullness (e.g., Bluemel et al., 2017), but also a noted bias in the ability of AN patients to perceive the size of food stimuli in comparison to HCs (e.g., Zitron-Emanuel et al., 2022). Specifically, there were two measures of prospective fullness, namely the Prior Prospective Fullness Belief and the Posterior Prospective Fullness Belief, both of which measured how much fuller participants predicted to be in relation to their (Post-Instructions) Pre-Sipping Baseline Fullness State Rating and their Post-Sipping Fullness State Rating, respectively. While the former measure was a prospective estimate based on prior knowledge and

experiences and the visual, exteroceptive information participants had by looking at the quantity they would have to sip on the bottle, when making their second prospective fullness estimate, participants could also use the experience they just had to predict how full they would be, as the quantity that was shown to them for the Second Fullness State Guess (once per trial) was the same as the one that was shown for their First Fullness State Guess (these two measures were then used to calculate the Prior and Posterior Prospective Fullness Beliefs; see Figure 4.1). In two separate analyses – one for each measure of prospective fullness beliefs – it was found that both Prior and Posterior Prospective Fullness Beliefs of the AN-WR group were higher than those of the HC group. Surprisingly and contrary to the prediction, although the AN patients reported an elevated state of fullness (at baseline, before the instructions although the group differences of Pre-Instructions Baseline Fullness State Rating were not significant) they did not significantly prospectively overestimate their fullness state in comparison to HCs. Participants in the present experiment and contrary to others (e.g., Brown et al., 2022) were not asked to fast for practical reasons around availability and testing, but also to not interfere with the clinical schedules of the AN group. As a result, it is likely that differences in fullness at the beginning of the experiment and related feelings influenced participants' prospective fullness beliefs. To assess this, the effects of feelings of fullness in the bladder and stomach on (prior and posterior) prospective fullness beliefs were added in control analyses. However, the results of the main analyses remained the same, suggesting that as tested here, other gastrointestinal feelings did not influence participants' answers. Moreover, the role of baseline fullness at the beginning of each trial and its effect on the predicted fullness state estimates was also considered. In particular, the pre-sipping fullness state of the AN population (when looking at the face value ratings of Post-Instructions Pre-Sipping Baseline Fullness State) was higher than that of the HCs (although this was a non-significant trend). Thus, it is possible that the change in fullness states was not going to be significantly greater in the AN group as they already reported higher baseline fullness than the other groups at the beginning of each trial. This comes in support of clinical observations that AN patients will report exaggerated levels of fullness post-consumption of small meals (Schalla & Stengel, 2019). Specifically, the AN group reported above average Pre-Instructions Baseline Fullness (both emotional and perceived) and their already high baselines may explain why their additional, predicted fullness is not that much higher – in line with the finding that AN patients exaggerate their fullness levels. However, no significant group differences were found in the Post-Instructions Pre-Sipping Baseline Fullness State Rating and the Pre-Instructions Baseline (Perceived and Emotional) Fullness State Rating and thus these measures were not added as covariates in the main

analyses. Although individual differences in traits did not explain the group effect on the prospective beliefs of fullness, these results call for further attention.

To look at what else could explain the overestimation in prospective fullness states in the AN-WR group in comparison to HCs, the role of illness duration was considered in light of the literature regarding the impact of AN and prolonged caloric restriction to the gastrointestinal tract (see Section 4.1). The effect of illness duration on Prior and Posterior Prospective Fullness Beliefs was considered in both the AN and AN-WR groups, separately, to explore how a longer duration of the acute AN state may have influenced one's beliefs about gastrointestinal states. In the AN group, patients' prior and posterior prospective fullness beliefs were predicted by illness duration. Interestingly this was not the case for the AN-WR group; yet this could be attributed to the fact that overall, the average duration of the illness in the AN-WR group was longer (as patients may relapse and recover and relapse again) and had less variance than that of the AN group who was still in the acute phase of AN. In other terms, the duration of AN (from the moment of first diagnosis until weight-restoration for at least six months, i.e., including potential relapses that occurred in between) may have contributed to the overall difficulties of the AN-WR group to make perceptual and metacognitive inferences about fullness states. It is likely that the difference in the results between the two clinical groups in comparison to the HCs are indicative of a trait-related characteristic reflecting the long-term consequences of AN on gastrointestinal interoception rather than a state resulting from acute malnutrition, medication, refeeding, and other treatments occurring during the acute stage. However, the lack of a significant finding in the AN-WR group should be further explored as the general limitations of this experiment (see below) could explain the lack of a non-significant result, namely a small effect size and power-related issues due to a small sample size especially in the AN-WR group.

Another challenge in drawing conclusions regarding this finding is the fact that there are few studies on AN and gastrointestinal interoception, and unfortunately most have either predominantly focused on individuals with acute AN, making it impossible to draw comparisons with weight-restored or recovered individuals, or the examination of weight-restoration effects was conducted shortly after patients were no longer in an acute AN state. For instance, in a longitudinal follow-up that found that after weight-restoration fullness increased and hunger reduced, the patients who were tested at baseline while in the acute phase of AN were followed-up after an average of 112 days post restoration (approximately 4 months;

Bluemel et al., 2017) in comparison to the present AN-WR group (weight-restored for at least 6 months). Another study also found a partial improvement in the perception of appetite, gastrointestinal symptoms and concentration of gastrointestinal hormones and peptides following a period of refeeding in an acute AN sample, supporting the role of trait-mechanisms in the pathogenesis and maintenance of AN symptoms and behaviours (Heruc et al., 2019; Nakahara et al., 2007). It is important to note that follow-ups in these studies were performed after one to two weeks after refeeding, whilst contrary findings in support of state and not trait effects also exist (e.g., Monteleone et al., 2016). All these studies, including the present one, point towards a trait mechanism. However, the current findings cannot be conclusive - further investigations are warranted to delineate the trajectory of illness especially considering that the participants of the AN and AN-WR groups were distinct individuals, and not acute AN patients who were followed-up after weight restoration. Nevertheless, these preliminary findings underscore the necessity to untangle the underlying mechanisms contributing to deficits in gastrointestinal interoception in AN. Particularly, exploration is warranted to understand how premorbid risk factors related to the overestimation and over-exaggeration of predicted fullness states may persist following weight-restoration, as indications from existing studies suggest an improvement in the perception of gastrointestinal signals during weight-restoration and the refeeding phase (e.g., Bluemel et al., 2017; Heruc et al., 2019). Another aim should be to elucidate how aspects of gastrointestinal interoception may evolve across different stages of the disorder and whether they persist over time or gradually improve with sustained recovery and maintained remission of AN.

An unexpected finding was that there was an overall condition effect in the opposite direction on participants' Prior Prospective Fullness Beliefs but also on their Posterior Prospective Fullness Beliefs. In other terms, although it was expected that the two clinical groups would misestimate their fullness even more than HCs in the High Quantity as opposed to the Low Quantity condition (i.e., when shown a larger quantity on the bottle vs. a lower quantity), a main effect of condition but no interaction with group was found on Prior Prospective Fullness Belief in that all participants reported to be fuller when shown the Low vs. High Quantities. This was also the case when looking at participants' Posterior Prospective Fullness Beliefs, however in this analysis it was also shown that the interaction between group and condition was driven by the AN-WR group's overestimation of fullness in the Low vs. High Quantity condition. All participants had to see the water quantity and make a judgement using the visual information from the bottle, and prior knowledge and beliefs when predicting how full they

would be after consuming that amount. It is possible that the low quantities are more ‘atypical’ to the quantities one will usually consume – e.g., the 2% trial corresponded to 0.43cm of the bottle’s height and it could be speculated that such low quantities lead to more arbitrary and exaggerated responses, as opposed to predictions in the 15% trial (which corresponded to 3.36cm of the bottle). Yet, the fact that a Condition main effect and no interaction with group was observed, may point towards a general difficulty in people’s ability to use visual, exteroceptive information and combine it with prior, interoceptive experiences and beliefs.

Further, when providing a posterior prospective estimate of fullness (i.e., Second Fullness State Guess), the AN-WR group showed greater overestimation of how they would be following consumption of a Low vs High Quantity, driving the Group x Condition interaction. However, this finding is difficult to interpret and requires further investigation as no other studies have adopted a similar design and studies on object estimation do not show a pattern of results similar to those in the present study (e.g., Glauert et al., 2009; Saramandi et al., 2022), i.e., object size perception abilities do not seem to differ between populations with eating disorders and matched HCs. For instance, AN, AN-WR and HC groups have previously shown no size estimation differences when asked to judge the size of a 2D printed balloon (i.e., a neutral object, not linked to the AN symptomatology) both before and after receiving feedback about the balloon’s size, whereas the two clinical groups were found to overestimate their body size more than HCs before and after related feedback, suggesting that the perception and feedback of one’s body size may be processed differently than feedback about another object (Saramandi et al., 2022). However, in the present experiment, participants were asked to combine exteroceptive with interoceptive knowledge and information to make an interoceptive inference (i.e., ‘*How full will you be?*’), whereas in the balloon-size estimation example participants only had to use the visual exteroceptive information to provide a size estimate. Although these present results point towards a generalised difficulty in combining exteroceptive and interoceptive information, this warrants further investigations to better understand what may explain the overestimation in lower as opposed to higher quantities of water.

Next, differences in the groups’ explicit retrospective beliefs about fullness were also examined (i.e., after drinking the indicated amount of water, also referred to as Experienced Retrospective Fullness Beliefs), expecting that retrospectively the clinical groups would report to be fuller than HCs. Contrary to this prediction, only the AN-WR group and not the AN group

retrospectively updated their predicted fullness beliefs less than the HCs groups. Specifically, the difference between how full the AN-WR group reported to be after ingesting the required water quantity and their prospective fullness state beliefs was significantly greater than the corresponding difference observed in the HC group, offering support for a perceptual distortion which may persist in individuals with AN post weight-restoration. However, the group effect was no longer significant when adding condition (i.e., Low vs. High quantity) as an interacting variable, in contrast to the prediction that the clinical groups would report greater Experienced Retrospective Fullness Beliefs than HCs especially after the High vs. Low quantity condition. It is possible that after a prolonged period of starvation and suppression of interoceptive signals, the AN-WR group had difficulty attending to and interpreting inner body states, as shown by their exaggerations of fullness (in comparison to HCs) even though they consumed the indicated water quantity and could use that embodied feedback to adjust their beliefs regarding fullness. In fact, an association has been found between poor interoceptive sensibility (i.e., where interoceptive sensibility is defined as the self-evaluation of subjective interoception, typically assessed via questionnaires; Garfinkel et al., 2015; Khalsa et al., 2018) and decreased consumption of water (Brown et al., 2022). In other words, AN individuals who were less able to identify interoceptive signals also consumed less water during the water load task (Brown et al., 2022). In this context, this could support the finding of the AN-WR group reporting to be fuller retrospectively, yet the three groups did not differ significantly in their interoceptive sensibility scores (as measured using the interoceptive deficits subscale of the EDI-3; Garner, 2004), and thus further research is needed to understand this result. Another explanation is that the discrepancy between the AN-WR and HC groups' reports on fullness was affected by the type of measure, consistent with a study examining perceived gastric sensations which also found no significant differences in the reported feelings of fullness between the AN and HC groups, but instead found that AN patients were less likely to report hunger (Kerr et al., 2022). Hence, it is likely that the perception of distinct gastrointestinal sensations and feelings such as hunger, fullness, thirst, and satiety differs; in other words, it is possible that the framing of the question (e.g., '*How full are you*' vs '*How hungry are you*' or '*How thirsty are you*') taps onto different aspects of interoception and while patients with AN may report similar levels of fullness than HCs, they will report reduced levels of hunger. Future research should further delineate between the differences in gastrointestinal perception and interoceptive inference when measuring aspects of hunger, thirst, fullness, and satiety to better understand the distinct roles of these signals in AN.

The updating of beliefs, retrospectively (as indexed by the Prediction Error) and prospectively (as indexed by the updating rate) was also examined, and it was hypothesised that the AN and AN-WR groups would have a greater Prediction Error and lower updating rate than HCs. Indeed, the AN-WR group had a greater Prediction Error than HCs; weight-restored individuals showed a greater discrepancy between how much fuller they reported being after the water-sipping trials and how much fuller they predicted they would be (i.e., difference between Experienced Retrospective and Prior Prospective Fullness Belief). Interestingly, and contrary to the prediction, this difference that was observed between the AN-WR and HC groups, was not observed between the AN and HC groups. Notwithstanding, and in line with findings in the cardiac modality (Chapter 3) and in another domain (body size estimation; Saramandi et al., 2022), patients with and weight-restored from AN showed lower updating rates than HCs. Notably, when looking at the rate in which the three groups updated their beliefs, the two clinical groups updated their beliefs regarding their fullness to a smaller degree than HCs when using the fixed amount of additional fullness they should have felt following ingestion of water to calculate their prediction error scores. The additional fullness values equal to the change the participant should have reported regarding their fullness if it matched the quantity of water that was given to them, measured as a percentage of the bottle's entire volume. Contrary to the prediction, only the AN-WR and not the AN group had a lower updating rate when the prediction error was calculated using participants' Experienced Retrospective Fullness Belief. Participants did not receive any explicit external feedback or information regarding the amount of water they actually drank; however, findings in another domain have shown that when receiving disconfirming information about one's emotional body size (e.g., a clinician's perspective regarding an AN patient's body size based on looks rather than objective parameters such as body weight and shape, BMI), acute AN patients and not AN-WR patients are more likely to rely on prior beliefs and subsequently update emotional body size estimates at a lower rate than HCs (Saramandi et al., 2022). However, differences between the AN, AN-WR and HC groups in updating rate were not observed when the feedback received was based on objective parameters, i.e., based on body weight and shape. This could further suggest that embodied, gastrointestinal feedback which is subjective to the individual state but also between individuals may also be considered less in individuals with AN when asked to make inferences about future interoceptive (e.g., fullness) states and this could be influenced by the prolonged period of starvation as updating rate differences were observed in both the acute and weight-restored groups.

The lack of, or reduced updating could also be due to a greater reliance on prior estimates of fullness or beliefs regarding how full they would be post-ingestion, perhaps due to a similar experience in the past or other factors, such as less precision about the new information relative to the precision of the prior beliefs (Friston, 2003; Mathys et al., 2011; Smith et al., 2020). The handful of trials of the present task did not allow to conduct full computational model comparisons and to explore which parameters best predict the three different groups' posterior prospective fullness estimates like those conducted in cardiac interoception (see Chapter 3). Hence, a simplified model was used which allowed to compare the slope differences between the groups and observe the extent to which the drinking experience during the task and the retrospective fullness estimates were taken into consideration by the subjects when making new, updated fullness estimates. Future research could add on to this finding by further exploring differences in updating rate in light of explicit feedback, e.g., either by comparing how updating rate may also differ when participants are told explicitly how much water they will be consuming or assess the differences between subjective gastrointestinal states (e.g., fullness) and perceptions of gastrointestinal states. Specifically, a recent paradigm uses a method that triggers mechanosensation via vibration (for details in mechanosensation in the gastrointestinal tract see Mercado-Perez & Beyder, 2022), in order to dynamically facilitate and alter the individual's conscious perception of gastrointestinal sensations and aid the exploration of abnormal sensations in conditions such as AN (Mayeli et al., 2023). A vibrating capsule that participants ingest is used to enhance the mechanosensations that typically occur in the gastrointestinal tract (e.g., distention and contractions) and participants can then attend to the sensations they feel in their stomach and rate their intensity. Results from a healthy sample show that individual perception of signals, their intensity and beliefs regarding gastrointestinal states can be studied under a computational modelling approach to help identify the elements which could contribute to the manifestation of aberrant interoceptive precision estimates, learning rates, and belief updating in clinical symptoms (Mayeli et al., 2023). Combining this method together with a water-sipping task (like the one used in the present experiment) could then allow the comparison between potential difficulties in detecting objective gastrointestinal sensations and interpreting the more 'subjective' gastrointestinal states such as the perception of fullness and updating of beliefs regarding fullness states, retrospectively and prospectively.

Furthermore, the precision of ratings of fullness and related beliefs using confidence as precision proxies was also examined. Specifically, individual subjective uncertainty was

examined by assessing how accurate participants thought their ratings and beliefs were, expecting the clinical groups to be underconfident in comparison to HCs. Instead, the AN group was significantly more confident in the accuracy of their First Fullness Guess State Rating in comparison to HCs (although lower confidence across all participants was observed regarding the accuracy of posterior prospective fullness beliefs in the high quantity condition in comparison to the low quantity condition). Previous theoretical frameworks have proposed that poorer gastrointestinal interoception could be in part attributed to a vicious circle in which intolerance of uncertainty, especially regarding interoceptive states leads to eating restriction as an attempt to minimise prediction errors (Barca & Pezzulo, 2020), offering support for the expected lower confidence in a clinical population. On the other hand, the greater levels of confidence could be explained by the rigidity sometimes observed in this population regarding their beliefs of how eating and drinking will make them feel and thus they restrict food and drink intake, as well as related delusional beliefs regarding aspects of their own body (e.g., body size; e.g., Konstantakopoulos et al., 2012, 2020). This conclusion however remains speculative as the confidence levels between the groups were not significantly different following the other fullness ratings and beliefs, and particularly participants' posterior prospective beliefs. It is noteworthy, that the significant group effect on the confidence levels was partially mediated by stress (as measured using the DASS-21 Stress subscale; Lovibond & Lovibond, 1995). Specifically, it was found that overall confidence decreased as the severity of stress traits increased. It has previously been shown that stress (as well as other AN comorbid symptoms including depression) affect participants' overall self-efficacy beliefs regarding future cardiac interoception abilities (Saramandi et al., 2023; Chapter 3). Similarly, it could be that higher or more severe stress traits could influence the confidence one may feel regarding their abilities to interpret gastrointestinal signals of fullness. This finding warrants further investigation given the widely recognised reciprocal relation between interoceptive dimensions, related parameters, and psychopathology.

The present study had several limitations which can be addressed in future studies. Firstly, due to recruitment difficulties, the size of the two clinical groups was smaller than that of the HC group ($N_{AN}=27$, $N_{AN-WR}=20$, $N_{HC}=47$) which may explain why not all models may have had sufficient power to detect effects for some of the predictors. Increasing the sample size, particularly of the AN and AN-WR groups could increase the statistical power of the analyses and improve the ability to detect true effects – especially for the control and exploratory analyses which in some instances were only performed in a subset of the participants due to

time constraints. Secondly, difficulties in cognitive flexibility and set-shifting were not assessed (like they were in the sample of the cardiac interoception task presented in Chapter 3) and thus any conclusions regarding difficulties in updating beliefs about fullness in AN relating to – to some extent – cognitive rigidity (e.g., the lower updating rate of the AN and AN-WR groups in comparison to HCs) remain speculative and require further testing. Another point to note is that the interest in this experiment was in examining the proportional changes in predicted fullness and how individuals adjusted their fullness states, rather than focussing on the quantities of water that was consumed. This was the case given the inherent variations in stomach size between the groups, leading to differential impacts on their perception of fullness following the water-sipping trials. In other terms, an individual with AN will have a smaller stomach than a weight-restored individual or a HC and even if all individuals were asked to follow the same pre-study fasting rules, upon consumption of the same quantity of water or food, a higher volume of the stomach of the individual with AN would be filled in comparison to the HC. Additionally, individual differences and heterogeneity within and between groups should be further considered. For instance, gastrointestinal signals, including fullness, satiety, and hunger, may fluctuate during a woman's menstrual cycle, however this was not recorded in the HC group (and in the AN-WR group, where appropriate) and thus was not considered in the analyses. Similarly, although contraceptive methods (e.g., pills and intra uterine devices) may contribute to gastrointestinal symptoms (e.g., constipation and delayed gastric emptying) and influence one's perception of gastrointestinal, interoceptive signals, their effects were not accounted for. That is, even though most participants, including amenorrhoeal women (from the AN and AN-WR groups), were on some type of contraceptive method, these data, and their role on the main dependent variables of interest could not be further investigated. Finally, although the acute AN group was (at the time of testing) receiving pharmacotherapy, as did some AN-WR participants when they were in the acute AN state, and medication can have secondary effects to the gastrointestinal signals, only the class of drugs was recorded and thus their effects remain speculative and require further investigation (e.g., effects of dosage, treatment duration, combination of drugs on gastrointestinal interoception). Despite these limitations, however, the present study offers preliminary insights into beliefs about and changes in (experienced) fullness during a water-sipping task and differences between clinical populations with and weight-restored from AN and matched HCs.

4.4.1 Conclusion

In conclusion, this study compared how groups with and weight-restored from AN differ from HCs in their (subjective) estimates of perceived fullness before and after completing a water-sipping task. The results offer novel insights in the role of AN in gastrointestinal interoception. Specifically, the results show that difficulties in perceptual and metacognitive inferences may endure post weight-restoration and be influenced by a prolonged period of illness duration. Moreover, results also showed that both acute and weight-restored AN patients show difficulty in updating predicted gastrointestinal states, possibly relying more on prior beliefs and expectations, or other parameters, than an in-the-moment experience of physiological state alteration (i.e., changes in fullness after drinking water during the task). Additionally, although the present design did not allow for a computational model comparison or the delineation of which parameters may best predict the belief updating about fullness of each group, parameters which can be used as precision-proxies (here, confidence) seemed to be affected by stress in AN, as indexed by reduced confidence levels regarding fullness in AN as compared to HC when severity of stress traits was higher. Furthermore, emotional aspects of fullness were found to differ to cognitive aspects of fullness (i.e., *How full do you feel* vs *How full are you*), however the lack of a group effect requires further testing in order to understand how in combination with other factors (e.g., reduced interoceptive sensibility) this could contribute to the psychopathology of AN at the various stages of the disorder (e.g., from premorbid and acute state, to weight-restoration and remission) and difficulty in discerning gastrointestinal states regarding fullness, hunger, thirst and satiety. Finally, the current observations can be used for comparisons in perception and related beliefs across modalities of interoception but also other AN symptomatology and behaviours (e.g., body image and body size; see above), to further delineate the complex relations between perception, beliefs, and clinical characteristics of AN across domains, modalities, and hierarchical levels of (meta)perceptual inferences.

Chapter 5: The Use of Heart Rate Biofeedback Treatment to Assist the Regulation of Embodied Feelings: A Clinical Audit on the Process and Effects of Biofeedback in Clinical Practice

5.1 Introduction

The enduring effects of Anorexia Nervosa (AN) include impaired health, psychiatric comorbidity, and poor quality of life for the person and those around them. Additionally, AN has the highest mortality and morbidity rates among psychiatric disorders (Arcelus et al., 2011; Petkova et al., 2019), with a crude mortality rate (CMR) being approximately 5.1 deaths per 1,000 person-years (National Guideline Alliance (UK), 2017). The prevalence of AN in young women in Western Europe and the US is estimated to be approximately 0.3% (ranging between 0.2% and 0.8%; Hoek & van Hoeken, 2003), and incidence rates are estimated to range between 4.2 and 12.6 per 100,000 person-years (Currin et al., 2005; Micali et al., 2013; Steinhausen & Jensen, 2015; see Section 1.2, Chapter 1) for more details on the epidemiology of AN). Moreover, these figures are likely underestimates of the true extent of AN. In the UK for instance, current estimates are between 10-15 years old, mainly derived from community-based primary care records and do not take into account data from secondary care records (Currin et al., 2005; Goodman et al., 2000; Hoek, 2006; House et al., 2012; Lask et al., 2005; Micali et al., 2013; National Guideline Alliance (UK), 2017; Nicholls et al., 2011; Smink et al., 2012). Hence, it is possible that the incidence rate and subsequently the relapse rates are higher, making the need for appropriate and effective treatments of utmost importance.

Despite the debilitating consequences of AN and the severity of the disorder, our understanding of the underlying mechanisms remains unclear. Existing studies have focused on the contributing (neuro)biological, perceptual, and social factors, such as identifying the neural substrates of impaired functions, the distortion of one's perceived body image, and the influence of social media (e.g., Kaye et al., 2013; Keery et al., 2004; Munro et al., 2017; Nunn et al., 2012; Trace et al., 2013; Uchôa et al., 2019). However, several recent accounts have come to suggest that individuals with AN also have a marked pessimism regarding their abilities (i.e., poor self-esteem and self-efficacy; e.g., Dimitropoulos et al., 2016; Halmi et al., 2005; Keshen et al., 2017; Pinto et al., 2008), and more recently, this pessimism was observed in patients' beliefs regarding their abilities to attend to internal bodily signals (Saramandi et al., 2022, 2023; see also Chapters 2, 3 and 4).

Despite recent findings that highlight individual ability to attend to internal bodily signals and one's beliefs regarding those abilities are affected in AN, current standard treatment recommendations do not cover these aspects of the disorder. The NICE 2004 guidelines (updated in 2017; National Collaborating Centre for Mental Health (UK), 2004; National Guideline Alliance (UK), 2017) specify that one of the key aims of AN treatment is to help people reach a healthy body weight for their age. Available first-line psychological interventions are the Maudsley Model of AN Treatment for Adults (MANTRA), general or eating-disorder specific Cognitive Behavioural Therapy (CBT or CBT-ED, respectively), and Specialist Supportive Clinical Management (SSCM). These treatments encourage weight restoration and typically involve a psychoeducational component (e.g., explaining the risks of malnutrition in CBT-ED, and recognising the link between symptoms and abnormal eating behaviours in SSCM). The national guidelines (NICE 2004, 2017) recommend that the individual is explained what each treatment involves to help them choose which they would prefer. If the selected first-line intervention is unacceptable, contraindicated or ineffective for the adult with AN, then the patient either needs to choose one of the other first-line interventions or opt for eating-disorder-focused focal psychodynamic therapy (FPT), as a second-line psychological intervention option (National Collaborating Centre for Mental Health (UK), 2004; National Guideline Alliance (UK), 2017). FPT is focussed on helping the individual understand what the symptom(s) mean to them, how they affect them and how they also influence their relationship with others and with the therapist.

To date, evidence has shown that CBT-ED in adults may be more effective at improving remission in comparison to any other of the aforementioned first-line psychological interventions (National Guideline Alliance (UK), 2017), however at 12 months follow-up the benefits were no longer evident. Similarly, the main comparisons between the outcomes of these interventions, such as BMI improvement and remission rates, did not show favourable results for one type of first-line intervention over another (e.g., when comparing treatment outcomes of CBT vs MANTRA; National Guideline Alliance (UK), 2017). Although certain outcomes, such as depression symptoms, may be showing an improvement following certain interventions (here, SSCM), these observations do not suffice to suggest that one intervention is more efficacious than another. Moreover, although CBT-ED aims to enhance patients' overall self-efficacy and self-esteem (see more below), CBT-ED mainly focuses on improving the patient's self-efficacy and self-esteem in their ability to carry out daily tasks independently, establishing balanced eating habits and developing strategies to effectively manage triggers,

and reducing body image concerns, rather than improving their self-regulation ability beliefs (i.e., their beliefs regarding their ability to manage/control internal bodily states, such as using breathing exercises to slow down a high heart rate as a result of a stressful state). MANTRA and SSCM also have components aimed at helping the patient develop a ‘non-anorexic identity’ and recognising the link between symptoms and abnormal eating behaviours (National Guideline Alliance (UK), 2017). Yet, all these first-line interventions lack helping the patient understand what bodily sensations such as a fast-beating heart, or sweaty palms when feeling stressed may mean, how they could regulate them, and improving their self-efficacy (where outcomes in a task are influenced by a person’s beliefs in their abilities to succeed in said task; Bandura, 1977).

Although these currently established first-line interventions target a wide range of symptoms and behaviours that are necessary for AN recovery, they do not seem to always address the perception and management of aspects targeted by third-wave behaviour therapies for eating disorders, such as a different approach to understanding one’s emotions and cognitions (Buerger et al., 2021; Hayes et al., 2011). Third-wave behaviour therapies for eating disorders may include Dialectical Behaviour therapy (DBT), pure or guided self-help and interpersonal psychotherapy (National Guideline Alliance (UK), 2017), which focus on meditation elements, mindfulness, emotions, and exploring one’s thoughts and beliefs regarding their cognitions (Linardon et al., 2017). Despite some local units/trusts using elements of third-wave approaches, data on their efficacy as a treatment option for AN remain scarce (see Linardon et al., 2017 for a first review and meta-analysis on third-wave approaches in eating disorders). Thus, until further empirical support for third-wave approaches is obtained, first-line treatments remain the primarily nationally recommended approach which may then be tailored to also comply with the local trust’s standards of treatment and care.

Specifically, at the Regional Eating Disorders Unit (REDU, NHS Lothian, Scotland) the main goals after admission, and in line with the current NICE guidelines, include medical stabilisation, providing nutritional support and refeeding, addressing the psychological and psychiatric instabilities influencing the patient’s physical health and providing psychological assessments and treatments (details on the local practice can be found here: <https://services.nhsllothian.scot/seeds/regional-eating-disorder-unit/>). Here, the interest was in assessing whether the psychological treatments in place meet the local standards of assistance in understanding and managing embodied feelings (e.g., elevated heart rate resulting from

anxiety during eating). In particular, the REDU multidisciplinary team (MDT) uses CBT-ED as a first-line intervention, which itself includes several crucial steps (Murphy et al., 2010). First, it is important to engage the patient and encourage them to take ‘ownership’ of the treatment, and then complete a case conceptualisation with their clinician. Next, the patient is also provided (psycho)education about their disorder and its treatment with a focus on the characteristic features of AN and their effects, body weight and its regulation, and the adverse effects of long-term starvation and malnutrition. The patient’s weight is also checked on a weekly basis which is an opportunity for the therapist to educate the patient about body weight, and how to interpret the numbers on the scale. CBT-ED guidelines also recommend that patients establish a pattern of regular eating – at REDU this is enhanced by involving patients in the meal planning twice a week. Once these initial steps have been implemented, the main body of the CBT-ED intervention is to address the key processes that are maintaining the patient’s eating disorder, such as identifying the overevaluation of one’s body shape and weight, addressing the experience of *‘feeling fat’*, addressing dietary rules (e.g., patients avoiding specific foods), and importantly also addressing low self-esteem and self-efficacy. This latter target is especially important as low self-esteem, and in general, a negative view of oneself in AN can be a barrier to treatment engagement and progress.

However, following an earlier service evaluation, the MDT at REDU observed that CBT-ED did not suffice to help patients manage and regulate their feelings when faced with triggering situations (e.g., during the weekly weighing, or meal planning) and patients’ levels of self-esteem and self-efficacy regarding recovery did not seem to improve. Thus, third-wave CBT approaches and additional therapy elements were incorporated. More specifically, patients could participate in optional weekly psychoeducation and mindfulness, and 10-minute relaxation sessions twice a week (e.g., progressive muscle relaxation, body scan technique), all aimed at helping the patients to learn more about ways in which they can relax and understand their own body and embodied feelings. Additionally, the MDT also introduced a body image group in which patients were encouraged to learn more about what body image is, its different components (for more see Gadsby, 2019; Saramandi et al., 2022) and improve their relationship with their body image based on acceptance and CBT principles. Finally, patients were offered decider and DBT skills where they practice mindfulness to notice their emotions, as well as anxiety management and assertiveness sessions based on CBT principles. Yet, these changes in the standard practice still did not sufficiently meet the local standard of care, as low participation rates were observed, and patients did not seem to trust the efficacy of these

optional treatment approaches. For instance, one reported that *“Often, with mindfulness I don’t know if it’s working, I don’t know if my heart is going down, and I am also not sure when to start the breathing exercises. I don’t notice the anxiety build up; it just hits me”*. Another patient also said *“I have always been very cynical about it [meditation/breathing exercises]. It’s very up in the air and not practical”*.

The MDT thus decided to consult individuals with lived AN experience (through Patient and Public Involvement (PPI) events, in line with the NHS PPI Strategy 2021-2026; NHS Accelerated Access Collaborative, 2022), and a team of academic (research) partners, including a Psychology Professor, a Research Fellow and a PhD student (AS) to identify how the standard practice that was in place could be enhanced in order to help patients with AN better understand and self-regulate embodied feelings. Following feedback from the PPI participants (individuals with lived AN experience, clinicians and researchers specialising in eating disorders), it was suggested to incorporate heart rate biofeedback (HRBF) as an add-on to the existing standard practice of mindfulness following experiences that are regarded as highly emotional (e.g., meal planning at the ward). Biofeedback interventions are a type of self-monitoring aid enabling patients to feel more in control of their health management (Heron & Smyth, 2010; Piwek et al., 2016; Yetisen et al., 2018). Biofeedback combines several techniques such as operant learning, using positive reinforcement to shape an individual’s behaviour (Babel, 2020; Skinner, 1965), and meditation- and mindfulness-inspired techniques in combination with real-time feedback on one’s physiological states, e.g., heart rate, electrodermal activity and respiratory rate (Goessl et al., 2017; Weerdmeester et al., 2020) keeping the participant engaged and letting them track their self-regulation abilities. In the latter, mindfulness and meditation techniques allow the individual to shift their attention to present, internal experiences in order to reach higher levels of wellbeing and emotional balance (Astin et al., 2003; Brown & Ryan, 2003; Khalsa et al., 2015). In eating disorders, such mindfulness-based interventions, meditation exercises and yoga interventions have shown positive results when used in conjunction with standard treatment approaches, for instance by improving psychological wellbeing (e.g., CBT, or family-based therapies; Beccia et al., 2018; Borden & Cook-Cottone, 2020).

Biofeedback is also based on various other key mechanisms, including the appraisal of self-efficacy and locus of control. Self-efficacy is a core predictor of outcomes in eating disorders (Berman, 2006; Glasofer et al., 2013; MacNeil et al., 2012). There are several accounts

suggesting that low-self efficacy is linked with higher treatment drop-out and end-of-treatment symptom severity in AN (Carter et al., 2012; Keshen et al., 2017). Yet, whilst available interventions for AN may target beliefs about the meaning of physiological signals (e.g., metacognitive therapies), or allow individuals to monitor their heart rate (via wearable devices, e.g., Fitbit or apple watch), they do not encourage individuals to learn more about their abilities to self-regulate, in turn enhancing their self-efficacy beliefs. Concurrently, increased self-efficacy may be linked to an internal locus of control, wherein individuals believe that outcomes are shaped by their own actions and outcomes, rather than by factors beyond their control (Bandura, 2006; Weerdmeester et al., 2020). With continuous biofeedback training for the self-regulation of physiological states, participants in training may start experiencing a shift in their locus of control and believe that they can influence their own bodily states.

The efficacy of biofeedback interventions has been piloted in several conditions such as anxiety, chronic pain, post-traumatic stress disorder (Alneyadi et al., 2021; Berry et al., 2014; Moss & Shaffer, 2017; Sielski et al., 2017; Tan et al., 2011; Weerdmeester et al., 2020), and more recently in eating disorders too (for review see Imperatori et al., 2018). Hitherto, there are very few pilot studies assessing biofeedback-based techniques in eating disorders and only two in AN (Pop-Jordanova, 2000; Scolnick et al., 2014; for review see: Imperatori et al., 2018). The authors showed that following training AN patients were able to self-regulate their heart rate variability, increasing treatment engagement (Scolnick et al., 2014), and electrodermal activity biofeedback allowed patients to feel more in control over physiological processes which were previously felt uncontrollable (e.g., stress when choosing a meal; Pop-Jordanova, 2000).

After the implementation of HRBF at REDU, the research team was invited to conduct a clinical audit (i.e., a quality improvement process that seeks to improve patient care and outcomes through systematic review of care against explicit criteria and the implementation of care; in line with the Principles for Best Practice in Clinical Audit (NICE/CHI, 2002); Scrivener et al., 2004) and assess the HRBF. Specifically, the following processes were implemented to the practice. Patients completed a guided breathing exercise following two stress-inducing exercises (meal planning and a mathematical test; see below). They also received real time HRBF to see their heart rate changes. During the audit, measures on patients' beliefs regarding their self-regulatory abilities, their heart rate and affect changes throughout the session, as well as qualitative feedback were obtained in order to assess whether this newly implemented

practice at the ward met the local standard of embodied feeling regulation and subsequent increase in related self-efficacy beliefs.

The aims were to (i) evaluate whether HRBF following guided breathing as an enhanced intervention at REDU for embodied feeling regulation (e.g., increased heart rate) and improvements in related self-efficacy beliefs regarding such self-regulatory abilities was implemented in a way that was feasible, and (ii) to evaluate if this new intervention did in fact bring improvements to the ward's practice. The former aim was achieved by gathering qualitative feedback from the MDT and patients on the feasibility of this practice in terms of time-duration of the process, staff training to deliver psychoeducation and monitor the HRBF, but also by observations made during the audit. The latter aim had four, more specific sub-aims, namely (1) to assess the acceptability of this new intervention as opposed to previous interventions (here, mindfulness) that were in place to help patients with self-regulation, (2) to assess whether there were actual heart rate changes following HRBF and the guided breathing patients completed, (3) to assess if HRBF improved patients' self-regulatory abilities and self-efficacy beliefs regarding such abilities, and (4) to collect feedback from patients to assess if and how HRBF may be preferred to other practices they may have previously tried. These aims were assessed by acquiring data on (i) patient participation, compliance, and completion rates, (ii) recording heart rate changes throughout the session, (iii) evidence that patients became more aware of changes in their heart rate following the HRBF as measured by data on patients' ability to feel their heartbeat before and after the HRBF and related belief/attitude changes, (iii) attitudes on the HRBF vs. other ward practices, (iv) qualitative feedback on the HRBF process. Finally, to assess if feeling more in control of one's embodied feelings via HRBF brought improvements to patients' overall mood, positive and negative affect was measured via items from a standardised self-report questionnaire before and after the HRBF. Adverse events were also recorded to assess the safety of the new practice. Given the aforementioned preliminary findings on the efficacy of biofeedback interventions, there could be promising avenues for the development of biofeedback-based interventions, to be used in conjunction with the current treatments for AN and eating disorders more generally (National Collaborating Centre for Mental Health (UK), 2004; National Guideline Alliance (UK), 2017).

5.2 Methods and Materials

5.2.1 Audit Setting and Recruitment

The clinical audit took place at the Regional Eating Disorders Unit (REDU), St John's Hospital, NHS Lothian, UK and was a collaboration between clinicians from REDU and academic partners (researchers) from University College London (UCL, UK). I would like to note that this clinical audit was conducted under the leadership of Dr Richard Taylor and Dr Susan Simpson. Audits do not constitute research, therefore patients gave written informed consent as per Trust audit procedures, and data protection came under the Trust's protection and sponsorship, and the forms were stored at the Trust. My involvement was to facilitate these clinicians to conduct this audit and my supervisors advised accordingly. At the advice of the Clinical, Educational, and Healthy Psychology Department at UCL we were informed that an audit can form part of a doctorate thesis, if it is treated as an audit which tests standards of practice in a given Trust, rather than a research project with generalisable results. This Chapter is written as such.

The audit included all patients admitted to REDU with a diagnosis of AN during May 2021. Treatment of AN at REDU follows the standard NICE guidelines (as outlined in Section 5.1). Specifically, when admitted to the 12-bed ward, the patients are given a tailored meal-plan to encourage higher caloric intake and for every 1kg they gain, they progress onto the next stage of the treatment. Once the patient has completed the nine stages of the treatment aimed at weight-gain, they proceed onto the (weight) maintenance stage (also known as the tenth stage). Participation was optional, and out of the total 13 women receiving in-patient treatment at REDU during the audit period, $n=6$ (46.15%) with clinically diagnosed AN volunteered (see Table 5.1 for clinical profiles). A full record of their medical and psychiatric treatment history was obtained, and comorbidities and other clinical information were also recorded as part of standard clinical notes. The clinical audit was carried out as part of a service evaluation and in accordance with the Department of Health, with all patients providing written consent prior to participation. All patients were made aware of their rights to withdraw without providing any explanation, and that withdrawing would not affect the quality of treatment they receive.

Table 5.1. Clinical Profiles of the REDU in-patient users.

AN Subtype	$n=4$: restrictive AN; $n=1$: binge-purge AN; $n=1$: purge AN
Age $M(SD)$	31.33 (9.35)
Age at Onset $M(SD)$	22.00(7.75)
Duration (years) – Total $M(SD)$	9.33(7.75)

Duration (months) – Current Relapse	9.33(3.88)
Admissions (incl. current)	n=3: ≤3 admissions; n=2: ≤5 admissions; n=1: 10 admissions
Treatment Stage	n=1 Stage 7; n=3: Stage 8; n=1 Stage 9; n=1 Maintenance Stage
Length of Current Admission (in months; M(SD))	2.5(2.35)
BMI – Current M(SD)	15.72(3.13)
Lowest BMI M(SD)	12.32(1.16)
Highest BMI M(SD)	23.95(5.08)
DASS-21 – Depression M(SD)	29.00(14.01)
DASS-21 – Anxiety M(SD)	21.33(12.69)
DASS-21 – Stress M(SD)	26.67(7.66)
EDI-3-ID M(SD)	20.00(9.14)
SAPAS M(SD)	4.00(1.79)
TAS-20 M(SD)	68.67(10.88)
GSE-10 M(SD)	18.17(5.53)
IUS-12 M(SD)	44.17(8.33)
Insight Error Score	-0.33(0.34)
Comorbidities	Multiple Comorbidities*: n=2 Borderline Personality Disorder (BPD): n=1 Depression Symptoms: n=1
Psychoactive Medications	SSRI (n=1) Benzodiazepines (n=1)

N.B. The Treatment Stages are used to determine the clinical state of the patient in terms of their caloric intake (10 stages in total), before entering the weight maintenance stage prior to discharge. Multiple Comorbidities include: BPD, Post-traumatic Stress Disorder (PTSD), Generalised Anxiety Disorder (GAD); Depression, Social Anxiety traits; Emotionally-Unstable Personality Disorder (EUPD). Abbreviations: DASS-21 (21-Item Depression, Anxiety, and Stress Scale; Lovibond & Lovibond, 1995); EDI-3-ID (Eating Disorders Inventory 3, Interceptive Deficits Scale; Garner, 2004); SAPAS (Symptoms of Personality Disorder; Moran et al., 2003); TAS-20 (20-Item Toronto Alexithymia Scale; Bagby et al., 1994); GSE-10 (10-Item General Self-Efficacy Scale; Schwarzer & Jerusalem, 1995); IUS-12 (12-Item Intolerance of Uncertainty Scale; Carleton et al., 2007); SSRI (Selective Serotonin Reuptake Inhibitor). The Insight Error Score was calculated using the difference between patient and clinician scores, weighted by the respective clinician score. Four scores were obtained, one for each item of the scale (i.e., insight into illness severity, thinness with respect to the average women in the country, assessing how pathological the eating behaviour and food approach are), which were then averaged. A negative insight error score suggests clinical insight deficits.

5.2.2 Development of the Heart Rate Biofeedback Practice: Feedback from the PPI Event

A Patient and Public Involvement (PPI) event was held during the new, HRBF practice development. At the first PPI event the MDT from REDU presented the feedback from patients regarding the existing standard practice that was in place to help individuals attend to and manage embodied feelings, while the research team introduced evidence-based insights regarding the efficacy of biofeedback interventions in health management. Individuals with lived AN experience provided insightful recommendations regarding ways the existing practice could be improved and shared ideas for how to best incorporate the heart rate biofeedback and relevant psychoeducation components in the ward. Input from PPI events is invaluable given that the input from the public, and in particular individuals with lived experience, may facilitate identifying what is at stake, suggest important outcome measures, and improve the quality and relevance of the audit given their own expertise and lived experience. Then, all attendees were invited to an open discussion, during which two key points were raised and relevant actions were taken and thus implemented in the HRBF practice protocol.

Key Point 1 and Action 1:

Individuals with lived experience suggested the biofeedback intervention should not be ‘over-medicalised’ given that patients already receive many charts and information about their health, and additional charts would not be recommended. Instead, they suggested to only show the summary of the E4 real-time application (and not the real-time data acquisition interface) and ask patients to only focus on the summary of their heart rate data (average beats per minute), which is what was implemented in the local, enhanced standard practice.

Key Point 2 and Action 2:

The second point addressed was the need to highlight during the familiarisation and psychoeducation phases that the biofeedback would not be another way to measure their health with specific expectations. Instead, it should be presented as a way in which one could explore ways to better understand their body and responses to stimuli. It was recommended that the clinician wore the watch first, to show the patient how their heart rate can vary, and the patient would also be given time to explore the watch and its features without the clinician looking at them or giving any feedback. It was also suggested the clinician used the metaphor to explain how the body is like a book and biofeedback could be used as another way to “read” the signals

coming in from one's own body, and breathing could be used to regulate some of the signals (e.g., heart rate). It was ensured that a lengthy familiarisation and psychoeducation phase were included during which patients were able to explore the apparatus in a safe space and ask questions, if they had any.

5.2.3 Summary of Process

To assess whether the implemented HRBF intervention met the local standards of understanding and self-regulating embodied feelings, patients completed a biofeedback session during which measures to assess if this new intervention met local standards of care were obtained. The session lasted approximately 90 minutes and although it was initially scheduled to be completed in a one-off session, in some instances it was separated into two sessions, due to mealtime interruptions organised by the ward (those receiving in-patient treatment at REDU are encouraged to maintain a strict eating schedule and we thus arranged a second session in instances in which not all tasks were completed before scheduled meals). Apart from the main intervention (comprising a psychoeducation component, stress induction and guided breathing, as outlined below), patients completed a heartbeat counting task (HCT; Schandry, 1981) to obtain a measure of their ability to accurately feel their heartbeat (i.e., interoceptive accuracy, IAcc) before and after engaging in the HRBF (which is aimed at increasing their awareness of such embodied feelings). Patients were briefed on the concept of self-regulation as part of the psychoeducation component of the intervention, and after getting familiarised with the equipment, they completed the stress induction and guided breathing components of the intervention. Feasibility measures were obtained throughout the session and participants gave feedback at the end of the session.

5.2.4 Baseline Measures Obtained for Clinical Audit Purposes

Before the main completing the HRBF, a series of baseline measures had to be obtained in order to use them to assess if the enhanced standard practice using HRBF led to process improvements and met local standards. Specifically, patients completed a series of baseline assessments and tasks, and familiarised themselves with the concepts and equipment introduced during the HRBF. First, patients completed the HCT (Schandry, 1981; for full instructions and design see Chapters 2 and 3) to measure their IAcc, their beliefs about their IAcc abilities, as well as their knowledge regarding their own average resting heart rate and in

the general population. Information on patients' clinical profiles and demographics was collected.

5.2.5 REDU Local Practice of HRBF: Process

5.2.5.1 Part 1: Psychoeducation and Familiarisation

When the HRBF was introduced, patients received psychoeducation from the MDT and were also given the opportunity to familiarise themselves with the equipment through which they would be receiving HRBF and practicing guided breathing. Specifically, patients were explained that heart rate tends to increase when doing activities or when one is stressed, and it tends to decrease when recovering from an activity or relaxing. Patients were also asked to imagine that their body is a book, and with the help of the watch and application interface they would be invited to explore different ways in which they can “read” this book. It was highlighted that experiences of heart rate fluctuation are normal and occur to everyone, and the interest of this enhanced practice was in exploring whether this could be a new way that may help, at least some people, control their heart rate. Patients were also explained that some individuals have trained themselves to breathe at a pace of 6 breaths/minute, which is considered to be the optimal breathing pace (the pace also used during the HRBF) to decrease one's heart rate. Patients were then invited to familiarise themselves with the watch and the application via which they would be receiving the biofeedback (Figure 5.1, Panel A) and then practiced guided breathing (familiarisation phase; Figure 5.1, Panels B and C). Patients who participated in the audit were told that this evaluation was performed to assess whether this new, enhanced practice led to certain improvements and how it compared to other regulation practices they may have tried previously at the ward. It was made clear that this was a service evaluation, with no expectations to perform a certain way or see a certain result, and instead several measures as well as feedback from them would be obtained.

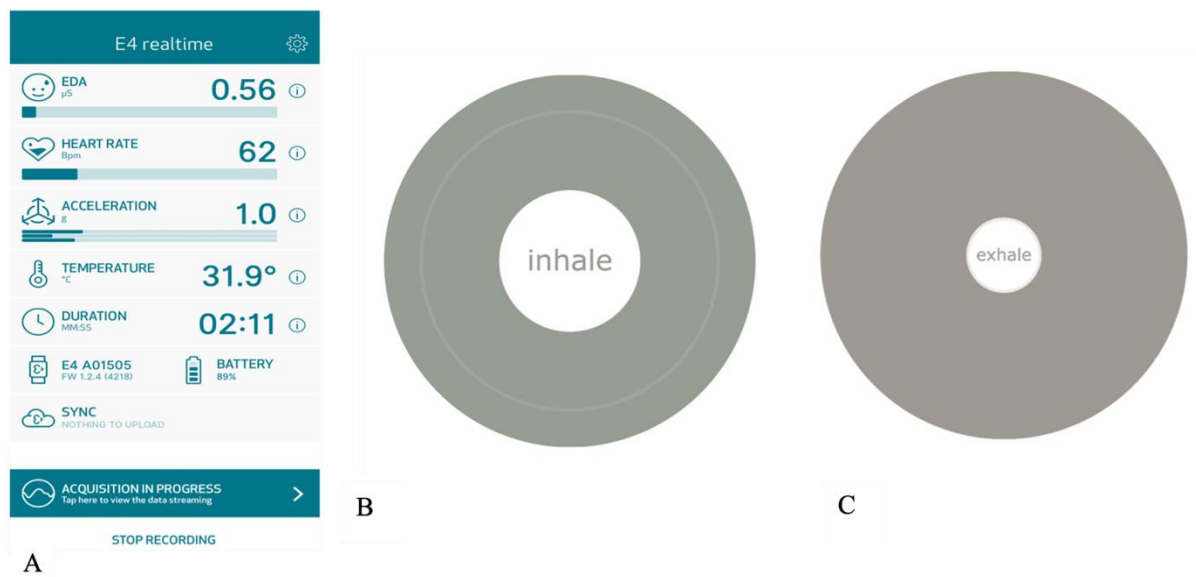


Figure 5.1. Interface of the Heart Rate Regulation Breathing Application and the Visual Aid for Guided Breathing.

Panel A shows the interface of the application that was used to give real-time biofeedback to patients. Patients were asked to only focus on the information displayed on the second row which is their heart rate (average beats/min). Patients were given this feedback before task instructions, after completing the task, and after completing the guided breathing intervention. Panels B and C show the guided breathing intervention. Patients were instructed to use the visual aid during the guided breathing intervention. Patients inhaled as the white circle on the left expanded to reach the first outer circle, and for as long as the word 'inhale' was visible in the centre of the circle (Panel A). Then, the circle started shrinking again, and patients had to exhale until the circle touched the pale, inner circle line, and for as long as the 'word' exhale was visible, as shown on Panel B.

5.2.5.2 Part 2: Heart Rate Biofeedback for Embodied Feeling Regulation

Patients at REDU were given HRBF before and after completing their meal-plan and after completing a guided breathing exercise, aimed at helping them regulate their heart rate (which typically increased during the meal-planning). To assess if HRBF worked better than other practices previously implemented at the ward to help patients self-regulate these embodied feelings and increase their self-efficacy beliefs, measures on the efficacy of the HRBF and changes in heart rate and self-efficacy beliefs regarding self-regulatory abilities were obtained during the meal planning. However, to assess whether indeed HRBF assisted patients beyond

situations directly linked with the AN symptomatology and behaviours (e.g., increased heart rate and anxiety during meal planning, or during weighing), the effect of HRBF on other daily activities was assessed. This was done by obtaining measures before and after patients completed a brief cognitive task in which patients were invited to complete subtractions on a laptop. The meal-planning (*Eating-Related Trigger Arm*) and the cognitive task (*Neutral Trigger Arm*) were completed in a counterbalanced order between patients.

i. Eating-Related Trigger Arm

AN patients often struggle with meal choice, and this was noticed when they had to select their meals at the ward. The patients are given a meal plan with fixed breakfast and snack options and different lunch and dinner options every three weeks. Making a meal decision increases their feelings of anxiety, subsequently leading to increased heart rate. Subsequently, the MDT introduced HRBF to assist with the self-regulation of embodied feelings which may arise during the meal planning. For the audit purposes, the heart rate of the patient before and after the meal planning and after the guided breathing was noted, and patients were also asked to rate their beliefs regarding their ability to control their heart rate by breathing before and after completing the meal planning. Patients had three minutes to make a lunch and mid-afternoon snack decision for the following day. Patients were not asked to plan their breakfast during the audit, as breakfast options are repeated every week, hence patients may be used to the choices and respond habitually and without much emotional engagement.

ii. Neutral Trigger Arm

To assess the improvement HRBF brings to patients' self-regulatory abilities of embodied feelings and related self-efficacy beliefs beyond symptom-specific contexts, the effects of HRBF were assessed in a neutral context. In the *Neutral Trigger Arm* patients' heart rate was increased in a neutral/control manner. Patients completed a short cognitive task, which was an adapted version of the short-mental arithmetic task used by Herbert and colleagues (Herbert et al., 2010). During the task, patients were shown a four-digit number on the screen and were asked to subtract '17' from the displayed number as quickly and as accurately as possible. Patients were shown four possible answers and were asked to indicate the correct answer by pressing one of the four respective keys placed in front of their right hand (left hand for left-handed participants). Selecting a response by pressing a key instead of saying it verbally reduced any variations in heart rate and HRV, which may have been caused by speaking and/or influenced by respiration (Tomaka et al., 1994). To match the duration of the *Eating-Related*

Trigger Arm patients were asked to stop the task after 3 minutes. Heart rate changes and related beliefs regarding self-regulatory abilities were also obtained here, as above.

5.2.5.3 Enhanced Standard Practice and Clinical Audit Measures: Heart Rate Monitoring, Guided Breathing and Belief Measurement

During the HRBF the clinician can visually monitor heart rate changes via the E4 realtime watch application. This was also done during the audit to ensure that the two stressors elicited an emotional response and in turn raised patients' heart rate, but also to observe changes in heart rate during the guided breathing. Patients were shown their heart rate before each of the task instructions (meal planning, and mathematical task). Then, they completed the task and were shown their heart rate again.

Next, patients completed the guided breathing: patients were asked to concentrate on the breathing visualisation on the laptop screen and follow the shrinking and expanding circles to regulate their breathing. The guided breathing lasted for 1.5 minutes, after which patients were shown their heart rate again.

Before and after the guided breathing patients were asked to rate their ability to regulate their heart rate by breathing. Specifically, patients were asked how well they thought they would be able to control their heart rate by breathing before engaging in the task (prior prospective belief), after being shown their heart rate upon completion of the guided breathing (retrospective belief), and they were also asked to report how well they would be able to control their heart rate in a similar, future task (posterior prospective belief), all on a scale of 0 (*not well at all*) to 100 (*extremely well*). These measures were obtained to assess whether the HRBF led to improvements in patients' beliefs regarding their self-regulatory abilities of embodied feelings, and specifically of elevated heart rate.

5.2.6 Part 3: Feedback on the process and the service improvements

Patients' views on the new practice and the potential efficacy of biofeedback-based interventions in comparison to the previous standard practice of weekly mindfulness sessions were assessed. Specifically, patients were asked to evaluate how well they think they are able to relax themselves or with the help of staff when feeling stressed at the ward, before and after engaging in the HRBF. Patients were also asked to give feedback on the guided breathing and

HRBF, share their thoughts on the difficulty of this enhanced practice, and discuss if they noticed and felt changes in their heartrate. Finally, they were asked if this enhanced practice is something they could see themselves practicing in the future, and something which could benefit them, and if they had any other comments or thoughts, they think could help improving the local standard practice. Patients were also given the option to provide further, anonymised, written feedback on their thoughts regarding the acceptability and feasibility of this new process as an add on to the previous standard practice of mindfulness at REDU.

5.2.7 Summary of Measures Used to Achieve the Audit's Aims

To assess if the feasibility of this new HRBF process, qualitative feedback from the ward's staff and patients was obtained, as well as observations during the audit. The focus was on evaluating whether the way the process had been implemented at the ward was sustainable and feasible. Next, to assess if there were improvements in the patients' treatment brought about the HRBF a few measures were used. Firstly, data on participation, compliance and completion rates were obtained and qualitatively compared against those of previous interventions aimed at meeting the same standards of care. Secondly, data on patients' heart rate change awareness were collected through (i) assessment of their interoceptive accuracy (via the HCT; Schandry, 1981) before and after receiving HRBF, (ii) recording and evaluation of heart rate changes throughout the HRBF session, (iii) comparing the change in patients' self-efficacy beliefs regarding self-regulatory abilities before and after completing the two stress-inducing tasks and HRBF, (iv) comparing attitudes on the HRBF vs mindfulness practice, (v) obtaining qualitative feedback from patients on their thoughts and opinions regarding the HRBF. Items from the Positive and Negative Affect Scale (PANAS; Watson et al., 1988) were used to measure changes in patients' affect before and after the HRBF.

5.2.8 Data Management and Analysis

This clinical audit could not use a blinding design for the patients or staff carrying out the audit. However, the data were analysed without access to information about the person, i.e., randomly generated patient IDs from Excel were used. Upon completion of the clinical audit, data were input into, analysed and plotted using Excel in a case-by-case manner.

Data analysis was primarily descriptive with summaries for each outcome measure, and without formal statistical comparison between arms (triggers). The main interest was in the

measures and establishing, whether the process was feasible and if it led to improvements in the local standards. First, there was an evaluation of the referral, recruitment, and completion rates by referral source. Reach and representativeness were described with regards to the proportions of patients who were contacted and screened for participation. Adherence through the proportion of patients who successfully completed the session and the success of the HRBF and guided breathing by comparing heart rate and self-efficacy beliefs prior to and post task were also evaluated. Behaviour change, in terms of mood, was measured by self-report questionnaires throughout the session. Feasibility of collecting outcome and resource use data was assessed by estimating the proportions of missing, incomplete, or incorrect data in each of the outcomes assessed. To monitor safety of the suggested intervention, any adverse events were noted and reported by study arm.

5.3 Findings and Observations

5.3.1 Assessing the Feasibility of the Process

Feasibility of the HRBF was deemed appropriate: one member of staff would be needed to lead the session with one patient at a time. The member of staff would need to be able to deliver the psychoeducation regarding changes in embodied feelings in response to contextual factors (e.g., increased heart rate during weighting, during an interview, etc.) and ways in which one can self-regulate (e.g., guided breathing). The reaction to getting HRBF after engaging in a guided breathing exercise to manage physiological changes in response to a stress-inducing experience (here, meal planning) was positive. The time frame in which the psychoeducation about self-regulation and HRBF were delivered, in line also with the ward's other practices and daily schedule of the patients (e.g., mealtimes, occupational therapy, etc.) was deemed appropriate. Some patients suggested the HRBF to be incorporated during group sessions with members of staff, however this needs to be further determined as the ward would then have to purchase enough watches for the patients and member of staff. This should be consulted with the NHS budget managers.

5.3.2 Assessing the Service Development

5.3.2.1 *Evaluation of Referral, Recruitment, and Completion Rates*

At the end of the first week of the audit, one patient was discharged, and another was admitted, bringing the total number of patients who could receive HRBF to 13. Out of N=13 patients at REDU, n=9 were eligible (n=2 were using nasogastric tube feeding, n=1 was bedbound, and n=1 had comorbidities which made them ineligible to receive the guided breathing

intervention). Critical health states or comorbidities which would have prevented the patients from seeking this enhanced treatment were identified as reasons for not taking part in the audit. One patient opted out due to an aversion towards heart rate monitoring practices. Out of the 13 patients, 69.23% were eligible. Out of the N=9 eligible patients, n=6 (66.6%) consented to participate in the audit by the end of recruitment period, and all participating patients adhered to and completed the session.

5.3.2.2 Documenting Heart Rate Changes

HRBF was given to six patients during the audit. The changes in heart rate were noted before and after completing each of the two stress-inducing tasks, and after the guided breathing. In the *Eating-Related Trigger Arm*, heart rate increased in 5 patients (83.3%) and decreased in one patient (16.6%) during the meal selection as seen by the difference in heart rate between the pre and post task heart rate scores (Figure 5.2). All 6 patients (100%) experienced a decrease in heart rate following the guided breathing as seen by the reduction in heart rate values at Post HRBF in comparison to Post Task (Figure 5.2). In the *Neutral Trigger Arm* the heart rate of 3 patients (50%) and four patients (66.6%) experienced a decrease in heart rate following the guided breathing. In two patients (33.3%) heart rate reduced during the task, and then decreased following the guided breathing in one and remained the same in the other one (Patient E, Figure 5.2). Finally, in one patient (16.6%) heart rate decreased during the task, and then decreased slightly more following the guided breathing (Figure 5.2). The physiological data of the patients are presented in Table 5.2.

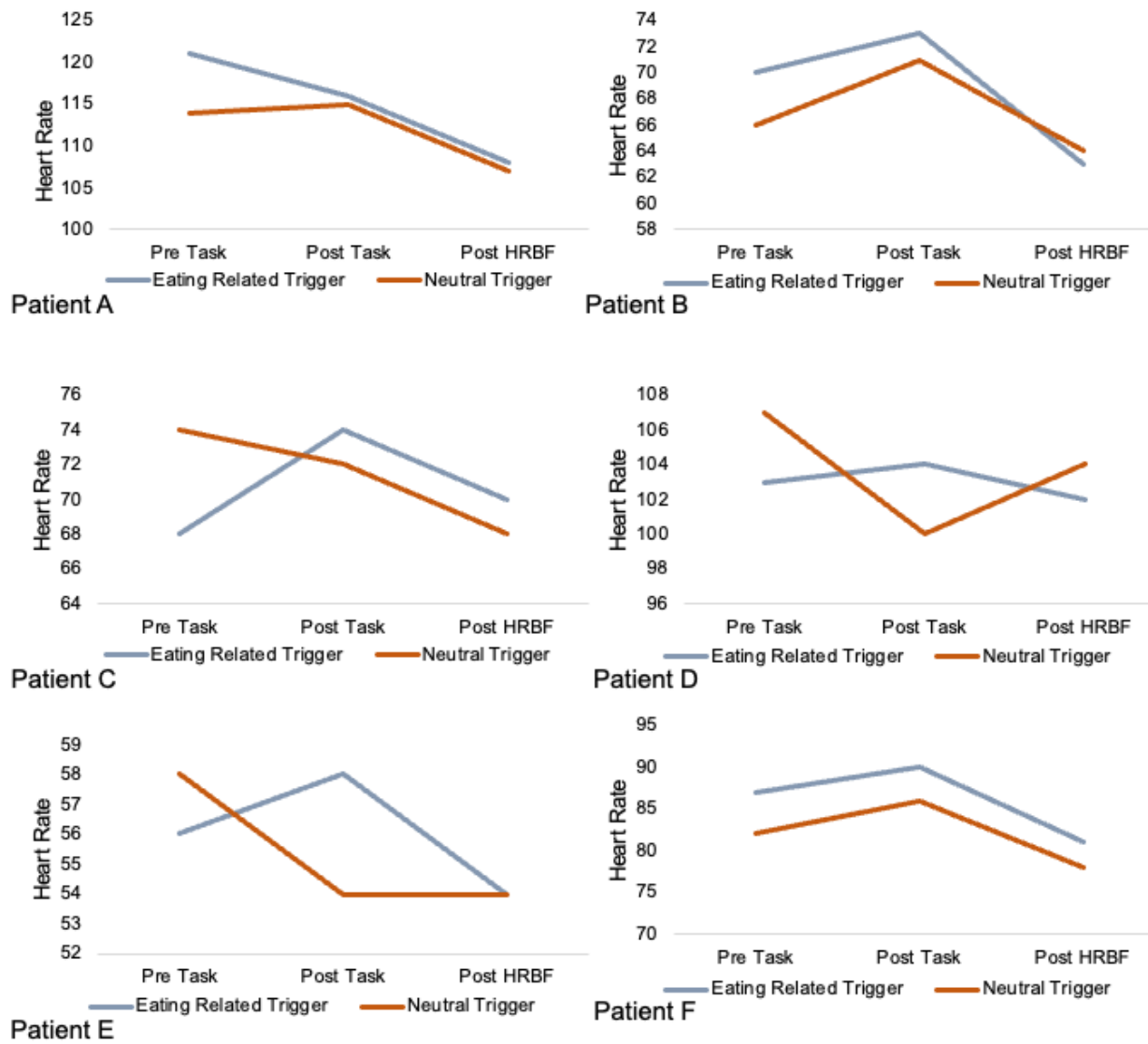


Figure 5.2. Heart rate changes per patient before and after the task.

The blue line represents heart rate changes for the Eating-Related Trigger Arm before and after the guided breathing and heart rate biofeedback (Pre Task, Post Task and Post HRBF, respectively), and the orange line represents the changes for the Neutral Trigger Arm.

Table 5.2. Physiological Data of the Patients.

	Patient A	Patient B	Patient C	Patient D	Patient E	Patient F
Blood Pressure on Admission Day	101-110/71-80	95/56	110-120/81-90	101/65	81-90/61-70	95/55
Pulse on Admission Day	76	73	49	57	45	83

Blood Pressure on Intervention Day	99/69	109/67	86/53	104/70	81-90/51-60	91/57
Pulse on Intervention Day	108	69	73	98	70	79
Patient Resting Heart rate at the Start of the Session (bpm)	112	73	66	87	56	96
Patient's Self-Reported Resting Heart Rate at the Start of the Session (bpm)	95	72	54	80	50	40
Patient's Self-Reported Resting Heart Rate at the End of the Session (bpm)	80	70	68	85	50	60

5.3.2.3 Evaluating Changes in Heart Rate Awareness (Interoceptive Accuracy)

Interoceptive Accuracy (IAcc) was measured using the Schandry (1981) formula (see Section 1.3, Chapter 1 for IAcc definition and Section 2.2, Chapter 2 for calculation formula). A cardiac IAcc score, as in this case, is used to assess how well one performed on a behavioural test of heartbeat detection, with a score of or close to 0 suggesting poor ability to feel one's own heartrate, while a score of 1 suggests that one was able to feel all the heartbeats they actually had during a given timeframe. To evaluate if the HRBF indeed helped patients become aware of their heart rate changes (and thus help them with practicing self-regulation), an IAcc score was obtained before the HRBF and compared against changes in the patient's IAcc score post-HRBF. Out of the 6 patients, 4 showed improvements in their abilities to notice their heartbeat after completing the HRBF, while the abilities of 1 did not change (Figure 5.3). Only 1 patient demonstrated a slightly worse performance and a lower score post-HRBF.

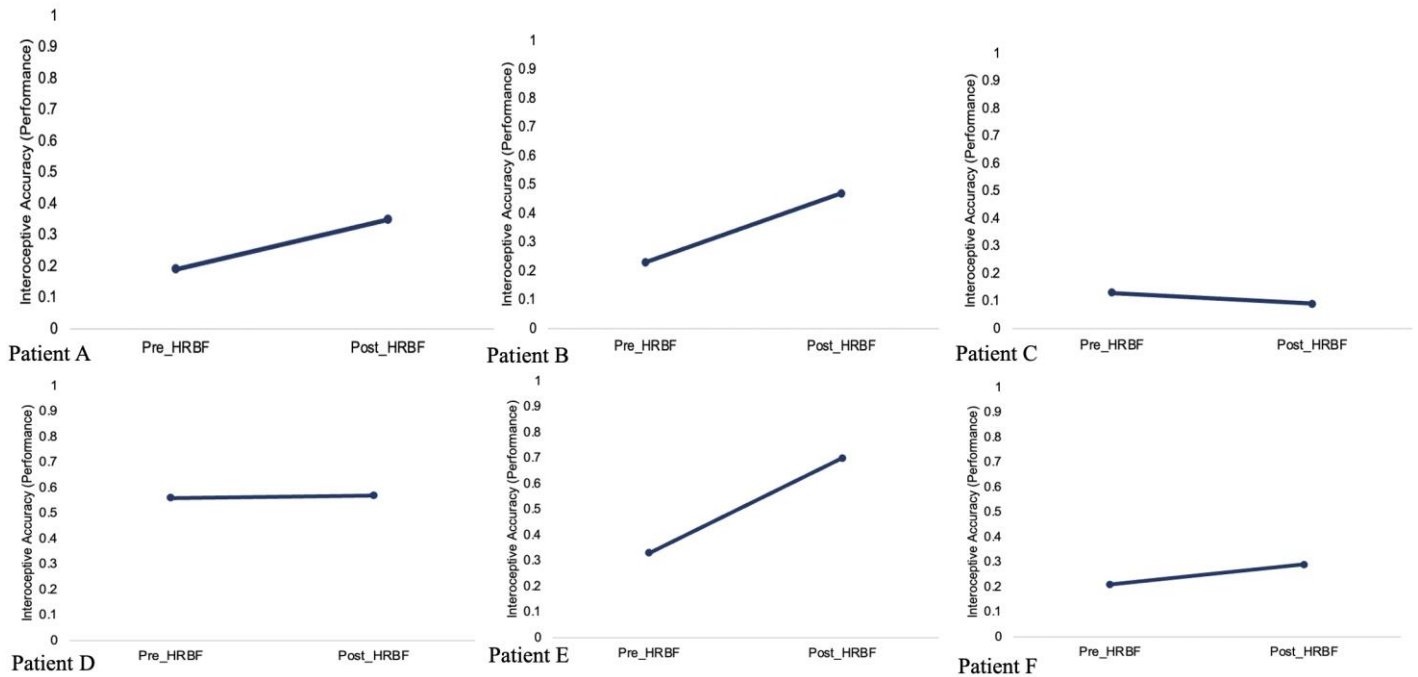


Figure 5.3. Interoceptive Accuracy before and after the heart rate biofeedback.

The changes in patients' Interoceptive Accuracy Scores (i.e., Performance in the Heartbeat Counting Task; Schandry, 1981) were measured before and after receiving heart rate Biofeedback (Pre_HRBF, and Post_HRBF, respectively).

5.3.2.4 Assessing Improvement in Self-Regulatory Ability (i.e., Self-Efficacy) Beliefs

When looking at patients' beliefs regarding their ability to control their heart rate by breathing and their beliefs regarding the difficulty of controlling their heart rate by breathing during the eating-related trigger arm, it was observed that: 4 patients (66.6%) retrospectively thought they controlled their heart rate better than they expected, 1 patient (16.6%) gave the same estimate before and after the intervention, and 1 patient (16.6%) thought she did not control her heart rate by breathing as well as she thought she would before the task (Figure 5.4). When asked to estimate how well they would be able to control their heart rate by breathing in a similar, future scenario, 5 patients (83.3%) thought they would do better than they did, and one (16.6%) thought she would do worse (Figure 5.4). Similarly, in the neutral trigger arm, 2 patients (33.3%) thought they did better than they expected, while 2 (33.3%) did the same than they expected, and 2 (33.3%) thought they did worse than they expected (Figure 5.4). However, 5 patients (83.3%) reported they thought they would do better than they did if the task was repeated in the future, under similar situations, whereas one patient (16.6%) thought they would do worse. Additionally, across both trigger arms, the majority of patients (100% in the neutral

trigger arm, and 83.3% in the eating-related trigger arm) reported they would be worse at controlling their heart rate by breathing in the future, if they did not have the breathing pace aid (in this case, the guided breathing via the expanding and shrinking circles displayed on the laptop screen; Figure 5.4). Overall, 5 patients (83.3%) reported they trusted the accuracy of the biofeedback they received regarding increases and decreases in their heart rate above chance ($M(SD)$: $76(14.75)$ and $79(14.32)$, respectively), and the same 1 patient (16.6%) reported she trusted the accuracy of the biofeedback on her increasing and decreasing heart rate below chance (18% in both instances, measured on a scale of 0 *not accurate at all* to 100 *extremely accurate*). This patient reported that they liked the breathing, and it was a helpful add-on to the standard practice but “*maybe without the feedback. It is more of what I feel in me. A lot of patients with eating disorders can get hooked on numbers. The breathing should be more than numbers. It is about how you feel. The science in me would like a graph over a number. I felt relaxed and calm, but your app said something different*”.

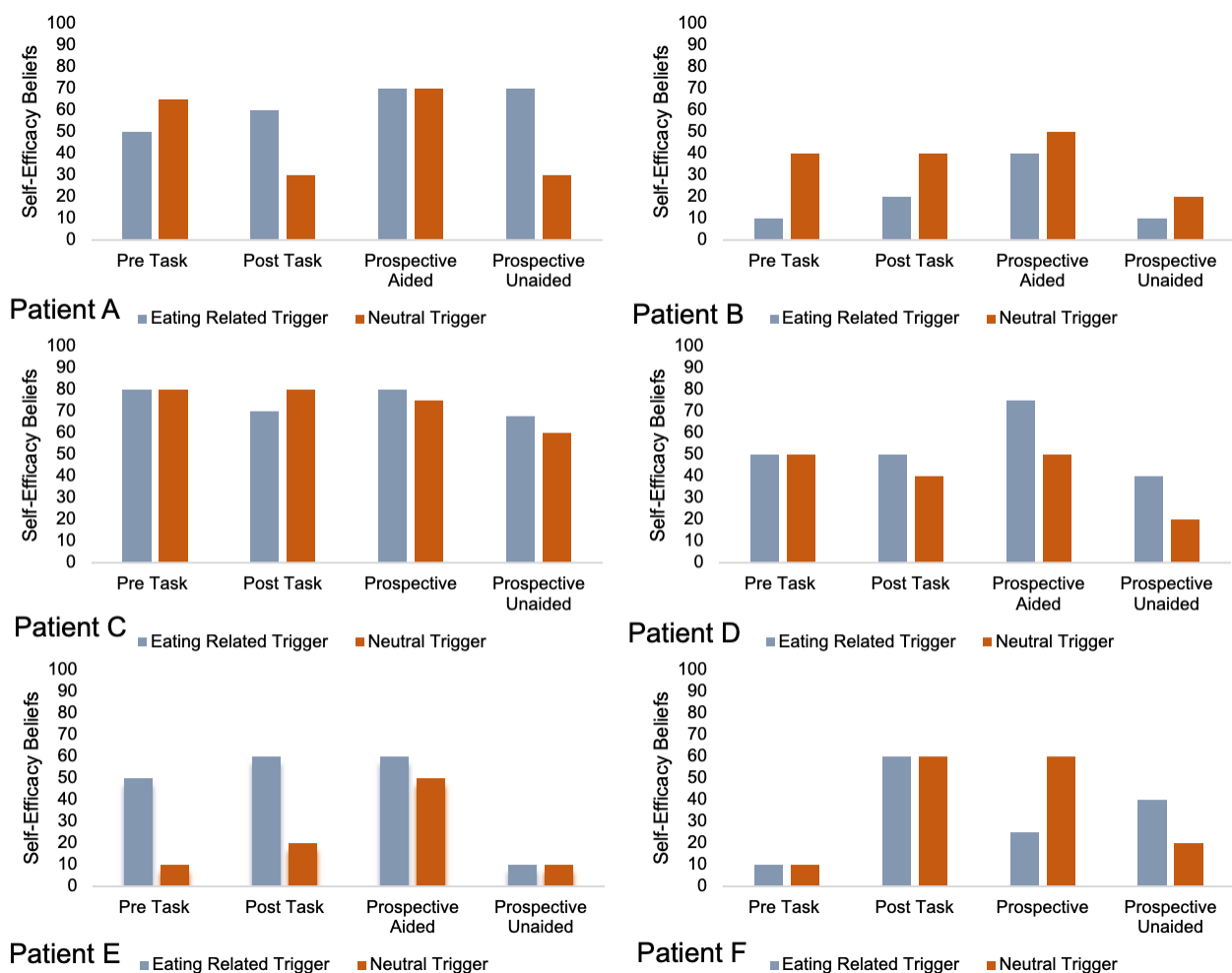


Figure 5.4. Patients’ self-efficacy beliefs before and after the tasks.

The figure demonstrates patients' self-efficacy beliefs regarding their ability to control their heart rate by breathing before starting the eating-related and neutral trigger arm tasks, respectively, post task where they had to provide a retrospective estimate on how well they were able to control their heart rate by breathing, and two prospective beliefs regarding their ability to control their heart rate by breathing in the future with and without the breathing pace aid.

5.3.2.5 Assessing Service Improvement: Comparing Attitudes Towards the Enhanced Practice (HRBF) vs Previous (Mindfulness) Standard Practice

Next, when looking at their global beliefs regarding their ability to relax themselves before attempting HRBF and after receiving HRBF on self-regulatory abilities, all 5 patients (83.3%) reported that after receiving visual HRBF to see how their heart rate changed before and after engaging in a stress-inducing task and after engaging in guided breathing, they would be better able to relax themselves the next time they were stressed at the ward, and 1 patient (16.6%) reported that they would be able to relax themselves the same way (in comparison to how well they could before the intervention). Additionally, 3 patients (50%) reported that they would be better, and 3 patients (50%) reported they would be the same at relaxing themselves with the help of the staff at the ward next time they were stressed after they had done guided breathing and received biofeedback as opposed to before trying this new, enhanced practice. More importantly, ratings from all patients (100%) showed that patients thought they would be better at relaxing themselves if doing the guided breathing together with the help of staff at the ward, as compared to doing it on their own (Figure 5.5). Finally, all patients (100%) were open to maintaining HRBF as a practice at the unit and they also said that completing the guided breathing during the meal-planning exercise helped them self-regulate and manage the feelings of stress associated with selecting meals they would need to consume.

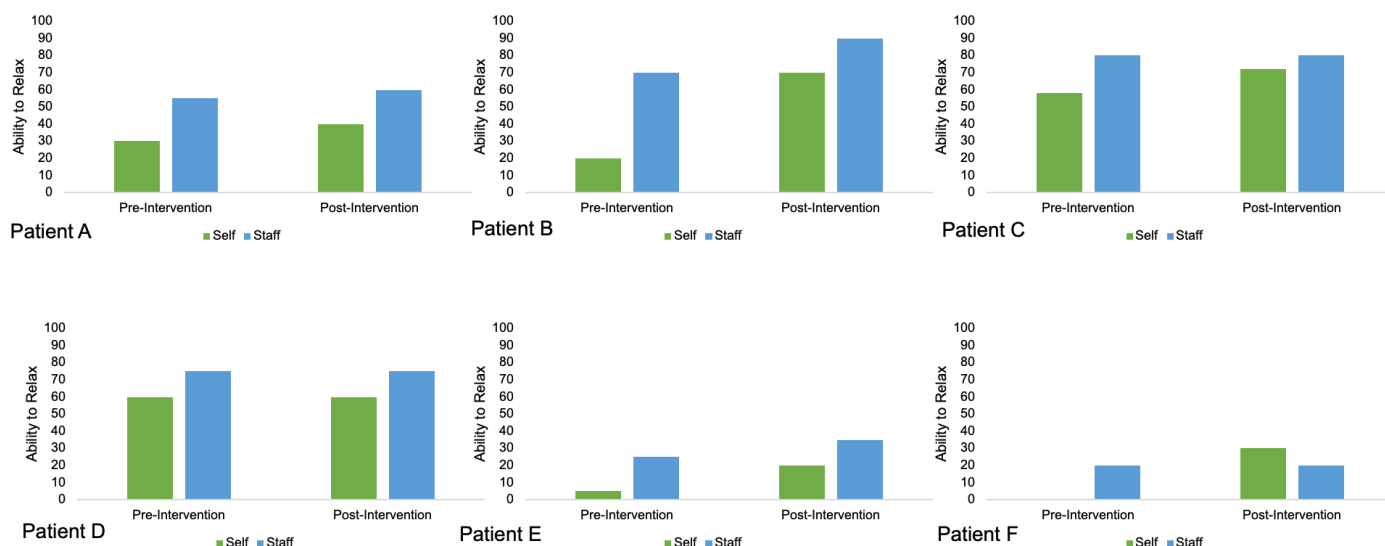


Figure 5.5. Beliefs about self-relaxation vs. assisted relaxation abilities.

The Figure shows patients' estimates in their ability to relax themselves on their own (green bars) and with the help of the staff (blue bars) when feeling stressed at the ward. N.B. Pre-Intervention is a measure of patients' estimates before trying the guided breathing with heart rate biofeedback (HRBF), and Post-Intervention is a measure of their estimates after engaging in guided breathing with HRBF.

5.3.3 Qualitative Feedback

Comments from the semi-structured interviews with the patients indicate the guided breathing with HRBF was a helpful add-on to the previous standard practice of mindfulness at REDU, but several factors which are explained in more detail in the discussion and action plan suggestions, were raised. For instance, 3 patients (50%) said that while they found the process helpful, they do not think it would be ideal for patients who are in a very acute AN state and just admitted in the hospital, and instead would be more beneficial throughout later stages of the remission and recovery process. Patients said that having the biofeedback helped them understand certain bodily signals *"I don't often know when I get stressed, so the biofeedback helped. I think seeing my pulse changes could give me warnings. I am bad at understanding my body, so biofeedback would help me know when I am starting to escalate"*. Another patient said *"Seeing the heart rate changes made me think it's [the biofeedback] useful. I was surprised my heart rate was not higher in the maths task and it changed my confidence and my ability to self-regulate"*, *"Now that I have seen the difference [with the biofeedback], I will do it [guided breathing and mindfulness]"*, and *"I guess you can control your heart rate by breathing"*. Also,

3 patients (50%) expressed they would have liked a shorter duration of the breathing exercise (the one in the audit was 3 minutes), but when asked if it was too tiring, they said it was acceptable. Patients also expressed that knowing about heart rate changes (e.g., during the weekly weighing, or mealtimes) was useful, as some expressed that *“it would be like a warning. I liked the [application] interface and circles (...). I found it quite calming and made changes more obvious”*, also adding that they would like the idea of having the guided breathing tools (here, expanding and shrinking circles) on the canteen flatscreens as a reminder to breathe at a certain pace when feeling overwhelmed during mealtimes, or other stressful instances. Finally, patients reported that they would like to practice HRBF all together and with the staff, as the HRBF was initially introduced during the 1:1 sessions with the clinician.

5.3.3.1 Affect (Mood) Changes

To assess if feeling more in control of embodied feelings as indexed by heart rate changes following guided breathing (via the HRBF) also improved patients' mood, the clinicians asked the patients to provide affect ratings before and after the HRBF. Positive and negative affect changes were measured using 4 items from the Positive Affect and Negative Affect Schedule (PANAS; Watson et al., 1988). Items were rated on a 1 (*very slightly* or *not at all*) to 5 (*extremely*) scale and scores were generated by calculating the sum on the positive and negative affect items separately. Positive affect scores remained the same in 3 patients (50%), they slightly increased in 1 patient (16.6%) and slightly decreased in 2 patients (33.3%; Figure 5.6). Negative affect scores remained the same in 2 patients (33.3%) and slightly decreased in 4 patients (66.6%).

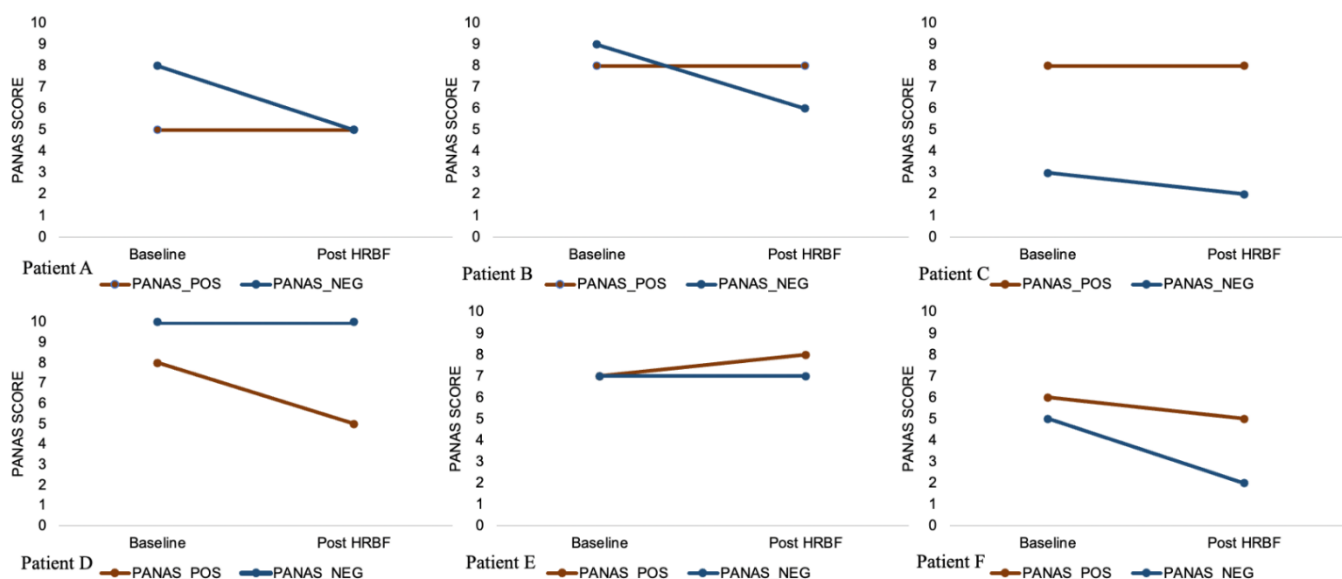


Figure 5.6. Positive Affect and Negative Affect Scores.

Patient positive (PANAS_POS) and negative (PANAS_NEG) PANAS scores before (Baseline) and after receiving heart rate biofeedback (Post HRBF). Abbreviations: HRBF (Heart Rate Biofeedback).

5.3.3.2 Adverse Events

One adverse event was recorded: one patient who had to split the session in two parts due to a coinciding mealtime reported extreme tiredness when they were invited to complete the second half of the session. The patient discussed the event and their feelings with a member of the team and was able to identify what triggered the tiredness. However, the patient did not wish to withdraw and asked to reschedule the remaining of the session for the next day, and successfully completed the intervention during the rescheduled slot. No other adverse physical responses or other events were noted during the audit.

5.4 Discussion

This audit had two primary and two secondary aims. The primary aims were to assess the process feasibility and service improvement after the REDU MDT introduced heart rate biofeedback (HRBF) with guided breathing to enhance their previous practice of mindfulness to meet the standard of aiding patients understand and self-regulate embodied feelings, such as heart rate signals. Using qualitative feedback and observations during the audit, it was shown that the new process was feasible and possible to implement as a practice at the ward. Next, measures to assess patients' acceptability of the new intervention (as measured by looking at participation, compliance, and completion rates), as well as measuring patients' ability to feel their heart rate before and after the HRBF, tracking actual heart rate changes during the session, and assessing changes in self-efficacy beliefs showed that this new, enhanced practice brought improvements to the patients' ability to understand and self-regulate embodied feelings. Moreover, qualitative feedback from the patients also shows their acceptability towards this new practice and comparisons to the previous standard practice which included optional group mindfulness sessions. The two secondary aims were to assess mood improvements brought by the improvement in patients' ability to self-regulate embodied feelings, and also assess the safety of the HRBF and guided breathing by recording any adverse events that occurred during the sessions. The observations, action points and recommendations, and limitations are discussed in turn below.

5.4.1 Primary Aim 1: Feasibility of the Heart Rate Biofeedback Process

During the audit it was observed that it would be feasible and effective to maintain HRBF as a standard practice. Staff members were appropriately trained to deliver psychoeducation to patients regarding ways in which one can notice and self-regulate embodied feelings such as heart rate. Moreover, the HRBF session fit within the ward's other practices and was delivered in a time-efficient manner and the process was simple and easy to follow for patients.

5.4.2 Primary Aim 2: Assessing the Service Improvements

Adding HRBF as an additional tool to the existing practices of mindfulness, guided breathing, and other techniques to improve understanding of one's body and feelings demonstrated to be effective in meeting the standards of care. Firstly, heart rate was reduced following the guided breathing and after seeing these changes with the HRBF most patients not only were able to better and more accurately feel their heartbeat (i.e., interoceptive accuracy improvements) but they also evaluated their ability to regulate their heart rate as better than they thought it would be and had higher self-efficacy beliefs for such abilities in the future. Importantly, there was an improvement in patients' self-efficacy beliefs regarding their ability to relax themselves at the ward after they engaged in the HRBF-enhanced practice, in comparison to the self-efficacy beliefs they had regarding self-regulatory abilities before the HRBF-enhanced practice was introduced (i.e., when such beliefs were measured after the mindfulness sessions – both quantitatively and qualitatively). Specifically, patients' post-HRBF self-efficacy beliefs were higher in comparison to their post-mindfulness self-efficacy beliefs (i.e., when asked how well they think they can relax themselves on their own/with the help of the staff when feeling stressed at the ward).

To the best of my knowledge, this was the first time that an eating disorders unit incorporated HRBF to enhance the local standards of care, and specifically facilitating the awareness and regulation of embodied feelings/'inner bodily sensations' in a population with AN. The existing local practices were evaluated as insufficient to meet these standards of care, as both first-line treatment options (here, CBT-ED) and alternative therapies (e.g., mindfulness, guided breathing) did not produce desirable effects in helping patients with self-regulation and increasing relevant self-efficacy beliefs, as seen by earlier REDU service evaluations, verbal patient reports, and attendance to the optional therapy sessions aimed at improving one's understanding of physiological responses and embodied feelings. It was shown that enhancing

evidence-based and third-wave CBT approaches with the use of HRBF was better than the CBT approach of safely exposing the individual to enhanced bodily sensations (e.g., hyperventilation following exercise) in order to facilitate appraisal and regulation of these sensations (Deacon et al., 2013; Heim et al., 2022). Others have also shown that biofeedback could be helpful in reducing common AN feelings (e.g., anxiety) or other eating disorder related behaviours (Imperatori et al., 2018; Meule et al., 2012; Pop-Jordanova, 2000; Rodríguez-Ruiz et al., 2009; Scolnick et al., 2014). Although the present audit showed that enhanced local practice led to improvements, research needs to be conducted to examine the extent to which HRBF may be better than other practices.

When evaluating patient response to the ward's newly implemented practice of HRBF, patients reported they preferred it to previous practices, and felt it was an effective way to help them learn how to regulate their heart rate and feelings of stress/anxiety they may experience. Patients were accepting of the HRBF, also in line with clinical pilot studies who have shown patients' openness to biofeedback-based interventions (e.g., Pop-Jordanova, 2000; Scolnick et al., 2014). Patients mentioned that there are instances during which they may not be familiar with physiological states and having a tool as an aid to prevent the sudden escalation was helpful. Importantly, several patients mentioned that although they had been encouraged to practice guided breathing or similar exercises in the past, the lack of insight into whether it was working or not was a preventative factor in them practicing more often. Patients said that after receiving HRBF and seeing how their heart rate changed "*I will do it*" (see Findings and Observations). This qualitative feedback, together with the quantitative changes in patients' self-efficacy beliefs are in line with the main aims of the audit, namely assessing patient receptivity of and response to the HRBF, becoming aware of heart rate changes, and evaluating whether it led to improvements in the quality-of-care patients receive at REDU.

5.4.3 Action Points and Recommendations

Taking the above observations into consideration, coupled with the patients' self-reported mood changes that occurred after they saw how one can self-regulate their embodied feelings, and the safe way in which the HRBF with guided breathing were administered, it was suggested that it would be feasible and effective to maintain HRBF with guided breathing as a standard practice. However, a few practical and logistical hurdles would have to be overcome in order to ensure proper implementation of this enhanced practice. First, the appropriate funding would

need to be secured to purchase the necessary biofeedback equipment for patients and staff at the ward and cover the training of practicing staff. Another point to eventually consider would be whether discharged patients will be discontinuing the practice of HRBF, or whether some sort of equipment could be provided to them. Yet, the long-term dependency effects on the biofeedback apparatus should also first be considered. The below action points are thus recommended:

1. HRBF to be maintained as a weekly treatment component in addition to the CBT-ED and mindfulness sessions.
2. HRBF to be delivered during meal-planning and consider the HRBF delivery during weekly weigh-ins, or other stress-inducing times at the ward.
3. HRBF to assist patients increase their self-efficacy beliefs regarding self-regulatory abilities with the aim to not be dependent on HRBF to recognise fluctuations in embodied feelings.
4. To determine the treatment stage during which HRBF should be introduced.
5. To ensure appropriate staff training to deliver the psychoeducation and monitor the HRBF delivery.
6. To ensure that all staff have received appropriate technical skills training to assist with and monitor the HRBF.
7. To consider simplifying the HRBF delivery by trying to use wearable devices that include an easy-to-follow guidance and feedback interface, such as the Polar wrist worn device.
8. To conduct a reaudit to provide an opportunity to reassess if this aspect of service meets local standards of care once formally implemented.
9. Future audits could test how many biofeedback sessions are actually needed to assist individuals in self-regulation skills and self-efficacy belief improvement.
10. Future audits also test how to best use the HRBF to see improvements in self-regulatory abilities beyond Anorexia Nervosa (AN)-related activities that may be perceived as stress inducing.
11. Further physiological specificity and calibration may be best in other conditions, but at the ward the wearables appeared sufficient for most patients.

5.4.4 Limitations

Despite the insightful snapshot data on a range of observations, this audit presented with certain limitations. Firstly, the sample size of patients who could receive the HRBF was relatively small and given this was not a research study, a sample size calculation and/or formal statistical analyses, were not performed. Moreover, less than 50% of patients' data were used which means the suitability of the intervention should be further evaluated, as factors such as patients' clinical state influenced eligibility. Clinical state and the suitability of the intervention was also received as feedback, wherein it was discussed that severely acute patients may be getting too fatigued to practice guided breathing and receive HRBF, and during the early admission stages other criteria, such as stabilising the patient's clinical status are at stake. However, it should be highlighted that the MDT noted that overall, more patients were willing to participate in the HRBF in comparison to the number of patients who used to participate in the mindfulness and guided breathing sessions. Additionally, the HRBF was practiced when patients were meal-planning, but this was not suitable for the two patients receiving feeding via a nasogastric tube. Thus, for these patients, the suitability of the HRBF could be assessed under other AN-relevant situations that are entailed in the ward's practices and which may be deemed as stressful (e.g., during the weekly weigh-ins patients must complete as part of their weight monitoring and therapy). Secondly, patients were shown their heart rate at specific timepoints, and it is not possible to know whether their beliefs about their ability to regulate their heart rate by breathing were influenced more by the biofeedback they received, or by a general knowledge of their heart rate and other factors which may not have been captured by a moment's heart rate value. Importantly, certain patients may not have been able to regulate their heart rate at their full potential given that the members of staff were present while they were completing the guided breathing with HRBF, and this social presence of another individual may have also influenced their physiological state. However, it is important to note that this audit was primarily concerned with assessing the process in which the new service was implemented and measure the improvements it brought in comparison to the standard practice of mindfulness and was not a study to investigate if and how HRBF worked, and which factors may have influenced the outcomes. Nonetheless, given the observations on the feasibility and acceptance of HRBF at a clinical setting, it was decided to formally test how a biofeedback-assisted intervention for eating and somatic symptom awareness could be implemented and how it differs to a standard mindfulness intervention (preregistration: <https://doi.org/10.1186/ISRCTN16762367>; [ISRCTN16762367](https://doi.org/10.1186/ISRCTN16762367)).

5.4.5 Conclusion

To conclude, the present audit evaluated whether adding HRBF to existing practices of mindfulness at a local eating disorders unit for individuals receiving treatment for AN was a feasible process that also led to service improvements by meeting local standards of facilitating embodied feeling awareness and self-regulation. The mood changes that patients experienced as a result of the HRBF were documented and the safety of the intervention was also considered. Observations and feedback from members of staff and patients reveal that for the most part the HRBF process was feasible in the ward, despite its technological nature, although some recommendations to increase its feasibility were formulated. It was also evident this enhanced practice helped patients become aware of their heart rate changes during stressful situations and guided breathing but also led to experienced changes in their self-efficacy beliefs regarding their self-regulatory abilities as noted by (i) the participation, compliance and completion rates, (ii) physiological changes in patients' heart rate, (iii) increased awareness in of their heartbeat post-HRBF, (iv) improvement in self-efficacy beliefs, (v) acceptability and attitude towards the HRBF in comparison to other proposed or available practices, and (vi) positive feedback regarding this service improvement. Taken together, these observations offer promising opportunities for the further evaluation of this practice development. The benefits of HRBF-enhanced interventions should be examined further before formally creating and implementing an intervention module in the local and/or national guidelines for the treatment of AN.

Chapter 6: General Discussion

6.1 Introduction

The overarching aim of this thesis was to study the formation and updating of perceptual and metacognitive beliefs about internal bodily signals (interoception) and their relation to eating disorders, particularly Anorexia Nervosa (AN). As discussed in more detail in Chapter 1 (see Section 1.3), interoception refers to the process of sensing, integrating, and interpreting physiological signals, such as signals of hunger and heart palpitations (Khalsa, Adolphs, et al., 2018). Disturbances in interoception have been associated with various psychopathologies, including AN where it is assumed that difficulties in sensing nutritional depletions may be associated with reduced food-seeking behaviours (Brown et al., 2022; Kerr et al., 2016; for reviews see (Jacquemot & Park, 2020; Khalsa et al., 2022). Moreover, efforts have been made towards untangling the different dimensions or levels of processing of interoception, ranging from lower-level abilities including the processing of neural and preconscious levels, to higher-order interpretations (for the initial three-dimensional model see Critchley & Garfinkel, 2017; Garfinkel et al., 2015; for more dimensions see Khalsa, Adolphs, et al., 2018; Suksasilp & Garfinkel, 2022). Yet, although the role of different levels of interoceptive processing in mental health (e.g., Bogaerts et al., 2010; Lavalley et al., 2023; Paulus et al., 2009; Paulus & Stein, 2010; Smith et al., 2020) and in AN more specifically (e.g., Crucianelli et al., 2016, 2021; Khalsa et al., 2015; Kinnaird et al., 2020) is being increasingly recognised, no previous study on AN had investigated deficits across domains and dimensions of interoception. Moreover, the higher-order aspects of interoception such as interoceptive awareness or insight (e.g., the correspondence between how accurately one performed in an interoceptive task and their subjective uncertainty ratings regarding that performance) is closely linked with the concept of metacognition (for more see Section 1.4, Chapter 1). Metacognition broadly refers to one's ability to reflect on, evaluate, monitor, and optimise the success of one's own cognitive processes across domains (Flavell, 1979; Fleming & Lau, 2014; Hoogervorst et al., 2023; Morales et al., 2018). Metacognition is also a multifaceted concept which is of high clinical relevance, especially when looking at the 'global' aspects of metacognition – global metacognition refers to one's overall sense of confidence about one's own abilities (also referred to as global self-performance estimates, or self-efficacy beliefs; Bandura, 1977; Rouault et al., 2019; Stephan et al., 2016), and has also been associated with the subjective experience of mental health symptoms (Seow et al., 2021), including explicit beliefs about one's cognition and related (dis)abilities, or its consequences (termed clinical insight; David,

2004). Yet, clinical research has mainly looked at the role of metacognition and such self-beliefs in treatment adherence and prognosis (see more in Section 1.4.2, Chapter 1), while research traditions have typically focussed on the ‘local’ aspect of metacognition, most often assessed by collecting trial-by-trial (retrospective) confidence ratings on tasks of perceptual or memory performance. Furthermore, research traditions have also focussed on assessing the domain-specificity or generality of mechanisms, by exploring if deficits in one domain (e.g., visual perception) also extend to deficits in another domain (e.g., memory; see Section 1.4.1, Chapter 1). Hence, while a wealth of insights into the way ‘local’ retrospective beliefs are formed in an experimental set-up and the effects of ‘global’ metacognition in clinical settings has been acquired, little is known about (i) the association between local and global metacognition in the formation and updating of interoceptive beliefs, (ii) their role in the AN psychopathology.

6.1.1 Summary of the Thesis’ Main Aims

Therefore, the present thesis had a main two-fold aim to address these gaps in interoception and metacognition: to experimentally and computationally investigate how AN patients and weight-restored patients (AN-WR) differ from age- and gender-matched healthy controls (HCs) in the metacognitive formation and updating of beliefs regarding their cardiac interoceptive abilities (Chapters 2 and 3). More specifically, the experiment presented in Chapter 2 was the first of the two studies on cardiac interoception; the preliminary findings presented in Chapter 2 were used to develop the cardiac interoception task further and preregister the experimental and computational study presented in Chapter 3. The experiment in Chapter 3 adds onto the findings of Chapter 2 as its design allowed to study how AN and AN-WR patients in comparison to HCs use nested, local and global retrospective beliefs about their ability to accurately feel their heartbeat and update them prospectively. Moreover, the addition of further measures (e.g., confidence ratings on accuracy of beliefs) allowed to create computational models to study which parameters best explain how participants update their beliefs (e.g., if they rely more on their actual performance in the task or their self-efficacy beliefs regarding that performance when prospectively predicting how well they would perform in a similar task in the future). Relatedly, to address the second part of the main aim, namely to characterise the domain-specificity of the predicted differences in forming and updating beliefs about cardiac interoception, Chapter 4 investigated how explicit beliefs about one’s abilities to predict future states of stomach fullness may change and update over time during a water-sipping task as well as which parameters may influence the prediction of a fullness state (e.g.,

exteroceptive, visual cues about the quantity they need to sip to make an interoceptive state inference). Inspired by certain clinical observations and the experimental findings from the aforementioned empirical studies (Chapters 2 – 4), Chapter 5 presented a clinical audit which was conducted after a local NHS eating disorders unit implemented new processes to their existing practices. The audit assessed the acceptability of the new heart rate biofeedback-enhanced intervention as opposed to previous, standard-practice interventions (mindfulness), and the effects on patients' self-regulatory abilities and changes in self-efficacy beliefs regarding such abilities.

In this concluding chapter, the findings of all the experimental studies and the clinical audit observations will be reviewed collectively, and methodological issues (including design and analysis limitations, and sampling issues), as well as the clinical impact of the findings will be discussed. Finally, future research opportunities will also be considered.

6.2 Summary of Main Findings

Overall, the main findings from this thesis revealed that patients with AN have a marked pessimism regarding their interoceptive abilities and often rely on these (pessimistic) prior beliefs regarding such abilities rather than in-the-moment (embodied) evidence from task performance to update such beliefs retrospectively or, prospectively. Additionally, it was found that the rate at which patients with AN use new information (e.g., evidence from cardiac perception task performance, or feelings of subjective stomach fullness) to update their beliefs is lower than that of HCs, which could be suggestive of a difficulty in using new information to update prior beliefs about an interoceptive state. Yet, while metacognitive deficits in AN were observed across both the cardiac and gastrointestinal modalities, the clinical AN groups showed comparable performance to HCs in their ability to accurately feel their heartbeats but showed differences in their perception of stomach fullness. These results suggest that interoceptive difficulties in AN may be dimension- and domain-specific. Additionally, the observations from the clinical audit suggest that using heart rate biofeedback to target self-efficacy beliefs and self-regulatory abilities in patients receiving AN treatment is a feasible process, met the local standards of practice, and numerous recommendations were formulated to facilitate the implementation of this intervention and increase the feasibility of the process. These results are extensively discussed in turn below.

6.3 Perceptual and Metacognitive Aspects of Interoception and their Domain-Specificity

6.3.1 Interoceptive Accuracy

Literature findings on the abilities of individuals with AN to perceive interoceptive signals have been mixed across the domains of interest, as was the case in this thesis. AN and AN-WR participants showed comparable performance to HCs when asked to feel their heartbeats without guessing or attempting any physical manipulation which could aid the detection of heartbeats (Chapters 2 and 3). These findings are consistent with other findings on cardiac interoception (e.g., Kinnaird et al., 2020; Lutz et al., 2019; though contrary findings also exist, e.g., Pollatos et al., 2008, 2016). These differences could be explained by sampling differences, and task instructions (for details see Section 2.4, Chapter 2 and Section 3.4, Chapter 3). For example, previous studies have looked at cardiac interoceptive accuracy differences across both subtypes of AN in comparison to controls (e.g., Pollatos et al., 2008), as opposed to the present thesis studies (Chapters 2 and 3) which only focussed on patients with the restrictive AN subtype. Moreover, a recent study showed that performance in the heartbeat counting task (HCT) is sensitive to task instructions (Desmedt et al., 2020). That is, participants' interoceptive accuracy scores are less influenced by confounds (e.g., time estimation abilities and knowledge about heart rate in the general population) when participants are encouraged to only report the heartbeats they felt rather than guess. Participants in the studies presented in Chapters 2 and 3 were given these adapted (no guessing allowed) instructions, and as tested here, possible confounds were not found to significantly affect participants' cardiac interoceptive accuracy scores.

Previous literature findings have also suggested that poor interoceptive accuracy in AN could be linked to the noted difficulty of AN patients in identifying and describing feelings, and externally-oriented thinking (also referred to as alexithymia; Kessler et al., 2006; Peres et al., 2020; for reviews see Lavender et al., 2015; Oldershaw et al., 2015). Although both AN and AN-WR groups across the experiments presented in Chapters 2 and 3 showed greater scores of alexithymia in comparison to the controls (as measured using the Toronto Alexithymia Scale, TAS-20; Bagby et al., 1994), as tested here, alexithymia did not significantly influence their HCT performance scores. It should be noted that assessments of alexithymia usually cover a range of global abilities (e.g., “*I am able to describe my feelings easily*”; “*I am often puzzled by sensations in my body*”), but in the task that was used in the experiments presented in

Chapters 2 and 3 participants were asked to tune into one specific sensation, namely their heartbeats, which could explain the non-significant differences in the performance of the groups.

Moreover, some findings suggest that aspects of alexithymia (e.g., difficulty describing feelings) in AN may be mediated by comorbid depression and anxiety (Peres et al., 2020; Torres et al., 2015), further highlighting a previously found association between anticipatory meal anxiety and heightened sensitivity to cardiac interoceptive signals in AN (Khalsa, Hassanpour, et al., 2018). Again, the AN and AN-WR groups in the present studies had greater levels of trait anxiety and stress than HCs (as measured using the Depression, Anxiety, and Stress Scale, DASS-21; Lovibond & Lovibond, 1995), but as tested here these two traits did not significantly influence cardiac interoceptive accuracy. The results from Chapters 2 and 3, and also taking into account evidence from previous studies suggest two things: (i) while anxiety and other comorbid AN traits may be influencing interoception more than AN itself, (ii) interoceptive accuracy may be contextual – in other words, the ability to accurately perceive interoceptive signals and changes may be – to some extent – influenced by factors which cause physiological changes. For example, in the experiments presented in Chapters 2 and 3, participants were asked to perceive heartbeats during trials of specific duration but their accuracy in doing so was not compared to their sensitivity during a situation which may be stressful to AN patients, such as before eating a meal (like in Khalsa, Hassanpour, et al., 2018), or when heartbeat was experimentally manipulated. While global difficulties relating to alexithymia and emotion regulation are a characteristic of AN, here it is proposed that interoceptive accuracy deficits are context- and domain-specific.

It should be noted that the reliability and validity of the findings regarding cardiac interoceptive accuracy should be regarded with caution given the numerous criticisms regarding the methodology of the HCT (e.g., Ainley et al., 2014; Brener & Ring, 2016; Desmedt et al., 2018, 2020; Murphy, Geary, et al., 2018; Shah et al., 2016; for details on the methodological criticisms regarding the HCT see Section 2.4, Chapter 2 and Section 3.4, Chapter 3). Despite the addition of control measures in the experiment presented in Chapter 3, the HCT is not an intuitive task; in other words, individuals are usually not consciously aware of their heartbeat (Khalsa et al., 2009). Even in a sample of typically developing adults, approximately 40% are not consciously aware of their heartbeat (Khalsa et al., 2009). One way to assess the reliability of the present findings in cardiac interoceptive accuracy would be to experimentally manipulate

the participants' heartbeat (e.g., by giving bolus isoproterenol infusions which alters the intensity of heart rate palpitations; Khalsa et al., 2009, 2015) and examine whether a similar pattern of results would be observed in the heart rate manipulation condition as opposed to the control condition (i.e., no manipulation of heart rate). Another option would be to use a second cardiac interoceptive task as a control task to examine whether the results between the groups would be similar to those observed in the HCT. For example, this control task could be the heart rate detection task (Whitehead et al., 1977), during which a two-alternative forced-choice (2AFC) task participants have to judge whether an external signal (e.g., a tone or a light) is presented in a synchronous or asynchronous manner with their heartbeat. Nevertheless, despite the acknowledged HCT limitations, as previously pointed out, it is an easy task for clinical populations to comprehend and for individuals to easily form beliefs about such abilities, which were also the focus throughout this thesis.

Furthermore, this thesis aimed to examine how results observed in the cardiac modality would generalise in the gastrointestinal modality. To do this, the experiment presented in Chapter 3 assessed subjective stomach fullness before and after participants drank different quantities of water. Even though the task design and constraints did not allow to obtain an actual measure of gastrointestinal interoceptive accuracy (e.g., by measuring physiological and structural dimensions of each participant's stomach, or by assessing how accurate participants were at noticing vibrations of a capsule in their stomach in comparison to controls; Anderson et al., 2023; Mayeli et al., 2021, 2023), differences in the perception of stomach fullness between the groups were noted. Specifically, the AN-WR group showed an overestimation of stomach fullness in comparison to HCs when they had to rate their perceived stomach fullness after drinking certain 'fixed' amounts of water that were shown to them by the experimenter. Previous studies have shown that patients with AN tend to report more exaggerated level of fullness after consuming a meal in comparison to HCs, and they also need to drink less water than HCs (during a Water Load Task) to feel completely full (Brown et al., 2022; Schalla & Stengel, 2019). Moreover, while healthy individuals engage in food consumption to address the dysphoria associated with hunger (Goldstone et al., 2009), AN patients often refuse to eat as food consumption is a dysphoric, and anxiogenic experience (Steinglass et al., 2010). There is little research on the eating behaviours of women remitted from AN, but some evidence points to reduced insular activation – a circuitry which motivates behaviour including eating (Rolls, 2016) – to taste stimulation during the hungry state in comparison to HCs (Kaye et al., 2020). Additionally, evidence has pointed to faulty reward processing and overactive cognitive

circuits during hunger and food-related somatosensory processing (Santel et al., 2006; Wierenga et al., 2015). One explanation for the results stemming from this thesis is that the fluctuations and differences in how one experiences subjective states of stomach fullness may be a subclinical trait, whose severity might in turn influence eating-related behaviours in individuals who are at a greater risk for developing an eating disorder, offering support for hunger suppression and extended satiation resulting from a suppression of interoceptive signals (Barca & Pezzulo, 2020). Indeed, previous studies in healthy population samples have shown that lower sensitivity to satiation signals is linked to more negative body image (referring to negative thoughts, perceptions, and feelings one may have about their body, also extending to appearance dissatisfaction, body surveillance, and body preoccupation; Cash, 2004; Cash & Deagle, 1997) and related body image disturbances, whilst greater sensitivity to satiation signals is linked to a more positive body image (which is distinct from negative body image and includes multiple facets such as body appreciation and body acceptance/love; Todd et al., 2020, 2021; Tylka & Wood-Barcalow, 2015; van Dyck et al., 2016). This conclusion regarding a trait-like mechanism of gastrointestinal sensitivity and risk for developing an eating disorder remains speculative and requires further investigations. Nevertheless, the findings from the present experiment could be used to further examine – in longitudinal designs – if a general bias in interpreting different aspects of stomach fullness could increase the risk of engaging in disordered eating behaviours and subsequently developing an eating disorder.

Unfortunately, the methods to study gastrointestinal interoception in AN vary, with most studies assessing sensibility to interoceptive signals via self-report questionnaires, rather than the perception of gastrointestinal signals (for review see Martin et al., 2019). Similarly, here, the perception of stomach fullness was assessed by a combination of beliefs regarding fullness states, visual information from the bottle, and the experience following the consumption of water rather than by obtaining objective measures of gastrointestinal differences across the three groups (AN, AN-WR and HC). Future studies should thus also aim to further investigate gastrointestinal interoception deficits in AN while accounting for factors such as stomach size, the role of hormones and peptides which regulate food intake and the motivation to eat, and other physiological parameters which may be contributing to individual differences in the perception of gastrointestinal states.

Collectively, the results from Chapters 2 – 4 offer support for the domain-specificity of interoceptive deficits at the perceptual level, and also support previous evidence regarding

similar performance in HCT between clinical AN samples and HCs. Additionally, despite the lack of an objective measure of gastrointestinal interoceptive accuracy, the results show aberrant processing of gastrointestinal information at the perceptual level. However, given the smaller sample sizes of the two clinical groups in Chapter 4 (in comparison to the HCs), more research is warranted to establish the validity and reliability of these results and further explore why in this case the AN group did not show significant differences in stomach fullness perception in comparison to HCs, despite noted deficits in the perception of hunger, satiety, and fullness in AN patients.

6.3.2 Interoceptive Metacognition

6.3.2.1 Prediction Errors and Retrospective Belief Updating

As discussed in detail in Chapter 1 (see Section 1.5), the judgements and subsequent actions in the external and internal worlds rely on the process of forming and adjusting predictions based on the combination of prior beliefs and new sensory information (Knill & Pouget, 2004). Nevertheless, adjusting beliefs in light of new information does not suffice when needing to create a more accurate representation of a state. Instead, the goal is to minimise the discrepancy between an actual and a predicted state (i.e., to minimise prediction errors). This can be achieved by engaging in a reciprocal exchange of bottom-up (low level) and top-down (high level) signals between the brain's hierarchical levels so as to provide corrective feedback for future predictions (or, belief updating; Aitchison & Lengyel, 2017; Barrett & Simmons, 2015; Friston et al., 2009; Owens et al., 2018; Seth, 2013).

Interestingly, the results from Chapters 2, 3 and 4 showed that individuals in the AN and AN-WR groups had difficulty minimising their prediction errors in both the cardiac and gastrointestinal interoception modalities. Specifically, AN patients predicted to perform significantly worse in the heartbeat counting task than they actually did (Chapter 2), and while the initial predictions of the AN and AN-WR groups (or, Prior Prospective Self-Efficacy Beliefs) were not significantly more pessimistic than those of the HCs, they misestimated their task performance retrospectively (or, Posterior Retrospective Self-Efficacy Beliefs; Chapter 3). In other words, the two clinical groups reported they did worse than they actually did. However, due to the opposite directionality of these differences in comparison to the HCs, who reported they did better than they actually did, one cannot claim that the clinical groups misestimated more or less than the HCs (see Appendix B, Table B.4). Nevertheless, in combination with the results on the discrepancy between performance and self-reported interoceptive abilities (or,

Interoceptive Trait Prediction Error (ITPE) z-score; see Appendix B, Table B.5), the results suggest that pessimistic beliefs in AN may be less malleable to change, maintaining a large discrepancy between expected and experienced afferent interoceptive signals from the body. Similarly, a greater prediction error in comparison to HCs was found between predicted states of stomach fullness and experienced stomach fullness in the AN-WR group following a post-sipping task (Chapter 4). These findings can be discussed in the context of (dis)confirmatory evidence and its role in belief adaptation, the role of pessimistic and negatively biased beliefs, and the role of large prediction errors and pessimistic ‘local’ retrospective metacognitive beliefs in the more global aspects of metacognition and mental health.

The large prediction errors as well as the poor retrospective evaluation could be attributed to the notion that people may discard new information if it disconfirms their existing (biased) beliefs (Lefebvre et al., 2017; Sharot, 2011; Sharot & Garrett, 2016). In schizophrenia for instance, delusions (defined as fixed false beliefs) are maintained despite evidence disconfirming them, suggesting that people with delusions may have greater belief inflexibility in comparison to HCs (Colbert et al., 2010; So et al., 2010). Interestingly, the degree of belief inflexibility seen in people with schizophrenia who experience delusions is largely dependent on the confidence attributed to prior beliefs – in other words, priors are less malleable if the individual holds a high degree of confidence on them (Woodward et al., 2008). Additionally, another recent study on schizophrenia showed that while patients had the tendency to ignore new information and instead persevere on previous responses during a modified predictive inference task, when they chose to update their beliefs, they relied completely on the new information (Nassar et al., 2021). This belief updating strategy reduces the amount of new information that can be integrated over time, subsequently limiting the flexibility and precision of beliefs patients may have (Nassar et al., 2021).

Importantly, however, recent work has also shown that while the updating of beliefs in healthy individuals depends on the valence of the disconfirming information (also known as optimism bias, e.g., Sharot, 2011; Sharot et al., 2012; Sharot & Garrett, 2016), this optimism bias is absent in clinical populations, who instead hold more strongly onto their negative beliefs (e.g., Korn et al., 2014; Kube et al., 2022; Woodward et al., 2008; for review on beliefs in mental health see Kube & Rozenkrantz, 2021). This stronger reliance on negative beliefs, in turn, leads to a lack of belief updating or the updating of beliefs without minimising prediction errors and thus the updating of future beliefs will rely on those biased priors. For example, people with

depression were given brief descriptions of an emotionally negative and relevant situation (e.g., “The company you are working for needs to lay off many employees, and you are called in to see your boss”), and then had to choose which of the four interpretations best fit the presented scenario (Everaert et al., 2018). Next, participants received further information on the scenario and were suggested an alternative interpretation (e.g., “Your boss shares how upset he is about having to lay off his employees and states that he wants you to stay because of your collegiality and achievements”). Interestingly, people with depressive symptoms showed more difficulty in abandoning their initial interpretation, despite the fact that the interpretation of the scenario they were subsequently given was more positive. Contrastingly, when positive interpretations were provided from the onset upon receiving novel negative information, depressive symptoms were not predictive of difficulties in adjusting positive interpretations. Adding to this, two other studies have shown that while healthy individuals and people with depression do not show differences in updating positive performance expectations following unexpectedly negative feedback (Kube, Kirchner, et al., 2019), people with depression, but not healthy control individuals, cannot update negative performance expectations despite receiving disconfirming information (in this case positive performance feedback; Kube, Rief, et al., 2019). Collectively, these findings suggest that the noted abnormal updating of negative prior beliefs (instead of generalised belief updating difficulties) in depression is in line with the cognitive model of depression (which posits that broadly, people’s perception of themselves within the world is centred on schemas, such as negative self-schemas; Beck, 1979; Beck & Haigh, 2014).

The belief updating strategy seen in schizophrenia and the suggestion of a specific difficulty in updating negative beliefs in depression could also apply to eating disorders where (i) AN patients also present with varying degrees of delusionality which in turn affects their perception of their body (for a review on the delusional dimension in AN see Behar et al., 2018), and (ii) therapy models attempt (but frequently fail; Abbate-Daga et al., 2013; Cooper, 2005; Hay et al., 2009; Murphy et al., 2010; Shapiro et al., 2007) to challenge the negative cognitions and beliefs individuals with eating disorders have about their self (Corstorphine, 2006; Fairburn et al., 2003; Murphy et al., 2010; Waller et al., 2002). In fact, as seen in Chapter 2, patients were able to use explicit, albeit false feedback regarding their performance in the HCT, to update their retrospective evaluation of their performance and adjust it accordingly to the valence of the feedback (e.g., if they were given positive feedback, then they evaluated their performance as better). On the other hand, in Chapter 3, patients with AN had lower confidence in the trial-by-trial accuracy of their responses during the heartbeat counting task in comparison to HCs,

which could explain why although their performance was not worse than that of HCs, their retrospective evaluation (in the absence of any other type of external or social feedback) of that performance was poorer – possibly relying onto their prior beliefs more. Coupled with the fact AN patients did not show significant differences in set-shifting abilities in comparison to HCs (as measured using the Wisconsin Card Sorting Test, WCST; Grant & Berg, 1948) these results suggest that: (i) patients with AN are able to use (dis)confirming information about a belief (e.g., here external, social feedback about interoceptive task performance) and update an isolated estimate, yet (ii) this could in fact be explained by the low confidence they hold on to that estimate (in this case, low confidence regarding their cardiac interoceptive accuracy) but this low confidence might prevent them from updating beliefs prospectively. Indeed, this was confirmed in Chapter 3 (see more below). As seen from the analyses on global, prospective metacognitive beliefs about interoception and the low rate at which these prospective beliefs are updated, the focus should also be placed on the ‘weight’ individuals place on their prior beliefs. This focus would allow for a better understanding of how the pessimistic beliefs and maladaptive belief updating observed in experimental settings could explain the poor AN treatment outcomes.

Another plausible interpretation of the results from Chapter 3, in light of the results presented in Chapter 2 and the observations from the clinical audit (presented in Chapter 5), is that AN patients are not unable to update their beliefs altogether, across all contexts and domains. Nevertheless, it is possible that AN patients have a difficulty – in comparison to HCs – in updating beliefs about their ‘interoceptive self, and interoceptive self-efficacy beliefs’ when their metacognitive evaluation of an interoceptive experience and interoceptive self-efficacy beliefs relies solely on their embodied experience (or, ‘evidence’) and existing beliefs (or, ‘priors’) and not on feedback (e.g., external, social feedback like in Chapter 2, or individual heart rate biofeedback like in the intervention presented in Chapter 5). Specifically, in Chapters 2 and 3, AN participants were pessimistic about their HCT performance when metacognitively evaluating / updating their beliefs after completing the task, but in Chapter 2 all participants, including those in the AN group, were able to use the social feedback from the experimenter (albeit arbitrary) to update their performance evaluation, and in the direction of the feedback. Moreover, in Chapter 3, AN patients were able to use the computerised feedback when the card matches they made were right or wrong, showing flexibility when exposed to changes in reinforcement in a similar way to HCs (as measured using the WCST; Grant & Berg, 1948). Further, in Chapter 5, patients were given heart rate biofeedback before and after engaging in

a guided breathing exercise, following a stressful situation (e.g., meal planning) and it was observed that in the majority of cases, the self-efficacy beliefs regarding one's ability to downregulate one's heart rate following a future, stressful situation increased after seeing heart rate changes in real time. Taken together, these results and observations suggest the need to differentiate between and further investigate (i) general belief updating mechanism differences in the absence of and the presence of feedback, and (ii) metacognitive belief updating (in the absence of non-embodied, self-feedback) mechanisms in AN patient populations and HCs. It was found that it is 'only' the latter mechanism which is affected in AN, future interventions could look into therapeutically targeting one's interoceptive metacognition with feedback training – this however warrants further investigations

6.3.3 Updating Prospective Beliefs, Learning Rates, and the Role of Precision

Although it has been hypothesised that the central nervous system makes inferences in a hierarchical and bidirectional manner (Friston, 2008), it has also been hypothesised that in AN the precision one places on incoming interoceptive, afferent signals is less than the precision that has been placed on priors (Paulus et al., 2019) – previous findings on the role of context-specific, hyper-precise priors which lead to strong predictions about visceral sensations support this hypothesis (e.g., Berner et al., 2018; Khalsa, Hassanpour, et al., 2018). However, it remained unclear how the brain can update prior beliefs over iterative observations (belief updating), and how quickly (i.e., learning rate). This thesis' results fill part of this gap by showing the role of pessimistic retrospective beliefs and prediction error minimisation difficulties in learning rates and belief updating, and the role of aberrant precision-weighting mechanisms.

Firstly, the updating of explicit, prospective self-efficacy beliefs in AN was influenced more by “pessimistic”, global retrospective beliefs about such performance (i.e., how accurate one thought they were after the task ends) rather than perceptual performance itself (i.e., how accurate one actually was), in line with the finding that prediction errors in AN are not minimised following embodied feedback and are more influenced by prior beliefs. Being able to minimise prediction errors helps the organism maintain its homeostatic balance (e.g., by adjusting prior expectations to match a current physiological state; Paulus et al., 2019). At the same time, in adaptive individuals, the brain should also be able to regulate the body and ensure a homeostatic balance is maintained by anticipating its future states and attempting to regulate them before they arise (or, allostasis; Sterling, 2014), but failure to do so can lead to the

development of mental health symptoms (Stephan et al., 2016). However, the results showed that the poor belief updating could be attributed to the greater reliance on prior beliefs, and the lower precision-weighting of the embodied, interoceptive feedback during task performance. In fact, aberrant precision-weighting has previously been theorised to be implicated in the maintenance of maladaptive beliefs, but this thesis was the first to show the role of faulty precision-weighting beyond retrospective judgements and also in the updating of posterior prospective beliefs.

For example, a recent study showed precision-weighting difficulties in AN patients during a multisensory integration task (Di Lernia et al., 2023). Specifically, AN patients and a HC group had to estimate their body size (using a well-validated virtual reality full-body illusion task; Serino et al., 2016). Patients estimated their body size before and after experiencing a body illusion, wherein they received tactile feedback on their abdomen in a synchronous and asynchronous manner as the avatar. The results showed that patients tended to ignore the new, bottom-up information received from the tasks and relied more on their prior, top-down beliefs regarding body size, leading to more misestimations in comparison to HCs, and failure to minimise the prediction error regarding body size estimation. Furthermore, two computational studies in transdiagnostic patient samples (with symptoms of depression, anxiety, substance use disorder and/or eating disorders; also briefly described in Section 1.5, Chapter 1) fit a Bayesian computational model of perception to behaviour to assess if distortions in precision-weighting had a link with misestimation of cardiac interoceptive states (Lavalley et al., 2023; Smith et al., 2020). The clinical samples showed a failure to adapt their beliefs regarding the precision of experienced interoceptive signals, and importantly did not place greater precision during the “breath-hold” condition which actually increases the magnitude of cardiac signals, unlike the HCs. This shows that patients may not be susceptible to in-the-moment physiological, interoceptive changes, hence relying more on prior beliefs regarding such states. These two studies were the first to computationally model behaviour on an interoceptive perception task, however, the eating disorders sample of the study was small ($N=14$ in Smith et al., 2020; $N=36$ in Lavalley et al., 2023) and the authors did not specify what eating disorder diagnosis the participants had. The results from the present thesis extend these findings by showing the role of this aberrant precision-weighting mechanism in the updating of prospective beliefs, and supports previous theories which suggest that this inappropriate weighting of prior beliefs versus new evidence, especially regarding interoception, may lead to the manifestation of several psychiatric symptoms as the result of faulty inference (Barca & Pezzulo, 2020;

Barrett & Simmons, 2015; Owens et al., 2018; Paulus et al., 2019; Petzschner et al., 2017; although see Stephan et al., 2016 for a discussion of such faulty updating at metacognitive and prospective cognition levels).

Similarly, beliefs about stomach fullness states in AN-WR patients were not only greater than those of HCs, but despite showing large prediction errors on experienced and predicted fullness states the AN-WR group did not update their beliefs prospectively. As explained in Section 1.5 (Chapter 1), if prediction errors are not corrected, and predictions about one's state are based on a misrepresentation of the interoceptive self, then this may result in uncertainty regarding bodily needs and in turn lead to maladaptive behaviours used to reduce this uncertainty. The present findings offer support for a previously theorised framework of active interoceptive inference in AN, wherein individuals may not 'know' what they are feeling and suppress the interoceptive signals by engaging in maladaptive eating behaviours, such as eating restriction (Barca & Pezzulo, 2020). However, the present results further extend this theory by showing that patients with and weight-restored from AN, in comparison to HCs, have a difficulty in not only anticipating a future state (here, '*How full will you be?*'), they also show a difficulty in using visual cues (here, water bottle quantity) and incorporating an interoceptive experience they just had (after drinking the water in the first trial) to update their beliefs regarding how full they would be in a similar scenario. However, when considering the above interpretation regarding metacognitive belief updating about interoception in AN vs. HC groups, one could further explore whether beliefs about a future gastrointestinal state would be updated differently in the presence of additional, external feedback as opposed to one relying only on their prior and present embodied, self-experience(s) and visual information by observing the water quantity on the bottle.

Moreover, two additional results call for further consideration. The AN-WR group consistently showed an overestimation of subjective fullness state and predictions about such state, both when making a first prediction and after drinking the indicated water quantity and asked to make a second prediction now that they had also drunk that quantity. Secondly, the AN group showed significantly greater confidence in the accuracy of their first stomach fullness prediction, and while their confidence in the accuracy of the second stomach fullness prediction was not significantly different to that of the HCs, when looking at the slope and standard error of these two measures (in comparison to HCs) they are similar [First Fullness State Guess Confidence: $\beta(SE)_{HC} = 70.74(2.88)$, $\beta(SE)_{AN} = 9.31(4.61)$; Second Fullness State Guess

Confidence: $\beta(SE)_{HC} = 72.97(3.00)$; $\beta(SE)_{AN} = 7.14(4.40)$]. The higher confidence of AN patients in predicted stomach fullness states as opposed to the HCs could support the rigidity in patients' beliefs regarding how eating will make them feel.

Moreover, these results could also be discussed in the context of the aforementioned role of (dis)confirming information in belief updating and how fast beliefs can be updated in light of new information. Specifically, the overestimation of fullness states in the AN-WR group in comparison to HCs and the overall lower rate at which the two clinical groups updated their fullness state predictions in comparison to the HC group could influence the previously noted difficulty in deciding when to eat and how much to eat which is precipitated by an interoceptive prediction error regarding satiety/hunger states (Frank et al., 2019). The low updating rate of the AN and AN-WR groups in comparison to that of HCs could – in part – be due to a difficulty in not only using an experienced interoceptive state to update priors, but also to a difficulty in integrating visual, exteroceptive information which would influence a future state (i.e., looking at the indicated water quantity and predicting how full they would be post-sipping). This would result in a mismatch between how full they thought they would be when looking at the water quantity on the bottle and when referring to their stomach fullness state after the first time they had consumed the indicated water quantity. In a previous study which used the same methodology on updating rates as the study presented in Chapter 4, patients with AN updated their body size estimates at a lower rate than HCs when asked to provide a retrospective estimate (of how large they felt their body was) following social feedback from the experimenter regarding their body size (Saramandi et al., 2022). Patients thus ignored the feedback from the experimenter which was disconfirming their overestimation of their body size and relied more on their priors. Additionally, as seen in the experiment by Di Lernia and colleagues, AN patients showed a preference towards their confidence on perceived interoceptive sensations, even when the new precision-weighted bottom-up multisensory information mismatched their metacognitive interoceptive beliefs (Di Lernia et al., 2023).

It should be noted however, that while AN patients typically overestimate how full meals can make them, estimating how full water quantities (which do not have a caloric value) will make them is different. Despite this, the results still offer novel preliminary evidence regarding gastrointestinal interoceptive belief updating in AN, suggesting that future studies should further examine gastrointestinal sensitivity in eating disorders vs. control populations and which parameters influence this sensitivity and beliefs regarding stomach states. Importantly,

altered anticipation and processing of aversive interoceptive experiences have also been noted in a sample of women remitted from bulimia nervosa, an eating disorder characterised by dysregulated intake of food, and homeostatic imbalance due to interoceptive disturbances (Berner et al., 2019; Stein et al., 2003; Walsh et al., 2003). Future studies should further explore the role of aversive interoceptive experiences and faulty precision-weighting in predictions to understand how a difficulty in expecting homeostatic changes may promote a range of disordered eating behaviours. For example, in restrictive eating observed in the restrictive subtype of AN, uncertainty about interoceptive states and low precision attributed to experienced states could promote starvation, whereas uncertainty and low precision may promote binge eating and purging behaviours in eating disorders such as bulimia nervosa (Berner et al., 2019).

Although precision-weighting mechanisms could not be computationally studied in Chapter 4, computational modelling analyses presented in Chapter 3 explored what explained the differences in precision-weighting learning rates between the AN and HC groups. When using Performance confidence (which was found to be poorer in the AN vs. HC group) as a precision proxy for the evidence, AN patients showed lower precision-weighted Learning Rates than the HC group. These findings add evidence to the previously theorised role of precision-weighting difficulties in AN and extend preliminary findings in other conditions.

Specifically, the previously theorised consequences of faulty inferences about interoceptive states, and prolonged dyshomeostatic states had been discussed in the context of substance use, depression and anxiety disorders. For instance, a model for drug craving had posited that as a result of alterations in dopaminergic modulation, frequent drug use results in over-precise expectations about physiological states (Gu & Filbey, 2017). This overprecision on expectations is coupled with underprecision in the interoceptive signals. Therefore, craving even after extended periods of abstinence can be seen as the brain's failure in updating beliefs about the substance-dependent physiological states of the body (Gu & Filbey, 2017). In another recently proposed framework for the role of faulty updating at metacognitive and prospective cognition levels, it was suggested that low self-efficacy beliefs regarding the brain's metacognitive abilities to regulate bodily states (also known as allostatic self-efficacy), may lead to a chronic dyshomeostatic bodily state and the development of fatigue and depression (Stephan et al., 2016). Taken together, these two examples suggest that a disruption in the brain's ability to accurately correct prediction errors about interoceptive states and

metacognitive evaluations of such beliefs could influence the subsequent decisions the individual makes.

Nevertheless, despite the different efforts to characterise the role of faulty interoceptive inference processing and predictions about bodily states in AN, no previous study had explicitly sampled patients' beliefs regarding interoceptive abilities before and after a cardiac perception task. Here, it was shown that patients 'ignore' the interoceptive information they receive during the task – potentially due to the high uncertainty regarding the embodied feedback they received during the HCT – and instead adhere to their pessimistic prior predictions of related abilities. In other words, patients expected to perform poorly, and despite performing better than they expected, when asked to predict their future performance in a similar HCT but under different trial durations they did not attempt to minimise the discrepancy between their prior prospective beliefs and their performance, and instead relied on their prior beliefs more. These results further suggest that therapeutic interventions should also look into targeting the 'precision-based' mechanism of belief updating as the weight AN patients place on their beliefs is likely to influence their everyday decisions, but also influencing their treatment outcomes and prognosis. For example, a patient who has low self-efficacy regarding their abilities to recover, may be less inclined to see and accept their progress and update their beliefs regarding their prognosis and thus not adhere to treatment and be more likely to relapse in the future. Importantly, however, this should be studied together with the effect of feedback on belief updating, as it is necessary to distinguish between the potential differences in general belief updating, and metacognitive belief updating about interoception.

Another finding which brings additional support for the need to target the precision-weighting mechanisms in belief updating was the association between the AN and AN-WR groups' Posterior Prospective Self-Efficacy Beliefs and their counterfactual beliefs about everyday life, such as fear of gaining weight, and less hope for improvement of clinical symptoms and behaviours. Clinical studies have previously shown that prospective self-efficacy beliefs at treatment onset (such as self-efficacy beliefs regarding one's abilities to change eating and weight control behaviours) can be predictive of eating disorder treatment dropout (Keshen et al., 2017). Moreover, patients with AN show higher levels of negative beliefs about worry (i.e., beliefs that thoughts are uncontrollable and dangerous; Cooper et al., 2007; Davenport et al., 2015; McDermott & Rushford, 2011; Olstad et al., 2015; Palomba et al., 2017) and higher positive beliefs about worry (e.g., beliefs that perseverative thinking is useful; Davenport et al.,

2015; McDermott & Rushford, 2011; for review on metacognition and eating disorders see (Palmieri et al., 2021; Sun et al., 2017) in comparison to the general population. Existing psychotherapeutic interventions such as Cognitive Behaviour Therapy- Enhanced (CBT-E; Atwood & Friedman, 2020) and the Maudsley Anorexia Nervosa Treatment for Adults (MANTRA; Schmidt et al., 2014) have been targeting the cognitions of patients, such as by identifying thinking and alternative thinking styles, they do not address the metacognitive beliefs that may arise, as these approaches do not directly alter the maladaptive prior beliefs but some of the factors which may influence their updating. At the same time, several other interventions have focussed on improving one's reflection upon and practices which involve interoceptive awareness, also known as interoceptive-based interventions (Farb et al., 2015). For example, interoceptive exposure uses strategies to help the individual attend to their feared sensations and challenge their catastrophic conditions in order to reduce the emotional distress resulting from sensations of physiological arousal (Stewart & Watt, 2008). Other interventions, such as mindfulness techniques focus on minimising the prediction errors about bodily states by encouraging the individual to shift their attention from a predicted to an observed body state (Farb et al., 2015). These interventions have been used to target behaviours across a wide range of diagnoses, but despite preliminary evidence for the benefits of such interoceptive-based interventions in autism spectrum disorders, and irritable bowel syndrome, the efficacy of interoceptive-based interventions in eating disorders remains inconclusive (for review on interoceptive-based interventions for mental health disorders see Heim et al., 2023). Finally, although metacognitive therapies have been used for psychiatric conditions where metacognitions are central (e.g., generalised anxiety disorder, Konstantellou et al., 2011; OCD, Halmi et al., 2005) in order to indirectly reduce worry, rumination, and maladaptive coping behaviours as patients improve (Solem et al., 2009), there exists no clinical trial on the effect of metacognitive therapy for eating disorders (Olstad et al., 2015).

Given the results from the present thesis, it is suggested that interventions targeting beliefs in AN should also target the metacognitive interoceptive beliefs and aberrant precision-weighting mechanisms that have been found in this population by targeting top-down pessimistic and strong prior beliefs regarding sensory signals. In fact, patients have often quoted not being able to recognise bodily sensations and thus struggling to self-regulate (e.g., *"I often don't know when I am feeling stressed...it just hits me"*; see Chapter 5), and as also seen from the discussions with clinicians and individuals with lived experience, interventions aimed at improving body awareness and self-regulatory abilities do not seem to be enough in AN as

patients report not trusting the process and not feeling the changes (see Chapter 5). The introduction of a heart rate biofeedback (HRBF) enhanced intervention as part of a local NHS eating disorder's unit standard practice showed improved outcomes in comparison to previous practices (Chapter 5). Patients' self-efficacy beliefs were improved after completing the HRBF as did most of patients' scores of interoceptive accuracy. Using the insights from Chapter 3 as well as other studies on the role of precision in predicting and updating bodily states (Di Lernia et al., 2023; Lavalley et al., 2023; Smith et al., 2020), it is possible that maximising the precision of the evidence (in this case, patients knew that the HRBF was a direct reflection of their actual heart rate changes) could help in altering the maladaptive precision-weighting on pessimistic interoceptive self-efficacy beliefs, potentially even when the new information is disconfirming patients priors (e.g., *"I felt my heart rate go up but your app did not show this."*, referring to the biofeedback interface). The empirical insights from this thesis, combined with the observations regarding the efficacy, feasibility and acceptability of HRBF for AN led to the design of a randomised clinical trial (RCT). The RCT examined the efficacy and mechanisms of action of HRBF during guided respiration exercises aimed at helping individuals with subclinical symptoms of disordered eating and somatic disorder down-regulate their own heart rate under different conditions of stress and to enhance related metacognitive self-efficacy beliefs. This HRBF intervention was compared against visual guided imagery exercises (RCT number: ISRCTN16762367). Although the outcomes from the audit are not generalisable as it was conducted at a local eating disorders unit, more local wards and NHS Trusts should assess the efficacy and evaluate the process of HRBF in comparison to other local standards of practice, in addition to research studies conducted in parallel to establish how to best target poor belief updating and precision-weighting about interoception and related abilities in AN.

Despite these results regarding precision-weighting differences between the AN and HC groups, it should be noted that precision-weighting mechanisms could be affected by the pharmacological medications that AN patients receive. Research has shown that neuromodulators (e.g., acetylcholine (ACh), serotonin, dopamine, N-methyl-d-aspartate) are involved in synaptic gain control mechanisms and could thus influence predictive and active inference processes; for example dopamine has been related to the encoding of prediction errors in reward learning (Aukstulewicz & Friston, 2016; Daw & Doya, 2006; Haarsma et al., 2021; Iglesias et al., 2013; Schultz et al., 1997). Moreover, a recent study in individuals with psychosis (where the dopamine system is overactive) and controls showed that dopamine modulates precision-weighting of prediction errors that signal surprise (i.e., a prediction error

which signals the degree of surprise without indicating the valence; Haarsma et al., 2021). This finding suggests that the predisposition to delusional beliefs could be explained by a failure to account for the precision of new information, therefore resulting in aberrant inferences (Haarsma et al., 2021). As seen in this thesis, patients with AN had pessimistic priors and posteriors, and failed to update them despite disconfirming embodied feedback, placing greater precision on top-down beliefs than bottom-up sensory signals during the tasks. Nevertheless, patients with AN, including those in the present sample are often heavily medicated and even though information on the sample's pharmacological treatment was collected in all studies, the effects of pharmacological medications on the beliefs of the patients in the current thesis sample cannot be characterised. Typically, by including a recovered sample, some of these effects can be accounted for, however several weight-restored patients in the present samples were also medicated for their comorbidities which endured beyond the acute AN state. Although the present thesis did not look into the causal relationship between AN and comorbidities, the individual differences of the participants and clinical complexities in clinical and subclinical samples should be considered further in future studies.

To summarise, the findings from the three studies on belief updating about cardiac and gastrointestinal interoceptive states showed that patients with AN and AN-WR participants did not update their beliefs regarding cardiac abilities and gastrointestinal states in the same way as controls. Evidence from the cardiac modality points to the role of precision-weighting issues influencing the amount of 'new' information participants will incorporate with their prior beliefs to update them retrospectively and prospectively. Moreover, the difficulties in updating stomach fullness states across both clinical groups could be linked to the general rigidity patients have regarding what to eat, when, and how much. Overall, the results point towards domain-general deficits in interoceptive metacognition in AN.

6.3.4 The role of clinical characteristics in the formation and updating of metacognitive interoceptive beliefs

Going beyond the main questions this thesis attempted to answer regarding the perceptual and metacognitive deficits in interoception and belief updating in AN, an important consideration was the transdiagnostic complexity of AN. To do this, this thesis also explored the role of comorbid clinical and cognitive traits in interoception as well as the role of other characteristics, such as state and trait mechanisms.

One notable finding stemming from this thesis' experimental studies is the role of depression and stress in self-efficacy, learning rates and confidence. Depression, anxiety, and stress are highly comorbid in eating disorders, including AN (Andrés-Pepiñá et al., 2020; Eskild-Jensen et al., 2020; Himmerich et al., 2019), and the results from the experiment presented in Chapter 3 offer evidence for a potential role of depression and stress in the updating of pessimistic self-efficacy beliefs regarding interoceptive abilities. These results are especially enlightening regarding the mechanisms by which depression and stress can affect one's way of updating beliefs; from a transdiagnostic and computational psychiatry perspective, the present findings not only show that depression and stress at severe levels can influence one's confidence levels and in turn result in lower precision-weighted learning rates (as seen in the AN vs HC groups), they also suggest that mood may relate to a dysfunctional precision-weighting mechanism that extends beyond local metacognition to global, prospective beliefs about one's performance.

Furthermore, the results from Chapter 3 regarding the role of depression traits on cardiac interoception self-efficacy beliefs and confidence point towards the possibility that pessimistic self-efficacy beliefs may be a more enduring trait of individuals with AN (given the more pessimistic prospective beliefs of the AN-WR group than the HC group's at a trend-level which were associated with depressive symptoms). Unfortunately, the design here was cross-sectional and thus the transdiagnostic complexity of symptoms and behaviours is not always clear. This is also the case in most other studies, and hence results regarding the state and trait effects of mental health symptoms in confidence remain mixed (e.g., for evidence of a trait effect see Eisenacher et al., 2015; Gawęda et al., 2018; but for evidence for state dependence see Moeller et al., 2016). However, by including individuals both during the acute and weight-restoration phase of AN and comparing how their behaviours and beliefs differed to those of a matched HC population allowed to examine if they were secondary consequences to the acute AN state or premorbid characteristics which endured post weight-restoration. For example, if differences between the AN but not the AN-WR group in comparison to HCs are significant, they would be suggestive of a state effect. Whereas a difference between both clinical groups in comparison to HCs would be suggestive of a trait effect. This latter inference requires careful interpretation given that the cross-sectional nature of the studies did not allow to differentiate between a 'true' trait marker which was present before the initial episode of acute AN or a possible 'scar' effect of AN resulting from potential changes in the brain network architecture enduring beyond the acute AN phase. Hence, future longitudinal studies are needed to further delineate this.

Additionally, as also mentioned in Section 1.2 (Chapter 1), there is still a lack of a consensus regarding the specific criteria for weight-restoration and recovery in AN and subsequent clinical terms of weight-restoration, remission and recovery are often used interchangeably even when referring to distinct phases of one's recovery process (Khalsa et al., 2017; also see Chapter 1, Section 1.2). Here, AN-WR patients were recruited if they met behavioural and physical signs of recovery for at least 6 months which could mean that while some individuals may have recovered for years, others had just entered the weight-restoration phase, which could contribute to large within-group variance. It is thus recommended that future research further explore changes in interoception, metacognition and belief updating throughout the different stages of AN (from the acute to the complete recovery phase) in parallel to ongoing efforts to establish clear definitions for each stage that will be used among those conducting relevant research in order to clarify which behaviours or characteristics are premorbid traits, states, or scar effects.

The results on cardiac belief updating and the difficulties in metacognitive inferences about stomach fullness states which were observed across both the AN and AN-WR samples in comparison to HCs are likely to be indicative of a trait-related characteristic. These results add to recent findings suggesting that poor levels of interoceptive sensibility (as measured using the self-report Multidimensional Assessment of Interoceptive Awareness (MAIA); Mehling et al., 2012) may represent a trait factor which increases the risk of developing AN, given that poor interoceptive awareness in comparison to HCs was found in a sample of individuals with a current diagnosis of AN, weight-restored from AN, but also in biological sisters of individuals with AN (Phillipou et al., 2022). Another study looking at the effects of Cognitive-Behavioural Therapy on interoceptive accuracy and sensibility in a group of acute AN patients (restricting type) and matched HCs also found that interoceptive deficits are present even as patients progress into the recovery stage (Fischer et al., 2016). Taken together, these results in combination with the results from the present thesis highlight that deficits across processing levels of interoception may be possible underlying mechanisms in the symptomatology of AN. Crucially, this trait role of interoception should be considered further, especially in light of criticisms regarding existing diagnostic criteria which largely focus on the disordered eating behaviours and not distortions in the experience and evaluation of interoception (see Section 1.2, Chapter 1; Broomfield et al., 2017; Stinson, 2019).

6.4 Methodological issues

The present thesis employed an interdisciplinary approach, combining theories and methods from numerous domains, including clinical and experimental psychology, (computational) psychiatry, and cognitive and computational neuroscience to address the literature gaps that were identified in the Introduction (Chapter 1) and investigated further in each chapter separately (Chapters 2-5). As such, adopting an interdisciplinary approach gives rise to several methodological challenges regarding the tasks used to study interoception and the analysis options. It is also noteworthy that several limitations can be attributed to sampling issues. All these challenges and limitations are discussed in detail in the subsections that follow.

6.4.1 Design Issues

Given the experimental constraints of existing tasks when it comes to assessing interoception (as discussed above and in Chapters 2 – 4), the focus in this thesis was placed on individuals' beliefs regarding their interoceptive abilities (cardiac modality) and belief updating regarding interoceptive states (gastrointestinal modality). This was done in an attempt to bridge some of the existing gaps between experimental and clinical research traditions (see Chapter 1, Section 1.4) and help our understanding of how (clinical) metacognition deficits could influence belief updating about inner bodily states and their impact on symptom development and maintenance. However, the studies used in this thesis were not studies that have been traditionally used to experimentally assess local and global metacognition (e.g., see Haddara & Rahnev, 2022). Consequently, aspects of metacognition such as metacognitive bias and sensitivity were not analysed using the traditional task-based measures and Type 2 Signal Detection Theory (SDT; Fleming & Lau, 2014; Galvin et al., 2003; Maniscalco & Lau, 2012) and instead the focus was placed on the updating of metacognitive (local and global retrospective and prospective beliefs) to characterise the associations between higher-order mechanisms of interoception and dimensions or levels of interoception. Future studies should continue using an interdisciplinary approach and applying Bayesian computational models across modalities of interoception to predict more specifically the relationship between computational parameters and physiological measures – an approach which could aid the understanding of individual differences and the role of neuropsychiatric symptoms in interoception (Garfinkel et al., 2022).

Importantly, however, the studies presented in this thesis only focused on self-efficacy beliefs regarding interoceptive abilities, which hinders the ability to generalise the findings in non-interoceptive modalities. Therefore, although the findings from this thesis can characterise the

pessimism regarding interoceptive self-efficacy beliefs in an AN population in comparison to controls across the cardiac and gastrointestinal interoceptive systems, one cannot establish whether pessimistic beliefs and lower updating (in comparison to controls) would also be observed in non-interoceptive domains. This could be achieved by employing an adapted non-interoceptive task (to also add questions about self-efficacy beliefs, such as “How well will you do/did you do?”), to assess metacognition and belief updating, such as the Metadots task (e.g., Fleming et al., 2014; Lund et al., 2023; see Section 1.4, Chapter 1 for details on metacognitive tasks in the perceptual domain). Further, although the present experiments assessed self-efficacy beliefs which were directly linked to the tasks (e.g., “How well will you do in the task?”, e.g., Chapters 2 and 3), general self-efficacy was not assessed via standardised questionnaires, such as the General Self-Efficacy Scale (GSE; Schwarzer & Jerusalem, 1995). By using such standardised questionnaires, one could assess if, for example, a generally poor sense of self-efficacy regarding one’s ability to cope with daily hassles and adapting after stressful life events, could be predictive of poor interoceptive self-efficacy beliefs, and/or correlated with everyday counterfactual beliefs (such as those assessed in Chapter 3, regarding fears and hopes towards improvement and recovery). At the time of the experimental conceptualisation and design the focus was placed on interoception and patterns across the cardiac and gastrointestinal modalities. In Chapter 3, the Wisconsin Card Sorting Task (WCST; Grant & Berger, 1948) was used to account for general difficulties in set shifting and cognitive rigidity, but even this measure is not the most appropriate to assess general self-efficacy. Therefore, future studies should consider employing a non-interoceptive, control task alongside a standardised measure, such as the GSE to examine if poor interoceptive self-efficacy and belief updating, forms part of a generalised difficulty in updating metacognitive self-efficacy beliefs. In a further step, a non-interoceptive task which measures self-efficacy and belief updating regarding AN behaviours in the exteroceptive domain (e.g., body size estimation) could be used to further disentangle whether pessimism regarding self-efficacy is only linked with and observed in AN-relevant symptoms and behaviours, or generalised across all levels of self-efficacy, tapping also onto non-directly relevant AN contexts and beliefs (e.g., beliefs regarding performance on a visual perception task).

6.4.2 Other Challenges and Limitations

One of the challenges of the present thesis was recruiting enough eligible participants for the studies, especially for the two clinical groups (AN and AN-WR groups). During the COVID-19 pandemic, the collaborating clinics in Italy and Greece noted that most new patients were

being diagnosed with the binge-purge subtype of AN instead of the restricting subtype and the age of illness onset was also lowered to less than 18 years old. Subsequently, most new acute AN patients receiving treatment were not eligible to participate and it was also not possible to recruit individuals with either subtype of AN. Specifically, patients with the restricting AN subtype display – to an extent – different behaviours and comorbidities than those with the binge-purge subtype (e.g., differences in decision-making performance; Cavedini et al., 2004; Danner et al., 2016; Garrido & Subirá, 2013; greater impulsivity and emotional dysregulation in the binge-purge subtype and overall greater psychopathology than in the restricting subtype; Carter et al., 2012; Lewis et al., 2019; Peat et al., 2009; Reas & Rø, 2018; Waxman, 2009). In another study, which looked at gastric emptying in patients with restricting and binge-purging subtypes of AN before and after four and 22 weeks of rehabilitation, gastric emptying had only improved in the restricting group (in comparison to controls) and not in the binge-purging group after the rehabilitation period (at both timepoints of measurement; Benini et al., 2004). These differences in behavioural, psychological, and physiological processes between the two AN subtypes are likely to influence interoception and related interoceptive abilities in different ways. For instance, perception of interoceptive signals and interoceptive beliefs (e.g., perception of fullness or beliefs about performance in a heartbeat counting task) could be influenced by the greater emotional dysregulation or delayed gastric emptying in the binge-purge AN group (in comparison to the restricting AN group). Therefore, there would need to be two separate acute AN groups to account for the neuropsychological and psychopathological differences relating to or influencing interoception. Unfortunately, it would not have been possible to recruit enough participants for both a binge-purge and restricting AN group in a timely manner for this thesis' completion and thus the focus was placed on those with the restricting subtype. This decision was mainly taken since most previous research on interoception has recruited individuals with the restricting subtype of AN.

In the same vein, it was challenging to recruit enough weight-restored individuals for the AN-WR group. Several individuals who were eligible to participate as part of the AN-WR group relapsed during the COVID-19 pandemic and did not maintain weight-restoration for a minimum of 6 months or show any other behavioural or physical signs of weight-restoration, thus were not able to participate. To tackle this issue, the testing periods for the experiments in Chapters 3 and 4 were extended, and eventually it was decided to stop recruitment when enough participants were recruited per group (and especially the clinical group) to analyse the data in a timely manner.

Unfortunately, although the samples in three out of the four main Chapters in this thesis (Chapters 2 – 4) had a multi-site approach which contributes to current efforts in neuroscience, psychology and psychiatry research practices for more representative and diverse samples (Naddaf, 2023), the overall diversity of the sample was rather limited as recruiting individuals below the age of 18 as well as individuals who were assigned male at birth would add an additional layer of complexity and variables to control for when interpreting the results which was not possible in the present studies. More specifically, beyond the ethical considerations and parental consent which would be required to recruit adolescent participants, one of the most important factors to consider when assessing interoception in adolescents with AN is the role of hormones and general changes that occur in puberty. Puberty is a crucial period of risk for the onset of eating pathology in adolescent girls, with changes in reproductive hormones often acting as one specific (psycho)pathogenic mechanism for symptom development (Harden et al., 2014). For instance, it has been established that women with AN (including adolescent girls) have lower concentrations of oestrogens and progesterone than matched healthy controls (Himmerich et al., 2010; Misra et al., 2004). Oestrogens are known to inhibit food intake, decrease meal size and advance the onset of satiety, while progesterone is an oestrogen antagonist and will increase food intake in the presence of oestrogens (Baker et al., 2012). When recruiting adolescent girls with AN one should not only account for the elevated risk for disordered eating (which remains unclear to a certain extent), but also adopt longitudinal study designs with multivariate measurement of pubertal development, also including direct measures of changes in these reproductive hormones, in order to discriminate the multiple, interconnected changes that occur during puberty due to changes in hormonal concentrations. This approach would help to delineate how these hormones may influence the course of the disorder and symptom manifestation and severity but also how such hormones may influence one's ability to perceive the various interoceptive signals. As a result, the cross-sectional design of the present thesis' experiments did not allow to determine if and how deficits observed in individuals with AN are particularly accentuated among women who experience hormonal changes during the different stages of puberty.

Similarly, sex differences in diagnostic criteria, hormones and interoception also meant that recruiting both men and women would require considerations of additional variables which would explain, to some extent, the predicted sex differences in interoception and belief updating. Initially, AN was thought to be a disorder primarily affecting women – even in the

literature, AN in males was not a topic of discussion until a century after AN was clinically described (see Bruch, 1971). A female-centric framework of assessment is likely to have hindered diagnosis and treatment of AN in males (Gorrell & Murray, 2019; Strother et al., 2012) since specific issues of eating disorders in males are not sufficiently covered in existing diagnostic manuals. For instance, one of the DSM-IV diagnostic criteria for AN was amenorrhea, but since a direct endocrine equivalent does not exist in male patients, it is possible that until the DSM-V changes were implemented (where amenorrhea is no longer a diagnostic criterion for AN), there was an underdiagnosis of AN in males (Gorrell & Murray, 2019). Another example is the drive for thinness in women with AN, which in men may also manifest as a drive for a muscularity-oriented body image (Murnen & Karazsia, 2017; Parent, 2013). Unfortunately, most studies on AN recruit women rather than men or both women and men with AN given that collaborating hospitals and clinics usually treat women with AN. For instance, out of the 1,009 patients who were hospitalised in St. Anne Hospital in Paris, France between 1988 and 2004 for AN treatment, only N=23 patients were anorectic males (Gueguen et al., 2012). This was also the case here – the collaborating hospitals only treated one adult male throughout the duration of data collection and testing period. Sex hormones (e.g., testosterone, estradiol, progesterone) seem to be leading contributors to differential risk for eating pathology in males and females, and there is also variation in the genetic contributions to eating pathology between the two sexes (e.g., adrenarche in males, and gonadarche in females; for review on the biological basis for sex differences in eating disorders see Culbert et al., 2021). Although men and women with AN may present with certain similar clinical features and no significant differences in the eating disorder inventory scores (Braun et al., 1999; Carlat & Camargo, 1997; Crisp et al., 2006; Gueguen et al., 2012; Sharp et al., 1994; Strober et al., 2006), the genetic and biological differences are theorised to lead to interoception differences between the two sexes. Although a systematic investigation of the effects of physical and hormonal changes on interoception is currently lacking, physical and hormonal change has been linked with fluctuations in physical systems from which interoceptive signals arise (e.g., increased resting heart rate in pregnancy, palpitations during pregnancy, menopause, and the luteal phase of the menstrual cycle; e.g., Hill & Pickinpaugh, 2008; Rosano et al., 1997). Subsequently, it is likely that such changes may result in numerous discrepancies between one's actual and expected bodily states, which in turn could contribute to the overall heightened attention to internal signals and increased somatic complaints reported by women (Murphy et al., 2019). More specifically, women – in comparison to men – show poorer objective interoceptive accuracy (e.g., cardiac, respiratory, gastric and sexual arousal;

Grabauskaitė et al., 2017; Harver et al., 1993; Suschinsky & Lalumière, 2012; Whitehead & Drescher, 1980), even after accounting for physiological differences, such as blood pressure, BMI, heart rate variability, and resting heart rate (Murphy, Brewer, et al., 2018). Additionally, it has been theorised that internal signals may be more difficult to perceive as a direct consequence of the frequent hormonal fluctuations that women experience (in comparison to men) which may lead them to rely more on external cues for gauging internal states (Murphy et al., 2019; Pennebaker & Roberts, 1992). Moreover, if interoceptive predictions are formed based on one's prior experiences (e.g., (Barrett & Simmons, 2015), frequent differences between actual and expected bodily states may also produce imprecise predictions and greater prediction errors regarding the expected interoceptive state of one's body. Hence, a study on the formation and updating of interoceptive beliefs in both men and women should account for the biological (including hormonal and physiological) factors which inherently lead to discrepancies in how the different sexes use priors and embodied feedback to gauge internal, interoceptive states.

Finally, most patients from the acute AN group were receiving inpatient treatment when the testing took place and due to clinical time constraints in certain instances the testing had to either take place in two separate sessions, or certain measures could not be completed, leading to missing data. The sessions consisted of several tasks beyond the completion of the main experimental task (e.g., HCT; Schandry, 1981, or the Water-Sipping Task), such as self-reported questionnaires and in some instances (e.g., Chapter 3), assessments of cognitive flexibility and set-shifting abilities. Due to the length of the session, some patients did not have time to complete all questionnaires at the end and hence in some instances there are missing data on self-reported questionnaires. This was also the case in some of the data from the AN-WR and HC groups who also did not complete some self-report questionnaires due to time constraints. Although there are no missing data on the main experimental measures, some data on comorbid psychological and cognitive traits and clinical measures were collected in only a subset of participants, and thus the overall power and effect size of the respective analyses is lower than what it could have been if there were no missing data. As a result, it is especially important for future studies to further examine the role of psychological (e.g., depression, anxiety) and cognitive (e.g., set-shifting, flexibility) traits in interoception and belief updating.

6.5 Clinical Impact

The studies and the clinical audit that were carried out and presented in this current thesis were experimental and thus have no direct clinical impact. However, the studies employed a mixed sample – there were two clinical groups of interest (an acute AN and a weight-restored AN group) and a gender- and age-matched HC group in each experiment. The clinical audit was run at a local NHS eating disorders unit providing inpatient treatment to individuals with AN. Hence, while the conclusions drawn from this thesis will need to be further assessed in feasibility and clinical trials before establishing therapeutic targets, it offers interdisciplinary insights directly relevant to AN.

To begin with, the studies described in Chapters 2 and 3, explored, for the first time, both local and global retrospective and prospective self-efficacy beliefs about cardiac interoception in AN, attempting to bridge some of the gaps between experimental and clinical research traditions on metacognition. More specifically, the key findings of these two Chapters highlight the pessimistic prior and posterior prospective self-efficacy beliefs in AN (in comparison to those of HCs) which fail to be updated despite a heartbeat counting task performance comparable to that of HCs. As discussed in Chapter 1, however, these metacognitive, self-efficacy beliefs in clinical tradition had mainly been assessed in relations to one's beliefs regarding their ability to regulate their eating, other bodily domains and cognition, but not regarding their interoceptive abilities. At the same time, experimental traditions had assessed local and global beliefs regarding non-interoceptive task performance and thus the role of such beliefs in the formation and updating of beliefs about one's interoceptive abilities was unclear. Here, across two studies, cardiac interoception self-efficacy beliefs were found to be poorer in AN in comparison to HCs, suggesting that a pessimism regarding one's understanding of their 'inner body' could have implications in the treatment and their ability to update counterfactual beliefs about symptoms and recovery.

Next, the study presented in Chapter 4, exploring the perception, prediction and updating of subjective fullness states in the gastrointestinal modality between AN, AN-WR and HC groups sheds light on the differences between clinical and healthy population regarding gastrointestinal states. More specifically, this study shows, for the first time, the perception and metacognitive beliefs about (subjective) gastrointestinal states are affected in AN and prior experiences regarding such states may be more influential in the formation and updating of

prospective state beliefs and inferences about fullness. Thus, these beliefs should be targeted therapeutically in order to help individuals in and post-recovery to attend to bodily states.

Across all experimental studies in this thesis, cognitive and psychological traits were taken into consideration. Interestingly, cognitive difficulties, such as in set-shifting and flexibility, were not found to be affecting cardiac interoception belief updating. The lack of an effect of cognitive difficulties should however be further explored as this was tested in only a subset of the sample in the cardiac modality and thus sample size and power-related issues may have hindered the results. However, and importantly, the role of depression and stress (comorbid AN traits) was noted across both the cardiac and gastrointestinal modalities. Specifically, greater subjective uncertainty and poorer belief updating were explained by more severe comorbid depression and stress.

Another noteworthy finding is one bridging some of the gaps between the clinical and experimental research observations in metacognition and self-efficacy. More specifically, in Chapter 3, an association between everyday counterfactual beliefs (such as fear of losing control over food and gaining weight) and prospective self-efficacy beliefs was found. This finding is of high clinical relevance as a compromised ability to update interoceptive self-efficacy beliefs could extend to a difficulty to update beliefs about recovery and a post-acute phase of everyday life and vice versa, although these results warrant further investigations as they were moderately significant ($p < 0.052$ and $p < 0.053$).

Finally, the audit showed the feasibility and effectiveness of incorporating heart rate biofeedback-based interventions to help patients with AN to improve their interoceptive self-efficacy beliefs and self-regulatory abilities to manage physiological responses (e.g., increased heart rate in response to a stressful stimulus or context) better than other standard processes that were used at a local NHS eating disorders unit. These observations are in line with the insights that stem from the main experimental Chapters of this thesis regarding the poor belief updating and self-efficacy beliefs in AN about interoception. Specifically, the findings from Chapters 2, 3, and 4 point towards the need to therapeutically target the pessimistic interoceptive self-efficacy beliefs and precision-weighting mechanisms underlying belief updating. Future studies are needed to examine how these could be formed into a validated and established intervention.

Overall, the results from this thesis suggest that precision-weighting difficulties in belief updating and aberrant beliefs regarding interoceptive states could be influencing the way in which patients make decisions about their bodies and in turn engage in maladaptive behaviours. Anecdotally, patients have reported not trusting the signals they receive from their body, or not knowing how to self-regulate or not feeling confidence about such abilities (see Chapter 5). Further efforts should be made into optimising the precision of new evidence as opposed to prior beliefs to help improve the global beliefs patients have about interoceptive abilities. Despite the design and sampling issues which are discussed above, this thesis offers novel insights into the interaction between the different dimensions of interoception and aspects of metacognition in AN.

6.6 Directions for future research

Based on the findings from this thesis and the methodological considerations which were discussed in Section 6.4, several directions for future research arise, but here just a few, initial outlines and suggestions will be provided.

Firstly, cross-modal studies should further assess the domain-specificity/generalizability of the interoceptive difficulties, also by examining interoception in other modalities, such as the respiratory axis, which could be affected when individuals with AN are stressed. This would be particularly important given the findings of a stress effect in belief updating and prospective self-efficacy beliefs in the cardiac modality.

Secondly, studies should continue focussing on the neural bases of metacognition and interoception and better map how the observed precision-weighting and belief updating deficits in AN correlate with altered activity in key neurotransmitter and anatomical systems. Moreover, the medication that the clinical population takes could also be accounted for to further examine how precision-weighting is affected in relation to HCs.

Thirdly, although studies have studied neural processing of interoceptive events in other eating disorders, such as bulimia nervosa, to my knowledge no study has examined metacognitive belief updating about interoception in bulimia nervosa. Future research could thus examine if and how the results in eating disorders marked by restriction of eating assimilate patterns observed in eating disorders marked by bingeing and purging.

Finally, future studies should consider addressing the sampling issues this thesis had. Namely, these include assessing interoceptive belief updating in men, as well as adolescents to further elucidate which patterns may be affected by biological and/or hormonal factors. Crucially, future studies should consider employing longitudinal designs, instead of only cross-sectional designs, in order to also better understand the transdiagnostic role and effects of comorbid traits and mechanisms in belief updating and self-efficacy beliefs in AN.

6.7 Concluding Remarks

To conclude, this thesis used an interdisciplinary approach to experimentally and computationally study interoceptive belief updating in the cardiac and gastrointestinal modalities in individuals with acute AN, weight-restored from AN, and HCs. Specifically, patients completed a HCT task to which self-efficacy beliefs about abilities to perform well (i.e., feel heartbeats) before and after the cardiac perception task were also sampled in order to study how these beliefs are formed and updated retrospectively, and prospectively. In the gastrointestinal modality, participants had to predict and update subjective states of stomach fullness. The results stemming from this thesis – despite the acknowledged methodological limitations – offer insights into the following. First, by assessing the different dimensions of interoception (e.g., interoceptive accuracy, self-reported beliefs on interoceptive sensibility, interoceptive insight) and across modalities, this thesis showed that while interoception in AN may not be compromised across all levels and modalities, they influence one another (e.g., poor confidence regarding interoceptive accuracy in the HCT may explain greater reliance on priors). Second, the results regarding the belief updating and learning rates across the cardiac and gastrointestinal modalities suggest that metacognitive deficits may be contributing to difficulties in (i) updating beliefs about abilities, influencing clinical treatment outcomes, and (ii) updating beliefs about bodily needs and responding to those needs (e.g., recognising a hunger cue and satisfying this hunger need). Third, the observations from the audit suggest that HRBF may be a promising add-on to existing practices of mindfulness in order to help patients improve their beliefs about self-regulatory abilities, but these results cannot be generalised unless more Trusts carry out a similar clinical audit. Moreover, to establish that HRBF is indeed better than other existing practices, clinical trials would have to assess this. Fourth, overall it was observed that AN patients were able to use non-embodied feedback to retrospectively and prospectively update their interoceptive self-efficacy beliefs, but did not update their beliefs in the same way as HCs when only relying on their prior beliefs and embodied, interoceptive feedback during task performance, suggesting a potential difference between the general ability

to update beliefs and metacognitive belief updating about interoception in AN. Future studies should further explore interoceptive belief updating across modalities to assess the reliability and validity of the present findings, as well as to continue assessing these across acute and weight-restored samples, to further delineate which observations are state or trait effects. Finally, for more generalisable results, future studies could also investigate how other parameters and factors may influence interoceptive belief updating and precision-weighting mechanisms, such as gender, age (e.g., adolescents vs. adults with AN), and pharmacological treatment.

References

- Abbate-Daga, G., Amianto, F., Delsedime, N., De-Bacco, C., & Fassino, S. (2013). Resistance to treatment and change in anorexia nervosa: A clinical overview. *BMC Psychiatry*, 13(1), 294. <https://doi.org/10.1186/1471-244X-13-294>
- Abbate-Daga, G., Buzzichelli, S., Amianto, F., Rocca, G., Marzola, E., McClintock, S. M., & Fassino, S. (2011). Cognitive flexibility in verbal and nonverbal domains and decision making in anorexia nervosa patients: A pilot study. *BMC Psychiatry*, 11(1), 162. <https://doi.org/10.1186/1471-244X-11-162>
- Adams, R. A. (2018). Chapter 7—Bayesian Inference, Predictive Coding, and Computational Models of Psychosis. In A. Anticevic & J. D. Murray (Eds.), *Computational Psychiatry* (pp. 175–195). Academic Press. <https://doi.org/10.1016/B978-0-12-809825-7.00007-9>
- Adams, R. A., Huys, Q. J. M., & Roiser, J. P. (2016). Computational Psychiatry: Towards a mathematically informed understanding of mental illness. *Journal of Neurology, Neurosurgery & Psychiatry*, 87(1), 53–63. <https://doi.org/10.1136/jnnp-2015-310737>
- Ainley, V., Brass, M., & Tsakiris, M. (2014). Heartfelt imitation: High interoceptive awareness is linked to greater automatic imitation. *Neuropsychologia*, 60, 21–28. <https://doi.org/10.1016/j.neuropsychologia.2014.05.010>
- Aitchison, L., & Lengyel, M. (2017). With or without you: Predictive coding and Bayesian inference in the brain. *Current Opinion in Neurobiology*, 46, 219–227. <https://doi.org/10.1016/j.conb.2017.08.010>
- Allen, M., Fardo, F., Dietz, M. J., Hillebrandt, H., Friston, K. J., Rees, G., & Roepstorff, A. (2016). Anterior insula coordinates hierarchical processing of tactile mismatch responses. *NeuroImage*, 127, 34–43. <https://doi.org/10.1016/j.neuroimage.2015.11.030>
- Allen, M., & Friston, K. (2018). From cognitivism to autopoiesis: Towards a computational framework for the embodied mind. *Synthese*, 195(6), 2459–2482. <https://doi.org/10.1007/s11229-016-1288-5>
- Allen, M., Legrand, N., Correa, C. M. C., & Fardo, F. (2020). Thinking through prior bodies: Autonomic uncertainty and interoceptive self-inference. *Behavioral and Brain Sciences*, 43, e91. <https://doi.org/10.1017/S0140525X19002899>

- Alneyadi, M., Drissi, N., Almeqbaali, M., & Ouhbi, S. (2021). Biofeedback-Based Connected Mental Health Interventions for Anxiety: Systematic Literature Review. *JMIR mHealth and uHealth*, 9(4), e26038. <https://doi.org/10.2196/26038>
- Ambrosecchia, M., Ardizzi, M., Russo, E., Ditaranto, F., Speciale, M., Vinai, P., Todisco, P., Maestro, S., & Gallese, V. (2017). Interoception and Autonomic Correlates during Social Interactions. Implications for Anorexia. *Frontiers in Human Neuroscience*, 11. <https://www.frontiersin.org/articles/10.3389/fnhum.2017.00219>
- American Psychiatric Association (Ed.). (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed). American Psychiatric Association.
- Anderluh, M. B., Tchanturia, K., Rabe-Hesketh, S., & Treasure, J. (2003). Childhood Obsessive-Compulsive Personality Traits in Adult Women With Eating Disorders: Defining a Broader Eating Disorder Phenotype. *American Journal of Psychiatry*, 160(2), 242–247. <https://doi.org/10.1176/appi.ajp.160.2.242>
- 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>
- Anderson, K., Watters, A., Dee, E., & Mehler, P. S. (2023). Can we predict the development of acute gastric dilatation in patients with anorexia nervosa? *Journal of Eating Disorders*, 11(1), 212. <https://doi.org/10.1186/s40337-023-00937-2>
- Andrés-Pepiñá, S., Plana, M. T., Flamarique, I., Romero, S., Borràs, R., Julià, L., Gárriz, M., & Castro-Fornieles, J. (2020). Long-term outcome and psychiatric comorbidity of adolescent-onset anorexia nervosa. *Clinical Child Psychology and Psychiatry*, 25(1), 33–44. <https://doi.org/10.1177/1359104519827629>
- Arbel, R., Koren, D., Klein, E., & Latzer, Y. (2013). The neurocognitive basis of insight into illness in anorexia nervosa: A pilot metacognitive study. *Psychiatry Research*, 209(3), 604–610. <https://doi.org/10.1016/j.psychres.2013.01.009>
- Arcelus, J., Mitchell, A. J., Wales, J., & Nielsen, S. (2011). Mortality Rates in Patients With Anorexia Nervosa and Other Eating Disorders: A Meta-analysis of 36 Studies. *Archives of General Psychiatry*, 68(7), 724–731. <https://doi.org/10.1001/archgenpsychiatry.2011.74>

- Arnaud, S., Sullivan, J., MacKinnon, A., & Bodell, L. P. (2023). Teasing Apart the Roles of Interoception, Emotion, and Self-Control in Anorexia Nervosa. *Review of Philosophy and Psychology*. <https://doi.org/10.1007/s13164-023-00706-9>
- Astin, J. A., Shapiro, S. L., Eisenberg, D. M., & Forsys, K. L. (2003). Mind-Body Medicine: State of the Science, Implications for Practice. *The Journal of the American Board of Family Practice*, 16(2), 131–147. <https://doi.org/10.3122/jabfm.16.2.131>
- Atwood, M. E., & Friedman, A. (2020). A systematic review of enhanced cognitive behavioral therapy (CBT-E) for eating disorders. *International Journal of Eating Disorders*, 53(3), 311–330. <https://doi.org/10.1002/eat.23206>
- Auksztulewicz, R., & Friston, K. (2016). Repetition suppression and its contextual determinants in predictive coding. *Cortex*, 80, 125–140. <https://doi.org/10.1016/j.cortex.2015.11.024>
- Avena, N. M., Rada, P., & Hoebel, B. G. (2008). Underweight rats have enhanced dopamine release and blunted acetylcholine response in the nucleus accumbens while bingeing on sucrose. *Neuroscience*, 156(4), 865–871. <https://doi.org/10.1016/j.neuroscience.2008.08.017>
- Aylward, J., Hales, C., Robinson, E., & Robinson, O. J. (2020). Translating a rodent measure of negative bias into humans: The impact of induced anxiety and unmedicated mood and anxiety disorders. *Psychological Medicine*, 50(2), 237–246. <https://doi.org/10.1017/S0033291718004117>
- Aziz, Q., Andersson, J., Valind, S., Sundin, A., Hamdy, S., Jones, A., Foster, E., Langstrom, B., & Thompson, D. (1997). Identification of human brain loci processing esophageal sensation using positron emission tomography. *Gastroenterology*, 113(1), 50–59. [https://doi.org/10.1016/S0016-5085\(97\)70079-9](https://doi.org/10.1016/S0016-5085(97)70079-9)
- Aziz, Q., Schnitzler, A., & Enck, P. (2000). Functional Neuroimaging of Visceral Sensation. *Journal of Clinical Neurophysiology*, 17(6), 604.
- Aziz, Q., Thompson, D. G., Ng, V. W. K., Hamdy, S., Sarkar, S., Brammer, M. J., Bullmore, E. T., Hobson, A., Tracey, I., Gregory, L., Simmons, A., & Williams, S. C. R. (2000). Cortical Processing of Human Somatic and Visceral Sensation. *The Journal of Neuroscience*, 20(7), 2657–2663. <https://doi.org/10.1523/JNEUROSCI.20-07-02657.2000>
- Babel, P. (2020). Operant conditioning as a new mechanism of placebo effects. *European Journal of Pain*, 24(5), 902–908. <https://doi.org/10.1002/ejp.1544>

- Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia scale—I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research*, 38(1), 23–32. [https://doi.org/10.1016/0022-3999\(94\)90005-1](https://doi.org/10.1016/0022-3999(94)90005-1)
- Baird, B., Smallwood, J., Gorgolewski, K. J., & Margulies, D. S. (2013). Medial and lateral networks in anterior prefrontal cortex support metacognitive ability for memory and perception. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 33(42), 16657–16665. <https://doi.org/10.1523/JNEUROSCI.0786-13.2013>
- Baker, A. (2020). Neurostimulation Approaches in the Treatment of Severe Anorexia Nervosa. *OBM Neurobiology*, 4(3), 1–19. <https://doi.org/10.21926/obm.neurobiol.2003069>
- Baker, J. H., Girdler, S. S., & Bulik, C. M. (2012). The role of reproductive hormones in the development and maintenance of eating disorders. *Expert Review of Obstetrics & Gynecology*, 7(6), 573–583. <https://doi.org/10.1586/eog.12.54>
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191–215. <https://doi.org/10.1037/0033-295X.84.2.191>
- Bandura, A. (2006). *Self-Efficacy Beliefs of Adolescents*. IAP.
- Bang, D., & Fleming, S. M. (2018). Distinct encoding of decision confidence in human medial prefrontal cortex. *Proceedings of the National Academy of Sciences*, 115(23), 6082–6087. <https://doi.org/10.1073/pnas.1800795115>
- Barca, L., & Pezzulo, G. (2020). Keep your interoceptive streams under control: An active inference perspective on anorexia nervosa. *Cognitive, Affective, & Behavioral Neuroscience*, 20(2), 427–440. <https://doi.org/10.3758/s13415-020-00777-6>
- Bardone, A. M., Perez, M., Abramson, L. Y., & Joiner Jr., T. E. (2003). Self-competence and self-liking in the prediction of change in bulimic symptoms. *International Journal of Eating Disorders*, 34(3), 361–369. <https://doi.org/10.1002/eat.10197>
- Barlow, D. H., Allen, L. B., & Choate, M. L. (2004). Toward a unified treatment for emotional disorders. *Behavior Therapy*, 35(2), 205–230. [https://doi.org/10.1016/S0005-7894\(04\)80036-4](https://doi.org/10.1016/S0005-7894(04)80036-4)
- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal*

- of Personality and Social Psychology*, 51(6), 1173–1182.
<https://doi.org/10.1037/0022-3514.51.6.1173>
- Barrett, L. F., & Simmons, W. K. (2015). Interoceptive predictions in the brain. *Nature Reviews Neuroscience*, 16(7), 419–429. <https://doi.org/10.1038/nrn3950>
- Barron, H. C., Auksztulewicz, R., & Friston, K. (2020). Prediction and memory: A predictive coding account. *Progress in Neurobiology*, 192, 101821.
<https://doi.org/10.1016/j.pneurobio.2020.101821>
- Barsky, A. J. (1992). Amplification, Somatization, and the Somatoform Disorders. *Psychosomatics*, 33(1), 28–34. [https://doi.org/10.1016/S0033-3182\(92\)72018-0](https://doi.org/10.1016/S0033-3182(92)72018-0)
- Beccia, A. L., Dunlap, C., Hanes, D. A., Courneene, B. J., & Zwickey, H. L. (2018). Mindfulness-based eating disorder prevention programs: A systematic review and meta-analysis. *Mental Health & Prevention*, 9, 1–12.
<https://doi.org/10.1016/j.mhp.2017.11.001>
- Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding Advantageously Before Knowing the Advantageous Strategy. *Science*, 275(5304), 1293–1295.
<https://doi.org/10.1126/science.275.5304.1293>
- Beck, A. T. (1979). *Cognitive Therapy of Depression*. Guilford Press.
- Beck, A. T., Baruch, E., Balter, J. M., Steer, R. A., & Warman, D. M. (2004). A new instrument for measuring insight: The Beck Cognitive Insight Scale. *Schizophrenia Research*, 68(2), 319–329. [https://doi.org/10.1016/S0920-9964\(03\)00189-0](https://doi.org/10.1016/S0920-9964(03)00189-0)
- Beck, A. T., & Haigh, E. A. P. (2014). Advances in Cognitive Theory and Therapy: The Generic Cognitive Model. *Annual Review of Clinical Psychology*, 10(1), 1–24.
<https://doi.org/10.1146/annurev-clinpsy-032813-153734>
- Behar, R., Arancibia, M., Gaete, M. I., Silva, H., & Meza-Concha, N. (2018). The delusional dimension of anorexia nervosa: Phenomenological, neurobiological and clinical perspectives. *Archives of Clinical Psychiatry (São Paulo)*, 45(1), 15–21.
<https://doi.org/10.1590/0101-608300000000148>
- Benini, L., Todesco, T., Dalle Grave, R., Deiorio, F., Salandini, L., & Vantini, I. (2004). Gastric Emptying in Patients with Restricting and Binge/Purging Subtypes of Anorexia Nervosa. *Official Journal of the American College of Gastroenterology | ACG*, 99(8), 1448.

- Benini, L., Todesco, T., Frulloni, L., Dalle Grave, R., Campagnola, P., Agugiario, F., Cusumano, C. D., Gabbrielli, A., & Vantini, I. (2010). Esophageal motility and symptoms in restricting and binge-eating/purging anorexia. *Digestive and Liver Disease*, 42(11), 767–772. <https://doi.org/10.1016/j.dld.2010.03.018>
- Benson, P. G., & Önköl, D. (1992). The effects of feedback and training on the performance of probability forecasters. *International Journal of Forecasting*, 8(4), 559–573. [https://doi.org/10.1016/0169-2070\(92\)90066-I](https://doi.org/10.1016/0169-2070(92)90066-I)
- Benwell, C. S. Y., Mohr, G., Wallberg, J., Kouadio, A., & Ince, R. A. A. (2022). Psychiatrically relevant signatures of domain-general decision-making and metacognition in the general population. *Npj Mental Health Research*, 1(1), Article 1. <https://doi.org/10.1038/s44184-022-00009-4>
- Berman, E. S. (2006). The relationship between eating self-efficacy and eating disorder symptoms in a non-clinical sample. *Eating Behaviors*, 7(1), 79–90. <https://doi.org/10.1016/j.eatbeh.2005.07.004>
- Bern, E. M., & O'Brien, R. F. (2013). Is it an eating disorder, gastrointestinal disorder, or both? *Current Opinion in Pediatrics*, 25(4), 463–470. <https://doi.org/10.1097/MOP.0b013e328362d1ad>
- Berner, L. A., Simmons, A. N., Wierenga, C. E., Bischoff-Grethe, A., Paulus, M. P., Bailer, U. F., Ely, A. V., & Kaye, W. H. (2018). Altered interoceptive activation before, during, and after aversive breathing load in women remitted from anorexia nervosa. *Psychological Medicine*, 48(1), 142–154. <https://doi.org/10.1017/S0033291717001635>
- Berner, L. A., Simmons, A. N., Wierenga, C. E., Bischoff-Grethe, A., Paulus, M. P., Bailer, U. F., & Kaye, W. H. (2019). Altered anticipation and processing of aversive interoceptive experience among women remitted from bulimia nervosa. *Neuropsychopharmacology*, 44(7), Article 7. <https://doi.org/10.1038/s41386-019-0361-4>
- Berntson, G. G., Cacioppo, J. T., & Quigley, K. S. (1993). Cardiac psychophysiology and autonomic space in humans: Empirical perspectives and conceptual implications. *Psychological Bulletin*, 114(2), 296–322. <https://doi.org/10.1037/0033-2909.114.2.296>
- Berry, M. E., Chapple, I. T., Ginsberg, J. P., Gleichauf, K. J., Meyer, J. A., & Nagpal, M. L. (2014). Non-pharmacological Intervention for Chronic Pain in Veterans: A Pilot

- Study of Heart Rate Variability Biofeedback. *Global Advances in Health and Medicine*, 3(2), 28–33. <https://doi.org/10.7453/gahmj.2013.075>
- Berthoud, H. (2006). Homeostatic and Non-homeostatic Pathways Involved in the Control of Food Intake and Energy Balance. *Obesity*, 14(S8).
<https://doi.org/10.1038/oby.2006.308>
- Bhome, R., McWilliams, A., Price, G., Poole, N. A., Howard, R. J., Fleming, S. M., & Huntley, J. D. (2022). Metacognition in functional cognitive disorder. *Brain Communications*, 4(2), fcac041. <https://doi.org/10.1093/braincomms/fcac041>
- Bischoff-Grethe, A., Wierenga, C. E., Berner, L. A., Simmons, A. N., Bailer, U., Paulus, M. P., & Kaye, W. H. (2018). Neural hypersensitivity to pleasant touch in women remitted from anorexia nervosa. *Translational Psychiatry*, 8(1), Article 1.
<https://doi.org/10.1038/s41398-018-0218-3>
- Bizeul, C., Sadowsky, N., & Rigaud, D. (2001). The prognostic value of initial EDI scores in anorexia nervosa patients: A prospective follow-up study of 5–10 years. *European Psychiatry*, 16(4), 232–238. [https://doi.org/10.1016/S0924-9338\(01\)00570-3](https://doi.org/10.1016/S0924-9338(01)00570-3)
- Bleuler, E. (1950). *Dementia praecox or the group of schizophrenias* (p. 548). International Universities Press.
- Bluemel, S., Menne, D., Milos, G., Goetze, O., Fried, M., Schwizer, W., Fox, M., & Steingoetter, A. (2017). Relationship of body weight with gastrointestinal motor and sensory function: Studies in anorexia nervosa and obesity. *BMC Gastroenterology*, 17(1), 4. <https://doi.org/10.1186/s12876-016-0560-y>
- Bogaerts, K., Van Eylen, L., Li, W., Bresseleers, J., Van Diest, I., De Peuter, S., Stans, L., Decramer, M., & Van Den Bergh, O. (2010). Distorted symptom perception in patients with medically unexplained symptoms. *Journal of Abnormal Psychology*, 119(1), 226–234. <https://doi.org/10.1037/a0017780>
- Bohon, C., Stice, E., & Spoor, S. (2009). Female emotional eaters show abnormalities in consummatory and anticipatory food reward: A functional magnetic resonance imaging study. *International Journal of Eating Disorders*, 42(3), 210–221.
<https://doi.org/10.1002/eat.20615>
- Bomba, M., Riva, A., Morzenti, S., Grimaldi, M., Neri, F., & Nacinovich, R. (2015). Global and regional brain volumes normalization in weight-recovered adolescents with anorexia nervosa: Preliminary findings of a longitudinal voxel-based morphometry

- study. *Neuropsychiatric Disease and Treatment*, 637.
<https://doi.org/10.2147/NDT.S73239>
- Borden, A., & Cook-Cottone, C. (2020). Yoga and eating disorder prevention and treatment: A comprehensive review and meta-analysis. *Eating Disorders*, 28(4), 400–437.
<https://doi.org/10.1080/10640266.2020.1798172>
- Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, 16(1), 5–13.
<https://doi.org/10.1002/wps.20375>
- Brannigan, M., Stevenson, R. J., & Francis, H. (2015). Thirst interoception and its relationship to a Western-style diet. *Physiology & Behavior*, 139, 423–429.
<https://doi.org/10.1016/j.physbeh.2014.11.050>
- Braun, D. L., Sunday, S. R., Huang, A., & Halmi, K. A. (1999). More males seek treatment for eating disorders. *International Journal of Eating Disorders*, 25(4), 415–424.
[https://doi.org/10.1002/\(SICI\)1098-108X\(199905\)25:4<415::AID-EAT6>3.0.CO;2-B](https://doi.org/10.1002/(SICI)1098-108X(199905)25:4<415::AID-EAT6>3.0.CO;2-B)
- 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>
- Brevers, D., Cleeremans, A., Bechara, A., Greisen, M., Kornreich, C., Verbanck, P., & Noël, X. (2014). Impaired Metacognitive Capacities in Individuals with Problem Gambling. *Journal of Gambling Studies*, 30(1), 141–152. <https://doi.org/10.1007/s10899-012-9348-3>
- Brewer, R., Murphy, J., & Bird, G. (2021). Atypical interoception as a common risk factor for psychopathology: A review. *Neuroscience & Biobehavioral Reviews*, 130, 470–508. <https://doi.org/10.1016/j.neubiorev.2021.07.036>
- Brooks, S. J., Barker, G. J., O'Daly, O. G., Brammer, M., Williams, S. C., Benedict, C., Schiöth, H. B., Treasure, J., & Campbell, I. C. (2011). Restraint of appetite and reduced regional brain volumes in anorexia nervosa: A voxel-based morphometric study. *BMC Psychiatry*, 11(1), 179. <https://doi.org/10.1186/1471-244X-11-179>
- Broomfield, C., Stedal, K., Touyz, S., & Rhodes, P. (2017). Labeling and defining severe and enduring anorexia nervosa: A systematic review and critical analysis. *International Journal of Eating Disorders*, 50(6), 611–623. <https://doi.org/10.1002/eat.22715>

- Brown, K. W., & Ryan, R. M. (2003). The benefits of being present: Mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology*, 84, 822–848. <https://doi.org/10.1037/0022-3514.84.4.822>
- Brown, T. A., Perry, T. R., Kaye, W. H., & Wierenga, C. E. (2022). Pilot study of a water load test as a measure of gastric interoception in anorexia nervosa. *Eating and Weight Disorders - Studies on Anorexia, Bulimia and Obesity*, 27(6), 2223–2228. <https://doi.org/10.1007/s40519-022-01376-9>
- Brown, T. A., Vanzhula, I. A., Reilly, E. E., Levinson, C. A., Berner, L. A., Krueger, A., Lavender, J. M., Kaye, W. H., & Wierenga, C. E. (2020). Body mistrust bridges interoceptive awareness and eating disorder symptoms. *Journal of Abnormal Psychology*, 129(5), 445–456. <https://doi.org/10.1037/abn0000516>
- Bruch, H. (1962). Perceptual and Conceptual Disturbances in Anorexia Nervosa. *Psychosomatic Medicine*, 24(2), 187.
- Bruch, H. (1971). Anorexia Nervosa in the Male. *Psychosomatic Medicine*, 33(1), 31.
- Bruch, H. (1978). *The Golden Cage: The Enigma of Anorexia Nervosa*. Harvard University Press.
- Buerger, A., Vloet, T. D., Haber, L., & Geissler, J. M. (2021). Third-wave interventions for eating disorders in adolescence – systematic review with meta-analysis. *Borderline Personality Disorder and Emotion Dysregulation*, 8, 20. <https://doi.org/10.1186/s40479-021-00158-6>
- Bulik, C. M., Tozzi, F., Anderson, C., Mazzeo, S. E., Aggen, S., & Sullivan, P. F. (2003). The Relation Between Eating Disorders and Components of Perfectionism. *American Journal of Psychiatry*, 160(2), 366–368. <https://doi.org/10.1176/appi.ajp.160.2.366>
- Buratti, S., Allwood, C. M., & Kleitman, S. (2013). First- and second-order metacognitive judgments of semantic memory reports: The influence of personality traits and cognitive styles. *Metacognition and Learning*, 8(1), 79–102. <https://doi.org/10.1007/s11409-013-9096-5>
- Cameron, O. G. (2001). Interoception: The Inside Story—A Model for Psychosomatic Processes. *Psychosomatic Medicine*, 63(5), 697.
- Capuron, L., & Miller, A. H. (2011). Immune system to brain signaling: Neuropsychopharmacological implications. *Pharmacology & Therapeutics*, 130(2), 226–238. <https://doi.org/10.1016/j.pharmthera.2011.01.014>

- Carey, M., Crucianelli, L., Preston, C., & Fotopoulou, A. (2021). The role of affective touch in whole-body embodiment remains equivocal. *Consciousness and Cognition*, 87, 103059. <https://doi.org/10.1016/j.concog.2020.103059>
- Carlat, D., & Camargo, C. A. (1997). Eating disorders in males: A report on 135 patients. *American Journal of Psychiatry*, 154(8), 1127–1132. <https://doi.org/10.1176/ajp.154.8.1127>
- Carleton, R. N., Norton, M. A. P. J., & Asmundson, G. J. G. (2007). Fearing the unknown: A short version of the Intolerance of Uncertainty Scale. *Journal of Anxiety Disorders*, 21(1), 105–117. <https://doi.org/10.1016/j.janxdis.2006.03.014>
- Carter, J. C., Blackmore, E., Sutandar-Pinnock, K., & Woodside, D. B. (2004). Relapse in anorexia nervosa: A survival analysis. *Psychological Medicine*, 34(4), 671–679. <https://doi.org/10.1017/S0033291703001168>
- Carter, J. C., Mercer-Lynn, K. B., Norwood, S. J., Bewell-Weiss, C. V., Crosby, R. D., Woodside, D. B., & Olmsted, M. P. (2012). A prospective study of predictors of relapse in anorexia nervosa: Implications for relapse prevention. *Psychiatry Research*, 200(2–3), 518–523. <https://doi.org/10.1016/j.psychres.2012.04.037>
- Cartwright-Hatton, S., & Wells, A. (1997). Beliefs about Worry and Intrusions: The Meta-Cognitions Questionnaire and its Correlates. *Journal of Anxiety Disorders*, 11(3), 279–296. [https://doi.org/10.1016/S0887-6185\(97\)00011-X](https://doi.org/10.1016/S0887-6185(97)00011-X)
- Cascino, G., Castellini, G., Stanghellini, G., Ricca, V., Cassioli, E., Ruzzi, V., Monteleone, P., & Monteleone, A. M. (2019). The Role of the Embodiment Disturbance in the Anorexia Nervosa Psychopathology: A Network Analysis Study. *Brain Sciences*, 9(10), Article 10. <https://doi.org/10.3390/brainsci9100276>
- Cash, T. F. (2004). Body image: Past, present, and future. *Body Image*, 1(1), 1–5. [https://doi.org/10.1016/S1740-1445\(03\)00011-1](https://doi.org/10.1016/S1740-1445(03)00011-1)
- Cash, T. F., & Deagle, E. A. (1997). The nature and extent of body-image disturbances in anorexia nervosa and bulimia nervosa: A meta-analysis. *International Journal of Eating Disorders*, 22(2), 107–126. [https://doi.org/10.1002/\(SICI\)1098-108X\(199709\)22:2<107::AID-EAT1>3.0.CO;2-J](https://doi.org/10.1002/(SICI)1098-108X(199709)22:2<107::AID-EAT1>3.0.CO;2-J)
- Castellini, G., Lo Sauro, C., Mannucci, E., Ravaldi, C., Rotella, C. M., Faravelli, C., & Ricca, V. (2011). Diagnostic Crossover and Outcome Predictors in Eating Disorders According to DSM-IV and DSM-V Proposed Criteria: A 6-Year Follow-Up Study. *Psychosomatic Medicine*, 73(3), 270. <https://doi.org/10.1097/PSY.0b013e31820a1838>

- Castro-Fornieles, J., Bargalló, N., Lázaro, L., Andrés, S., Falcon, C., Plana, M. T., & Junqué, C. (2009). A cross-sectional and follow-up voxel-based morphometric MRI study in adolescent anorexia nervosa. *Journal of Psychiatric Research*, 43(3), 331–340.
<https://doi.org/10.1016/j.jpsychires.2008.03.013>
- Cavedini, P., Bassi, T., Ubbiali, A., Casolari, A., Giordani, S., Zorzi, C., & Bellodi, L. (2004). Neuropsychological investigation of decision-making in anorexia nervosa. *Psychiatry Research*, 127(3), 259–266.
<https://doi.org/10.1016/j.psychres.2004.03.012>
- Ceunen, E., Vlaeyen, J. W. S., & Van Diest, I. (2016). On the Origin of Interoception. *Frontiers in Psychology*, 7.
<https://www.frontiersin.org/articles/10.3389/fpsyg.2016.00743>
- Chapman, S., Colvin, L. E., Vuorre, M., Cocchini, G., Metcalfe, J., Huey, E. D., & Cosentino, S. (2018). Cross domain self-monitoring in anosognosia for memory loss in Alzheimer's disease. *Cortex*, 101, 221–233.
<https://doi.org/10.1016/j.cortex.2018.01.019>
- Christoff, K., Ream, J. M., Geddes, L. P. T., & Gabrieli, J. D. E. (2003). Evaluating Self-Generated Information: Anterior Prefrontal Contributions to Human Cognition. *Behavioral Neuroscience*, 117(6), 1161–1168. <https://doi.org/10.1037/0735-7044.117.6.1161>
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3), 181–204.
<https://doi.org/10.1017/S0140525X12000477>
- Clark, L., Bechara, A., Damasio, H., Aitken, M. R. F., Sahakian, B. J., & Robbins, T. W. (2008). Differential effects of insular and ventromedial prefrontal cortex lesions on risky decision-making. *Brain*, 131(5), 1311–1322.
<https://doi.org/10.1093/brain/awn066>
- Coddington, D. R., & Bruch, H. (1970). Gastric Perceptivity in Normal, Obese and Schizophrenic Subjects. *Psychosomatics*, 11(6), 571–579.
[https://doi.org/10.1016/S0033-3182\(70\)71578-8](https://doi.org/10.1016/S0033-3182(70)71578-8)
- Coniglio, K. A., Becker, K. R., Franko, D. L., Zayas, L. V., Plessow, F., Eddy, K. T., & Thomas, J. J. (2017). Won't stop or can't stop? Food restriction as a habitual behavior among individuals with anorexia nervosa or atypical anorexia nervosa. *Eating Behaviors*, 26, 144–147. <https://doi.org/10.1016/j.eatbeh.2017.03.005>

- Cooper, M., Collison, A. O., Collica, S. C., Pan, I., Tamashiro, K. L., Redgrave, G. W., Schreyer, C. C., & Guarda, A. S. (2021). Gastrointestinal symptomatology, diagnosis, and treatment history in patients with underweight avoidant/restrictive food intake disorder and anorexia nervosa: Impact on weight restoration in a meal-based behavioral treatment program. *International Journal of Eating Disorders*, 54(6), 1055–1062. <https://doi.org/10.1002/eat.23535>
- Cooper, M. J. (2005). Cognitive theory in anorexia nervosa and bulimia nervosa: Progress, development and future directions. *Clinical Psychology Review*, 25(4), 511–531. <https://doi.org/10.1016/j.cpr.2005.01.003>
- Cooper, M. J., Deepak, K., Grocutt, E., & Bailey, E. (2007). The experience of ‘feeling fat’ in women with anorexia nervosa, dieting and non-dieting women: An exploratory study. *European Eating Disorders Review*, 15(5), 366–372. <https://doi.org/10.1002/erv.785>
- Corcoran, A. W., & Hohwy, J. (2017). *Allostasis, interoception, and the free energy principle: Feeling our way forward*. PsyArXiv. <https://doi.org/10.31234/osf.io/zbqnx>
- Cornwell, B. R., Garrido, M. I., Overstreet, C., Pine, D. S., & Grillon, C. (2017). The Unpredictive Brain Under Threat: A Neurocomputational Account of Anxious Hypervigilance. *Biological Psychiatry*, 82(6), 447–454. <https://doi.org/10.1016/j.biopsych.2017.06.031>
- Corstorphine, E. (2006). Cognitive–Emotional–Behavioural Therapy for the eating disorders: Working with beliefs about emotions. *European Eating Disorders Review*, 14(6), 448–461. <https://doi.org/10.1002/erv.747>
- Cosentino, S., Metcalfe, J., Cary, M. S., De Leon, J., & Karlawish, J. (2011). Memory Awareness Influences Everyday Decision Making Capacity about Medication Management in Alzheimer’s Disease. *International Journal of Alzheimer’s Disease*, 2011, e483897. <https://doi.org/10.4061/2011/483897>
- Couturier, J., & Lock, J. (2006). What is recovery in adolescent anorexia nervosa? *International Journal of Eating Disorders*, 39(7), 550–555. <https://doi.org/10.1002/eat.20309>
- Cowdrey, F. A., & Park, R. J. (2012). The role of experiential avoidance, rumination and mindfulness in eating disorders. *Eating Behaviors*, 13(2), 100–105. <https://doi.org/10.1016/j.eatbeh.2012.01.001>

- Craig, A. D. (2003). Interoception: The sense of the physiological condition of the body. *Current Opinion in Neurobiology*, 13(4), 500–505. [https://doi.org/10.1016/S0959-4388\(03\)00090-4](https://doi.org/10.1016/S0959-4388(03)00090-4)
- Craig, A. D. (2004). Human feelings: Why are some more aware than others? *Trends in Cognitive Sciences*, 8(6), 239–241. <https://doi.org/10.1016/j.tics.2004.04.004>
- Crane, A. M., Roberts, M. E., & Treasure, J. (2007). Are obsessive-compulsive personality traits associated with a poor outcome in anorexia nervosa? A systematic review of randomized controlled trials and naturalistic outcome studies. *International Journal of Eating Disorders*, 40(7), 581–588. <https://doi.org/10.1002/eat.20419>
- Crisp, A., Gowers, S., Joughin, N., McClelland, L., Rooney, B., Nielsen, S., Bowyer, C., Halek, C., & Hartman, D. (2006). 1.5. Anorexia nervosa in males: Similarities and differences to anorexia nervosa in females. *European Eating Disorders Review*, 14(3), 163–167. <https://doi.org/10.1002/erv.703>
- Critchley, H. D., & Garfinkel, S. N. (2017). Interoception and emotion. *Current Opinion in Psychology*, 17, 7–14. <https://doi.org/10.1016/j.copsyc.2017.04.020>
- Crucianelli, L., Cardi, V., Treasure, J., Jenkinson, P. M., & Fotopoulou, A. (2016). The perception of affective touch in anorexia nervosa. *Psychiatry Research*, 239, 72–78. <https://doi.org/10.1016/j.psychres.2016.01.078>
- Crucianelli, L., Demartini, B., Goeta, D., Nisticò, V., Saramandi, A., Bertelli, S., Todisco, P., Gambini, O., & Fotopoulou, A. (2021). The Anticipation and Perception of Affective Touch in Women with and Recovered from Anorexia Nervosa. *Neuroscience*, 464, 143–155. <https://doi.org/10.1016/j.neuroscience.2020.09.013>
- Crucianelli, L., & Ehrsson, H. H. (2023). The Role of the Skin in Interoception: A Neglected Organ? *Perspectives on Psychological Science*, 18(1), 224–238. <https://doi.org/10.1177/17456916221094509>
- Culbert, K. M., Racine, S. E., & Klump, K. L. (2015). Research Review: What we have learned about the causes of eating disorders – a synthesis of sociocultural, psychological, and biological research. *Journal of Child Psychology and Psychiatry*, 56(11), 1141–1164. <https://doi.org/10.1111/jcpp.12441>
- Culbert, K. M., Sisk, C. L., & Klump, K. L. (2021). A Narrative Review of Sex Differences in Eating Disorders: Is There a Biological Basis? *Clinical Therapeutics*, 43(1), 95–111. <https://doi.org/10.1016/j.clinthera.2020.12.003>

- Curran, L., Schmidt, U., Treasure, J., & Jick, H. (2005). Time trends in eating disorder incidence. *The British Journal of Psychiatry*, 186(2), 132–135.
<https://doi.org/10.1192/bjp.186.2.132>
- Dalton, B., Bartholdy, S., Campbell, I. C., & Schmidt, U. (2018). Neurostimulation in Clinical and Sub-clinical Eating Disorders: A Systematic Update of the Literature. *Current Neuropsychopharmacology*, 16(8), 1174–1192.
<https://doi.org/10.2174/1570159X166666180108111532>
- Damasio, A. R. (1994). Descartes' Error and the Future of Human Life. *Scientific American*, 271(4), 144–144. <https://doi.org/10.1038/scientificamerican1094-144>
- Danner, U. N., Sternheim, L., Bijsterbosch, J. M., Dingemans, A. E., Evers, C., & van Elburg, A. A. (2016). Influence of negative affect on decision making in women with restrictive and binge-purge type anorexia nervosa. *Psychiatry Research*, 239, 39–46.
<https://doi.org/10.1016/j.psychres.2016.02.054>
- Davenport, E., Rushford, N., Soon, S., & McDermott, C. (2015). Dysfunctional metacognition and drive for thinness in typical and atypical anorexia nervosa. *Journal of Eating Disorders*, 3(1), 24. <https://doi.org/10.1186/s40337-015-0060-4>
- David, A. S. (2004a). *Insight and Psychosis: Awareness of Illness in Schizophrenia and Related Disorders*. OUP Oxford.
- David, A. S. (2004b). *Insight and Psychosis: Awareness of Illness in Schizophrenia and Related Disorders*. OUP Oxford.
- David, A. S., Bedford, N., Wiffen, B., & Gilleen, J. (2012). Failures of metacognition and lack of insight in neuropsychiatric disorders. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1594), 1379–1390.
<https://doi.org/10.1098/rstb.2012.0002>
- Davidovic, M., Karjalainen, L., Starck, G., Wentz, E., Björnsdotter, M., & Olausson, H. (2018). Abnormal brain processing of gentle touch in anorexia nervosa. *Psychiatry Research: Neuroimaging*, 281, 53–60.
<https://doi.org/10.1016/j.pscychresns.2018.08.007>
- Davis, B. J., Lysaker, P. H., Salyers, M. P., & Minor, K. S. (2020). The insight paradox in schizophrenia: A meta-analysis of the relationship between clinical insight and quality of life. *Schizophrenia Research*, 223, 9–17.
<https://doi.org/10.1016/j.schres.2020.07.017>

- Davis, C., & Woodside, D. B. (2002). Sensitivity to the rewarding effects of food and exercise in the eating disorders. *Comprehensive Psychiatry*, 43(3), 189–194. <https://doi.org/10.1053/comp.2002.32356>
- Daw, N. D., & Doya, K. (2006). The computational neurobiology of learning and reward. *Current Opinion in Neurobiology*, 16(2), 199–204. <https://doi.org/10.1016/j.conb.2006.03.006>
- Dayan, P., & Berridge, K. C. (2014). Model-based and model-free Pavlovian reward learning: Revaluation, revision, and revelation. *Cognitive, Affective, & Behavioral Neuroscience*, 14(2), 473–492. <https://doi.org/10.3758/s13415-014-0277-8>
- De Gardelle, V., Le Corre, F., & Mamassian, P. (2016). Confidence as a Common Currency between Vision and Audition. *PLOS ONE*, 11(1), e0147901. <https://doi.org/10.1371/journal.pone.0147901>
- Deacon, B. J., Lickel, J. J., Farrell, N. R., Kemp, J. J., & Hipol, L. J. (2013). Therapist perceptions and delivery of interoceptive exposure for panic disorder. *Journal of Anxiety Disorders*, 27(2), 259–264. <https://doi.org/10.1016/j.janxdis.2013.02.004>
- Desender, K., Boldt, A., & Yeung, N. (2018). Subjective Confidence Predicts Information Seeking in Decision Making. *Psychological Science*, 29(5), 761–778. <https://doi.org/10.1177/0956797617744771>
- 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., 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., Walentynowicz, M., & Corneille, O. (2023). *The New Measures of Interoceptive Accuracy: A Systematic Review and Assessment*. <https://doi.org/10.31219/osf.io/f56xe>
- 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>

- Dhopatkar, N., Keeler, J. L., Mutwalli, H., Whelan, K., Treasure, J., & Himmerich, H. (2023). Gastrointestinal symptoms, gut microbiome, probiotics and prebiotics in anorexia nervosa: A review of mechanistic rationale and clinical evidence. *Psychoneuroendocrinology*, 147, 105959. <https://doi.org/10.1016/j.psyneuen.2022.105959>
- Di Lernia, D., Serino, S., Tuena, C., Cacciatore, C., Polli, N., & Riva, G. (2023). Mental health meets computational neuroscience: A predictive Bayesian account of the relationship between interoception and multisensory bodily illusions in anorexia nervosa. *International Journal of Clinical and Health Psychology*, 23(4), 100383. <https://doi.org/10.1016/j.ijchp.2023.100383>
- Dimitropoulos, G., McCallum, L., Colasanto, M., Freeman, V. E., & Gadalla, T. (2016). The Effects of Stigma on Recovery Attitudes in People With Anorexia Nervosa in Intensive Treatment. *The Journal of Nervous and Mental Disease*, 204(5), 370. <https://doi.org/10.1097/NMD.0000000000000480>
- 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>
- Donoso, M., Collins, A. G. E., & Koechlin, E. (2014). Foundations of human reasoning in the prefrontal cortex. *Science*, 344(6191), 1481–1486. <https://doi.org/10.1126/science.1252254>
- Dörsam, A. F., Mack, I., Kögel, L., Zipfel, S., & Giel, K. E. (2020). How do patients with eating disorders perceive and evaluate food portion sizes? A systematic review. *European Eating Disorders Review*, 28(4), 398–409. <https://doi.org/10.1002/erv.2743>
- Dunn, B. D., Dalgleish, T., Ogilvie, A. D., & Lawrence, A. D. (2007). Heartbeat perception in depression. *Behaviour Research and Therapy*, 45(8), 1921–1930. <https://doi.org/10.1016/j.brat.2006.09.008>
- Dunn, B. D., Evans, D., Makarova, D., White, J., & Clark, L. (2012). Gut feelings and the reaction to perceived inequity: The interplay between bodily responses, regulation, and perception shapes the rejection of unfair offers on the ultimatum game. *Cognitive, Affective, & Behavioral Neuroscience*, 12(3), 419–429. <https://doi.org/10.3758/s13415-012-0092-z>

- 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. *Psychological Science*, 21(12), 1835–1844. <https://doi.org/10.1177/0956797610389191>
- 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>
- Eisenacher, S., Rausch, F., Ainser, F., Mier, D., Veckenstedt, R., Schirmbeck, F., Lewien, A., Englisch, S., Andreou, C., Moritz, S., Meyer-Lindenberg, A., Kirsch, P., & Zink, M. (2015). Investigation of metamemory functioning in the at-risk mental state for psychosis. *Psychological Medicine*, 45(15), 3329–3340. <https://doi.org/10.1017/S0033291715001373>
- Emanuelson, L., Drew, R., & Köteles, F. (2015). Interoceptive sensitivity, body image dissatisfaction, and body awareness in healthy individuals. *Scandinavian Journal of Psychology*, 56(2), 167–174. <https://doi.org/10.1111/sjop.12183>
- Eshkevari, E., Rieger, E., Musiat, P., & Treasure, J. (2014). An investigation of interoceptive sensitivity in eating disorders using a heartbeat detection task and a self-report measure. *European Eating Disorders Review: The Journal of the Eating Disorders Association*, 22(5), 383–388. <https://doi.org/10.1002/erv.2305>
- Eskild-Jensen, M., Støving, R. K., Flindt, C. F., & Sjogren, M. (2020). Comorbid depression as a negative predictor of weight gain during treatment of anorexia nervosa: A systematic scoping review. *European Eating Disorders Review*, 28(6), 605–619. <https://doi.org/10.1002/erv.2787>
- Everaert, J., Bronstein, M. V., Cannon, T. D., & Joormann, J. (2018). Looking Through Tinted Glasses: Depression and Social Anxiety Are Related to Both Interpretation Biases and Inflexible Negative Interpretations. *Clinical Psychological Science*, 6(4), 517–528. <https://doi.org/10.1177/2167702617747968>
- 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>

- Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders*, 16(4), 363–370. [https://doi.org/10.1002/1098-108X\(199412\)16:4<363::AID-EAT2260160405>3.0.CO;2-#](https://doi.org/10.1002/1098-108X(199412)16:4<363::AID-EAT2260160405>3.0.CO;2-#)
- Fairburn, C. G., Cooper, Z., & Shafran, R. (2003). Cognitive behaviour therapy for eating disorders: A “transdiagnostic” theory and treatment. *Behaviour Research and Therapy*, 41(5), 509–528. [https://doi.org/10.1016/S0005-7967\(02\)00088-8](https://doi.org/10.1016/S0005-7967(02)00088-8)
- 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://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2015.00763>
- Fassino, S., Amianto, F., Gramaglia, C., Facchini, F., & Abbate Daga, G. (2004). Temperament and character in eating disorders: Ten years of studies. *Eating and Weight Disorders - Studies on Anorexia, Bulimia and Obesity*, 9(2), 81–90. <https://doi.org/10.1007/BF03325050>
- Fassino, S., Pieró, A., Daga, G. A., Leombruni, P., Mortara, P., & Rovera, G. G. (2002). Attentional biases and frontal functioning in anorexia nervosa: Neuropsychological Tests in AN. *International Journal of Eating Disorders*, 31(3), 274–283. <https://doi.org/10.1002/eat.10028>
- Fassino, S., Pierò, A., Gramaglia, C., & Abbate-Daga, G. (2004). Clinical, Psychopathological and Personality Correlates of Interoceptive Awareness in Anorexia nervosa, Bulimia nervosa and Obesity. *Psychopathology*, 37(4), 168–174. <https://doi.org/10.1159/000079420>
- Fatt, S. J., Mitchison, D., Bussey, K., & Mond, J. (2022). Methods used to assess insight in individuals with eating disorders: A scoping review. *Journal of Mental Health*, 1–12. <https://doi.org/10.1080/09638237.2022.2069696>
- Fealey, R. D. (2013). Interoception and autonomic nervous system reflexes thermoregulation. In *Handbook of Clinical Neurology* (Vol. 117, pp. 79–88). Elsevier. <https://doi.org/10.1016/B978-0-444-53491-0.00007-9>
- Ferentzi, E., Wilhelm, O., & Köteles, F. (2022). What counts when heartbeats are counted. *Trends in Cognitive Sciences*, 26(10), 832–835. <https://doi.org/10.1016/j.tics.2022.07.009>

- Fernández-Aranda, F., Poyastro Pinheiro, A., Tozzi, F., La Via, M., Thornton, L. M., Plotnicov, K. H., Kaye, W. H., Fichter, M. M., Halmi, K. A., Kaplan, A. S., Woodside, D. B., Klump, K. L., Strober, M., Crow, S., Mitchell, J., Rotondo, A., Keel, P., Berrettini, W. H., Rickels, K. E., ... Bulik, C. M. (2007). Symptom Profile of Major Depressive Disorder in Women with Eating Disorders. *Australian & New Zealand Journal of Psychiatry*, 41(1), 24–31.
<https://doi.org/10.1080/00048670601057718>
- Fernández-Aranda, F., Treasure, J., Paslakis, G., Agüera, Z., Giménez, M., Granero, R., Sánchez, I., Serrano-Troncoso, E., Gorwood, P., Herpertz-Dahlmann, B., Bonin, E.-M., Monteleone, P., & Jiménez-Murcia, S. (2021). The impact of duration of illness on treatment nonresponse and drop-out: Exploring the relevance of enduring eating disorder concept. *European Eating Disorders Review*, 29(3), 499–513.
<https://doi.org/10.1002/erv.2822>
- Fichter, M. M., & Quadflieg, N. (1999). Six-Year course and outcome of anorexia nervosa. *International Journal of Eating Disorders*, 26(4), 359–385.
[https://doi.org/10.1002/\(SICI\)1098-108X\(199912\)26:4<359::AID-EAT2>3.0.CO;2-7](https://doi.org/10.1002/(SICI)1098-108X(199912)26:4<359::AID-EAT2>3.0.CO;2-7)
- Fischer, D., Berberich, G., Zaudig, M., Krauseneck, T., Weiss, S., & Pollatos, O. (2016). Interoceptive Processes in Anorexia Nervosa in the Time Course of Cognitive-Behavioral Therapy: A Pilot Study. *Frontiers in Psychiatry*, 7.
<https://doi.org/10.3389/fpsy.2016.00199>
- Fitz-Clarke, J. R. (2007). Computer simulation of human breath-hold diving: Cardiovascular adjustments. *European Journal of Applied Physiology*, 100(2), 207–224.
<https://doi.org/10.1007/s00421-007-0421-z>
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34(10), 906–911.
<https://doi.org/10.1037/0003-066X.34.10.906>
- Fleck, M. S., Daselaar, S. M., Dobbins, I. G., & Cabeza, R. (2006). Role of prefrontal and anterior cingulate regions in decision-making processes shared by memory and nonmemory tasks. *Cerebral Cortex (New York, N.Y.: 1991)*, 16(11), 1623–1630.
<https://doi.org/10.1093/cercor/bhj097>
- Fleming, S. M. (2020). Awareness as inference in a higher-order state space. *Neuroscience of Consciousness*, 2020(1), niz020. <https://doi.org/10.1093/nc/niz020>

- Fleming, S. M. (2024). Metacognition and Confidence: A Review and Synthesis. *Annual Review of Psychology*, 75(1), annurev-psych-022423-032425.
<https://doi.org/10.1146/annurev-psych-022423-032425>
- Fleming, S. M., & Daw, N. D. (2017). Self-evaluation of decision-making: A general Bayesian framework for metacognitive computation. *Psychological Review*, 124(1), 91–114. <https://doi.org/10.1037/rev0000045>
- Fleming, S. M., & Dolan, R. J. (2012). The neural basis of metacognitive ability. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1594), 1338–1349. <https://doi.org/10.1098/rstb.2011.0417>
- Fleming, S. M., Huijgen, J., & Dolan, R. J. (2012). Prefrontal contributions to metacognition in perceptual decision making. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 32(18), 6117–6125.
<https://doi.org/10.1523/JNEUROSCI.6489-11.2012>
- Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience*, 8. <https://www.frontiersin.org/articles/10.3389/fnhum.2014.00443>
- Fleming, S. M., Massoni, S., Gajdos, T., & Vergnaud, J.-C. (2016). Metacognition about the past and future: Quantifying common and distinct influences on prospective and retrospective judgments of self-performance. *Neuroscience of Consciousness*, 2016(1), niw018. <https://doi.org/10.1093/nc/niw018>
- Fleming, S. M., Ryu, J., Golfinos, J. G., & Blackmon, K. E. (2014). Domain-specific impairment in metacognitive accuracy following anterior prefrontal lesions. *Brain*, 137(10), 2811–2822. <https://doi.org/10.1093/brain/awu221>
- Fleming, S. M., Weil, R. S., Nagy, Z., Dolan, R. J., & Rees, G. (2010). Relating introspective accuracy to individual differences in brain structure. *Science (New York, N.Y.)*, 329(5998), 1541–1543. <https://doi.org/10.1126/science.1191883>
- Fletcher, P. C., & Frith, C. D. (2009). Perceiving is believing: A Bayesian approach to explaining the positive symptoms of schizophrenia. *Nature Reviews Neuroscience*, 10(1), Article 1. <https://doi.org/10.1038/nrn2536>
- Foa, E. B., Huppert, J. D., Leiberg, S., Langner, R., Kichic, R., Hajcak, G., & Salkovskis, P. M. (2002). The Obsessive-Compulsive Inventory: Development and validation of a short version. *Psychological Assessment*, 14(4), 485–496.
<https://doi.org/10.1037/1040-3590.14.4.485>

- Fogelson, N., Litvak, V., Peled, A., Fernandez-del-Olmo, M., & Friston, K. (2014). The functional anatomy of schizophrenia: A dynamic causal modeling study of predictive coding. *Schizophrenia Research*, 158(1), 204–212.
<https://doi.org/10.1016/j.schres.2014.06.011>
- Fotopoulou, A. (2015). The virtual bodily self: Mentalisation of the body as revealed in anosognosia for hemiplegia. *Consciousness and Cognition*, 33, 500–510.
<https://doi.org/10.1016/j.concog.2014.09.018>
- Frank, G. K., Shott, M. E., Hagman, J. O., & Mittal, V. A. (2013). Taste Reward Circuitry Related Brain Structures Characterize Ill and Recovered Anorexia Nervosa and Bulimia Nervosa. *The American Journal of Psychiatry*, 170(10), 1152–1160.
<https://doi.org/10.1176/appi.ajp.2013.12101294>
- Frank, G. K. W. (2016). The Perfect Storm—A Bio-Psycho-Social Risk Model for Developing and Maintaining Eating Disorders. *Frontiers in Behavioral Neuroscience*, 10. <https://www.frontiersin.org/articles/10.3389/fnbeh.2016.00044>
- Frank, G. K. W. (2021). From Desire to Dread—A Neurocircuitry Based Model for Food Avoidance in Anorexia Nervosa. *Journal of Clinical Medicine*, 10(11), Article 11.
<https://doi.org/10.3390/jcm10112228>
- Frank, G. K. W., DeGuzman, M. C., & Shott, M. E. (2019). Motivation to eat and not to eat – The psycho-biological conflict in anorexia nervosa. *Physiology & Behavior*, 206, 185–190. <https://doi.org/10.1016/j.physbeh.2019.04.007>
- Frank, G. K. W., Oberndorfer, T. A., Simmons, A. N., Paulus, M. P., Fudge, J. L., Yang, T. T., & Kaye, W. H. (2008). Sucrose activates human taste pathways differently from artificial sweetener. *NeuroImage*, 39(4), 1559–1569.
<https://doi.org/10.1016/j.neuroimage.2007.10.061>
- Franko, D. L., Tabri, N., Keshaviah, A., Murray, H. B., Herzog, D. B., Thomas, J. J., Coniglio, K., Keel, P. K., & Eddy, K. T. (2018). Predictors of long-term recovery in anorexia nervosa and bulimia nervosa: Data from a 22-year longitudinal study. *Journal of Psychiatric Research*, 96, 183–188.
<https://doi.org/10.1016/j.jpsychires.2017.10.008>
- Franz, M., Popp, K., Schaefer, R., Sitte, W., Schneider, C., Hardt, J., Decker, O., & Braehler, E. (2008). Alexithymia in the German general population. *Social Psychiatry and Psychiatric Epidemiology*, 43(1), 54–62. <https://doi.org/10.1007/s00127-007-0265-1>

- Friederich, H.-C., & Herzog, W. (2010). Cognitive-Behavioral Flexibility in Anorexia Nervosa. In R. A. H. Adan & W. H. Kaye (Eds.), *Behavioral Neurobiology of Eating Disorders* (Vol. 6, pp. 111–123). Springer Berlin Heidelberg.
https://doi.org/10.1007/7854_2010_83
- Friederich, H.-C., Walther, S., Bendszus, M., Biller, A., Thomann, P., Zeigermann, S., Katus, T., Brunner, R., Zastrow, A., & Herzog, W. (2012). Grey matter abnormalities within cortico-limbic-striatal circuits in acute and weight-restored anorexia nervosa patients. *NeuroImage*, 59(2), 1106–1113. <https://doi.org/10.1016/j.neuroimage.2011.09.042>
- Friston, K. (2003). Learning and inference in the brain. *Neural Networks*, 16(9), 1325–1352.
<https://doi.org/10.1016/j.neunet.2003.06.005>
- Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1456), 815–836.
<https://doi.org/10.1098/rstb.2005.1622>
- Friston, K. (2008). Hierarchical Models in the Brain. *PLoS Computational Biology*, 4(11), e1000211. <https://doi.org/10.1371/journal.pcbi.1000211>
- Friston, K. (2009). The free-energy principle: A rough guide to the brain? *Trends in Cognitive Sciences*, 13(7), 293–301. <https://doi.org/10.1016/j.tics.2009.04.005>
- Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2), Article 2. <https://doi.org/10.1038/nrn2787>
- Friston, K. (2017). Precision Psychiatry. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 2(8), 640–643. <https://doi.org/10.1016/j.bpsc.2017.08.007>
- Friston, K., Daunizeau, J., Kilner, J., & Kiebel, S. J. (2010). Action and behavior: A free-energy formulation. *Biological Cybernetics*, 102(3), 227–260.
<https://doi.org/10.1007/s00422-010-0364-z>
- Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., O'Doherty, J., & Pezzulo, G. (2016). Active inference and learning. *Neuroscience & Biobehavioral Reviews*, 68, 862–879. <https://doi.org/10.1016/j.neubiorev.2016.06.022>
- Friston, K. J., Daunizeau, J., & Kiebel, S. J. (2009). Reinforcement Learning or Active Inference? *PLoS ONE*, 4(7), e6421. <https://doi.org/10.1371/journal.pone.0006421>
- Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1521), 1211–1221. <https://doi.org/10.1098/rstb.2008.0300>

- Friston, K., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. *Journal of Physiology-Paris*, 100(1–3), 70–87. <https://doi.org/10.1016/j.jphysparis.2006.10.001>
- Friston, K., Rigoli, F., Ognibene, D., Mathys, C., Fitzgerald, T., & Pezzulo, G. (2015). Active inference and epistemic value. *Cognitive Neuroscience*, 6(4), 187–214. <https://doi.org/10.1080/17588928.2015.1020053>
- Friston, K., Schwartenbeck, P., Fitzgerald, T., Moutoussis, M., Behrens, T., & Dolan, R. (2013). The anatomy of choice: Active inference and agency. *Frontiers in Human Neuroscience*, 7. <https://www.frontiersin.org/articles/10.3389/fnhum.2013.00598>
- Frömer, R., Nassar, M. R., Bruckner, R., Stürmer, B., Sommer, W., & Yeung, N. (2021). Response-based outcome predictions and confidence regulate feedback processing and learning. *eLife*, 10, e62825. <https://doi.org/10.7554/eLife.62825>
- Fu, Z., Beam, D., Chung, J. M., Reed, C. M., Mamelak, A. N., Adolphs, R., & Rutishauser, U. (2022). The geometry of domain-general performance monitoring in the human medial frontal cortex. *Science*, 376(6593), eabm9922. <https://doi.org/10.1126/science.abm9922>
- Fuglset, T. S., Endestad, T., Hilland, E., Bang, L., Tamnes, C. K., Landrø, N. I., & Rø, Ø. (2016). Brain volumes and regional cortical thickness in young females with anorexia nervosa. *BMC Psychiatry*, 16(1), 404. <https://doi.org/10.1186/s12888-016-1126-9>
- Fujisawa, T. X., Yatsuga, C., Mabe, H., Yamada, E., Masuda, M., & Tomoda, A. (2015). Anorexia Nervosa during Adolescence Is Associated with Decreased Gray Matter Volume in the Inferior Frontal Gyrus. *PLOS ONE*, 10(6), e0128548. <https://doi.org/10.1371/journal.pone.0128548>
- Füstös, J., Gramann, K., Herbert, B. M., & Pollatos, O. (2013). On the embodiment of emotion regulation: Interoceptive awareness facilitates reappraisal. *Social Cognitive and Affective Neuroscience*, 8(8), 911–917. <https://doi.org/10.1093/scan/nss089>
- Gadsby, S. (2019). Body representations and cognitive ontology: Drawing the boundaries of the body image. *Consciousness and Cognition*, 74, 102772. <https://doi.org/10.1016/j.concog.2019.102772>
- Galvin, S. J., Podd, J. V., Drga, V., & Whitmore, J. (2003). Type 2 tasks in the theory of signal detectability: Discrimination between correct and incorrect decisions. *Psychonomic Bulletin & Review*, 10(4), 843–876. <https://doi.org/10.3758/BF03196546>

- García-Cordero, I., Sedeño, L., De La Fuente, L., Slachevsky, A., Forno, G., Klein, F., Lillo, P., Ferrari, J., Rodriguez, C., Bustin, J., Torralva, T., Baez, S., Yoris, A., Esteves, S., Melloni, M., Salamone, P., Huepe, D., Manes, F., García, A. M., & Ibañez, A. (2016). Feeling, learning from and being aware of inner states: Interoceptive dimensions in neurodegeneration and stroke. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1708), 20160006. <https://doi.org/10.1098/rstb.2016.0006>
- Garfinkel, S. N., Schulz, A., & Tsakiris, M. (2022). Addressing the need for new interoceptive methods. *Biological Psychology*, 170, 108322. <https://doi.org/10.1016/j.biopsycho.2022.108322>
- 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. (2004). *Eating Disorder Inventory—3 (EDI-3). Professional Manual*. Odessa, FL: Psychological Assessment Resources, 1.
- Garrido, I., & Subirá, S. (2013). Decision-making and impulsivity in eating disorder patients. *Psychiatry Research*, 207(1), 107–112. <https://doi.org/10.1016/j.psychres.2012.10.004>
- Gárriz, M., Andrés-Perpiñá, S., Plana, M. T., Flamarique, I., Romero, S., Julià, L., & Castro-Fornieles, J. (2021). Personality disorder traits, obsessive ideation and perfectionism 20 years after adolescent-onset anorexia nervosa: A recovered study. *Eating and Weight Disorders - Studies on Anorexia, Bulimia and Obesity*, 26(2), 667–677. <https://doi.org/10.1007/s40519-020-00906-7>
- Gaudio, S., Wiemerslage, L., Brooks, S. J., & Schiöth, H. B. (2016). A systematic review of resting-state functional-MRI studies in anorexia nervosa: Evidence for functional connectivity impairment in cognitive control and visuospatial and body-signal integration. *Neuroscience & Biobehavioral Reviews*, 71, 578–589. <https://doi.org/10.1016/j.neubiorev.2016.09.032>
- Gawęda, Ł., Li, E., Lavoie, S., Whitford, T. J., Moritz, S., & Nelson, B. (2018). Impaired action self-monitoring and cognitive confidence among ultra-high risk for psychosis

- and first-episode psychosis patients. *European Psychiatry*, 47, 67–75.
<https://doi.org/10.1016/j.eurpsy.2017.09.003>
- Geeraerts, B., Van Oudenhove, L., Dupont, P., Vanderghinste, D., Bormans, G., Van Laere, K., & Tack, J. (2011). Different regional brain activity during physiological gastric distension compared to balloon distension: A H215O-PET study. *Neurogastroenterology & Motility*, 23(6), 533–e203. <https://doi.org/10.1111/j.1365-2982.2010.01642.x>
- Gendall, K. A., Joyce, P. R., Carter, F. A., McIntosh, V. V., & Bulik, C. M. (2005). Childhood gastrointestinal complaints in women with bulimia nervosa. *International Journal of Eating Disorders*, 37(3), 256–260. <https://doi.org/10.1002/eat.20088>
- Gershman, S. J. (2019). *What does the free energy principle tell us about the brain?* <https://doi.org/10.48550/ARXIV.1901.07945>
- Geurts, L. S., Cooke, J. R. H., Van Bergen, R. S., & Jehee, J. F. M. (2021). *Subjective confidence reflects representation of Bayesian probability in cortex* [Preprint]. Neuroscience. <https://doi.org/10.1101/2021.04.10.439272>
- Gillan, C. M., & Seow, T. X. F. (2020). Carving Out New Transdiagnostic Dimensions for Research in Mental Health. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 5(10), 932–934. <https://doi.org/10.1016/j.bpsc.2020.04.013>
- Gillberg, I. C., Billstedt, E., Wentz, E., Anckarsäter, H., Råstam, M., & Gillberg, C. (2010). Attention, executive functions, and mentalizing in anorexia nervosa eighteen years after onset of eating disorder. *Journal of Clinical and Experimental Neuropsychology*, 32(4), 358–365. <https://doi.org/10.1080/13803390903066857>
- Glaser, D. R., Haaga, D. A. F., Hannallah, L., Field, S. E., Kozlosky, M., Reynolds, J., Yanovski, J. A., & Tanofsky-Kraff, M. (2013). Self-efficacy beliefs and eating behavior in adolescent girls at-risk for excess weight gain and binge eating disorder: Self-Efficacy Beliefs and Eating Behavior. *International Journal of Eating Disorders*, 46(7), 663–668. <https://doi.org/10.1002/eat.22160>
- Glauert, R., Rhodes, G., Byrne, S., Fink, B., & Grammer, K. (2009). Body dissatisfaction and the effects of perceptual exposure on body norms and ideals. *International Journal of Eating Disorders*, 42(5), 443–452. <https://doi.org/10.1002/eat.20640>
- Gloster, A. T., Rhoades, H. M., Novy, D., Klotsche, J., Senior, A., Kunik, M., Wilson, N., & Stanley, M. A. (2008). Psychometric properties of the Depression Anxiety and Stress

- Scale-21 in older primary care patients. *Journal of Affective Disorders*, 110(3), 248–259. <https://doi.org/10.1016/j.jad.2008.01.023>
- Goddard, E., Hibbs, R., Raenker, S., Salerno, L., Arcelus, J., Boughton, N., Connan, F., Goss, K., Laszlo, B., Morgan, J., Moore, K., Robertson, D., S, S., Schreiber-Kounine, C., Sharma, S., Whitehead, L., Schmidt, U., & Treasure, J. (2013). A multi-centre cohort study of short term outcomes of hospital treatment for anorexia nervosa in the UK. *BMC Psychiatry*, 13(1), 287. <https://doi.org/10.1186/1471-244X-13-287>
- Goessl, V. C., Curtiss, J. E., & Hofmann, S. G. (2017). The effect of heart rate variability biofeedback training on stress and anxiety: A meta-analysis. *Psychological Medicine*, 47(15), 2578–2586. <https://doi.org/10.1017/S0033291717001003>
- Goldstein, R. Z., Craig, A. D., Bechara, A., Garavan, H., Childress, A. R., Paulus, M. P., & Volkow, N. D. (2009). The Neurocircuitry of Impaired Insight in Drug Addiction. *Trends in Cognitive Sciences*, 13(9), 372–380. <https://doi.org/10.1016/j.tics.2009.06.004>
- Goldstone, A. P., Prechtl de Hernandez, C. G., Beaver, J. D., Muhammed, K., Croese, C., Bell, G., Durighel, G., Hughes, E., Waldman, A. D., Frost, G., & Bell, J. D. (2009). Fasting biases brain reward systems towards high-calorie foods. *European Journal of Neuroscience*, 30(8), 1625–1635. <https://doi.org/10.1111/j.1460-9568.2009.06949.x>
- González-Arias, M., Martínez-Molina, A., Galdames, S., & Urzúa, A. (2018). Psychometric Properties of the 20-Item Toronto Alexithymia Scale in the Chilean Population. *Frontiers in Psychology*, 9, 963. <https://doi.org/10.3389/fpsyg.2018.00963>
- Goodman, R., Ford, T., Simmons, H., Gatward, R., & Meltzer, H. (2000). Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *The British Journal of Psychiatry*, 177(6), 534–539. <https://doi.org/10.1192/bjp.177.6.534>
- Goodrick, G. K., Pendleton, V. R., Kimball, K. T., Carlos Poston, W. S., Reeves, R. S., & Foreyt, J. P. (1999). Binge eating severity, self-concept, dieting self-efficacy and social support during treatment of binge eating disorder. *International Journal of Eating Disorders*, 26(3), 295–300. [https://doi.org/10.1002/\(SICI\)1098-108X\(199911\)26:3<295::AID-EAT7>3.0.CO;2-7](https://doi.org/10.1002/(SICI)1098-108X(199911)26:3<295::AID-EAT7>3.0.CO;2-7)
- Gordon, C. M., Dougherty, D. D., Fischman, A. J., Emans, S. J., Grace, E., Lamm, R., Alpert, N. M., Majzoub, J. A., & Rauch, S. L. (2001). Neural substrates of anorexia nervosa:

- A behavioral challenge study with positron emission tomography. *The Journal of Pediatrics*, 139(1), 51–57. <https://doi.org/10.1067/mpd.2001.114768>
- Gorrell, S., & Murray, S. B. (2019). Eating Disorders in Males. *Child and Adolescent Psychiatric Clinics of North America*, 28(4), 641–651. <https://doi.org/10.1016/j.chc.2019.05.012>
- 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>
- Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*, 38(4), 404–411. <https://doi.org/10.1037/h0059831>
- Greenfeld, D. G., Anyan, W. R., Hobart, M., Quinlan, D. M., & Plantès, M. (1991). Insight into illness and outcome in anorexia nervosa. *International Journal of Eating Disorders*, 10(1), 101–109. [https://doi.org/10.1002/1098-108X\(199101\)10:1<101::AID-EAT2260100111>3.0.CO;2-7](https://doi.org/10.1002/1098-108X(199101)10:1<101::AID-EAT2260100111>3.0.CO;2-7)
- Gregertsen, E. C., Mandy, W., & Serpell, L. (2017). The Egocentric Nature of Anorexia: An Impediment to Recovery in Anorexia Nervosa Treatment. *Frontiers in Psychology*, 8. <https://www.frontiersin.org/articles/10.3389/fpsyg.2017.02273>
- Gu, X., Eilam-Stock, T., Zhou, T., Anagnostou, E., Kolevzon, A., Soorya, L., Hof, P. R., Friston, K., & Fan, J. (2015). Autonomic and brain responses associated with empathy deficits in autism spectrum disorder. *Human Brain Mapping*, 36(9), 3323–3338. <https://doi.org/10.1002/hbm.22840>
- Gu, X., & Filbey, F. (2017). A Bayesian Observer Model of Drug Craving. *JAMA Psychiatry*, 74(4), 419–420. <https://doi.org/10.1001/jamapsychiatry.2016.3823>
- Gu, X., Lohrenz, T., Salas, R., Baldwin, P. R., Soltani, A., Kirk, U., Cinciripini, P. M., & Montague, P. R. (2015). Belief about nicotine selectively modulates value and reward prediction error signals in smokers. *Proceedings of the National Academy of Sciences*, 112(8), 2539–2544. <https://doi.org/10.1073/pnas.1416639112>
- Gueguen, J., Godart, N., Chambry, J., Brun-Eberentz, A., Foulon, C., Divac, PhD, S. M., Guelfi, J., Rouillon, F., Falissard, B., & Huas, C. (2012). Severe anorexia nervosa in men: Comparison with severe AN in women and analysis of mortality. *International Journal of Eating Disorders*, 45(4), 537–545. <https://doi.org/10.1002/eat.20987>

- Haarsma, J., Fletcher, P. C., Griffin, J. D., Taverne, H. J., Ziauddeen, H., Spencer, T. J., Miller, C., Katthagen, T., Goodyer, I., Diederer, K. M. J., & Murray, G. K. (2021). Precision weighting of cortical unsigned prediction error signals benefits learning, is mediated by dopamine, and is impaired in psychosis. *Molecular Psychiatry*, 26(9), Article 9. <https://doi.org/10.1038/s41380-020-0803-8>
- Haas, S., Brock, C., Krogh, K., Gram, M., Nissen, T. D., Lundby, L., Laurberg, S., & Drewes, A. M. (2014). Cortical evoked potentials in response to rapid balloon distension of the rectum and anal canal. *Neurogastroenterology & Motility*, 26(6), 862–873. <https://doi.org/10.1111/nmo.12341>
- Haddara, N., & Rahnev, D. (2022). The Impact of Feedback on Perceptual Decision-Making and Metacognition: Reduction in Bias but No Change in Sensitivity. *Psychological Science*, 33(2), 259–275. <https://doi.org/10.1177/09567976211032887>
- Halmi, K. A., Agras, W. S., Crow, S., Mitchell, J., Wilson, G. T., Bryson, S. W., & Kraemer, H. C. (2005). Predictors of Treatment Acceptance and Completion in Anorexia Nervosa: Implications for Future Study Designs. *Archives of General Psychiatry*, 62(7), 776–781. <https://doi.org/10.1001/archpsyc.62.7.776>
- Halmi, K. A., & Sunday, S. R. (1991). Temporal patterns of hunger and fullness ratings and related cognitions in anorexia and bulimia. *Appetite*, 16(3), 219–237. [https://doi.org/10.1016/0195-6663\(91\)90060-6](https://doi.org/10.1016/0195-6663(91)90060-6)
- Halmi, K. A., Sunday, S. R., Klump, K. L., Strober, M., Leckman, J. F., Fichter, M., Kaplan, A., Woodside, B., Treasure, J., Berrettini, W. H., Al Shabboat, M., Bulik, C. M., & Kaye, W. H. (2003). Obsessions and compulsions in anorexia nervosa subtypes. *International Journal of Eating Disorders*, 33(3), 308–319. <https://doi.org/10.1002/eat.10138>
- Ham, T. E., Bonnelle, V., Hellyer, P., Jilka, S., Robertson, I. H., Leech, R., & Sharp, D. J. (2014). The neural basis of impaired self-awareness after traumatic brain injury. *Brain*, 137(2), 586–597. <https://doi.org/10.1093/brain/awt350>
- Harden, K. P., Kretsch, N., Moore, S. R., & Mendle, J. (2014). Descriptive review: Hormonal influences on risk for eating disorder symptoms during puberty and adolescence. *International Journal of Eating Disorders*, 47(7), 718–726. <https://doi.org/10.1002/eat.22317>
- Harkness, D. L., & Keshava, A. L. (2017). Moving from the What to the How and Where – Bayesian Models and Predictive ProcessingMoving from the What to the How and

- Where – Bayesian Models and Predictive Processing. *Philosophy and Predictive Processing*. <https://doi.org/10.15502/9783958573178>
- Harris, C., & Barraclough, B. (1998). Excess mortality of mental disorder. *British Journal of Psychiatry*, 173(1), 11–53. <https://doi.org/10.1192/bjp.173.1.11>
- Harrison, O. K., Garfinkel, S. N., Marlow, L., Finnegan, S. L., Marino, S., Köchli, L., Allen, M., Finnemann, J., Keur-Huizinga, L., Harrison, S. J., Stephan, K. E., Pattinson, K. T. S., & Fleming, S. M. (2021). The Filter Detection Task for measurement of breathing-related interoception and metacognition. *Biological Psychology*, 165, 108185. <https://doi.org/10.1016/j.biopsycho.2021.108185>
- Harrop, E. N., Mensinger, J. L., Moore, M., & Lindhorst, T. (2021). Restrictive eating disorders in higher weight persons: A systematic review of atypical anorexia nervosa prevalence and consecutive admission literature. *International Journal of Eating Disorders*, 54(8), 1328–1357. <https://doi.org/10.1002/eat.23519>
- Hartmann, A. S., Thomas, J. J., Wilson, A. C., & Wilhelm, S. (2013). Insight impairment in body image disorders: Delusionality and overvalued ideas in anorexia nervosa versus body dysmorphic disorder. *Psychiatry Research*, 210(3), 1129–1135. <https://doi.org/10.1016/j.psychres.2013.08.010>
- Harver, A., Katkin, E. S., & Bloch, E. (1993). Signal-detection outcomes on heartbeat and respiratory resistance detection tasks in male and female subjects. *Psychophysiology*, 30(3), 223–230. <https://doi.org/10.1111/j.1469-8986.1993.tb03347.x>
- Hassanpour, M. S., Simmons, W. K., Feinstein, J. S., Luo, Q., Lapidus, R. C., Bodurka, J., Paulus, M. P., & Khalsa, S. S. (2018). The Insular Cortex Dynamically Maps Changes in Cardiorespiratory Interoception. *Neuropsychopharmacology*, 43(2), Article 2. <https://doi.org/10.1038/npp.2017.154>
- Hatch, A., Madden, S., Kohn, M., Clarke, S., Touyz, S., & Williams, L. M. (2010). Anorexia nervosa: Towards an integrative neuroscience model. *European Eating Disorders Review*, 18(3), 165–179. <https://doi.org/10.1002/erv.974>
- Hay, P. P., Bacaltchuk, J., Stefano, S., & Kashyap, P. (2009). Psychological treatments for bulimia nervosa and bingeing. *Cochrane Database of Systematic Reviews*. <https://doi.org/10.1002/14651858.CD000562.pub3>
- Hayes, S. C., Villatte, M., Levin, M., & Hildebrandt, M. (2011). Open, Aware, and Active: Contextual Approaches as an Emerging Trend in the Behavioral and Cognitive

- Therapies. *Annual Review of Clinical Psychology*, 7(1), 141–168.
<https://doi.org/10.1146/annurev-clinpsy-032210-104449>
- Hedman, A., Breithaupt, L., Hübel, C., Thornton, L. M., Tillander, A., Norring, C., Birgegård, A., Larsson, H., Ludvigsson, J. F., Sävendahl, L., Almqvist, C., & Bulik, C. M. (2019). Bidirectional relationship between eating disorders and autoimmune diseases. *Journal of Child Psychology and Psychiatry*, 60(7), 803–812.
<https://doi.org/10.1111/jcpp.12958>
- Heereman, J., Walter, H., & Heekeren, H. R. (2015). A task-independent neural representation of subjective certainty in visual perception. *Frontiers in Human Neuroscience*, 9. <https://www.frontiersin.org/articles/10.3389/fnhum.2015.00551>
- 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>
- Heim, N., Bobou, M., Tanzer, M., Jenkinson, P., Steinert, C., & Fotopoulou, A. (2022). *Interventions for interoception in mental health disorders: A systematic review of randomised-controlled trials*. OSF Preprints. <https://doi.org/10.31219/osf.io/rgm62>
- Helverskov, J. L., Clausen, L., Mors, O., Frydenberg, M., Thomsen, P. H., & Rokkedal, K. (2010). Trans-diagnostic outcome of eating disorders: A 30-month follow-up study of 629 patients. *European Eating Disorders Review*, 18(6), 453–463.
<https://doi.org/10.1002/erv.1025>
- Herbert, B. M. (2021). Interoception and Its Role for Eating, Obesity, and Eating Disorders. *European Journal of Health Psychology*.
<https://econtent.hogrefe.com/doi/10.1027/2512-8442/a000062>
- Herbert, B. M., Herbert, C., & Pollatos, O. (2011). On the Relationship Between Interoceptive Awareness and Alexithymia: Is Interoceptive Awareness Related to Emotional Awareness?: Interoceptive Awareness and Alexithymia. *Journal of Personality*, 79(5), 1149–1175. <https://doi.org/10.1111/j.1467-6494.2011.00717.x>
- 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. (2014). Attenuated interoceptive sensitivity in overweight and obese individuals. *Eating Behaviors*, 15(3), 445–448.
<https://doi.org/10.1016/j.eatbeh.2014.06.002>
- 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>
- Herbert, B. M., Pollatos, O., & Schandry, R. (2007). Interoceptive sensitivity and emotion processing: An EEG study. *International Journal of Psychophysiology*, 65(3), 214–227. <https://doi.org/10.1016/j.ijpsycho.2007.04.007>
- Heron, K. E., & Smyth, J. M. (2010). Ecological momentary interventions: Incorporating mobile technology into psychosocial and health behaviour treatments. *British Journal of Health Psychology*, 15(1), 1–39. <https://doi.org/10.1348/135910709X466063>
- Herpertz-Dahlmann, B., Seitz, J., & Konrad, K. (2011). Aetiology of anorexia nervosa: From a “psychosomatic family model” to a neuropsychiatric disorder? *European Archives of Psychiatry and Clinical Neuroscience*, 261(2), 177. <https://doi.org/10.1007/s00406-011-0246-y>
- Heruc, G. A., Little, T. J., Kohn, M., Madden, S., Clarke, S., Horowitz, M., & Feinle-Bisset, C. (2019). Appetite Perceptions, Gastrointestinal Symptoms, Ghrelin, Peptide YY and State Anxiety Are Disturbed in Adolescent Females with Anorexia Nervosa and Only Partially Restored with Short-Term Refeeding. *Nutrients*, 11(1), Article 1.
<https://doi.org/10.3390/nu11010059>
- Hilgenstock, R., Weiss, T., & Witte, O. W. (2014). You’d better think twice: Post-decision perceptual confidence. *NeuroImage*, 99, 323–331.
<https://doi.org/10.1016/j.neuroimage.2014.05.049>
- Hill, C. C., & Pickinpaugh, J. (2008). Physiologic Changes in Pregnancy. *Surgical Clinics of North America*, 88(2), 391–401. <https://doi.org/10.1016/j.suc.2007.12.005>
- Himmerich, H., Hotopf, M., Shetty, H., Schmidt, U., Treasure, J., Hayes, R. D., Stewart, R., & Chang, C.-K. (2019). Psychiatric comorbidity as a risk factor for mortality in people with anorexia nervosa. *European Archives of Psychiatry and Clinical Neuroscience*, 269(3), 351–359. <https://doi.org/10.1007/s00406-018-0937-8>
- Himmerich, H., Schönknecht, P., Heitmann, S., & Sheldrick, A. J. (2010). Laboratory Parameters and Appetite Regulators in Patients with Anorexia Nervosa. *Journal of*

- Psychiatric Practice*, 16(2), 82–92.
<https://doi.org/10.1097/01.pra.0000369969.87779.1c>
- Hobson, A. R., Sarkar, S., Furlong, P. L., Thompson, D. G., & Aziz, Q. (2000). A cortical evoked potential study of afferents mediating human esophageal sensation. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, 279(1), G139–G147.
<https://doi.org/10.1152/ajpgi.2000.279.1.G139>
- Hoek, H. W. (2006). Incidence, prevalence and mortality of anorexia nervosa and other eating disorders. *Current Opinion in Psychiatry*, 19(4), 389.
<https://doi.org/10.1097/01.yco.0000228759.95237.78>
- Hoek, H. W., & van Hoeken, D. (2003). Review of the prevalence and incidence of eating disorders. *International Journal of Eating Disorders*, 34(4), 383–396.
<https://doi.org/10.1002/eat.10222>
- Hölzl, R., Erasmus, L.-P., & Möltner, A. (1996). Detection, discrimination and sensation of visceral stimuli. *Biological Psychology*, 42(1), 199–214. [https://doi.org/10.1016/0301-0511\(95\)05155-4](https://doi.org/10.1016/0301-0511(95)05155-4)
- Honkalampi, K., Hintikka, J., Laukkanen, E., & Viinamäki, J. L. H. (2001). Alexithymia and Depression: A Prospective Study of Patients With Major Depressive Disorder. *Psychosomatics*, 42(3), 229–234. <https://doi.org/10.1176/appi.psy.42.3.229>
- Hoogervorst, K., Banellis, L., & Allen, M. (2023). *Domain-specific updating of metacognitive self-beliefs*. PsyArXiv. <https://doi.org/10.31234/osf.io/knsb4>
- House, J., Schmidt, U., Craig, M., Landau, S., Simic, M., Nicholls, D., Hugo, P., Berelowitz, M., & Eisler, I. (2012). Comparison of specialist and nonspecialist care pathways for adolescents with anorexia nervosa and related eating disorders. *International Journal of Eating Disorders*, 45(8), 949–956. <https://doi.org/10.1002/eat.22065>
- Hoven, M., Lebreton, M., Engelmann, J. B., Denys, D., Luigjes, J., & Van Holst, R. J. (2019). Abnormalities of confidence in psychiatry: An overview and future perspectives. *Translational Psychiatry*, 9(1), 268. <https://doi.org/10.1038/s41398-019-0602-7>
- Hoven, M., Luigjes, J., Denys, D., Rouault, M., & van Holst, R. J. (2023). How do confidence and self-beliefs relate in psychopathology: A transdiagnostic approach. *Nature Mental Health*, 1(5), Article 5. <https://doi.org/10.1038/s44220-023-00062-8>

- Huang, Y., & Rao, R. P. N. (2011). Predictive coding. *WIREs Cognitive Science*, 2(5), 580–593. <https://doi.org/10.1002/wcs.142>
- Huys, Q. J. M., Browning, M., Paulus, M. P., & Frank, M. J. (2021). Advances in the computational understanding of mental illness. *Neuropsychopharmacology*, 46(1), 3–19. <https://doi.org/10.1038/s41386-020-0746-4>
- Huys, Q. J. M., Maia, T. V., & Frank, M. J. (2016). Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature Neuroscience*, 19(3), Article 3. <https://doi.org/10.1038/nn.4238>
- Iglesias, S., Mathys, C., Brodersen, K. H., Kasper, L., Piccirelli, M., den Ouden, H. E. M., & Stephan, K. E. (2013). Hierarchical Prediction Errors in Midbrain and Basal Forebrain during Sensory Learning. *Neuron*, 80(2), 519–530. <https://doi.org/10.1016/j.neuron.2013.09.009>
- Imperator, C., Mancini, M., Della Marca, G., Valenti, E. M., & Farina, B. (2018). Feedback-Based Treatments for Eating Disorders and Related Symptoms: A Systematic Review of the Literature. *Nutrients*, 10(11), Article 11. <https://doi.org/10.3390/nu10111806>
- Irwin, M. R., & Cole, S. W. (2011). Reciprocal regulation of the neural and innate immune systems. *Nature Reviews Immunology*, 11(9), Article 9. <https://doi.org/10.1038/nri3042>
- Jacobi, C., Hayward, C., De Zwaan, M., Kraemer, H. C., & Agras, W. S. (2004). Coming to Terms With Risk Factors for Eating Disorders: Application of Risk Terminology and Suggestions for a General Taxonomy. *Psychological Bulletin*, 130(1), 19–65. <https://doi.org/10.1037/0033-2909.130.1.19>
- Jacobi, C., Paul, T., De Zwaan, M., Nutzinger, D. O., & Dahme, B. (2004). Specificity of self-concept disturbances in eating disorders. *International Journal of Eating Disorders*, 35(2), 204–210. <https://doi.org/10.1002/eat.10240>
- Jacquemot, A. M. M. C., & Park, R. (2020). The Role of Interoception in the Pathogenesis and Treatment of Anorexia Nervosa: A Narrative Review. *Frontiers in Psychiatry*, 11. <https://www.frontiersin.org/articles/10.3389/fpsyt.2020.00281>
- Jenkinson, P. M., Koukoutsakis, A., Panagiotopoulou, E., Vagnoni, E., Demartini, B., Nisticò, V., Gambini, O., Christakou, A., & Fotopoulou, A. (2020). *Body appearance values modulate risk aversion in eating restriction* [Preprint]. PsyArXiv. <https://doi.org/10.31234/osf.io/rnzgs>

- Jenkinson, P. M., Koukoutsakis, A., Panagiotopoulou, E., Vagnoni, E., Demartini, B., Nistico, V., Gambini, O., Christakou, A., & Fotopoulou, A. (2023). Body appearance values modulate risk aversion in eating restriction. *Journal of Experimental Psychology: General*. <https://doi.org/10.1037/xge0001445>
- Jenkinson, P. M., Taylor, L., & Laws, K. R. (2018). Self-reported interoceptive deficits in eating disorders: A meta-analysis of studies using the eating disorder inventory. *Journal of Psychosomatic Research*, 110, 38–45. <https://doi.org/10.1016/j.jpsychores.2018.04.005>
- Johnson, P. M., & Kenny, P. J. (2010). Dopamine D2 receptors in addiction-like reward dysfunction and compulsive eating in obese rats. *Nature Neuroscience*, 13(5), Article 5. <https://doi.org/10.1038/nn.2519>
- Karpowicz, E., Skärsäter, I., & Nevenon, L. (2009). Self-esteem in patients treated for anorexia nervosa. *International Journal of Mental Health Nursing*, 18(5), 318–325. <https://doi.org/10.1111/j.1447-0349.2009.00621.x>
- Kass, R. E., & Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association*, 90(430), 773–795. <https://doi.org/10.1080/01621459.1995.10476572>
- Kästner, D., Löwe, B., & Gumz, A. (2019). The role of self-esteem in the treatment of patients with anorexia nervosa – A systematic review and meta-analysis. *International Journal of Eating Disorders*, 52(2), 101–116. <https://doi.org/10.1002/eat.22975>
- Katkin, E. S., Morell, M. A., Goldband, S., Bernstein, G. L., & Wise, J. A. (1982). Individual Differences in Heartbeat Discrimination. *Psychophysiology*, 19(2), 160–166. <https://doi.org/10.1111/j.1469-8986.1982.tb02538.x>
- Katkin, E. S., Wiens, S., & Öhman, A. (2001). Nonconscious Fear Conditioning, Visceral Perception, and the Development of Gut Feelings. *Psychological Science*, 12(5), 366–370. <https://doi.org/10.1111/1467-9280.00368>
- Katyal, S., Huys, Q., Dolan, R., & Fleming, S. M. (2023). *How underconfidence is maintained in anxiety and depression*. PsyArXiv. <https://doi.org/10.31234/osf.io/qcg92>
- Kaye, W. H., Fudge, J. L., & Paulus, M. (2009). New insights into symptoms and neurocircuit function of anorexia nervosa. *Nature Reviews Neuroscience*, 10(8), Article 8. <https://doi.org/10.1038/nrn2682>

- Kaye, W. H., Wierenga, C. E., Bailer, U. F., Simmons, A. N., & Bischoff-Grethe, A. (2013). Nothing tastes as good as skinny feels: The neurobiology of anorexia nervosa. *Trends in Neurosciences*, 36(2), 110–120. <https://doi.org/10.1016/j.tins.2013.01.003>
- Kaye, W. H., Wierenga, C. E., Bischoff-Grethe, A., Berner, L. A., Ely, A. V., Bailer, U. F., Paulus, M. P., & Fudge, J. L. (2020). Neural Insensitivity to the Effects of Hunger in Women Remitted From Anorexia Nervosa. *American Journal of Psychiatry*, 177(7), 601–610. <https://doi.org/10.1176/appi.ajp.2019.19030261>
- Keery, H., van den Berg, P., & Thompson, J. K. (2004). An evaluation of the Tripartite Influence Model of body dissatisfaction and eating disturbance with adolescent girls. *Body Image*, 1(3), 237–251. <https://doi.org/10.1016/j.bodyim.2004.03.001>
- Kelley, A. E. (2004). Ventral striatal control of appetitive motivation: Role in ingestive behavior and reward-related learning. *Neuroscience & Biobehavioral Reviews*, 27(8), 765–776. <https://doi.org/10.1016/j.neubiorev.2003.11.015>
- Kelley, A. E., Baldo, B. A., Pratt, W. E., & Will, M. J. (2005). Corticostriatal-hypothalamic circuitry and food motivation: Integration of energy, action and reward. *Physiology & Behavior*, 86(5), 773–795. <https://doi.org/10.1016/j.physbeh.2005.08.066>
- Kennerley, S. W., Walton, M. E., Behrens, T. E. J., Buckley, M. J., & Rushworth, M. F. S. (2006). Optimal decision making and the anterior cingulate cortex. *Nature Neuroscience*, 9(7), Article 7. <https://doi.org/10.1038/nn1724>
- Kepecs, A., Uchida, N., Zariwala, H. A., & Mainen, Z. F. (2008). Neural correlates, computation and behavioural impact of decision confidence. *Nature*, 455(7210), Article 7210. <https://doi.org/10.1038/nature07200>
- Kerr, K. L., DeVille, D. C., Moseman, S. E., & Simmons, W. K. (2022). Aberrant gastric functioning in weight-restored but not acute restricting-type anorexia nervosa: An electrogastrography study. *International Journal of Eating Disorders*, 55(10), 1384–1389. <https://doi.org/10.1002/eat.23802>
- Kerr, K. L., Moseman, S. E., Avery, J. A., Bodurka, J., & Simmons, W. K. (2017). Influence of visceral interoceptive experience on the brain's response to food images in anorexia nervosa. *Psychosomatic Medicine*, 79(7), 777–784. <https://doi.org/10.1097/PSY.0000000000000486>
- 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), Article 2.
<https://doi.org/10.1038/npp.2015.174>
- Keshen, A., Helson, T., Town, J., & Warren, K. (2017). Self-efficacy as a predictor of treatment outcome in an outpatient eating disorder program. *Eating Disorders*, 25(5), 406–419. <https://doi.org/10.1080/10640266.2017.1324073>
- Keski-Rahkonen, A., Hoek, H. W., Susser, E. S., Linna, M. S., Sihvola, E., Raevuori, A., Bulik, C. M., Kaprio, J., & Rissanen, A. (2007). Epidemiology and Course of Anorexia Nervosa in the Community. *American Journal of Psychiatry*, 164(8), 1259–1265. <https://doi.org/10.1176/appi.ajp.2007.06081388>
- Kessler, H., Schwarze, M., Filipic, S., Traue, H. C., & Von Wietersheim, J. (2006). Alexithymia and facial emotion recognition in patients with eating disorders. *International Journal of Eating Disorders*, 39(3), 245–251.
<https://doi.org/10.1002/eat.20228>
- Kever, A., Pollatos, O., Vermeulen, N., & Grynberg, D. (2015). Interoceptive sensitivity facilitates both antecedent- and response-focused emotion regulation strategies. *Personality and Individual Differences*, 87, 20–23.
<https://doi.org/10.1016/j.paid.2015.07.014>
- 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., Berner, L. A., & Anderson, L. M. (2022). Gastrointestinal Interoception in Eating Disorders: Charting a New Path. *Current Psychiatry Reports*, 24(1), 47–60.
<https://doi.org/10.1007/s11920-022-01318-3>
- 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., Hassanpour, M. S., Strober, M., Craske, M. G., Arevian, A. C., & Feusner, J. D. (2018). Interoceptive Anxiety and Body Representation in Anorexia Nervosa.

Frontiers in Psychiatry, 9.

<https://www.frontiersin.org/articles/10.3389/fpsy.2018.00444>

- Khalsa, S. S., Portnoff, L. C., McCurdy-McKinnon, D., & Feusner, J. D. (2017). What happens after treatment? A systematic review of relapse, remission, and recovery in anorexia nervosa. *Journal of Eating Disorders*, 5(1), 20.
<https://doi.org/10.1186/s40337-017-0145-3>
- Khalsa, S. S., Rudrauf, D., Damasio, A. R., Davidson, R. J., Lutz, A., & Tranel, D. (2008). Interoceptive awareness in experienced meditators. *Psychophysiology*, 45(4), 671–677. <https://doi.org/10.1111/j.1469-8986.2008.00666.x>
- Khalsa, S. S., Rudrauf, D., Davidson, R., & Tranel, D. (2015). The effect of meditation on regulation of internal body states. *Frontiers in Psychology*, 6.
<https://www.frontiersin.org/articles/10.3389/fpsyg.2015.00924>
- Khalsa, S. S., Rudrauf, D., Sandesara, C., Olshansky, B., & Tranel, D. (2009). Bolus isoproterenol infusions provide a reliable method for assessing interoceptive awareness. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, 72(1), 34–45.
<https://doi.org/10.1016/j.ijpsycho.2008.08.010>
- Kiani, R., & Shadlen, M. N. (2009). Representation of Confidence Associated with a Decision by Neurons in the Parietal Cortex. *Science*, 324(5928), 759–764.
<https://doi.org/10.1126/science.1169405>
- Kim, M., Kim, S., Lee, K.-U., & Jeong, B. (2020). Pessimistically biased perception in panic disorder during risk learning. *Depression and Anxiety*, 37(7), 609–619.
<https://doi.org/10.1002/da.23007>
- Kim, Y.-R., Kim, J.-H., Kim, M. J., & Treasure, J. (2014). Differential Methylation of the Oxytocin Receptor Gene in Patients with Anorexia Nervosa: A Pilot Study. *PLoS ONE*, 9(2), e88673. <https://doi.org/10.1371/journal.pone.0088673>
- Kinnaird, E., Stewart, C., & Tchanturia, K. (2020). Interoception in Anorexia Nervosa: Exploring Associations With Alexithymia and Autistic Traits. *Frontiers in Psychiatry*, 11. <https://www.frontiersin.org/articles/10.3389/fpsy.2020.00064>
- Kirsch, L. P., Mathys, C., Papadaki, C., Talelli, P., Friston, K., Moro, V., & Fotopoulou, A. (2021). Updating beliefs beyond the here-and-now: The counter-factual self in anosognosia for hemiplegia. *Brain Communications*, 3(2), fcab098.
<https://doi.org/10.1093/braincomms/fcab098>

- Kjelsås, E., Bjørnstrøm, C., & Gøtestam, K. G. (2004). Prevalence of eating disorders in female and male adolescents (14–15 years). *Eating Behaviors*, 5(1), 13–25.
[https://doi.org/10.1016/S1471-0153\(03\)00057-6](https://doi.org/10.1016/S1471-0153(03)00057-6)
- Klein, S. B., Altinyazar, V., & Metz, M. A. (2013). Facets of Self in Schizophrenia: The Reliability and Accuracy of Trait Self-Knowledge. *Clinical Psychological Science*, 1(3), 276–289. <https://doi.org/10.1177/2167702612474263>
- 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>
- Knill, D. C., & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences*, 27(12), 712–719.
<https://doi.org/10.1016/j.tins.2004.10.007>
- Kojima, S., Nagai, N., Nakabeppu, Y., Muranaga, T., Deguchi, D., Nakajo, M., Masuda, A., Nozoe, S., & Naruo, T. (2005). Comparison of regional cerebral blood flow in patients with anorexia nervosa before and after weight gain. *Psychiatry Research: Neuroimaging*, 140(3), 251–258. <https://doi.org/10.1016/j.psychresns.2005.08.002>
- Konstantakopoulos, G., Georgantopoulos, G., Gonidakis, F., Michopoulos, I., Stefanatou, P., & David, A. S. (2020). Development and validation of the schedule for the assessment of insight in eating disorders (SAI-ED). *Psychiatry Research*, 292, 113308. <https://doi.org/10.1016/j.psychres.2020.113308>
- Konstantakopoulos, G., Ioannidi, N., Patrikelis, P., & Gonidakis, F. (2020). The impact of theory of mind and neurocognition on delusionality in anorexia nervosa. *Journal of Clinical and Experimental Neuropsychology*, 42(6), 611–621.
<https://doi.org/10.1080/13803395.2020.1786504>
- Konstantakopoulos, G., Tchanturia, K., Surguladze, S. A., & David, A. S. (2011). Insight in eating disorders: Clinical and cognitive correlates. *Psychological Medicine*, 41(9), 1951–1961. <https://doi.org/10.1017/S0033291710002539>
- Konstantakopoulos, G., Varsou, E., Dikeos, D., Ioannidi, N., Gonidakis, F., Papadimitriou, G., & Oulis, P. (2012). Delusionality of body image beliefs in eating disorders. *Psychiatry Research*, 200(2), 482–488.
<https://doi.org/10.1016/j.psychres.2012.03.023>

- Kopp, B., Lange, F., & Steinke, A. (2021). The Reliability of the Wisconsin Card Sorting Test in Clinical Practice. *Assessment*, 28(1), 248–263.
<https://doi.org/10.1177/1073191119866257>
- Kordy, H., Krämer, B., Palmer, R. L., Papezova, H., Pellet, J., Richard, M., Treasure, J., & Cost Action B6. (2002). Remission, recovery, relapse, and recurrence in eating disorders: Conceptualization and illustration of a validation strategy. *Journal of Clinical Psychology*, 58(7), 833–846. <https://doi.org/10.1002/jclp.2013>
- Koreki, A., Funayama, M., Terasawa, Y., Onaya, M., & Mimura, M. (2021). Aberrant Interoceptive Accuracy in Patients With Schizophrenia Performing a Heartbeat Counting Task. *Schizophrenia Bulletin Open*, 2(1), sgaa067.
<https://doi.org/10.1093/schizbullopen/sgaa067>
- Korn, C. W., Sharot, T., Walter, H., Heekeren, H. R., & Dolan, R. J. (2014). Depression is related to an absence of optimistically biased belief updating about future life events. *Psychological Medicine*, 44(3), 579–592.
<https://doi.org/10.1017/S0033291713001074>
- Kringelbach, M. L., & Rolls, E. T. (2004). The functional neuroanatomy of the human orbitofrontal cortex: Evidence from neuroimaging and neuropsychology. *Progress in Neurobiology*, 72(5), 341–372. <https://doi.org/10.1016/j.pneurobio.2004.03.006>
- Kube, T., Kirchner, L., Lemmer, G., & Glombiewski, J. A. (2022). How the Discrepancy Between Prior Expectations and New Information Influences Expectation Updating in Depression—The Greater, the Better? *Clinical Psychological Science*, 10(3), 430–449. <https://doi.org/10.1177/21677026211024644>
- Kube, T., Kirchner, L., Rief, W., Gärtner, T., & Glombiewski, J. A. (2019). Belief updating in depression is not related to increased sensitivity to unexpectedly negative information. *Behaviour Research and Therapy*, 123, 103509.
<https://doi.org/10.1016/j.brat.2019.103509>
- Kube, T., Rief, W., Gollwitzer, M., Gärtner, T., & Glombiewski, J. A. (2019). Why dysfunctional expectations in depression persist – Results from two experimental studies investigating cognitive immunization. *Psychological Medicine*, 49(9), 1532–1544. <https://doi.org/10.1017/S0033291718002106>
- Kube, T., & Rozenkrantz, L. (2021). When Beliefs Face Reality: An Integrative Review of Belief Updating in Mental Health and Illness. *Perspectives on Psychological Science*, 16(2), 247–274. <https://doi.org/10.1177/1745691620931496>

- Kucharska-Pietura, K., Nikolaou, V., Masiak, M., & Treasure, J. (2004). The recognition of emotion in the faces and voice of anorexia nervosa. *International Journal of Eating Disorders*, 35(1), 42–47. <https://doi.org/10.1002/eat.10219>
- Kurth, F., Zilles, K., Fox, P. T., Laird, A. R., & Eickhoff, S. B. (2010). A link between the systems: Functional differentiation and integration within the human insula revealed by meta-analysis. *Brain Structure and Function*, 214(5), 519–534. <https://doi.org/10.1007/s00429-010-0255-z>
- Lang, K., Lopez, C., Stahl, D., Tchanturia, K., & Treasure, J. (2014). Central coherence in eating disorders: An updated systematic review and meta-analysis. *The World Journal of Biological Psychiatry*, 15(8), 586–598. <https://doi.org/10.3109/15622975.2014.909606>
- Lang, K., Stahl, D., Espie, J., Treasure, J., & Tchanturia, K. (2014). Set shifting in children and adolescents with anorexia nervosa: An exploratory systematic review and meta-analysis. *International Journal of Eating Disorders*, 47(4), 394–399. <https://doi.org/10.1002/eat.22235>
- Lapidus, R. C., Puhl, M., Kuplicki, R., Stewart, J. L., Paulus, M. P., Rhudy, J. L., Feinstein, J. S., Khalsa, S. S., & on behalf of the Tulsa 1000 Investigators. (2020). Heightened affective response to perturbation of respiratory but not pain signals in eating, mood, and anxiety disorders. *PLOS ONE*, 15(7), e0235346. <https://doi.org/10.1371/journal.pone.0235346>
- Lask, B., Bryant-Waugh, R., Wright, F., Campbell, M., Willoughby, K., & Waller, G. (2005). Family physician consultation patterns indicate high risk for early-onset anorexia nervosa. *International Journal of Eating Disorders*, 38(3), 269–272. <https://doi.org/10.1002/eat.20163>
- Lavalley, C. A., Hakimi, N., Taylor, S., Kuplicki, R., Forthman, K. L., Stewart, J. L., Paulus, M. P., Khalsa, S. S., & Smith, R. (2023). *Transdiagnostic failure to adapt interoceptive precision estimates across affective, substance use, and eating disorders: A replication study* [Preprint]. *Psychiatry and Clinical Psychology*. <https://doi.org/10.1101/2023.10.11.23296870>
- Lavender, J. M., Wonderlich, S. A., Engel, S. G., Gordon, K. H., Kaye, W. H., & Mitchell, J. E. (2015). Dimensions of emotion dysregulation in anorexia nervosa and bulimia nervosa: A conceptual review of the empirical literature. *Clinical Psychology Review*, 40, 111–122. <https://doi.org/10.1016/j.cpr.2015.05.010>

- Lebow, J., Sim, L. A., & Accurso, E. C. (2018). Is there clinical consensus in defining weight restoration for adolescents with anorexia nervosa? *Eating Disorders*, 26(3), 270–277. <https://doi.org/10.1080/10640266.2017.1388664>
- Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian Cognitive Modeling: A Practical Course*. Cambridge University Press.
- Lefebvre, G., Lebreton, M., Meyniel, F., Bourgeois-Gironde, S., & Palminteri, S. (2017). Behavioural and neural characterization of optimistic reinforcement learning. *Nature Human Behaviour*, 1(4), Article 4. <https://doi.org/10.1038/s41562-017-0067>
- 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>
- Lewis, Y. D., Gilon Mann, T., Enoch-Levy, A., Dubnov-Raz, G., Gothelf, D., Weizman, A., & Stein, D. (2019). Obsessive–compulsive symptomatology in female adolescent inpatients with restrictive compared with binge–purge eating disorders. *European Eating Disorders Review*, 27(3), 224–235. <https://doi.org/10.1002/erv.2638>
- Lilenfeld, L. R. R., Wonderlich, S., Riso, L. P., Crosby, R., & Mitchell, J. (2006). Eating disorders and personality: A methodological and empirical review. *Clinical Psychology Review*, 26(3), 299–320. <https://doi.org/10.1016/j.cpr.2005.10.003>
- Lin, J. A., Woods, E. R., & Bern, E. M. (2021). Common and Emergent Oral and Gastrointestinal Manifestations of Eating Disorders. *Gastroenterology & Hepatology*, 17(4), 157–167.
- Linardon, J., Fairburn, C. G., Fitzsimmons-Craft, E. E., Wilfley, D. E., & Brennan, L. (2017). The empirical status of the third-wave behaviour therapies for the treatment of eating disorders: A systematic review. *Clinical Psychology Review*, 58, 125–140. <https://doi.org/10.1016/j.cpr.2017.10.005>
- Lipsman, N., Woodside, D. B., & Lozano, A. M. (2015). Neurocircuitry of limbic dysfunction in anorexia nervosa. *Cortex*, 62, 109–118. <https://doi.org/10.1016/j.cortex.2014.02.020>
- 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](https://doi.org/10.1016/0005-7967(94)00075-U)

- Löwe, B., Zipfel, S., Buchholz, C., Dupont, Y., Reas, D. L., & Herzog, W. (2001). Long-term outcome of anorexia nervosa in a prospective 21-year follow-up study. *Psychological Medicine*, 31(5), 881–890. <https://doi.org/10.1017/S003329170100407X>
- Lund, A. E., Correa, C., Fardo, F., Fleming, S. M., & Allen, M. (2023). *Domain Generality in Metacognitive Ability: A Confirmatory Study Across Visual Perception, Memory, and General Knowledge*. PsyArXiv. <https://doi.org/10.31234/osf.io/p97s5>
- Lutz, A. P. C., Schulz, A., Voderholzer, U., Koch, S., van Dyck, Z., & Vögele, C. (2019). Enhanced cortical processing of cardio-afferent signals in anorexia nervosa. *Clinical Neurophysiology*, 130(9), 1620–1627. <https://doi.org/10.1016/j.clinph.2019.06.009>
- Lysaker, P. H., Carcione, A., Dimaggio, G., Johannesen, J. K., Nicolò, G., Procacci, M., & Semerari, A. (2005). Metacognition amidst narratives of self and illness in schizophrenia: Associations with neurocognition, symptoms, insight and quality of life. *Acta Psychiatrica Scandinavica*, 112(1), 64–71. <https://doi.org/10.1111/j.1600-0447.2005.00514.x>
- Lysaker, P. H., Chernov, N., Moiseeva, T., Sozinova, M., Dmitryeva, N., Alyoshin, V., Luther, L., Karpenko, O., & Kostyuk, G. (2021). Clinical insight, cognitive insight and metacognition in psychosis: Evidence of mediation. *Journal of Psychiatric Research*, 140, 1–6. <https://doi.org/10.1016/j.jpsychires.2021.05.030>
- Lysaker, P. H., Gagen, E., Wright, A., Vohs, J. L., Kukla, M., Yanos, P. T., & Hasson-Ohayon, I. (2019). Metacognitive Deficits Predict Impaired Insight in Schizophrenia Across Symptom Profiles: A Latent Class Analysis. *Schizophrenia Bulletin*, 45(1), 48–56. <https://doi.org/10.1093/schbul/sby142>
- Lysaker, P. H., Vohs, J., Minor, K. S., Irrarázaval, L., Leonhardt, B., Hamm, J. A., Kukla, M., Popolo, R., Luther, L., Buck, K. D., Wasmuth, S., & Dimaggio, G. (2015). Metacognitive Deficits in Schizophrenia: Presence and Associations With Psychosocial Outcomes. *Journal of Nervous & Mental Disease*, 203(7), 530–536. <https://doi.org/10.1097/NMD.0000000000000323>
- Machado, P. P. P., Grilo, C. M., & Crosby, R. D. (2017). Evaluation of the DSM-5 Severity Indicator for Anorexia Nervosa. *European Eating Disorders Review*, 25(3), 221–223. <https://doi.org/10.1002/erv.2508>
- MacNeil, L., Esposito-Smythers, C., Mehlenbeck, R., & Weismore, J. (2012). The effects of avoidance coping and coping self-efficacy on eating disorder attitudes and behaviors:

- A stress-diathesis model. *Eating Behaviors*, 13(4), 293–296.
<https://doi.org/10.1016/j.eatbeh.2012.06.005>
- Maguire, S., O'Dell, A., Touyz, L., & Russell, J. (2013). Oxytocin and Anorexia Nervosa: A Review of the Emerging Literature. *European Eating Disorders Review*, 21(6), 475–478. <https://doi.org/10.1002/erv.2252>
- Mainz, V., Schulte-Rüther, M., Fink, G. R., Herpertz-Dahlmann, B., & Konrad, K. (2012). Structural Brain Abnormalities in Adolescent Anorexia Nervosa Before and After Weight Recovery and Associated Hormonal Changes. *Psychosomatic Medicine*, 74(6), 574–582. <https://doi.org/10.1097/PSY.0b013e31824ef10e>
- Maniscalco, B., & Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and Cognition*, 21(1), 422–430. <https://doi.org/10.1016/j.concog.2011.09.021>
- Maniscalco, B., McCurdy, L. Y., Odegaard, B., & Lau, H. (2017). Limited Cognitive Resources Explain a Trade-Off between Perceptual and Metacognitive Vigilance. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 37(5), 1213–1224. <https://doi.org/10.1523/JNEUROSCI.2271-13.2016>
- Maniscalco, J. W., & Rinaman, L. (2018). Vagal Interoceptive Modulation of Motivated Behavior. *Physiology*, 33(2), 151–167. <https://doi.org/10.1152/physiol.00036.2017>
- Mårild, K., Størdal, K., Bulik, C. M., Rewers, M., Ekbom, A., Liu, E., & Ludvigsson, J. F. (2017). Celiac Disease and Anorexia Nervosa: A Nationwide Study. *Pediatrics*, 139(5), e20164367. <https://doi.org/10.1542/peds.2016-4367>
- Martin, E., Dourish, C. T., Rotshtein, P., Spetter, M. S., & Higgs, S. (2019). Interoception and disordered eating: A systematic review. *Neuroscience & Biobehavioral Reviews*, 107, 166–191. <https://doi.org/10.1016/j.neubiorev.2019.08.020>
- Martini, M., Marzola, E., Brustolin, A., & Abbate-Daga, G. (2021). Feeling imperfect and imperfectly feeling: A network analysis on perfectionism, interoceptive sensibility, and eating symptomatology in anorexia nervosa. *European Eating Disorders Review*, 29(6), 893–909. <https://doi.org/10.1002/erv.2863>
- Mascolo, M., Geer, B., Feuerstein, J., & Mehler, P. S. (2017). Gastrointestinal comorbidities which complicate the treatment of anorexia nervosa. *Eating Disorders*, 25(2), 122–133. <https://doi.org/10.1080/10640266.2016.1255108>

- Mathys, C. D., Lomakina, E. I., Daunizeau, J., Iglesias, S., Brodersen, K. H., Friston, K., & Stephan, K. E. (2014). Uncertainty in perception and the Hierarchical Gaussian Filter. *Frontiers in Human Neuroscience*, 8. <https://www.frontiersin.org/articles/10.3389/fnhum.2014.00825>
- Mathys, C., Daunizeau, J., Friston, K., & Stephan, K. (2011). A Bayesian Foundation for Individual Learning Under Uncertainty. *Frontiers in Human Neuroscience*, 5. <https://www.frontiersin.org/articles/10.3389/fnhum.2011.00039>
- 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), Article 1. <https://doi.org/10.1038/s41467-023-39058-4>
- Mayeli, A., Al Zoubi, O., White, E. J., Chappelle, S., Kuplicki, R., Smith, R., Feinstein, J. S., Bodurka, J., Paulus, M. P., & Khalsa, S. S. (2021). *Neural indicators of human gut feelings* [Preprint]. Neuroscience. <https://doi.org/10.1101/2021.02.11.430867>
- Mayer, E. A., Naliboff, B. D., & Craig, A. D. B. (2006). Neuroimaging of the Brain-Gut Axis: From Basic Understanding to Treatment of Functional GI Disorders. *Gastroenterology*, 131(6), 1925–1942. <https://doi.org/10.1053/j.gastro.2006.10.026>
- Mazancieux, A., Fleming, S. M., Souchay, C., & Moulin, C. J. A. (2020). Is there a G factor for metacognition? Correlations in retrospective metacognitive sensitivity across tasks. *Journal of Experimental Psychology: General*, 149(9), 1788–1799. <https://doi.org/10.1037/xge0000746>
- McCurdy, L. Y., Maniscalco, B., Metcalfe, J., Liu, K. Y., de Lange, F. P., & Lau, H. (2013). Anatomical coupling between distinct metacognitive systems for memory and visual perception. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 33(5), 1897–1906. <https://doi.org/10.1523/JNEUROSCI.1890-12.2013>
- McDermott, C. J., & Rushford, N. (2011). Dysfunctional metacognitions in anorexia nervosa. *Eating and Weight Disorders - Studies on Anorexia, Bulimia and Obesity*, 16(1), e49–e55. <https://doi.org/10.1007/BF03327521>
- McEwen, B. S., & Wingfield, J. C. (2003). The concept of allostasis in biology and biomedicine. *Hormones and Behavior*, 43(1), 2–15. [https://doi.org/10.1016/S0018-506X\(02\)00024-7](https://doi.org/10.1016/S0018-506X(02)00024-7)
- McNally, R. J. (2016). Can network analysis transform psychopathology? *Behaviour Research and Therapy*, 86, 95–104. <https://doi.org/10.1016/j.brat.2016.06.006>

- 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>
- Mercado-Perez, A., & Beyder, A. (2022). Gut feelings: Mechanosensing in the gastrointestinal tract. *Nature Reviews. Gastroenterology & Hepatology*, 19(5), 283–296. <https://doi.org/10.1038/s41575-021-00561-y>
- Meule, A., Freund, R., Skirde, A. K., Vögele, C., & Kübler, A. (2012). Heart Rate Variability Biofeedback Reduces Food Cravings in High Food Cravers. *Applied Psychophysiology and Biofeedback*, 37(4), 241–251. <https://doi.org/10.1007/s10484-012-9197-y>
- Meyer, R., Kröner-Herwig, B., & Spörkel, H. (1990). The effect of exercise and induced expectations on visceral perception in asthmatic patients. *Journal of Psychosomatic Research*, 34(4), 455–460. [https://doi.org/10.1016/0022-3999\(90\)90069-G](https://doi.org/10.1016/0022-3999(90)90069-G)
- Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the penn state worry questionnaire. *Behaviour Research and Therapy*, 28(6), 487–495. [https://doi.org/10.1016/0005-7967\(90\)90135-6](https://doi.org/10.1016/0005-7967(90)90135-6)
- Micali, N., Hagberg, K. W., Petersen, I., & Treasure, J. L. (2013). The incidence of eating disorders in the UK in 2000–2009: Findings from the General Practice Research Database. *BMJ Open*, 3(5), e002646. <https://doi.org/10.1136/bmjopen-2013-002646>
- Miles, S., Gnat, I., Phillipou, A., & Nedeljkovic, M. (2020). Cognitive flexibility in acute anorexia nervosa and after recovery: A systematic review. *Clinical Psychology Review*, 81, 101905. <https://doi.org/10.1016/j.cpr.2020.101905>
- Miles, S., Phillipou, A., Sumner, P., & Nedeljkovic, M. (2022). Cognitive flexibility and the risk of anorexia nervosa: An investigation using self-report and neurocognitive assessments. *Journal of Psychiatric Research*, 151, 531–538. <https://doi.org/10.1016/j.jpsychires.2022.05.043>
- Milos, G., Kuenzli, C., Soelch, C. M., Schumacher, S., Moergeli, H., & Mueller-Pfeiffer, C. (2013). How much should I eat? Estimation of meal portions in anorexia nervosa. *Appetite*, 63, 42–47. <https://doi.org/10.1016/j.appet.2012.12.016>

- Mintz, A. R., Dobson, K. S., & Romney, D. M. (2003). Insight in schizophrenia: A meta-analysis. *Schizophrenia Research*, 61(1), 75–88. [https://doi.org/10.1016/S0920-9964\(02\)00316-X](https://doi.org/10.1016/S0920-9964(02)00316-X)
- Miskovic-Wheatley, J., Bryant, E., Ong, S. H., Vatter, S., Le, A., Aouad, P., Barakat, S., Boakes, R., Brennan, L., Bryant, E., Byrne, S., Caldwell, B., Calvert, S., Carroll, B., Castle, D., Caterson, I., Chelius, B., Chiem, L., Clarke, S., ... National Eating Disorder Research Consortium. (2023). Eating disorder outcomes: Findings from a rapid review of over a decade of research. *Journal of Eating Disorders*, 11(1), 85. <https://doi.org/10.1186/s40337-023-00801-3>
- Misra, M., Aggarwal, A., Miller, K. K., Almazan, C., Worley, M., Soyka, L. A., Herzog, D. B., & Klibanski, A. (2004). Effects of Anorexia Nervosa on Clinical, Hematologic, Biochemical, and Bone Density Parameters in Community-Dwelling Adolescent Girls. *Pediatrics*, 114(6), 1574–1583. <https://doi.org/10.1542/peds.2004-0540>
- Misra, M., & Klibanski, A. (2014). Endocrine consequences of anorexia nervosa. *The Lancet Diabetes & Endocrinology*, 2(7), 581–592. [https://doi.org/10.1016/S2213-8587\(13\)70180-3](https://doi.org/10.1016/S2213-8587(13)70180-3)
- Moeller, S. J., Fleming, S. M., Gan, G., Zilverstand, A., Malaker, P., d'Oleire Uquillas, F., Schneider, K. E., Preston-Campbell, R. N., Parvaz, M. A., Maloney, T., Alia-Klein, N., & Goldstein, R. Z. (2016). Metacognitive impairment in active cocaine use disorder is associated with individual differences in brain structure. *European Neuropsychopharmacology*, 26(4), 653–662. <https://doi.org/10.1016/j.euroneuro.2016.02.009>
- Monteleone, A. M., Castellini, G., Volpe, U., Ricca, V., Lelli, L., Monteleone, P., & Maj, M. (2018). Neuroendocrinology and brain imaging of reward in eating disorders: A possible key to the treatment of anorexia nervosa and bulimia nervosa. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 80, 132–142. <https://doi.org/10.1016/j.pnpbp.2017.02.020>
- Monteleone, A. M., Monteleone, P., Dalle Grave, R., Nigro, M., El Ghoch, M., Calugi, S., Cimino, M., & Maj, M. (2016). Ghrelin response to hedonic eating in underweight and short-term weight restored patients with anorexia nervosa. *Psychiatry Research*, 235, 55–60. <https://doi.org/10.1016/j.psychres.2015.12.001>

- Monteleone, P., Castaldo, E., & Maj, M. (2008). Neuroendocrine dysregulation of food intake in eating disorders. *Regulatory Peptides*, 149(1–3), 39–50.
<https://doi.org/10.1016/j.regpep.2007.10.007>
- Morales, J., Lau, H., & Fleming, S. M. (2018). Domain-General and Domain-Specific Patterns of Activity Supporting Metacognition in Human Prefrontal Cortex. *Journal of Neuroscience*, 38(14), 3534–3546. <https://doi.org/10.1523/JNEUROSCI.2360-17.2018>
- Moran, P., Leese, M., Lee, T., Walters, P., Thornicroft, G., & Mann, A. (2003). Standardised Assessment of Personality – Abbreviated Scale (SAPAS): Preliminary validation of a brief screen for personality disorder. *The British Journal of Psychiatry*, 183(3), 228–232. <https://doi.org/10.1192/bjp.183.3.228>
- Moran, R. J., Campo, P., Symmonds, M., Stephan, K. E., Dolan, R. J., & Friston, K. (2013). Free Energy, Precision and Learning: The Role of Cholinergic Neuromodulation. *Journal of Neuroscience*, 33(19), 8227–8236.
<https://doi.org/10.1523/JNEUROSCI.4255-12.2013>
- Moritz, S., Ramdani, N., Klass, H., Andreou, C., Jungclaussen, D., Eifler, S., Englisch, S., Schirmbeck, F., & Zink, M. (2014). Overconfidence in incorrect perceptual judgments in patients with schizophrenia. *Schizophrenia Research: Cognition*, 1(4), 165–170.
<https://doi.org/10.1016/j.scog.2014.09.003>
- Moses-Payne, M. E., Rollwage, M., Fleming, S. M., & Roiser, J. P. (2019). Postdecision Evidence Integration and Depressive Symptoms. *Frontiers in Psychiatry*, 10, 639.
<https://doi.org/10.3389/fpsy.2019.00639>
- Moss, D., & Shaffer, F. (2017). The Application of Heart Rate Variability Biofeedback to Medical and Mental Health Disorders. *Biofeedback*, 45(1), 2–8.
<https://doi.org/10.5298/1081-5937-45.1.03>
- Munro, C., Randell, L., & Lawrie, S. M. (2017). An Integrative Bio-Psycho-Social Theory of Anorexia Nervosa. *Clinical Psychology & Psychotherapy*, 24(1), 1–21.
<https://doi.org/10.1002/cpp.2047>
- Murnen, S. K., & Karazsia, B. T. (2017). A review of research on men's body image and drive for muscularity. In *The psychology of men and masculinities* (pp. 229–257). American Psychological Association. <https://doi.org/10.1037/0000023-009>

- 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., 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., 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., Geary, H., Millgate, E., Catmur, C., & Bird, G. (2018). Direct and indirect effects of age on interoceptive accuracy and awareness across the adult lifespan. *Psychonomic Bulletin & Review*, 25(3), 1193–1202. <https://doi.org/10.3758/s13423-017-1339-z>
- 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>
- Murphy, J., Viding, E., & Bird, G. (2019). Does atypical interoception following physical change contribute to sex differences in mental illness? *Psychological Review*, 126(5), 787–789. <https://doi.org/10.1037/rev0000158>
- Murphy, R., Straebler, S., Cooper, Z., & Fairburn, C. G. (2010). Cognitive Behavioral Therapy for Eating Disorders. *The Psychiatric Clinics of North America*, 33(3), 611–627. <https://doi.org/10.1016/j.psc.2010.04.004>
- Naddaf, M. (2023). ‘Intention is not action’: Brain-research centre steps up quest for equality. *Nature*. <https://doi.org/10.1038/d41586-023-01425-y>
- Nakahara, T., Kojima, S., Tanaka, M., Yasuhara, D., Harada, T., Sagiya, K., Muranaga, T., Nagai, N., Nakazato, M., Nozoe, S., Naruo, T., & Inui, A. (2007). Incomplete restoration of the secretion of ghrelin and PYY compared to insulin after food ingestion following weight gain in anorexia nervosa. *Journal of Psychiatric Research*, 41(10), 814–820. <https://doi.org/10.1016/j.jpsychires.2006.07.021>

- Naliboff, B. D., Berman, S., Suyenobu, B., Labus, J. S., Chang, L., Stains, J., Mandelkern, M. A., & Mayer, E. A. (2006). Longitudinal Change in Perceptual and Brain Activation Response to Visceral Stimuli in Irritable Bowel Syndrome Patients. *Gastroenterology*, 131(2), 352–365. <https://doi.org/10.1053/j.gastro.2006.05.014>
- Naqvi, N. H., & Bechara, A. (2010). The insula and drug addiction: An interoceptive view of pleasure, urges, and decision-making. *Brain Structure and Function*, 214(5–6), 435–450. <https://doi.org/10.1007/s00429-010-0268-7>
- Nassar, M. R., Waltz, J. A., Albrecht, M. A., Gold, J. M., & Frank, M. J. (2021). All or nothing belief updating in patients with schizophrenia reduces precision and flexibility of beliefs. *Brain*, 144(3), 1013–1029. <https://doi.org/10.1093/brain/awaa453>
- National Collaborating Centre for Mental Health (UK). (2004). *Eating Disorders: Core Interventions in the Treatment and Management of Anorexia Nervosa, Bulimia Nervosa and Related Eating Disorders*. British Psychological Society (UK). <http://www.ncbi.nlm.nih.gov/books/NBK49304/>
- National Guideline Alliance (UK). (2017). *Eating Disorders: Recognition and Treatment*. National Institute for Health and Care Excellence (NICE). <http://www.ncbi.nlm.nih.gov/books/NBK436876/>
- NHS Accelerated Access Collaborative. (2022). *NHS Accelerated Access Collaborative » Accelerated Access Collaborative patient and public involvement strategy*. <https://www.england.nhs.uk/aac/publication/accelerated-access-collaborative-patient-and-public-involvement-strategy/>
- Nicholls, D. E., Lynn, R., & Viner, R. M. (2011). Childhood eating disorders: British national surveillance study. *The British Journal of Psychiatry*, 198(4), 295–301. <https://doi.org/10.1192/bjp.bp.110.081356>
- Nicolò, G., Dimaggio, G., Popolo, R., Carcione, A., Procacci, M., Hamm, J., Buck, K. D., Pompili, E., Buccione, I., Lagrotteria, B., & Lysaker, P. H. (2012). Associations of Metacognition With Symptoms, Insight, and Neurocognition in Clinically Stable Outpatients With Schizophrenia. *Journal of Nervous & Mental Disease*, 200(7), 644–647. <https://doi.org/10.1097/NMD.0b013e31825bfb10>
- Nikolova, N., Harrison, O., Toohey, S., Brændholt, M., Legrand, N., Correa, C., Vejlø, M., Jensen, M. S., Fardo, F., & Allen, M. (2022). The respiratory resistance sensitivity

- task: An automated method for quantifying respiratory interoception and metacognition. *Biological Psychology*, 170, 108325.
<https://doi.org/10.1016/j.biopsycho.2022.108325>
- Nord, C. L., Lawson, R. P., & Dalgleish, T. (2021). Disrupted Dorsal Mid-Insula Activation During Interoception Across Psychiatric Disorders. *American Journal of Psychiatry*, 178(8), 761–770. <https://doi.org/10.1176/appi.ajp.2020.20091340>
- Norris, M. L., Harrison, M. E., Isserlin, L., Robinson, A., Feder, S., & Sampson, M. (2016). Gastrointestinal complications associated with anorexia nervosa: A systematic review. *International Journal of Eating Disorders*, 49(3), 216–237.
<https://doi.org/10.1002/eat.22462>
- Nunn, K., Frampton, I., Fuglset, T. S., Törzsök-Sonnevend, M., & Lask, B. (2011). Anorexia nervosa and the insula. *Medical Hypotheses*, 76(3), 353–357.
<https://doi.org/10.1016/j.mehy.2010.10.038>
- Nunn, K., Frampton, I., Gordon, I., & Lask, B. (2008). The fault is not in her parents but in her insula—A neurobiological hypothesis of anorexia nervosa. *European Eating Disorders Review*, 16(5), 355–360. <https://doi.org/10.1002/erv.890>
- Nunn, K., Frampton, I., & Lask, B. (2012). Anorexia nervosa – A noradrenergic dysregulation hypothesis. *Medical Hypotheses*, 78(5), 580–584.
<https://doi.org/10.1016/j.mehy.2012.01.033>
- Olatunji, B. O., Levinson, C., & Calebs, B. (2018). A network analysis of eating disorder symptoms and characteristics in an inpatient sample. *Psychiatry Research*, 262, 270–281. <https://doi.org/10.1016/j.psychres.2018.02.027>
- Oldershaw, A., Hambrook, D., Stahl, D., Tchanturia, K., Treasure, J., & Schmidt, U. (2011). The socio-emotional processing stream in Anorexia Nervosa. *Neuroscience & Biobehavioral Reviews*, 35(3), 970–988.
<https://doi.org/10.1016/j.neubiorev.2010.11.001>
- Oldershaw, A., Hambrook, D., Tchanturia, K., Treasure, J., & Schmidt, U. (2010). Emotional Theory of Mind and Emotional Awareness in Recovered Anorexia Nervosa Patients. *Psychosomatic Medicine*, 72(1), 73–79.
<https://doi.org/10.1097/PSY.0b013e3181c6c7ca>
- Oldershaw, A., Lavender, T., Sallis, H., Stahl, D., & Schmidt, U. (2015). Emotion generation and regulation in anorexia nervosa: A systematic review and meta-analysis of self-

- report data. *Clinical Psychology Review*, 39, 83–95.
<https://doi.org/10.1016/j.cpr.2015.04.005>
- O’Leary, A. (1985). Self-efficacy and health. *Behaviour Research and Therapy*, 23(4), 437–451. [https://doi.org/10.1016/0005-7967\(85\)90172-X](https://doi.org/10.1016/0005-7967(85)90172-X)
- Oliva, R., Baiano, M., Salvo, P., Cereser, L., Castiello, U., & Begliomini, C. (2020). Metacognition in individuals recovered from anorexia nervosa: A voxel-based morphometry study. *Psychiatry Research: Neuroimaging*, 304, 111138.
<https://doi.org/10.1016/j.psychresns.2020.111138>
- Olstad, S., Solem, S., Hjemdal, O., & Hagen, R. (2015). Metacognition in eating disorders: Comparison of women with eating disorders, self-reported history of eating disorders or psychiatric problems, and healthy controls. *Eating Behaviors*, 16, 17–22.
<https://doi.org/10.1016/j.eatbeh.2014.10.019>
- Oppenheimer, S., & Cechetto, D. (2016). The Insular Cortex and the Regulation of Cardiac Function. In Y. S. Prakash (Ed.), *Comprehensive Physiology* (1st ed., pp. 1081–1133). Wiley. <https://doi.org/10.1002/cphy.c140076>
- O’Reilly, J. X., Jbabdi, S., & Behrens, T. E. J. (2012). How can a Bayesian approach inform neuroscience? *European Journal of Neuroscience*, 35(7), 1169–1179.
<https://doi.org/10.1111/j.1460-9568.2012.08010.x>
- Owens, A. P., Allen, M., Ondobaka, S., & Friston, K. (2018). Interoceptive inference: From computational neuroscience to clinic. *Neuroscience & Biobehavioral Reviews*, 90, 174–183. <https://doi.org/10.1016/j.neubiorev.2018.04.017>
- Owens, A. P., Friston, K., Low, D. A., Mathias, C. J., & Critchley, H. D. (2018). Investigating the relationship between cardiac interoception and autonomic cardiac control using a predictive coding framework. *Autonomic Neuroscience*, 210, 65–71.
<https://doi.org/10.1016/j.autneu.2018.01.001>
- Palmer, C., Ainley, V., & Tsakiris, M. (2019). *Fine Tuning Your Heart: A novel method for measuring interoceptive accuracy*. PsyArXiv. <https://doi.org/10.31234/osf.io/qz7r9>
- Palmer, C. J., Seth, A. K., & Hohwy, J. (2015). The felt presence of other minds: Predictive processing, counterfactual predictions, and mentalising in autism. *Consciousness and Cognition*, 36, 376–389. <https://doi.org/10.1016/j.concog.2015.04.007>
- Palmer-Cooper, E. C., Wright, A. C., McGuire, N., Montagnese, M., Dlugunovych, V., Cella, M., & Wykes, T. (2023). Metacognition and psychosis-spectrum experiences: A study

- of objective and subjective measures. *Schizophrenia Research*, S0920996422004613. <https://doi.org/10.1016/j.schres.2022.12.014>
- Palmieri, S., Mansueto, G., Ruggiero, G. M., Caselli, G., Sassaroli, S., & Spada, M. M. (2021). Metacognitive beliefs across eating disorders and eating behaviours: A systematic review. *Clinical Psychology & Psychotherapy*, 28(5), 1254–1265. <https://doi.org/10.1002/cpp.2573>
- Palomba, D., Venturini, M., Rausa, M., Contin, S. A., Penolazzi, B., Schumann, R., & Ballardini, D. (2017). Reduced sympathetic activity and dysfunctional metacognition in patients with anorexia nervosa: A preliminary study. *Journal of Evidence-Based Psychotherapies*, 17(1), 1–20. <https://doi.org/10.24193/jebp.2017.1.1>
- Pannu, J. K., & Kaszniak, A. W. (2005). Metamemory Experiments in Neurological Populations: A Review. *Neuropsychology Review*, 15(3), 105–130. <https://doi.org/10.1007/s11065-005-7091-6>
- Parent, M. C. (2013). Clinical considerations in etiology, assessment, and treatment of men's muscularity-focused body image disturbance. *Psychology of Men & Masculinity*, 14(1), 88–100. <https://doi.org/10.1037/a0025644>
- Paulus, M. P. (2007). Decision-Making Dysfunctions in Psychiatry—Altered Homeostatic Processing? *Science*, 318(5850), 602–606. <https://doi.org/10.1126/science.1142997>
- 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 and Function*, 214(5–6), 451–463. <https://doi.org/10.1007/s00429-010-0258-9>
- Paulus, M. P., Tapert, S. F., & Schulteis, G. (2009). The role of interoception and alliesthesia in addiction. *Pharmacology Biochemistry and Behavior*, 94(1), 1–7. <https://doi.org/10.1016/j.pbb.2009.08.005>
- Peat, C., Mitchell, J. E., Hoek, H. W., & Wonderlich, S. A. (2009). Validity and utility of subtyping anorexia nervosa. *International Journal of Eating Disorders*, 42(7), 590–594. <https://doi.org/10.1002/eat.20717>

- Pennebaker, J. W., & Roberts, T.-A. (1992). Toward a His and Hers Theory of Emotion: Gender Differences in Visceral Perception. *Journal of Social and Clinical Psychology, 11*(3), 199–212. <https://doi.org/10.1521/jscp.1992.11.3.199>
- Peres, V., Corcos, M., Robin, M., & Pham-Scottez, A. (2020). Emotional intelligence, empathy and alexithymia in anorexia nervosa during adolescence. *Eating and Weight Disorders - Studies on Anorexia, Bulimia and Obesity, 25*(1), 1–8. <https://doi.org/10.1007/s40519-018-0482-5>
- Peters, M., Charles, L., & Maniscalco, B. (2023). Optimal metacognitive decision strategies in Signal Detection Theory. *Journal of Vision, 23*(9), 5465. <https://doi.org/10.1167/jov.23.9.5465>
- Petkova, H., Simic, M., Nicholls, D., Ford, T., Prina, A. M., Stuart, R., Livingstone, N., Kelly, G., Macdonald, G., Eisler, I., Gowers, S., Barrett, B. M., & Byford, S. (2019). Incidence of anorexia nervosa in young people in the UK and Ireland: A national surveillance study. *BMJ Open, 9*(10), e027339. <https://doi.org/10.1136/bmjopen-2018-027339>
- Petzschner, F. H., Garfinkel, S. N., Paulus, M. P., Koch, C., & Khalsa, S. S. (2021). Computational Models of Interoception and Body Regulation. *Trends in Neurosciences, 44*(1), 63–76. <https://doi.org/10.1016/j.tins.2020.09.012>
- Petzschner, F. H., Weber, L. A. E., Gard, T., & Stephan, K. E. (2017). Computational Psychosomatics and Computational Psychiatry: Toward a Joint Framework for Differential Diagnosis. *Biological Psychiatry, 82*(6), 421–430. <https://doi.org/10.1016/j.biopsych.2017.05.012>
- Pezzulo, G., Rigoli, F., & Friston, K. (2015). Active Inference, homeostatic regulation and adaptive behavioural control. *Progress in Neurobiology, 134*, 17–35. <https://doi.org/10.1016/j.pneurobio.2015.09.001>
- Phillipou, A., Rossell, S. L., Castle, D. J., & Gurvich, C. (2022). Interoceptive awareness in anorexia nervosa. *Journal of Psychiatric Research, 148*, 84–87. <https://doi.org/10.1016/j.jpsychires.2022.01.051>
- Pike, K. M. (1998). Long-term course of anorexia nervosa: Response, relapse, remission, and recovery. *Clinical Psychology Review, 18*(4), 447–475. [https://doi.org/10.1016/S0272-7358\(98\)00014-2](https://doi.org/10.1016/S0272-7358(98)00014-2)
- Pike, K. M., So, M., Hilbert, A., Maekawa, H., Shimanouchi, T., Wilfley, D., Dohm, F., Fairburn, C. G., & Weissman, R. S. (2021). Risk factors for anorexia nervosa and

- bulimia nervosa in Japan and compared to a U.S. sample. *International Journal of Eating Disorders*, 54(2), 155–167. <https://doi.org/10.1002/eat.23442>
- Pinto, A. M., Guarda, A. S., Heinberg, L. J., & DiClemente, C. C. (2006). Development of the eating disorder recovery self-efficacy questionnaire. *International Journal of Eating Disorders*, 39(5), 376–384. <https://doi.org/10.1002/eat.20256>
- Pinto, A. M., Heinberg, L. J., Coughlin, J. W., Fava, J. L., & Guarda, A. S. (2008). The Eating Disorder Recovery Self-Efficacy Questionnaire (EDRSQ): Change with treatment and prediction of outcome. *Eating Behaviors*, 9(2), 143–153. <https://doi.org/10.1016/j.eatbeh.2007.07.001>
- Piwek, L., Ellis, D. A., Andrews, S., & Joinson, A. (2016). The Rise of Consumer Health Wearables: Promises and Barriers. *PLOS Medicine*, 13(2), e1001953. <https://doi.org/10.1371/journal.pmed.1001953>
- 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>
- Pollatos, O., Gramann, K., & Schandry, R. (2007). Neural systems connecting interoceptive awareness and feelings. *Human Brain Mapping*, 28(1), 9–18. <https://doi.org/10.1002/hbm.20258>
- Pollatos, O., Herbert, B. M., Berberich, G., Zaudig, M., Krauseneck, T., & Tsakiris, M. (2016). Atypical Self-Focus Effect on Interoceptive Accuracy in Anorexia Nervosa. *Frontiers in Human Neuroscience*, 10. <https://www.frontiersin.org/articles/10.3389/fnhum.2016.00484>
- Pollatos, O., Herbert, B. M., Matthias, E., & Schandry, R. (2007). Heart rate response after emotional picture presentation is modulated by interoceptive awareness. *International Journal of Psychophysiology*, 63(1), 117–124. <https://doi.org/10.1016/j.ijpsycho.2006.09.003>
- Pollatos, O., Kurz, A.-L., Albrecht, J., Schreder, T., Kleemann, A. M., Schöpf, V., Kopietz, R., Wiesmann, M., & Schandry, R. (2008). Reduced perception of bodily signals in anorexia nervosa. *Eating Behaviors*, 9(4), 381–388. <https://doi.org/10.1016/j.eatbeh.2008.02.001>

- Pollatos, O., Matthias, E., & Keller, J. (2015). When interoception helps to overcome negative feelings caused by social exclusion. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.00786>
- Pollatos, O., Schandry, R., Auer, D. P., & 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>
- Pop-Jordanova, N. (2000). Psychological characteristics and biofeedback mitigation in preadolescents with eating disorders. *Pediatrics International*, 42(1), 76–81. <https://doi.org/10.1046/j.1442-200x.2000.01176.x>
- Porges, S. W. (1993). *Body perception questionnaire*. Laboratory of Developmental Assessment: University of Maryland.
- Pouget, A., Drugowitsch, J., & Kepecs, A. (2016). Confidence and certainty: Distinct probabilistic quantities for different goals. *Nature Neuroscience*, 19(3), Article 3. <https://doi.org/10.1038/nn.4240>
- Pyne, J. M., Bean, D., & Sullivan, G. (2001). Characteristics of Patients with Schizophrenia Who Do Not Believe They Are Mentally Ill. *The Journal of Nervous and Mental Disease*, 189(3), 146.
- Quadt, L., Critchley, H. D., & Garfinkel, S. N. (2018). The neurobiology of interoception in health and disease. *Annals of the New York Academy of Sciences*, 1428(1), 112–128. <https://doi.org/10.1111/nyas.13915>
- Raevuori, A., Keski-Rahkonen, A., & Hoek, H. W. (2014). A review of eating disorders in males. *Current Opinion in Psychiatry*, 27(6), 426. <https://doi.org/10.1097/YCO.0000000000000113>
- Ralph, A. F., Brennan, L., Byrne, S., Caldwell, B., Farmer, J., Hart, L. M., Heruc, G. A., Maguire, S., Piya, M. K., Quin, J., Trobe, S. K., Wallis, A., Williams-Tchen, A., & Hay, P. (2022). Management of eating disorders for people with higher weight: Clinical practice guideline. *Journal of Eating Disorders*, 10(1), 121. <https://doi.org/10.1186/s40337-022-00622-w>
- Reas, D. L., & Rø, Ø. (2018). Less symptomatic, but equally impaired: Clinical impairment in restricting versus binge-eating/purging subtype of anorexia nervosa. *Eating Behaviors*, 28, 32–37. <https://doi.org/10.1016/j.eatbeh.2017.12.004>

- Remes-Troche, J. M., Attaluri, A., Chahal, P., & Rao, S. S. C. (2012). Barostat or dynamic balloon distention test: Which technique is best suited for esophageal sensory testing? *Diseases of the Esophagus*, 25(7), 584–589. <https://doi.org/10.1111/j.1442-2050.2011.01294.x>
- Richard, A., Meule, A., Georgii, C., Voderholzer, U., Cuntz, U., Wilhelm, F. H., & Blechert, J. (2019). Associations between interoceptive sensitivity, intuitive eating, and body mass index in patients with anorexia nervosa and normal-weight controls. *European Eating Disorders Review*, 27(5), 571–577. <https://doi.org/10.1002/erv.2676>
- Riedlinger, C., Schmidt, G., Weiland, A., Stengel, A., Giel, K. E., Zipfel, S., Enck, P., & Mack, I. (2020). Which Symptoms, Complaints and Complications of the Gastrointestinal Tract Occur in Patients With Eating Disorders? A Systematic Review and Quantitative Analysis. *Frontiers in Psychiatry*, 11. <https://www.frontiersin.org/articles/10.3389/fpsyt.2020.00195>
- Rief, W., Hiller, W., & Margraf, J. (1998). Cognitive aspects of hypochondriasis and the somatization syndrome. *Journal of Abnormal Psychology*, 107(4), 587–595. <https://doi.org/10.1037/0021-843X.107.4.587>
- Rigaud, D., Bedig, G., Merrouche, M., Vulpillat, M., Bonfils, S., & Apfelbaum, M. (1988). Delayed gastric emptying in anorexia nervosa is improved by completion of a renutrition program. *Digestive Diseases and Sciences*, 33(8), 919–925. <https://doi.org/10.1007/BF01535985>
- Rigaud, D., Pennacchio, H., Bizeul, C., Reveillard, V., & Vergès, B. (2011). Outcome in AN adult patients: A 13-year follow-up in 484 patients. *Diabetes & Metabolism*, 37(4), 305–311. <https://doi.org/10.1016/j.diabet.2010.11.020>
- Robinson, P. H. (1989). Perceptivity and Paraceptivity During Measurement of Gastric Emptying in Anorexia and Bulimia Nervosa. *British Journal of Psychiatry*, 154(3), 400–405. <https://doi.org/10.1192/bjp.154.3.400>
- Rodríguez-Ruiz, S., Ruiz-Padial, E., Vera, N., Fernández, C., Anllo-Vento, L., & Vila, J. (2009). Effect of Heart Rate Variability on Defensive Reaction and Eating Disorder Symptomatology in Chocolate Cravers. *Journal of Psychophysiology*, 23(3), 95–103. <https://doi.org/10.1027/0269-8803.23.3.95>
- Rolls, E. T. (2011). Taste, olfactory and food texture reward processing in the brain and obesity. *International Journal of Obesity*, 35(4), Article 4. <https://doi.org/10.1038/ijo.2010.155>

- Rolls, E. T. (2016). Reward Systems in the Brain and Nutrition. *Annual Review of Nutrition*, 36(1), 435–470. <https://doi.org/10.1146/annurev-nutr-071715-050725>
- Rollwage, M., Dolan, R. J., & Fleming, S. M. (2018). Metacognitive Failure as a Feature of Those Holding Radical Beliefs. *Current Biology*, 28(24), 4014–4021.e8. <https://doi.org/10.1016/j.cub.2018.10.053>
- Romero, L. M., Dickens, M. J., & Cyr, N. E. (2009). The reactive scope model—A new model integrating homeostasis, allostasis, and stress. *Hormones and Behavior*, 55(3), 375–389. <https://doi.org/10.1016/j.yhbeh.2008.12.009>
- Rosano, G. M., Rillo, M., Leonardo, F., Pappone, C., & Chierchia, S. L. (1997). Palpitations: What is the mechanism, and when should we treat them? *International Journal of Fertility and Women's Medicine*, 42(2), 94–100.
- Rosenberg, M. (2015). *Society and the Adolescent Self-Image*. Princeton University Press.
- Rouault, M., Dayan, P., & Fleming, S. M. (2019). Forming global estimates of self-performance from local confidence. *Nature Communications*, 10(1), Article 1. <https://doi.org/10.1038/s41467-019-09075-3>
- Rouault, M., & Fleming, S. M. (2020). Formation of global self-beliefs in the human brain. *Proceedings of the National Academy of Sciences*, 117(44), 27268–27276. <https://doi.org/10.1073/pnas.2003094117>
- Rouault, M., McWilliams, A., Allen, M. G., & Fleming, S. M. (2018). Human Metacognition Across Domains: Insights from Individual Differences and Neuroimaging. *Personality Neuroscience*, 1, e17. <https://doi.org/10.1017/pen.2018.16>
- Rouault, M., Seow, T., Gillan, C. M., & Fleming, S. M. (2018). Psychiatric Symptom Dimensions Are Associated With Dissociable Shifts in Metacognition but Not Task Performance. *Biological Psychiatry*, 84(6), 443–451. <https://doi.org/10.1016/j.biopsych.2017.12.017>
- 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>
- Roux, H., Ali, A., Lambert, S., Radon, L., Huas, C., Curt, F., Berthoz, S., Godart, N., Godart, N., Berthoz, S., Lalanne, C., Duclos, J., Mattar, L., Roux, H., Thiébaud, M. R., Vibert, S., Hubert, T., Courty, A., Ringuenet, D., ... the EVHAN Group. (2016). Predictive

- factors of dropout from inpatient treatment for anorexia nervosa. *BMC Psychiatry*, 16(1), 339. <https://doi.org/10.1186/s12888-016-1010-7>
- Rupprechter, S., Stankevicius, A., Huys, Q. J. M., Steele, J. D., & Seriès, P. (2018). Major Depression Impairs the Use of Reward Values for Decision-Making. *Scientific Reports*, 8(1), Article 1. <https://doi.org/10.1038/s41598-018-31730-w>
- Russell, T. A., Schmidt, U., Doherty, L., Young, V., & Tchanturia, K. (2009). Aspects of social cognition in anorexia nervosa: Affective and cognitive theory of mind. *Psychiatry Research*, 168(3), 181–185. <https://doi.org/10.1016/j.psychres.2008.10.028>
- Safdari, S., Khoramdel, K., & Kamranian, E. (2013). *The Role of Metacognitive Beliefs in Eating Disorders*. 1(2), 96–99.
- Saini, F., Ponzo, S., Silvestrin, F., Fotopoulou, A., & David, A. S. (2022). Depersonalization disorder as a systematic downregulation of interoceptive signals. *Scientific Reports*, 12(1), 22123. <https://doi.org/10.1038/s41598-022-22277-y>
- Salbach-Andrae, H., Lenz, K., Simmendinger, N., Klinkowski, N., Lehmkuhl, U., & Pfeiffer, E. (2008). Psychiatric Comorbidities among Female Adolescents with Anorexia Nervosa. *Child Psychiatry and Human Development*, 39(3), 261–272. <https://doi.org/10.1007/s10578-007-0086-1>
- Santel, S., Baving, L., Krauel, K., Münte, T. F., & Rotte, M. (2006). Hunger and satiety in anorexia nervosa: fMRI during cognitive processing of food pictures. *Brain Research*, 1114(1), 138–148. <https://doi.org/10.1016/j.brainres.2006.07.045>
- Santonicola, A., Gagliardi, M., Guarino, M. P. L., Siniscalchi, M., Ciacchi, C., & Iovino, P. (2019). Eating Disorders and Gastrointestinal Diseases. *Nutrients*, 11(12), Article 12. <https://doi.org/10.3390/nu11123038>
- Sapra, A., Malik, A., & Bhandari, P. (2023). Vital Sign Assessment. In *StatPearls*. StatPearls Publishing. <http://www.ncbi.nlm.nih.gov/books/NBK553213/>
- Saramandi, A., Crucianelli, L., Koukoutsakis, A., Nisticò, V., Baiza, A., Goeta, D., Demartini, B., Gambini, O., Jenkinson, P., & Fotopoulou, A. (2022). *Belief updating about Interoception and Body Size Estimation in Anorexia Nervosa*. PsyArXiv. <https://doi.org/10.31234/osf.io/rntsf>
- Saramandi, A., Crucianelli, L., Koukoutsakis, A., Nisticò, V., Mavromara, L., Goeta, D., Boido, G., Gonidakis, F., Demartini, B., Bertelli, S., Gambini, O., Jenkinson, P., &

- Fotopoulou, A. (2023). *Updating Prospective, Self-Efficacy Beliefs About Cardiac Interoception in Anorexia Nervosa: An Experimental and Computational Study*. PsyArXiv. <https://doi.org/10.31234/osf.io/c8mjv>
- Sato, Y., & Fukudo, S. (2015). Gastrointestinal symptoms and disorders in patients with eating disorders. *Clinical Journal of Gastroenterology*, 8(5), 255–263. <https://doi.org/10.1007/s12328-015-0611-x>
- Schaefer, M., Egloff, B., & Witthöft, M. (2012). Is interoceptive awareness really altered in somatoform disorders? Testing competing theories with two paradigms of heartbeat perception. *Journal of Abnormal Psychology*, 121(3), 719–724. <https://doi.org/10.1037/a0028509>
- Schalla, M. A., & Stengel, A. (2019). Gastrointestinal alterations in anorexia nervosa—A systematic review. *European Eating Disorders Review*. <https://doi.org/10.1002/erv.2679>
- 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>
- Schebendach, J. E., Mayer, L. E., Devlin, M. J., Attia, E., Contento, I. R., Wolf, R. L., & Walsh, B. T. (2008). Dietary energy density and diet variety as predictors of outcome in anorexia nervosa. *The American Journal of Clinical Nutrition*, 87(4), 810–816. <https://doi.org/10.1093/ajcn/87.4.810>
- Schmidt, U., & Treasure, J. (2006). Anorexia nervosa: Valued and visible. A cognitive-interpersonal maintenance model and its implications for research and practice. *British Journal of Clinical Psychology*, 45(3), 343–366. <https://doi.org/10.1348/014466505X53902>
- Schmidt, U., Wade, T. D., & Treasure, J. (2014). The Maudsley Model of Anorexia Nervosa Treatment for Adults (MANTRA): Development, Key Features, and Preliminary Evidence. *Journal of Cognitive Psychotherapy*, 28(1), 48–71.
- Schnyer, D. M., Verfaellie, M., Alexander, M. P., LaFleche, G., Nicholls, L., & Kaszniak, A. W. (2004). A role for right medial prefrontal cortex in accurate feeling-of-knowing judgements: Evidence from patients with lesions to frontal cortex. *Neuropsychologia*, 42(7), 957–966. <https://doi.org/10.1016/j.neuropsychologia.2003.11.020>
- Schultz, W., Dayan, P., & Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*, 275(5306), 1593–1599. <https://doi.org/10.1126/science.275.5306.1593>

- Schwarzer, R., & Jerusalem, M. (1995). Generalised Self-Efficacy Scale. In *In Weinman, J., Wright, S., & Johnston, M. Measures in health psychology: A user's portfolio. Causal and control beliefs* (pp. 35–37). NFER-NELSON.
- Scolnick, B., Mostofsky, D. I., & Keane, R. J. (2014). Pilot study employing heart rate variability biofeedback training to decrease anxiety in patients with eating disorders. *Journal of Eating Disorders*, 2(1), 17. <https://doi.org/10.1186/2050-2974-2-17>
- Scrivener, R., National Institute for Clinical Excellence, Royal College of Nursing, University of Leicester, & Commission for Health Improvement (Eds.). (2004). *Principles for best practice in clinical audit* (repr). Radcliffe Medical.
- Seidel, M., Petermann, J., Diestel, S., Ritschel, F., Boehm, I., King, J. A., Geisler, D., Bernardoni, F., Roessner, V., Goschke, T., & Ehrlich, S. (2016). A naturalistic examination of negative affect and disorder-related rumination in anorexia nervosa. *European Child & Adolescent Psychiatry*, 25(11), 1207–1216. <https://doi.org/10.1007/s00787-016-0844-3>
- Seitz, J., Herpertz-Dahlmann, B., & Konrad, K. (2016). Brain morphological changes in adolescent and adult patients with anorexia nervosa. *Journal of Neural Transmission*, 123(8), 949–959. <https://doi.org/10.1007/s00702-016-1567-9>
- Seow, T. X. F., & Gillan, C. M. (2020). Transdiagnostic Phenotyping Reveals a Host of Metacognitive Deficits Implicated in Compulsivity. *Scientific Reports*, 10(1), Article 1. <https://doi.org/10.1038/s41598-020-59646-4>
- Seow, T. X. F., Rouault, M., Gillan, C. M., & Fleming, S. M. (2021). How Local and Global Metacognition Shape Mental Health. *Biological Psychiatry*, 90(7), 436–446. <https://doi.org/10.1016/j.biopsych.2021.05.013>
- Serino, S., Pedroli, E., Keizer, A., Triberti, S., Dakanalis, A., Pallavicini, F., Chirico, A., & Riva, G. (2016). Virtual Reality Body Swapping: A Tool for Modifying the Allocentric Memory of the Body. *Cyberpsychology, Behavior, and Social Networking*, 19(2), 127–133. <https://doi.org/10.1089/cyber.2015.0229>
- Serpell, L., Treasure, J., Teasdale, J., & Sullivan, V. (1999). Anorexia nervosa: Friend or foe? *International Journal of Eating Disorders*, 25(2), 177–186. [https://doi.org/10.1002/\(SICI\)1098-108X\(199903\)25:2<177::AID-EAT7>3.0.CO;2-D](https://doi.org/10.1002/(SICI)1098-108X(199903)25:2<177::AID-EAT7>3.0.CO;2-D)
- 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>

- Seth, A. K., & Friston, K. (2016). Active interoceptive inference and the emotional brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1708), 20160007. <https://doi.org/10.1098/rstb.2016.0007>
- Seth, A. K., Suzuki, K., & Critchley, H. D. (2012). An Interoceptive Predictive Coding Model of Conscious Presence. *Frontiers in Psychology*, 2. <https://doi.org/10.3389/fpsyg.2011.00395>
- 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>
- Shapiro, J. R., Berkman, N. D., Brownley, K. A., Sedway, J. A., Lohr, K. N., & Bulik, C. M. (2007). Bulimia nervosa treatment: A systematic review of randomized controlled trials. *International Journal of Eating Disorders*, 40(4), 321–336. <https://doi.org/10.1002/eat.20372>
- Sharot, T. (2011). *The optimism bias*. 21, R941–R945.
- Sharot, T., & Garrett, N. (2016). Forming Beliefs: Why Valence Matters. *Trends in Cognitive Sciences*, 20(1), 25–33. <https://doi.org/10.1016/j.tics.2015.11.002>
- Sharot, T., Kanai, R., Marston, D., Korn, C. W., Rees, G., & Dolan, R. J. (2012). Selectively altering belief formation in the human brain. *Proceedings of the National Academy of Sciences*, 109(42), 17058–17062. <https://doi.org/10.1073/pnas.1205828109>
- Sharp, C. W., Clark, S. A., Dunan, J. R., Blackwood, D. H. R., & Shapiro, C. M. (1994). Clinical presentation of anorexia nervosa in males: 24 new cases. *International Journal of Eating Disorders*, 15(2), 125–134. [https://doi.org/10.1002/1098-108X\(199403\)15:2<125::AID-EAT2260150204>3.0.CO;2-O](https://doi.org/10.1002/1098-108X(199403)15:2<125::AID-EAT2260150204>3.0.CO;2-O)
- Shekriladze, I., & Tchanturia, K. (2017). Acculturation to Western Culture in the Context of Eating Disorders. In T. Wade (Ed.), *Encyclopedia of Feeding and Eating Disorders* (pp. 1–4). Springer Singapore. https://doi.org/10.1007/978-981-287-104-6_204
- Sherrington, C. (1906). *The Integrative Action of the Nervous System*. Yale University Press.
- 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
- Shimamura, A. P., & Squire, L. R. (1986). Memory and metamemory: A study of the feeling-of-knowing phenomenon in amnesic patients. *Journal of Experimental Psychology*.

- Learning, Memory, and Cognition*, 12(3), 452–460. <https://doi.org/10.1037//0278-7393.12.3.452>
- Shivkumar, K., Ajjola, O. A., Anand, I., Armour, J. A., Chen, P., Esler, M., De Ferrari, G. M., Fishbein, M. C., Goldberger, J. J., Harper, R. M., Joyner, M. J., Khalsa, S. S., Kumar, R., Lane, R., Mahajan, A., Po, S., Schwartz, P. J., Somers, V. K., Valderrabano, M., ... Zipes, D. P. (2016). Clinical neurocardiology defining the value of neuroscience-based cardiovascular therapeutics. *The Journal of Physiology*, 594(14), 3911–3954. <https://doi.org/10.1113/JP271870>
- Sielski, R., Rief, W., & Glombiewski, J. A. (2017). Efficacy of Biofeedback in Chronic back Pain: A Meta-Analysis. *International Journal of Behavioral Medicine*, 24(1), 25–41. <https://doi.org/10.1007/s12529-016-9572-9>
- Simons, L. E., Elman, I., & Borsook, D. (2014). Psychological processing in chronic pain: A neural systems approach. *Neuroscience & Biobehavioral Reviews*, 39, 61–78. <https://doi.org/10.1016/j.neubiorev.2013.12.006>
- Skinner, B. F. (1965). *Science And Human Behavior*. Simon and Schuster.
- Smink, F. R. E., van Hoeken, D., & Hoek, H. W. (2012). Epidemiology of Eating Disorders: Incidence, Prevalence and Mortality Rates. *Current Psychiatry Reports*, 14(4), 406–414. <https://doi.org/10.1007/s11920-012-0282-y>
- Smink, F. R. E., van Hoeken, D., & Hoek, H. W. (2013). Epidemiology, course, and outcome of eating disorders. *Current Opinion in Psychiatry*, 26(6), 543. <https://doi.org/10.1097/YCO.0b013e328365a24f>
- Smith, K. E., Mason, T. B., & Lavender, J. M. (2018). Rumination and eating disorder psychopathology: A meta-analysis. *Clinical Psychology Review*, 61, 9–23. <https://doi.org/10.1016/j.cpr.2018.03.004>
- Smith, R., Badcock, P., & Friston, K. J. (2021). Recent advances in the application of predictive coding and active inference models within clinical neuroscience. *Psychiatry and Clinical Neurosciences*, 75(1), 3–13. <https://doi.org/10.1111/pcn.13138>
- Smith, R., Feinstein, J. S., Kuplicki, R., Forthman, K. L., Stewart, J. L., Paulus, M. P., & Khalsa, S. S. (2021). Perceptual insensitivity to the modulation of interoceptive signals in depression, anxiety, and substance use disorders. *Scientific Reports*, 11(1), Article 1. <https://doi.org/10.1038/s41598-021-81307-3>

- Smith, R., Kirlic, N., Stewart, J. L., Touthang, J., Kuplicki, R., Khalsa, S. S., Feinstein, J., Paulus, M. P., & Aupperle, R. L. (2021). Greater decision uncertainty characterizes a transdiagnostic patient sample during approach-avoidance conflict: A computational modelling approach. *Journal of Psychiatry and Neuroscience*, 46(1), E74–E87. <https://doi.org/10.1503/jpn.200032>
- Smith, R., Kuplicki, R., Feinstein, J., Forthman, K. L., Stewart, J. L., Paulus, M. P., Investigators, T. 1000, & Khalsa, S. S. (2020). A Bayesian computational model reveals a failure to adapt interoceptive precision estimates across depression, anxiety, eating, and substance use disorders. *PLOS Computational Biology*, 16(12), e1008484. <https://doi.org/10.1371/journal.pcbi.1008484>
- Smith, R., Mayeli, A., Taylor, S., Al Zoubi, O., Naegle, J., & Khalsa, S. S. (2021). Gut inference: A computational modelling approach. *Biological Psychology*, 164, 108152. <https://doi.org/10.1016/j.biopsycho.2021.108152>
- Smith, R., Thayer, J. F., Khalsa, S. S., & Lane, R. D. (2017). The hierarchical basis of neurovisceral integration. *Neuroscience & Biobehavioral Reviews*, 75, 274–296. <https://doi.org/10.1016/j.neubiorev.2017.02.003>
- Solmi, M., Collantoni, E., Meneguzzo, P., Degortes, D., Tenconi, E., & Favaro, A. (2018). Network analysis of specific psychopathology and psychiatric symptoms in patients with eating disorders. *International Journal of Eating Disorders*, 51(7), 680–692. <https://doi.org/10.1002/eat.22884>
- Souchay, C., Isingrini, M., & Gil, R. (2006). Metamemory Monitoring and Parkinson's Disease. *Journal of Clinical and Experimental Neuropsychology*, 28(4), 618–630. <https://doi.org/10.1080/13803390590935453>
- Spada, M. M., Caselli, G., Nikčević, A. V., & Wells, A. (2015). Metacognition in addictive behaviors. *Addictive Behaviors*, 44, 9–15. <https://doi.org/10.1016/j.addbeh.2014.08.002>
- Spielberger, C. D., Gonzalez-Reigosa, F., Martinez-Urrutia, A., Natalicio, L. F. S., & Natalicio, D. S. (1971). *The State-Trait Anxiety Inventory*. 5. <https://journal.sipsych.org/index.php/IJP/article/view/620>
- Stange, J. P., Alloy, L. B., & Fresco, D. M. (2017). Inflexibility as a vulnerability to depression: A systematic qualitative review. *Clinical Psychology: Science and Practice*, 24(3), 245–276. <https://doi.org/10.1037/h0101744>

- Stankevicius, A., Huys, Q. J. M., Kalra, A., & Seriès, P. (2014). Optimism as a Prior Belief about the Probability of Future Reward. *PLOS Computational Biology*, 10(5), e1003605. <https://doi.org/10.1371/journal.pcbi.1003605>
- Startup, H., Lavender, A., Oldershaw, A., Stott, R., Tchanturia, K., Treasure, J., & Schmidt, U. (2013). Worry and Rumination in Anorexia Nervosa. *Behavioural and Cognitive Psychotherapy*, 41(3), 301–316. <https://doi.org/10.1017/S1352465812000847>
- Steele, A. L., Bergin, J., & Wade, T. D. (2011). Self-efficacy as a robust predictor of outcome in guided self-help treatment for broadly defined bulimia nervosa. *International Journal of Eating Disorders*, 44(5), 389–396. <https://doi.org/10.1002/eat.20830>
- Stein, D., Kaye, W. H., Matsunaga, H., Myers, D., Orbach, I., Har-Even, D., Frank, G., & Rao, R. (2003). Pain perception in recovered bulimia nervosa patients. *International Journal of Eating Disorders*, 34(3), 331–336. <https://doi.org/10.1002/eat.10164>
- Steinglass, J. E., Berner, L. A., & Attia, E. (2019). Cognitive Neuroscience of Eating Disorders. *Psychiatric Clinics of North America*, 42(1), 75–91. <https://doi.org/10.1016/j.psc.2018.10.008>
- Steinglass, J. E., Eisen, J. L., Attia, E., Mayer, L., & Walsh, B. T. (2007). Is Anorexia Nervosa a Delusional Disorder? An Assessment of Eating Beliefs in Anorexia Nervosa. *Journal of Psychiatric Practice®*, 13(2), 65. <https://doi.org/10.1097/01.pra.0000265762.79753.88>
- Steinglass, J. E., Sysko, R., Mayer, L., Berner, L. A., Schebendach, J., Wang, Y., Chen, H., Albano, A. M., Simpson, H. B., & Walsh, B. T. (2010). Pre-meal anxiety and food intake in anorexia nervosa. *Appetite*, 55(2), 214–218. <https://doi.org/10.1016/j.appet.2010.05.090>
- Steinhausen, H.-C., & Jensen, C. M. (2015). Time trends in lifetime incidence rates of first-time diagnosed anorexia nervosa and bulimia nervosa across 16 years in a Danish nationwide psychiatric registry study. *The International Journal of Eating Disorders*, 48(7), 845–850. <https://doi.org/10.1002/eat.22402>
- Stephan, K. E., Manjaly, Z. M., Mathys, C. D., Weber, L. A. E., Paliwal, S., Gard, T., Tittgemeyer, M., Fleming, S. M., Haker, H., Seth, A. K., & Petzschner, F. H. (2016). Allostatic Self-efficacy: A Metacognitive Theory of Dyshomeostasis-Induced Fatigue and Depression. *Frontiers in Human Neuroscience*, 10. <https://doi.org/10.3389/fnhum.2016.00550>

- Stephan, K. E., Schlagenhaut, F., Huys, Q. J. M., Raman, S., Aponte, E. A., Brodersen, K. H., Rigoux, L., Moran, R. J., Daunizeau, J., Dolan, R. J., Friston, K., & Heinz, A. (2017). Computational neuroimaging strategies for single patient predictions. *NeuroImage*, 145, 180–199. <https://doi.org/10.1016/j.neuroimage.2016.06.038>
- Sterling, P. (2014). Homeostasis vs Allostatics: Implications for Brain Function and Mental Disorders. *JAMA Psychiatry*, 71(10), 1192–1193. <https://doi.org/10.1001/jamapsychiatry.2014.1043>
- Stern, E. R. (2014). Neural Circuitry of Interoception: New Insights into Anxiety and Obsessive-Compulsive Disorders. *Current Treatment Options in Psychiatry*, 1(3), 235–247. <https://doi.org/10.1007/s40501-014-0019-0>
- Sterzer, P., Adams, R. A., Fletcher, P., Frith, C., Lawrie, S. M., Muckli, L., Petrovic, P., Uhlhaas, P., Voss, M., & Corlett, P. R. (2018). The Predictive Coding Account of Psychosis. *Biological Psychiatry*, 84(9), 634–643. <https://doi.org/10.1016/j.biopsych.2018.05.015>
- Steward, T., Menchon, J. M., Jiménez-Murcia, S., Soriano-Mas, C., & Fernandez-Aranda, F. (2018). Neural Network Alterations Across Eating Disorders: A Narrative Review of fMRI Studies. *Current Neuropharmacology*, 16(8), 1150–1163. <https://doi.org/10.2174/1570159X15666171017111532>
- Stewart, S. H., & Watt, M. C. (2008). Introduction to the Special Issue on Interoceptive Exposure in the Treatment of Anxiety and Related Disorders: Novel Applications and Mechanisms of Action. *Journal of Cognitive Psychotherapy*, 22(4), 291–302. <https://doi.org/10.1891/0889-8391.22.4.291>
- Stinson, C. (2019). The absent body in psychiatric diagnosis, treatment, and research. *Synthese*, 196(6), 2153–2176. <https://doi.org/10.1007/s11229-017-1507-8>
- Stone, J., Pal, S., Blackburn, D., Reuber, M., Thekkumpurath, P., & Carson, A. (2015). Functional (Psychogenic) Cognitive Disorders: A Perspective from the Neurology Clinic. *Journal of Alzheimer's Disease*, 48(s1), S5–S17. <https://doi.org/10.3233/JAD-150430>
- Strigo, I. A., & Craig, A. D. (Bud). (2016). Interoception, homeostatic emotions and sympathovagal balance. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1708), 20160010. <https://doi.org/10.1098/rstb.2016.0010>
- Strober, M., Freeman, R., Lampert, C., Diamond, J., & Kaye, W. (2000). Controlled Family Study of Anorexia Nervosa and Bulimia Nervosa: Evidence of Shared Liability and

- Transmission of Partial Syndromes. *American Journal of Psychiatry*, 157(3), 393–401. <https://doi.org/10.1176/appi.ajp.157.3.393>
- Strober, M., Freeman, R., Lampert, C., Diamond, J., Teplinsky, C., & DeAntonio, M. (2006). Are there gender differences in core symptoms, temperament, and short-term prospective outcome in anorexia nervosa? *International Journal of Eating Disorders*, 39(7), 570–575. <https://doi.org/10.1002/eat.20293>
- Strober, M., Freeman, R., & Morrell, W. (1997). The long-term course of severe anorexia nervosa in adolescents: Survival analysis of recovery, relapse, and outcome predictors over 10-15 years in a prospective study. *International Journal of Eating Disorders*, 22(4), 339–360. [https://doi.org/10.1002/\(SICI\)1098-108X\(199712\)22:4<339::AID-EAT1>3.0.CO;2-N](https://doi.org/10.1002/(SICI)1098-108X(199712)22:4<339::AID-EAT1>3.0.CO;2-N)
- Strother, E., Lemberg, R., Stanford, S. C., & Turberville, D. (2012). Eating disorders in men: Underdiagnosed, undertreated, and misunderstood. *Eating Disorders*, 20(5), 346–355. <https://doi.org/10.1080/10640266.2012.715512>
- 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>
- Sun, W. M., Read, N. W., Prior, A., Daly, J.-A., Cheah, S. K., & Grundy, D. (1990). Sensory and motor responses to rectal distention vary according to rate and pattern of balloon inflation. *Gastroenterology*, 99(4), 1008–1015. [https://doi.org/10.1016/0016-5085\(90\)90620-G](https://doi.org/10.1016/0016-5085(90)90620-G)
- Sun, X., Zhu, C., & So, S. H. W. (2017). Dysfunctional metacognition across psychopathologies: A meta-analytic review. *European Psychiatry*, 45, 139–153. <https://doi.org/10.1016/j.eurpsy.2017.05.029>
- Surgenor, L. J., Maguire, S., Russell, J., & Touyz, S. (2007). Self-liking and self-competence: Relationship to symptoms of anorexia nervosa. *European Eating Disorders Review*, 15(2), 139–145. <https://doi.org/10.1002/erv.734>
- Suschinsky, K. D., & Lalumière, M. L. (2012). Is Sexual Concordance Related to Awareness of Physiological States? *Archives of Sexual Behavior*, 41(1), 199–208. <https://doi.org/10.1007/s10508-012-9931-9>
- Swinbourne, J. M., & Touyz, S. W. (2007). The co-morbidity of eating disorders and anxiety disorders: A review. *European Eating Disorders Review*, 15(4), 253–274. <https://doi.org/10.1002/erv.784>

- Szmukler, G. I., Young, G. P., Lichtenstein, M., & Andrews, J. T. (1990). A serial study of gastric emptying in anorexia nervosa and bulimia. *Australian and New Zealand Journal of Medicine*, 20(3), 220–225. <https://doi.org/10.1111/j.1445-5994.1990.tb01023.x>
- Tan, G., Dao, T. K., Farmer, L., Sutherland, R. J., & Gevirtz, R. (2011). Heart Rate Variability (HRV) and Posttraumatic Stress Disorder (PTSD): A Pilot Study. *Applied Psychophysiology and Biofeedback*, 36(1), 27–35. <https://doi.org/10.1007/s10484-010-9141-y>
- Tchanturia, K., Davies, H., Harrison, A., Fox, J. R. E., Treasure, J., & Schmidt, U. (2012). Altered social hedonic processing in eating disorders. *International Journal of Eating Disorders*, 45(8), 962–969. <https://doi.org/10.1002/eat.22032>
- Tchanturia, K., Davies, H., Roberts, M., Harrison, A., Nakazato, M., Schmidt, U., Treasure, J., & Morris, R. (2012). Poor Cognitive Flexibility in Eating Disorders: Examining the Evidence using the Wisconsin Card Sorting Task. *PLOS ONE*, 7(1), e28331. <https://doi.org/10.1371/journal.pone.0028331>
- Tchanturia, K., Lloyd, S., & Lang, K. (2013). Cognitive remediation therapy for anorexia nervosa: Current evidence and future research directions. *International Journal of Eating Disorders*, 46(5), 492–495. <https://doi.org/10.1002/eat.22106>
- Tenconi, E., Meregalli, V., Buffa, A., Collantoni, E., Cavallaro, R., Meneguzzo, P., & Favaro, A. (2023). Belief Inflexibility and Cognitive Biases in Anorexia Nervosa—The Role of the Bias against Disconfirmatory Evidence and Its Clinical and Neuropsychological Correlates. *Journal of Clinical Medicine*, 12(5), 1746. <https://doi.org/10.3390/jcm12051746>
- Titova, O. E., Hjorth, O. C., Schiöth, H. B., & Brooks, S. J. (2013). Anorexia nervosa is linked to reduced brain structure in reward and somatosensory regions: A meta-analysis of VBM studies. *BMC Psychiatry*, 13(1), 110. <https://doi.org/10.1186/1471-244X-13-110>
- Todd, J. (2020). *Bodies from the inside out: An investigation of the relationships between interoception and body image* [Doctoral, Anglia Ruskin University]. <https://arro.anglia.ac.uk/id/eprint/706752/>
- Todd, J., Aspell, J. E., Barron, D., & Swami, V. (2019). Multiple dimensions of interoceptive awareness are associated with facets of body image in British adults. *Body Image*, 29, 6–16. <https://doi.org/10.1016/j.bodyim.2019.02.003>

- Todd, J., Aspell, J. E., Barron, D., Toh, E. K. L., Zahari, H. S., Mohd. Khatib, N. A., Laughton, R., & Swami, V. (2020). Greater gastric interoception is associated with more positive body image: Evidence from adults in Malaysia and the United Kingdom. *Body Image*, 34, 101–111. <https://doi.org/10.1016/j.bodyim.2020.05.011>
- Todd, J., Cardellicchio, P., Swami, V., Cardini, F., & Aspell, J. E. (2021). Weaker implicit interoception is associated with more negative body image: Evidence from gastric-alpha phase amplitude coupling and the heartbeat evoked potential. *Cortex*, 143, 254–266. <https://doi.org/10.1016/j.cortex.2021.07.006>
- Toma, T. P., Zighelboim, J., Phillips, S. F., & Talley, N. J. (1996). Methods for studying intestinal sensitivity and compliance: In vitro studies of balloons and a barostat. *Neurogastroenterology & Motility*, 8(1), 19–28. <https://doi.org/10.1111/j.1365-2982.1996.tb00238.x>
- Tomaka, J., Blascovich, J., & Swart, L. (1994). Effects of vocalization on cardiovascular and electrodermal responses during mental arithmetic. *International Journal of Psychophysiology*, 18(1), 23–33. [https://doi.org/10.1016/0167-8760\(84\)90012-6](https://doi.org/10.1016/0167-8760(84)90012-6)
- Toray, T., & Cooley, E. (1997). Weight Fluctuation, Bulimic Symptoms, and Self-Efficacy for Control of Eating. *The Journal of Psychology*, 131(4), 383–392. <https://doi.org/10.1080/00223989709603524>
- Torres, S., Guerra, M. P., Lencastre, L., Miller, K., Vieira, F. M., Roma-Torres, A., Brandão, I., & Costa, P. (2015). Alexithymia in anorexia nervosa: The mediating role of depression. *Psychiatry Research*, 225(1), 99–107. <https://doi.org/10.1016/j.psychres.2014.10.023>
- Trace, S. E., Baker, J. H., Peñas-Lledó, E., & Bulik, C. M. (2013). The Genetics of Eating Disorders. *Annual Review of Clinical Psychology*, 9(1), 589–620. <https://doi.org/10.1146/annurev-clinpsy-050212-185546>
- Treasure, J., & Schmidt, U. (2013). The cognitive-interpersonal maintenance model of anorexia nervosa revisited: A summary of the evidence for cognitive, socio-emotional and interpersonal predisposing and perpetuating factors. *Journal of Eating Disorders*, 1(1), 13. <https://doi.org/10.1186/2050-2974-1-13>
- Tsakiris, M., & Critchley, H. (2016). Interoception beyond homeostasis: Affect, cognition and mental health. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1708), 20160002. <https://doi.org/10.1098/rstb.2016.0002>

- Tylka, T. L., & Wood-Barcalow, N. L. (2015). What is and what is not positive body image? Conceptual foundations and construct definition. *Body Image, 14*, 118–129.
<https://doi.org/10.1016/j.bodyim.2015.04.001>
- Uchôa, F. N. M., Uchôa, N. M., Daniele, T. M. da C., Lustosa, R. P., Garrido, N. D., Deana, N. F., Aranha, Á. C. M., & Alves, N. (2019). Influence of the Mass Media and Body Dissatisfaction on the Risk in Adolescents of Developing Eating Disorders. *International Journal of Environmental Research and Public Health, 16*(9), Article 9.
<https://doi.org/10.3390/ijerph16091508>
- Uher, R., Brammer, M. J., Murphy, T., Campbell, I. C., Ng, V. W., Williams, S. C. R., & Treasure, J. (2003). Recovery and chronicity in anorexia nervosa: Brain activity associated with differential outcomes. *Biological Psychiatry, 54*(9), 934–942.
[https://doi.org/10.1016/S0006-3223\(03\)00172-0](https://doi.org/10.1016/S0006-3223(03)00172-0)
- Uher, R., Murphy, T., Brammer, M. J., Dalgleish, T., Phillips, M. L., Ng, V. W., Andrew, C. M., Williams, S. C. R., Campbell, I. C., & Treasure, J. (2004). Medial Prefrontal Cortex Activity Associated With Symptom Provocation in Eating Disorders. *American Journal of Psychiatry, 161*(7), 1238–1246.
<https://doi.org/10.1176/appi.ajp.161.7.1238>
- Unal, A., Altug, F., Erden, A., Cavlak, U., & Senol, H. (2021). Validity and reliability of the Body Awareness Questionnaire in patients with non-specific chronic low back pain. *Acta Neurologica Belgica, 121*(3), 701–705. <https://doi.org/10.1007/s13760-020-01399-y>
- Van Camp, L. S. C., Sabbe, B. G. C., & Oldenburg, J. F. E. (2017). Cognitive insight: A systematic review. *Clinical Psychology Review, 55*, 12–24.
<https://doi.org/10.1016/j.cpr.2017.04.011>
- Van der Ploeg, H. M. (1980). Validity of the Zelf-Beoordelings-Vragenlijst (A Dutch version of the Spielberger State-Trait Anxiety Inventory). *Nederlands Tijdschrift Voor de Psychologie En Haar Grensgebieden, 35*, 243–249.
- van Dyck, Z., Schulz, A., Blechert, J., Herbert, B. M., Lutz, A. P. C., & Vögele, C. (2021). Gastric interoception and gastric myoelectrical activity in bulimia nervosa and binge-eating disorder. *International Journal of Eating Disorders, 54*(7), 1106–1115.
<https://doi.org/10.1002/eat.23291>
- van Dyck, Z., Vögele, C., Blechert, J., Lutz, A. P. C., Schulz, A., & Herbert, B. M. (2016). The Water Load Test As a Measure of Gastric Interoception: Development of a Two-

- Stage Protocol and Application to a Healthy Female Population. *PLOS ONE*, 11(9), e0163574. <https://doi.org/10.1371/journal.pone.0163574>
- Van Rijsoort, S., Emmelkamp, P., & Vervaeke, G. (1999). The Penn State Worry Questionnaire and the Worry Domains Questionnaire: Structure, reliability and validity. *Clinical Psychology & Psychotherapy*, 6(4), 297–307. [https://doi.org/10.1002/\(SICI\)1099-0879\(199910\)6:4<297::AID-CPP206>3.0.CO;2-E](https://doi.org/10.1002/(SICI)1099-0879(199910)6:4<297::AID-CPP206>3.0.CO;2-E)
- Verdejo-Garcia, A., Clark, L., & Dunn, B. D. (2012). The role of interoception in addiction: A critical review. *Neuroscience & Biobehavioral Reviews*, 36(8), 1857–1869. <https://doi.org/10.1016/j.neubiorev.2012.05.007>
- Vervaeke, M., Puttevils, L., Hoekstra, R. H. A., Fried, E., & Vanderhasselt, M. (2021). Transdiagnostic vulnerability factors in eating disorders: A network analysis. *European Eating Disorders Review*, 29(1), 86–100. <https://doi.org/10.1002/erv.2805>
- Viltart, O., Duriez, P., & Tolle, V. (2018). Metabolic and neuroendocrine adaptations to undernutrition in anorexia nervosa: From a clinical to a basic research point of view. *Hormone Molecular Biology and Clinical Investigation*, 36(1). <https://doi.org/10.1515/hmbci-2018-0010>
- Vohs, J. L., George, S., Leonhardt, B. L., & Lysaker, P. H. (2016). An integrative model of the impairments in insight in schizophrenia: Emerging research on causal factors and treatments. *Expert Review of Neurotherapeutics*, 16(10), 1193–1204. <https://doi.org/10.1080/14737175.2016.1199275>
- Vohs, J. L., Lysaker, P. H., Liffick, E., Francis, M. M., Leonhardt, B. L., James, A., Buck, K. D., Hamm, J. A., Minor, K. S., Mehdiyou, N., & Breier, A. (2015). Metacognitive Capacity as a Predictor of Insight in First-Episode Psychosis. *Journal of Nervous & Mental Disease*, 203(5), 372–378. <https://doi.org/10.1097/NMD.0000000000000291>
- von Helmholtz, H. (1867). *Handbuch der physiologischen Optik*. Voss.
- Von Leupoldt, A., Chan, P.-Y. S., Esser, R. W., & Davenport, P. W. (2013). Emotions and Neural Processing of Respiratory Sensations Investigated With Respiratory-Related Evoked Potentials. *Psychosomatic Medicine*, 75(3), 244–252. <https://doi.org/10.1097/PSY.0b013e31828251cf>
- von Mohr, M., & Fotopoulou, A. (2018). *The cutaneous borders of interoception: Active and Social inference of pain and pleasure on the skin. The interoceptive mind: From homeostasis to awareness*. Oxford University Press.

- Wagner, A., Aizenstein, H., Mazurkewicz, L., Fudge, J., Frank, G. K., Putnam, K., Bailer, U. F., Fischer, L., & Kaye, W. H. (2008). Altered Insula Response to Taste Stimuli in Individuals Recovered from Restricting-Type Anorexia Nervosa. *Neuropsychopharmacology*, 33(3), Article 3. <https://doi.org/10.1038/sj.npp.1301443>
- Waller, G., Dickson, C., & Ohanian, V. (2002). Cognitive content in bulimic disorders: Core beliefs and eating attitudes. *Eating Behaviors*, 3(2), 171–178. [https://doi.org/10.1016/S1471-0153\(01\)00056-3](https://doi.org/10.1016/S1471-0153(01)00056-3)
- Walsh, B. T., Zimmerli, E., Devlin, M. J., Guss, J., & Kissileff, H. R. (2003). A disturbance of gastric function in bulimia nervosa. *Biological Psychiatry*, 54(9), 929–933. [https://doi.org/10.1016/S0006-3223\(03\)00176-8](https://doi.org/10.1016/S0006-3223(03)00176-8)
- Wang, G.-J., Tomasi, D., Backus, W., Wang, R., Telang, F., Geliebter, A., Korner, J., Bauman, A., Fowler, J. S., Thanos, P. K., & Volkow, N. D. (2008). Gastric distention activates satiety circuitry in the human brain. *NeuroImage*, 39(4), 1824–1831. <https://doi.org/10.1016/j.neuroimage.2007.11.008>
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54, 1063–1070. <https://doi.org/10.1037/0022-3514.54.6.1063>
- Watson, H. J., Yilmaz, Z., Thornton, L. M., Hübel, C., Coleman, J. R. I., Gaspar, H. A., Bryois, J., Hinney, A., Leppä, V. M., Mattheisen, M., Medland, S. E., Ripke, S., Yao, S., Giusti-Rodríguez, P., Hanscombe, K. B., Purves, K. L., Adan, R. A. H., Alfredsson, L., Ando, T., ... Bulik, C. M. (2019). Genome-wide association study identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa. *Nature Genetics*, 51(8), Article 8. <https://doi.org/10.1038/s41588-019-0439-2>
- Waxman, S. E. (2009). A systematic review of impulsivity in eating disorders. *European Eating Disorders Review*, 17(6), 408–425. <https://doi.org/10.1002/erv.952>
- Weerdmeester, J., Rooij, M. M. van, Engels, R. C., & Granic, I. (2020). An Integrative Model for the Effectiveness of Biofeedback Interventions for Anxiety Regulation: Viewpoint. *Journal of Medical Internet Research*, 22(7), e14958. <https://doi.org/10.2196/14958>
- Wells, A. (2001). *Emotional disorders and metacognition: Innovative cognitive therapy*. John Wiley & Sons.

- Wells, A. (2013). Advances in Metacognitive Therapy. *International Journal of Cognitive Therapy*, 6(2), 186–201. <https://doi.org/10.1521/ijct.2013.6.2.186>
- Wells, A., & Cartwright-Hatton, S. (2004). A short form of the metacognitions questionnaire: Properties of the MCQ-30. *Behaviour Research and Therapy*, 42(4), 385–396. [https://doi.org/10.1016/S0005-7967\(03\)00147-5](https://doi.org/10.1016/S0005-7967(03)00147-5)
- Wells, A., Fisher, P., Myers, S., Wheatley, J., Patel, T., & Brewin, C. R. (2012). Metacognitive therapy in treatment-resistant depression: A platform trial. *Behaviour Research and Therapy*, 50(6), 367–373. <https://doi.org/10.1016/j.brat.2012.02.004>
- Wells, A., & Matthews, G. (1994). *Attention and Emotion: A clinical perspective*. Lawrence Erlbaum Associates.
- Wells, A., & Matthews, G. (1996). Modelling cognition in emotional disorder: The S-REF model. *Behaviour Research and Therapy*, 34(11–12), 881–888. [https://doi.org/10.1016/S0005-7967\(96\)00050-2](https://doi.org/10.1016/S0005-7967(96)00050-2)
- West, M., McMaster, C. M., Staudacher, H. M., Hart, S., Jacka, F. N., Stewart, T., Loughman, A., Rocks, T., & Ruusunen, A. (2021). Gastrointestinal symptoms following treatment for anorexia nervosa: A systematic literature review. *International Journal of Eating Disorders*, 54(6), 936–951. <https://doi.org/10.1002/eat.23469>
- West, R. K., Harrison, W. J., Matthews, N., Mattingley, J. B., & Sewell, D. K. (2023). Modality independent or modality specific? Common computations underlie confidence judgements in visual and auditory decisions. *PLOS Computational Biology*, 19(7), e1011245. <https://doi.org/10.1371/journal.pcbi.1011245>
- Westmoreland, P., Krantz, M. J., & Mehler, P. S. (2016). Medical Complications of Anorexia Nervosa and Bulimia. *The American Journal of Medicine*, 129(1), 30–37. <https://doi.org/10.1016/j.amjmed.2015.06.031>
- Westwater, M. L., Mancini, F., Shapleske, J., Serfontein, J., Ernst, M., Ziauddeen, H., & Fletcher, P. C. (2021). Dissociable hormonal profiles for psychopathology and stress in anorexia and bulimia nervosa. *Psychological Medicine*, 51(16), 2814–2824. <https://doi.org/10.1017/S0033291720001440>
- Westwood, H., Stahl, D., Mandy, W., & Tchanturia, K. (2016). The set-shifting profiles of anorexia nervosa and autism spectrum disorder using the Wisconsin Card Sorting Test: A systematic review and meta-analysis. *Psychological Medicine*, 46(9), 1809–1827. <https://doi.org/10.1017/S0033291716000581>

- 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., Heiman, P., & Blackwell, B. (1977). Relation of heart rate control to heartbeat perception. *Biofeedback and Self-Regulation*, 2(4), 371–392.
<https://doi.org/10.1007/BF00998623>
- Wickham, H. (2016). Data Analysis. In H. Wickham (Ed.), *Ggplot2: Elegant Graphics for Data Analysis* (pp. 189–201). Springer International Publishing.
https://doi.org/10.1007/978-3-319-24277-4_9
- Wiens, S. (2005). Interoception in emotional experience: *Current Opinion in Neurology*, 18(4), 442–447. <https://doi.org/10.1097/01.wco.0000168079.92106.99>
- Wierenga, C. E., Bischoff-Grethe, A., Melrose, A. J., Irvine, Z., Torres, L., Bailer, U. F., Simmons, A., Fudge, J. L., McClure, S. M., Ely, A., & Kaye, W. H. (2015). Hunger Does Not Motivate Reward in Women Remitted from Anorexia Nervosa. *Biological Psychiatry*, 77(7), 642–652. <https://doi.org/10.1016/j.biopsych.2014.09.024>
- Wiklund, C. A., Rania, M., Kuja-Halkola, R., Thornton, L. M., & Bulik, C. M. (2021). Evaluating disorders of GUT-BRAIN interaction in eating disorders. *International Journal of Eating Disorders*, 54(6), 925–935. <https://doi.org/10.1002/eat.23527>
- Williams, D. (2020). Predictive coding and thought. *Synthese*, 197(4), 1749–1775.
<https://doi.org/10.1007/s11229-018-1768-x>
- Woodward, T. S., Moritz, S., Cuttler, C., & Whitman, J. C. (2006). The Contribution of a Cognitive Bias Against Disconfirmatory Evidence (BADE) to Delusions in Schizophrenia. *Journal of Clinical and Experimental Neuropsychology*, 28(4), 605–617. <https://doi.org/10.1080/13803390590949511>
- Woodward, T. S., Moritz, S., Menon, M., & Klinge, R. (2008). Belief inflexibility in schizophrenia. *Cognitive Neuropsychiatry*, 13(3), 267–277.
<https://doi.org/10.1080/13546800802099033>
- Wootton, B. M., Diefenbach, G. J., Bragdon, L. B., Steketee, G., Frost, R. O., & Tolin, D. F. (2015). A contemporary psychometric evaluation of the Obsessive Compulsive Inventory—Revised (OCI-R). *Psychological assessment*, 27(3), 874
- World Health Organization. (2022). *ICD-11: International classification of diseases (11th revision)*. <https://icd.who.int/>

- Wright, A. C., Lysaker, P. H., Fowler, D., & Greenwood, K. (2023). Clinical insight in first episode psychosis: The role of metacognition. *Journal of Mental Health*, 32(1), 78–86. <https://doi.org/10.1080/09638237.2021.1922629>
- Xiao, Y., Liu, D., Cline, M. A., & Gilbert, E. R. (2020). Chronic stress and adipose tissue in the anorexic state: Endocrine and epigenetic mechanisms. *Adipocyte*, 9(1), 472–483. <https://doi.org/10.1080/21623945.2020.1803643>
- Yang, M., Li, Z., Xu, X., Fang, D., Zou, D., Xu, G., Sun, Z., & Tu, Z. (2006). Characterization of cortical potentials evoked by oesophageal balloon distention and acid perfusion in patients with functional heartburn. *Neurogastroenterology & Motility*, 18(4), 292–299. <https://doi.org/10.1111/j.1365-2982.2006.00761.x>
- Ye, Q., Zou, F., Lau, H., Hu, Y., & Kwok, S. C. (2018). Causal Evidence for Mnemonic Metacognition in Human Precuneus. *The Journal of Neuroscience*, 38(28), 6379–6387. <https://doi.org/10.1523/JNEUROSCI.0660-18.2018>
- Yetisen, A. K., Martinez-Hurtado, J. L., Ünal, B., Khademhosseini, A., & Butt, H. (2018). Wearables in Medicine. *Advanced Materials*, 30(33), 1706910. <https://doi.org/10.1002/adma.201706910>
- Yokoyama, O., Miura, N., Watanabe, J., Takemoto, A., Uchida, S., Sugiura, M., Horie, K., Sato, S., Kawashima, R., & Nakamura, K. (2010). Right frontopolar cortex activity correlates with reliability of retrospective rating of confidence in short-term recognition memory performance. *Neuroscience Research*, 68(3), 199–206. <https://doi.org/10.1016/j.neures.2010.07.2041>
- Yon, D., & Frith, C. (2021). *Precision and the Bayesian Brain*. 31, 1017–1034.
- 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>
- Zerwas, S., Lund, B. C., Von Holle, A., Thornton, L. M., Berrettini, W. H., Brandt, H., Crawford, S., Fichter, M. M., Halmi, K. A., Johnson, C., Kaplan, A. S., La Via, M., Mitchell, J., Rotondo, A., Strober, M., Woodside, D. B., Kaye, W. H., & Bulik, C. M. (2013). Factors associated with recovery from anorexia nervosa. *Journal of Psychiatric Research*, 47(7), 972–979. <https://doi.org/10.1016/j.jpsychires.2013.02.011>

- Zipfel, S., Giel, K. E., Bulik, C. M., Hay, P., & Schmidt, U. (2015). Anorexia nervosa: Aetiology, assessment, and treatment. *The Lancet Psychiatry*, 2(12), 1099–1111. [https://doi.org/10.1016/S2215-0366\(15\)00356-9](https://doi.org/10.1016/S2215-0366(15)00356-9)
- Zipfel, S., Sammet, I., Rapps, N., Herzog, W., Herpertz, S., & Martens, U. (2006). Gastrointestinal disturbances in eating disorders: Clinical and neurobiological aspects. *Autonomic Neuroscience*, 129(1), 99–106. <https://doi.org/10.1016/j.autneu.2006.07.023>
- Zipfel, S., Seibel, M. J., Löwe, B., Beumont, P. J., Kasperk, C., & Herzog, W. (2001). Osteoporosis in Eating Disorders: A Follow-Up Study of Patients with Anorexia and Bulimia Nervosa. *The Journal of Clinical Endocrinology & Metabolism*, 86(11), 5227–5233. <https://doi.org/10.1210/jcem.86.11.8050>
- Zitron-Emanuel, N., Ganel, T., Albin, E., Abbate-Daga, G., & Marzola, E. (2022). The perception of food size and food shape in anorexia nervosa. *Appetite*, 169, 105858. <https://doi.org/10.1016/j.appet.2021.105858>
- Zucker, N. L., & Bulik, C. M. (2020). On bells, saliva, and abdominal pain or discomfort: Early aversive visceral conditioning and vulnerability for anorexia nervosa. *International Journal of Eating Disorders*, 53(4), 508–512. <https://doi.org/10.1002/eat.23255>
- Zucker, N. L., Kragel, P. A., Wagner, H. R., Keeling, L., Mayer, E., Wang, J., Kang, M. S., Merwin, R., Simmons, W. K., & LaBar, K. S. (2017). The Clinical Significance of Posterior Insular Volume in Adolescent Anorexia Nervosa. *Psychosomatic Medicine*, 79(9), 1025–1035. <https://doi.org/10.1097/PSY.0000000000000510>
- Zucker, N., Moskovich, A., Bulik, C. M., Merwin, R., Gaddis, K., Losh, M., Piven, J., Wagner, H. R., & LaBar, K. S. (2013). Perception of affect in biological motion cues in anorexia nervosa. *International Journal of Eating Disorders*, 46(1), 12–22. <https://doi.org/10.1002/eat.22062>
- Zulkosky, K. (2009). Self-Efficacy: A Concept Analysis. *Nursing Forum*, 44(2), 93–102. <https://doi.org/10.1111/j.1744-6198.2009.00132.x>

Appendix A

Appendix A includes the detailed instructions participants were given at the beginning of the Heartbeat Counting Task and supplementary material from control and exploratory analyses for the experiment presented in Chapter 2.

Section 1: Supplemental Methods

Task Instructions

Experimenter to Participant: “In the task that follows I will ask you a lot of questions about heartbeats. I will ask you to estimate how many heartbeats you felt without looking at any watches or clocks, or by touching your chest/taking your pulse. Please try to stay concentrated and answer all questions truthfully. I will guide you through the process. Do you have any questions?” If the participant has no questions, proceed to the detailed instructions. If the participant has questions, first answer them and then proceed with the following: “Great, I am now going to explain the task in detail. Please sit comfortably in your chair. Once you feel comfortable, try to not move during the task. Please sit with your feet flat on the floor (i.e., legs uncrossed) and with both hands on the table in front of you and your eyes open or closed, whatever feels more comfortable to you. The instructions for this task are simple, but it is important to follow them carefully. In a moment, I will ask you to try to sense your heartbeat, without actively feeling (touching) your wrists or neck pulsations, or attempting any other physical manipulation. You may feel your heartbeats, you may not feel them at all, or you may occasionally feel some of them. You will be asked to count from the moment the experimenter prompts you with the ‘Go’ signal until you hear the ‘Stop’ signal, which will be your cue to stop counting. After you hear ‘Stop’ you will be asked to verbally report the exact number of heartbeats you felt. We will repeat this exercise several times. For this task, it is very important that you only count the heartbeats you really feel, without trying to guess. I really insist that you please only count what you feel. It could mean reporting no beat at all, some beats, or all beats that actually happened. It is also important to not talk during the trials or move. You will be given rest breaks in between the trials. Do you have any questions about these instructions?”. If the participant has no questions, proceed with taking their Prospective Performance Estimate and completing the HCT.

Section 2: Supplemental Results

Table A.1. Table for Control Analyses for Hypotheses 1 and 2 (Chapter 2).

Predictors	$\beta(SE)$	t	95% CI	p
Group Differences on Prior Prospective Performance Estimates. Controlling for Age and Average Heartrate				
Intercept	58.23(12.60)	4.62	33.25 - 83.21	<0.001
AN	-16.89(5.54)	-3.05	-27.88 - -5.91	0.003
AN-WR	-5.95(5.49)	-1.09	-16.83 - 4.92	0.28
AGE	0.15(0.29)	0.5	-0.44 - 0.73	0.618
Average Heartrate	-0.12(0.14)	-0.83	-0.40 - 0.16	0.411
R ² / Adj R ²	0.09 / 0.06			
Group Differences on Performance. Controlling for Age and Average Heartrate.				
Intercept	69.41(12.39)	5.6	44.86 - 93.97	<0.001
AN	-2.17(5.45)	-0.4	-12.97 - 8.63	0.692
AN-WR	5.01(5.42)	0.92	-5.73 - 15.76	0.357
AGE	0.29(0.29)	1.01	-0.28 - 0.87	0.314
Average Heartrate	-0.26(0.14)	-1.86	-0.54 - 0.02	0.065
R ² / Adj R ²	0.08 / 0.04			
Group Differences on Performance (25-second trial). Controlling for Age.				
Intercept	55.42(7.65)	7.25	40.30 - 70.54	<0.001
AN	1.36(4.69)	0.29	-7.91 - 10.64	0.772
AN-WR	3.38(4.84)	0.7	-6.19 - 12.94	0.486
AGE	0.26(0.27)	0.94	-0.28 - 0.79	0.350
R ² / Adj R ²	0.01 / -0.01			
Group Differences on Performance (45-second trial). Controlling for Age.				
Intercept	55.62(7.62)	7.3	40.54 - 70.69	<0.001
AN	4.48(4.66)	0.96	-4.73 - 13.70	0.338
AN-WR	5.54(4.83)	1.15	-4.02 - 15.10	0.254
AGE	0.20(0.27)	0.75	-0.33 - 0.74	0.456
R ² / Adj R ²	0.02 / -0.006			
Group Differences on Performance (65-second trial). Controlling for Age.				
Intercept	49.01(7.54)	6.5	34.09 - 63.93	<0.001
AN	1.83(4.61)	0.4	-7.28 - 10.94	0.692
AN-WR	6.61(4.77)	1.39	-2.83 - 16.05	0.168
AGE	0.34(0.27)	1.28	-0.19 - 0.87	0.204

R^2 / Adj R^2 0.03 / 0.004

Abbreviations: AN (Acute Anorexia Nervosa Group); AN-WR (Weight-Restored from Anorexia Nervosa Group); HC (Healthy Controls Group). Bolded values denote statistical significance ($p < 0.05$).

Table A.2. Analyses with Psychometric Traits as Interacting Predictors (Chapter 2).

Predictors	β (SE)	t	95% CI	p
Effect of Depression x Group on Prospective Performance Estimates				
Intercept	62.21(5.99)	10.39	50.24 - 74.17	<0.001
AN	-19.45(12.05)	-1.61	-43.54 - 4.64	0.112
AN-WR	-0.26(10.02)	-1.02	-30.29 - 9.77	0.310
Depression	-0.02(1.45)	-0.02	-2.92 - 2.87	0.987
AN x Depression	0.03(1.63)	0.02	-3.24 - 3.29	0.988
AN-WR x Depression	-0.33(1.65)	-0.2	-3.62 - 2.96	0.842
R^2 / Adj R^2	0.14 / 0.08			
Effect of Depression x Group on Performance				
Intercept	58.32(6.33)	9.22	45.67 - 70.96	<0.001
AN	19.91(12.74)	1.56	-5.55 - 45.37	0.123
AN-WR	8.60(10.59)	0.81	-12.57 - 29.77	0.420
Depression	-1.37(1.53)	-0.9	-4.44 - 1.49	0.373
AN x Depression	0.34(1.73)	0.2	-3.11 - 3.79	0.845
AN-WR x Depression	1.54(1.74)	0.89	-1.94 - 5.02	0.379
R^2 / Adj R^2	0.10 / 0.03			
Effect of Anxiety x Group on Prospective Performance Estimates				
Intercept	60.25(6.37)	9.46	47.52 - 72.98	<0.001
AN	-17.89(10.04)	-1.78	-37.96 - 2.19	0.080
AN-WR	-7.43(9.79)	-0.76	-27.00 - 12.13	0.45
Anxiety	0.99(2.34)	0.43	-3.68 - 5.66	0.672
AN x Anxiety	-0.94(2.51)	-0.37	-5.95 - 4.07	0.71
AN-WR x Anxiety	-1.60(2.51)	-0.64	-6.63 - 3.42	0.526
R^2 / Adj R^2	0.15 / 0.08			
Effect of Anxiety x Group on Performance				

Intercept	59.94(6.76)	8.87	46.43 - 73.45	<0.001
AN	12.49(10.66)	1.17	-8.82 - 33.80	0.246
AN-WR	9.62(10.39)	0.93	-11.15 - 30.39	0.358
Anxiety	-2.78(2.48)	-1.12	-7.73 - 2.18	0.267
AN x Anxiety	1.75(2.66)	0.66	-3.57 - 7.07	0.513
AN-WR x Anxiety	2.58(2.67)	0.97	-2.75 - 7.91	0.337
R ² / Adj R ²	0.098 / 0.03			

Effect of Stress x Group on Prospective Performance Estimates				
Intercept	59.83(8.28)	7.23	43.29 - 76.38	<0.001
AN	-11.77(14.82)	-0.79	-41.38 - 17.85	0.430
AN-WR	-15.43(12.65)	-1.22	-40.72 - 9.87	0.227
Stress	0.39(1.17)	0.33	-1.95 - 2.73	0.739
AN x Stress	-0.80(1.47)	-0.54	-3.74 - 2.14	0.590
AN-WR x Stress	0.11	0.07	-2.85 - 3.06	0.943
R ² / Adj R ²	0.12 / 0.08			

Effect of Anxiety x Group on Performance				
Intercept	55.75(8.92)	6.25	37.92 - 73.57	<0.001
AN	18.29(15.96)	1.15	-13.61 - 50.19	0.256
AN-WR	12.40(13.63)	0.91	-14.85 - 39.65	0.366
Stress	-0.19(1.26)	-0.15	-2.71 - 2.34	0.884
AN x Stress	-0.51(1.58)	-0.32	-3.68 - 2.65	0.747
AN-WR x Stress	0.20(1.59)	0.13	-2.98 - 3.39	0.898
R ² / Adj R ²	0.07 / -0.004			

Effect of Body Awareness (BAQ) x Group on Prospective Performance Estimates				
Intercept (HC)	17.92(24.24)	0.74	-30.52 - 66.35	0.463
AN	3.56(30.55)	0.12	-57.49 - 62.62	0.908
AN-WR	42.26(31.51)	1.34	-20.72 - 105.24	0.185
BAQ	0.34(0.29)	1.18	-0.24 - 0.93	0.244
AN x BAQ	-0.14(0.39)	-0.36	-0.91 - 0.63	0.721
AN-WR x BAQ	-0.51(0.39)	-1.28	-1.29 - 0.28	0.204
R ² / Adj R ²	0.09 / 0.02			

Effect of Body Awareness (BAQ) x Group on Performance				
Intercept (HC)	55.52(22.63)	2.45	10.30 - 100.73	0.017

AN	2.07(28.74)	0.07	-55.37 - 59.51	<i>0.943</i>
AN-WR	3.42(29.98)	0.11	-56.48 - 63.33	<i>0.909</i>
BAQ	0.11(0.27)	0.41	-0.43 - 0.66	<i>0.682</i>
AN x BAQ	-0.07(0.36)	-0.18	-0.79 - 0.66	<i>0.858</i>
AN-WR x BAQ	-0.05(0.37)	-0.14	-0.80 - 0.69	<i>0.886</i>
R ² / Adj R ²	0.01 / -0.07			
Effect of Alexithymia (TAS-20) x Group on Prospective Performance Estimates				
Intercept (HC)	73.73(14.71)	5.01	44.64 - 102.82	<i><0.001</i>
AN	-9.86(18.74)	-0.53	-46.93 - 27.21	<i>0.600</i>
AN-WR	-11.74(17.74)	-0.66	-46.83 - 23.35	<i>0.509</i>
TAS-20	-0.53(0.38)	-1.41	-1.27 - 0.21	<i>0.162</i>
AN x TAS-20	0.14(0.42)	0.33	-0.69 - 0.96	<i>0.743</i>
AN-WR x TAS	0.25(0.42)	0.59	-0.59 - 1.09	<i>0.558</i>
R ² / Adj R ²	0.13 / 0.09			
Effect of Alexithymia (TAS-20) x Group on Performance				
Intercept (HC)	81.23(14.38)	5.65	52.78 - 109.68	<i><0.001</i>
AN	-8.01(18.46)	-0.43	-44.52 - 28.50	<i>0.665</i>
AN-WR	-3.51(17.48)	-0.2	-38.08 - 31.07	<i>0.841</i>
TAS-20	-0.55(0.37)	-1.5	-1.27 - 0.17	<i>0.135</i>
AN x TAS-20	0.38(0.41)	0.939	-0.42 - 1.19	<i>0.349</i>
AN-WR x TAS-20	0.30(0.41)	0.733	-0.52 - 1.12	<i>0.465</i>
R ² / Adj R ²	0.05 / 0.01			

N.B. Depression, Anxiety, and Stress were measured using the DASS-21 Scale; body awareness was measured using the BAQ (Body Awareness Questionnaire; Shields et al., 1989); alexithymia was measured using the TAS-20 (20-Item Toronto Alexithymia Scale; Bagby et al., 1994). Abbreviations: AN (Acute Anorexia Nervosa Group); AN-WR (Weight-Restored from Anorexia Nervosa Group); HC (Healthy Controls Group). Bolded values denote statistical significance ($p < 0.05$).

Table A.3. Effects of Illness on Prospective Performance Estimates and Performance (Chapter 2).

Predictors	β (SE)	t	95% CI	p

Effect of Illness Duration on Prospective Performance Estimates (in the AN group only)				
Intercept	36.58(4.60)	7.95	27.32 - 45.84	<0.001
Illness Duration	0.23(0.37)	0.629	-0.51 - 0.98	0.532
R ² / Adj R ²	0.008 / -0.013			
Effect of Illness Duration on Performance (in the AN group only)				
Intercept	61.77(35.1)	17.59	54.71 - 68.84	<0.001
Illness Duration	0.12(0.28)	0.45	-0.43 - 0.68	0.656
R ² / Adj R ²	0.004 / -0.02			
Effect of Illness Severity on Prospective Performance Estimates (in the AN group only)				
Intercept	26.56(7.75)	3.43	54.71 - 68.84	0.001
Illness Severity	0.25(0.13)	1.95	-0.43 - 0.68	0.058
R ² / Adj R ²	0.08 / 0.06			
Effect of Illness Severity on Performance (in the AN group only)				
Intercept	62.94(5.92)	10.62	51.02 - 74.86	<0.001
Illness Severity	0.01(0.10)	0.14	-0.18 - 0.21	0.888
R ² / Adj R ²	0.00 / -0.02			
Effect of Illness Duration on Prospective Performance Estimates (in the AN-WR group only)				
Intercept	41.64(5.24)	7.95	31.03 - 52.25	<0.001
Illness Duration	1.38(0.88)	1.57	-0.41 - 3.16	0.126
R ² / Adj R ²	0.06 / 0.04			
Effect of Illness Duration on Performance (in the AN-WR group only)				
Intercept	65.57(5.24)	12.52	54.94 - 76.19	<0.001
Illness Duration	0.31(0.88)	0.36	-1.47 - 2.10	0.724
R ² / Adj R ²	0.004 / -0.024			
Effect of Illness Severity on Prospective Performance Estimates (in the AN-WR group only)				
Intercept	44.89(7.81)	7.45	29.05 - 60.74	<0.001
Illness Severity	0.05(0.11)	0.46	-0.17 - 0.27	0.650
R ² / Adj R ²	0.006 / -0.022			
Effect of Illness Severity on Performance (in the AN-WR group only)				
Intercept	70.27(8.07)	8.70	53.87 - 86.66	<0.001
Illness Severity	-0.03(0.11)	-0.30	-0.26 - 0.19	0.763
R ² / Adj R ²	0.003 / -0.03			

Abbreviations: AN (Acute Anorexia Nervosa Group); AN-WR (Weight-Restored from Anorexia Nervosa Group). Bolded values denote statistical significance ($p < 0.05$).

Table A.4. Effects of Insight into Illness on Prospective Performance Estimates and Performance (Chapter 2).

Predictors	$\beta(SE)$	t	95% CI	p
Effect of Illness Severity Error on Prospective Performance Estimates				
Intercept	52.52(8.00)	6.57	36.05 - 68.99	<0.001
Severity Error Score	-3.37(2.25)	-1.5	-8.01 - 1.27	0.147
R ² / Adj R ²	0.08 / 0.05			
Effect of Illness Severity Error on Performance				
Intercept	0.63(0.06)	10.18	0.50 - 0.76	<0.001
Severity Error Score	0.01(0.02)	0.35	-0.03 - 0.04	0.733
R ² / Adj R ²	0.005 / -0.035			
Effect of Health Consequences Error Score on Prospective Performance Estimates				
Intercept	42.96(5.51)	7.71	31.12 - 53.80	<0.001
Health Consequences	0.26(2.10)	0.12	-4.08 - 4.59	0.903
R ² / Adj R ²	0.001 / -0.039			
Effect of Health Consequences Error Score on Performance				
Intercept	0.68(0.04)	17.86	0.61 - 0.76	<0.001
Health Consequences	-0.03(0.01)	-1.92	-0.06 - 0.00	0.066
R ² / Adj R ²	0.13 / 0.09			
Effect of Future Perspectives on Prospective Performance Estimates				
Intercept	57.57(16.90)	3.41	22.76 - 92.37	0.002
Fear	-1.84(2.02)	-0.01	-5.99 - 2.31	0.370
R ² / Adj R ²	0.03 / -0.006			
Effect of Future Perspectives on Performance				
Intercept	0.82(0.12)	6.67	0.57 - 1.07	<0.001
Fear	-0.02(0.01)	-1.45	-0.05 - 0.01	0.160
R ² / Adj R ²	0.08 / 0.04			
N.B. For illness severity and health consequences an error score was computed by subtracting patients' scores from the clinician's respective score on that item. Error scores could range from				

-10 to +10, with a score of -10 indicating negative bias (i.e., patients see their condition and the effects as more severe than the clinician), a score of 0 indicating no bias between patient and clinician perspective, and a score of +10 indicating positive bias (i.e., patients see their condition and the effects as less severe than the clinician). In the future perspectives the average of the two (fear of losing control over eating behaviour in the future and hopes for illness severity improving in the future) items was used (patient scores only used here, and the rating on the fear item (item number 5, see Table 2.2 in Chapter 2 was reverse scored). For the 'Health Consequences' dimension an average of the two error scores from the two items in that dimension was calculated.

Appendix B

Appendix B includes additional information for the Methods and Results sections control for the experiment presented in Chapter 3. Specifically, the first section provides details on the secondary aims and hypotheses and respective data analyses plan. The second section presents the findings from the additional analyses on the main measures, the output from the analyses on the secondary hypotheses and measures, and the output from the exploratory analyses.

Section 1: Supplemental Methods

Secondary Hypotheses, Measures and Analysis

In a secondary set of hypotheses and analyses, the interest was in examining how the three groups differed in their beliefs on their ability prior to completing the HCT. To do this, the groups' Prior Prospective Self-Efficacy Beliefs were compared, expecting that the beliefs of the AN and AN-WR groups would be more 'pessimistic' about their abilities to perform well in the Heartbeat Counting Task (HCT; Schandry, 1981) as indexed by their lower values as compared to the HCs, and as seen in Chapter 2. This was examined by running a multilevel modelling analysis (MLM) to assess whether Group predicted differences in Prior Prospective Self-Efficacy Beliefs, while controlling for Age, and adding Study Site as a random effect. Although no significant group differences were found on Prior Prospective Self-Efficacy Beliefs (see below), I explored whether traits and behaviours often encountered in populations with AN explained the pattern of the results, in order to also better understand the prospective pessimistic beliefs of the clinical populations (i.e., Posterior Prospective Self-Efficacy Beliefs).

After participants completed the HCT, they also provided a Posterior Retrospective Self-Efficacy Belief, where they were asked to assess how well they were able to feel their heartbeats. The respective analysis examined whether there would be group differences in Performance estimation retrospectively, expecting that the two clinical groups would misestimate their Performance in the HCT more than HCs as indexed by the greater discrepancy between their Posterior Retrospective Self-Efficacy Beliefs relative to their Performance. To do this, a Performance Error Percentage score was calculated by multiplying by 100 the Performance-weighted difference between Posterior Retrospective Self-Efficacy Beliefs and Performance. To evaluate between-group differences the logarithm of the Performance Error Percentage was used as the outcome variable (although non-preregistered initially, values were logarithmically transformed because the outcome variable followed a lognormal distribution; see below). Performance Error Percentage data were not normally

distributed and upon evaluation, the discrepancy between participants' beliefs and their actual performance was better distributed (Figures B.4-B.6). Hence, although not initially preregistered it was decided to also examine the difference between Beliefs and Performance (outcome variable) between the three groups (predictor variable) and within each group, separately.

Then, the interest was in evaluating whether the two clinical (AN and AN-WR) groups underestimated Performance in the cardiac interoception task, and specifically whether they would show poorer interoceptive awareness (correspondence between Performance and Performance Confidence Estimates) in comparison to HCs. It was predicted that the two clinical groups would show poorer interoceptive awareness in comparison to HCs. These differences were examined by running Pearson's correlations between participants' Performance and Performance Confidence – correlations were run across groups, and within each group separately. However, these analyses could not indicate the extent to which bottom-up interoceptive signals (e.g., cardiac signals during HCT) with regards to top-down prior beliefs, and interoceptive traits influence the precision of metacognitive, prospective beliefs (here, Posterior Prospective Self-Efficacy Belief). Hence, to elucidate to what extent the discrepancy between Performance and interoceptive sensibility (subjective measure of one's ability to attend to interoceptive signals; here, measured via the interoceptive deficits subscale from the EDI-3; (EDI-3-ID; Garner, 2004) explain interoceptive awareness, an interoceptive trait prediction error (ITPE) z-score was obtained (here, calculated as the difference between the Performance z-score and the EDI-3-ID z-score; see Garfinkel et al., 2016 for details). It was predicted that the AN and AN-WR groups would underestimate their Performance in the cardiac interoception task in comparison to HCs as indexed by their more negative ITPE values. The ITPE score was then converted into a z-score, with a positive value suggesting that participants overestimated their interoceptive abilities, whilst a negative one suggested an underestimation of one's own interoceptive abilities. Then, a linear regression was run to examine whether group predicted ITPE z-score differences.

After each belief (e.g., Prior Prospective Self-Efficacy Belief) and heartbeat counting trial, participants were asked to report how confident they were in the accuracy of their answer. I then examined whether the three groups differed in confidence throughout different experimental stages, and also used the confidence estimates after each heartbeat counting trial (Performance Confidence) in the aforementioned Bayesian Belief Updating and interoceptive

awareness analyses (see Sections 3.2 and 3.3, Chapter 3). Given findings of low confidence on the HCT in a previously reported AN sample (Kinnaird et al., 2020) it was expected that the clinical groups would be less confident in the accuracy of their answers throughout the experiment, especially during the HCT task and in their Posterior Retrospective Self-Efficacy Beliefs. To examine this, the same analyses that were performed for Prior and Posterior Prospective Self-Efficacy Beliefs were also performed for the Prior Prospective, Performance, and Posterior Prospective Confidence to assess between-group differences.

Then, a series of control analyses on the primary measures were performed. Given criticisms on the effects of time awareness in HCT performance (Brener & Ring, 2016), a control task was added; during this task participants, for the same time intervals as in the HCT, estimated how much time they thought had passed, rather than how many heartbeats they felt. Additional control measures included a 5-min baseline measurement of heart rate variability (HRV; calculated as the root mean square (RMSSD) of the difference in the time elapsed between two consecutive inter-beat intervals), and beliefs about typical heartbeats of oneself and of the general population were also sampled (by asking them to report how many heartbeats they think they have/the general population has per minute when at rest), for the same control purposes based on previous literature showing that individual differences in both time estimation and knowledge about heartbeats may influence performance on heartbeat counting (Brener & Ring, 2016; Knapp-Kline & Kline, 2005). Then an analysis was conducted using Performance as the outcome variable, Group as the predictor variable, and Time Accuracy, knowledge about heartbeats in the general population and in the self as control variables. Time accuracy was calculated using the same formula that used for Performance (see equation 3.1, Chapter 3) and replacing actual recorded heartbeats with the duration of the respective time trial (in sec) and the estimated, felt heartbeats with participants' estimation of the time elapsed (in sec). The Self-Heartbeat Accuracy score was calculated as the difference between the samples' average heartbeat and the participant's response regarding their own average heart rate when at rest, weighted by the sample average. The General-Heartbeat Accuracy score was calculated in the same way but replacing the sample average heartbeat with the general population average heartbeat scores when at rest (80) and using participants' response on the average heart rate of the population when at rest, instead of their own. An analysis was also conducted on group differences in HRV but given the lack of significant findings, HRV was not added in any further analyses (see below). Then, additional, control analyses for Performance in each time interval separately were run to ensure that overall Performance

results were not exclusively driven by any one of the intervals. Preliminary analyses to see if BMI was different between groups were also carried out, as BMI was a group-categorisation criterion. As expected, BMI was significantly lower in the AN vs. HC group, the differences between the BMI of the AN-WR and HC groups were not significant. Given criticisms on the effects of BMI on interoceptive abilities (Richard et al., 2019), within-group tests were run to examine potential effects of BMI on Performance (i.e., interoceptive accuracy).

Furthermore, to exclude the possibility that difficulties in updating prospective self-efficacy beliefs relate to a domain-general cognitive inflexibility, I also tested cognitive flexibility and particularly set-shifting using the Wisconsin Card Sorting Task (WCST; Grant & Berg, 1948) in a subset of the participants ($N_{AN}=11$; $N_{AN-WR}=20$; $N_{HC}=34$). Moreover, to account for potential effects of cognitive inflexibility on learning rates and beliefs, first a between-group differences analysis was performed on the percentage of correct WCST scores, and an MLM was run with learning rate (the one that used Performance Confidence as a precision proxy) as the outcome variable, Group as the predictor variable and the WCST as the random effect. This analysis was repeated using participants' percentage of preservative responses; that is, arranging the cards on the basis of former trial criteria (see Arbel et al., 2013).

Additionally, various exploratory analyses were carried out to examine the effect of clinical measures and traits on the main variables. To test the validity of the experimental measure of Posterior Prospective Self-Efficacy Beliefs about interoception, in the clinical populations only (both AN and AN-WR groups), their relation to prospective, everyday beliefs about patients' AN symptoms was also assessed. Specifically, I explored how everyday fears and hopes (e.g., *"Do you believe that the severity of your condition will improve in the future?"*; *"Are you fearful of gaining weight?"*) from a questionnaire developed by the present research team (see Table 3.3), predicted (Prior and Posterior) Prospective Self-Efficacy Beliefs in the AN and AN-WR groups (separately). The first set of analyses used subscale scores (hopes and fears, separately) as standalone IVs, and in the second set of analyses the total score (average of the four items from the hopes and fears subscales) was used. I used patient scores only on future hopes and fears as the predictor variables and self-efficacy beliefs as the outcome variables. Given that these future hopes and fears items are similar to some EDI-3 items (Garner, 2004), exploratory correlations between standalone item scores of the EDI-3 and Posterior Prospective Self-Efficacy Beliefs were also carried out (only in the two clinical groups, separately: Table B.12).

Finally, deficits into insight have often been reported in the AN population; items from the insight into illness scale developed by the present research team (see Table 3.3; Chapter 3) were used to explore the potential correlations between aspects of insight and self-efficacy beliefs, Performance and learning rates in the AN group only. The scoring of the items, the analyses and results are reported in more detail below.

Section 2: Supplemental Results

The output from the main analyses, as well as the output of some key secondary measures (namely Prior and Retrospective Self-Efficacy Beliefs, and Interoceptive Awareness) is presented below.

Main Measures (Performance and Posterior Prospective Self-Efficacy Beliefs) – Control and Exploratory Analyses

Additional analyses to examine whether Performance was influenced by participant ID and study site, BMI, time awareness, knowledge about heartbeats and heart-rate variability (HRV) are presented in Table B1

A multiple linear regression was run, using Performance as the outcome variable, Group, Time Accuracy, and Self- and General-Heartbeat Accuracy scores as the predictor variables. However, no significant effects were found (Table B.1a). I then examined whether Performance differed in each time interval separately but found no significant differences (Table B.1b). Next, BMI was not added as an IV in the main MLM for between-group difference on Performance given that it was a criterion for grouping participants (as expected the AN group had a significantly lower BMI than the HCs; Table B.1c), but three separate regressions were conducted to examine whether BMI affected Performance within each group, separately. None of the three within-group analyses yielded significant results, suggesting that in the present cohort, Performance was not significantly affected by BMI (Table B.1d). Finally, no significant between-group HRV differences were found (Table B.1e), thus HRV was not added as a control variable in the Performance analysis.

Table B.1. Additional and Control Analyses for Performance (Chapter 3).

Table B.1a: Performance (IAcc) with participant ID and Study Site as Random Effects				
Predictors	$\beta(SE)$	t	CI	p
(Intercept)	56.18(5.77)	9.74	44.78 – 67.59	<0.001
AN	-0.89(6.98)	-0.13	-14.68 – 12.91	0.899

AN-WR	-2.48(0.08)	-0.41	-14.49 – 9.53	0.684
Time Accuracy	-0.13(0.07)	-1.92	-0.26 – 0.00	0.057
Self-Heartbeat Accuracy	-20.15(14.45)	-1.40	-48.72 – 8.41	0.156
General-Heartbeat Accuracy	-5.99(12.18)	-0.49	-30.08 – 18.10	0.624

Table B.1b: Group Differences in Performance Per Trial Duration

(Intercept) ¹	0.40(0.03)	13.41	0.34 – 0.46	<0.001
AN	0.10(0.06)	1.59	-0.02 – 0.22	0.113
AN-WR	-0.02(0.06)	-0.40	-0.14 – 0.10	0.693

¹This analysis is for the 25sec long trial

(Intercept) ²	0.37(0.03)	11.49	0.30 – 0.43	<0.001
AN	-0.09(0.09)	-0.92	-0.27 – 0.10	0.360
AN-WR	0.10(0.07)	1.35	-0.05 – 0.24	0.181

²This analysis is for the 30sec long trial

(Intercept) ³	0.40(0.03)	14.38	0.35 – 0.46	<0.001
AN	0.07(0.06)	1.17	-0.05 – 0.18	0.245
AN-WR	0.03(0.06)	0.47	-0.09 – 0.14	0.641

³This analysis is for the 45sec long trial

(Intercept) ⁴	0.41(0.03)	14.59	0.35 – 0.46	<0.001
AN	0.06(0.06)	1.12	-0.05 – 0.18	0.264
AN-WR	-0.004(0.06)	-0.06	-0.12 – 0.11	0.950

⁴This analysis is for the 65sec long trial

Table B.1c: BMI Differences Between Groups

HC (Intercept)	20.48(0.53)	38.30	19.41 - 21.54	<0.001
AN	-4.35(0.93)	-4.69	-6.19 - -2.51	<.001
AN-WR	-0.54(1.05)	-0.52	-2.62 - 1.53	0.604

Table B.1d: Performance (IAcc) within each group

AN	15.13(60.53)	0.25	-110.41 – 140.66	0.805
BMI	2.32(3.74)	0.62	-5.43 – 10.06	0.542
AN-WR	64.82(78.53)	0.83	-104.84 - 234.48	0.424
BMI	-1.18(3.93)	-0.3	-9.67 - 7.32	0.770
HC	42.58(18.86)	2.26	4.61 – 80.56	0.029
BMI	0.26(0.90)	0.28	-1.55 - 2.06	0.777

Table B.1e: Heart Rate Variability (HRV) Differences Between Groups

HC (Intercept)	5.92(0.27)	21.83	5.39 - 6.46	<0.001
AN	0.72(0.72)	1.01	-0.70 - 2.14	0.316

AN-WR	0.07(0.62)	-0.12	-1.30 - 1.16	0.909

N.B. For Self-Heartbeat Knowledge, participants were asked to report how many heartbeats they think they have per minute. A self-accuracy score was obtained per participant, by weighing the difference between the participant's actual heartbeats per minute and their answer with the number of heartbeats the participant had in a minute when at rest. For General-Heartbeat Knowledge participants were asked to report how many heartbeats they think the average person has per minute when at rest. A score on General-Heartbeat accuracy was obtained by weighing the difference between the actual heartbeat average in the general population and the participant's answer with the average in the general population. Given average heartbeats per minute in the general population may range from 60 to 100 beats per minute (bpm; (Sapra et al., 2023)) I used the average (i.e., 80bpm). Participant ID refers to each participant's response per trial. Abbreviations: AN (Acute Anorexia Nervosa Group); AN-WR (Weight-Restored from Anorexia Nervosa Group); HC (Healthy Controls Group). Bolded values denote statistical significance ($p < 0.05$).

Additional analyses to explore what drives the lower Posterior Prospective Self-Efficacy Beliefs of the AN group.

Another aim was to further examine the role of traits and symptoms seen in AN and their effect on Posterior Prospective Self-Efficacy Beliefs. To do that, first a correlation between the Depression, Anxiety, and Stress subscale scores from DASS-21, participants' IUS-12, OCI-R, TAS-20, and EDI-3-ID scores was run. Given the high correlation among all those measures, a principal component analysis (PCA) was performed, and I used the 7 resulting principal components in 2 separate linear regressions to examine their effect on self-efficacy beliefs. Two separate linear regressions had to be performed (one with the all 7 principal components and one with the IUS-12, OCI-R, TAS-20, and EDI-3-ID components only) given that participants from one of the sites did not complete the DASS-21 questionnaire. To further explore what drives the lower Posterior Prospective Self-Efficacy Beliefs of the clinical groups (in comparison to the HCs'), an MLM using the principal component scores as predictors and Study Side as a random effect was run. The results showed that Depression and Stress scores as measured using the DASS-21 questionnaire (Lovibond & Lovibond, 1995), but not the other psychometric traits had a significant effect on Posterior Self-Efficacy Beliefs (Table B.2a). This analysis was complemented by running non-preregistered linear regressions on Posterior Self-Efficacy Beliefs, to explore whether Depression and Stress would explain the group effects previously observed. Indeed, a significant effect was found in both analyses, suggesting that part of the observed pessimism in self-efficacy beliefs could be attributed to depression and stress. To further examine whether this was a state or trait effect Holm-corrected linear regressions within each clinical group separately were run; the Depression and Stress principal

component scores were used as the two separate predictors and the results showed a significant effect of Depression on the AN-WR group's Posterior Prospective Self-Efficacy Beliefs, but not on the AN group's, and no significant effect of Stress on neither group's Posterior Prospective Self-Efficacy Beliefs. To better understand the pattern, the data were plotted and show that as Depression and Stress scores increase, Posterior Self-Efficacy Beliefs tend to decrease (Figures B.1 and B.2). However, after a certain threshold of Depression and Stress, the Posterior Prospective Self-Efficacy Beliefs plateau.

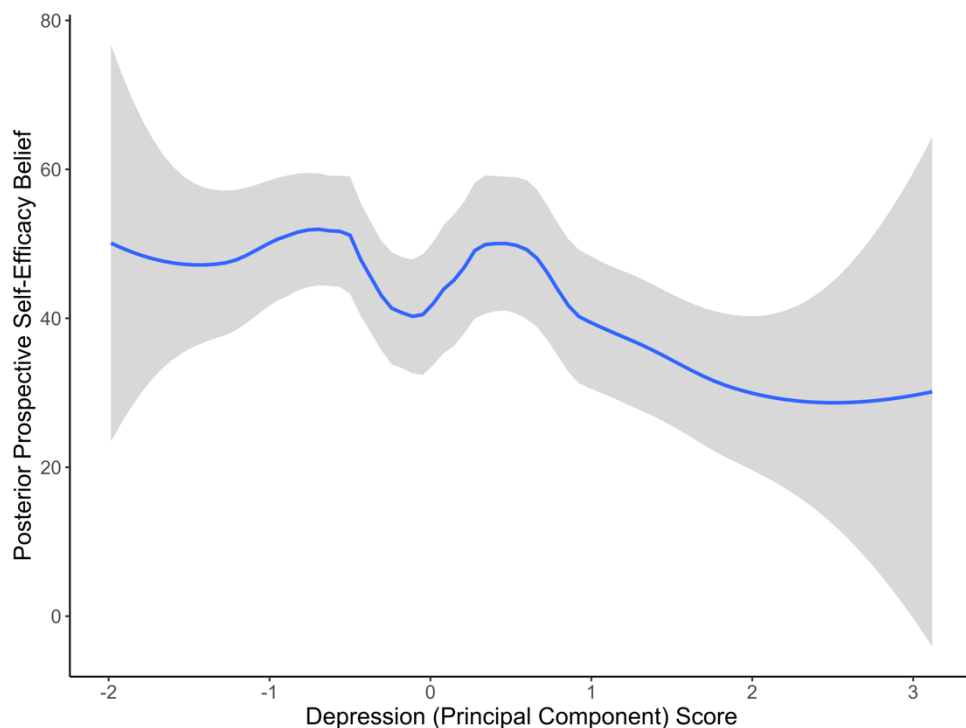


Figure B.1. Effect of Depression on Posterior Prospective Self-Efficacy Beliefs (Chapter 3).

Visual representation of the Posterior Prospective Self-Efficacy Beliefs decreasing as the Depression Scores increase, and after a certain threshold (2), the Self-Efficacy Belief values reach a plateau.

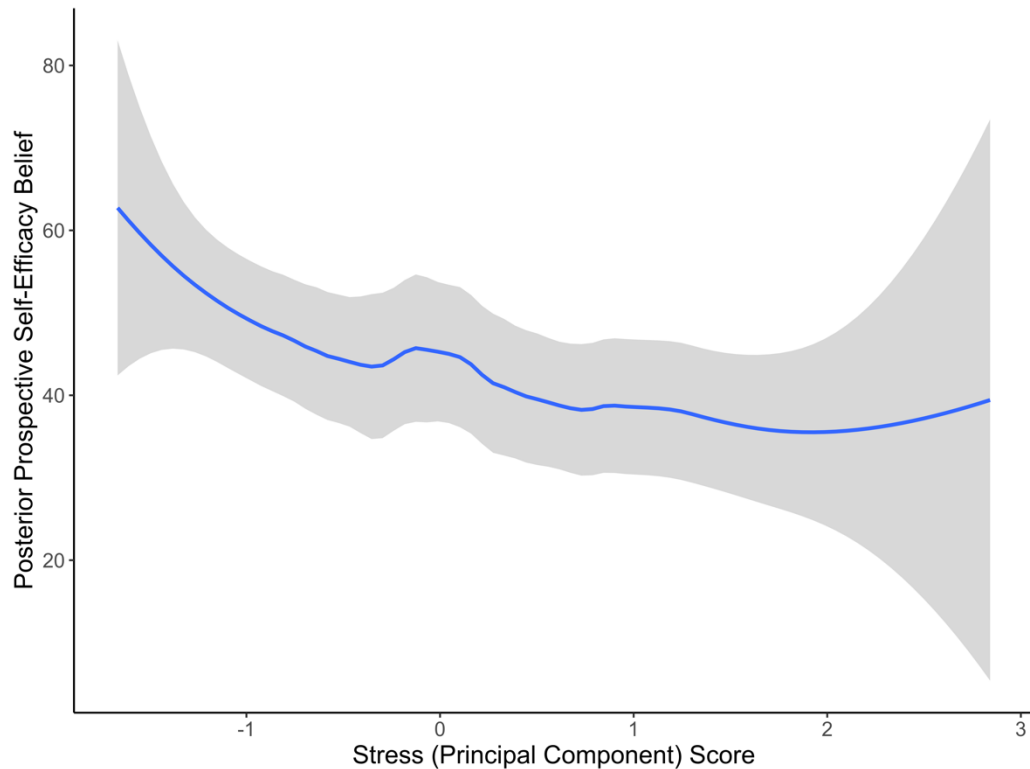


Figure B.2. Effect of Stress on Posterior Prospective Self-Efficacy Beliefs (Chapter 3).

Visual representation of the Posterior Prospective Self-Efficacy Beliefs decreasing as the Stress Scores increase, and after a certain threshold (1), the Self-Efficacy Belief values reach a plateau.

Then, non-preregistered, exploratory analyses were run to explore whether set-shifting difficulties explained the difference on the groups' Posterior Prospective Self-Efficacy Beliefs. However, as seen on Table B.2b, the effect of the WCST percentage of correct and preservative scores on participants' Posterior Prospective Self-Efficacy Beliefs was non-significant. These findings and their implications are discussed in detail in Section 3.4 (Chapter 3).

Table B.2. Additional Analyses on Posterior Prospective Self-Efficacy Beliefs (Chapter 3).

Table B.2a: Analysis Using the Depression, Stress, Anxiety, IUS-12, OCI-R, TAS-20 and EDI-3-ID Principal Components as Predictors				
Predictors	β(SE)	t	95% CI	p
(Intercept)	43.93(1.93)	22.74	40.15 – 47.72	<0.001
Depression	-4.96(1.89)	-2.63	-8.66 – -1.26	0.009
Stress	-5.64(1.92)	-2.94	-9.40 – -1.88	0.003
Anxiety	0.77(1.90)	0.40	-2.96 – 4.50	0.687
IUS-12	0.97(1.80)	0.54	-2.55 – 4.49	0.589

OCI-R	-1.02(1.91)	-0.54	-4.77 – 2.72	0.592
TAS-20	2.04(1.95)	1.05	-1.77 – 5.86	0.294
EDI-3-ID	2.07(2.10)	0.98	-2.06 – 6.19	0.326
Random Effects				
σ^2	431.35			
τ_{00} STUDY SITE	0.00			
N STUDY SITE	3			
Observations	120			
Marginal R ² / Conditional R ²	0.127/NA			
Analysis Using the Depression Principal Component as the Predictor				
(Intercept)	45.24(2.37)	19.08	40.55 – 49.94	<0.001
AN	-1.79(5.34)	-0.33	-12.37 – 8.79	0.739
AN-WR	-5.88(5.98)	-0.98	-17.73 – 5.97	0.328
Depression	-4.77(2.08)	-2.29	-8.89 – -0.64	0.024
Holm-corrected linear regressions in each clinical group separately, using the Depression Principal Component as the Predictor				
(Intercept: AN)	40.61(5.32)	7.63	29.54 – 51.68	<0.001
Depression	-0.81(4.01)	-0.20	-9.15 – 7.53	0.841
(Intercept: AN-WR)	39.97(6.39)	6.26	26.16 – 53.77	<0.001
Depression	-14.49(5.31)	-2.73	-25.95 – -3.02	0.034
Analysis Using the Stress Principal Component as the Predictor				
(Intercept)	45.04(2.37)	18.99	40.35 – 49.74	<0.001
AN	-3.16(5.13)	-0.62	-13.32 – 7.00	0.539
AN-WR	-3.03(6.16)	-0.49	-15.23 – 9.17	0.624
Stress	-5.01(2.04)	-2.46	-9.05 – -0.97	0.016
Holm-corrected linear regressions in each clinical group separately, using the Stress Principal Component as the Predictor				
(Intercept: AN)	41.00(4.79)	8.55	31.03 – 50.97	<0.001
Stress	-2.63(4.84)	-0.54	-12.70 – 7.44	0.593
(Intercept: AN-WR)	47.93(8.78)	5.46	28.95 – 66.90	<0.001
Stress	-15.05(8.56)	-1.76	-33.54 – 3.44	0.204
Analysis Using the IUS-12, OCI-R, TAS-20 and EDI-3-ID Principal Components				
(Intercept)	45.90(2.19)	20.93	41.60 – 50.20	<0.001
IUS-12	-1.04(1.66)	-0.63	-4.30 – 2.22	0.532
OCI-R	-2.23(1.68)	-1.32	-5.53 – 1.07	0.186
TAS-20	-0.93(1.68)	-0.55	-4.22 – 2.36	0.580
EDI-3-ID	-1.36(1.75)	-0.78	-4.78 – 2.06	0.436

Random Effects				
σ^2	509.33			
τ_{00} STUDY SITE	7.28			
ICC	0.01			
N STUDY SITE	4			
Observations	187			
Marginal R^2 / Conditional R^2	0.017/0.031			
Table B.2b: Analyses on Posterior Prospective Self-Efficacy Beliefs with the WCST Scores				
Analysis with Group and WCST Percentage Score as the Predictor Variables				
(Intercept)	49.06(11.77)	4.17	25.69 – 72.43	<0.001
AN	-16.36(7.54)	-2.17	-31.32 - -1.40	0.032
AN-WR	-7.80(6.00)	-1.30	-19.71 – 4.11	0.197
WCST_PCT	0.04(0.02)	0.26	-0.27 – 0.35	0.799
Analysis with Group and WCST Preservative Percentage Score as the Predictor Variables				
Intercept	49.02(7.45)	6.58	34.22 – 63.82	<0.001
AN	-16.66(7.69)	-2.17	-31.93 - -1.39	0.033
AN-WR	-7.92(6.22)	-1.28	-20.30 – 4.42	0.205
WCST_PRES_PCT	0.09(0.23)	0.40	-0.37 – 0.55	0.693
N.B. The principal component values were used in the analyses above. Abbreviations: AN (Acute Anorexia Nervosa Group); AN-WR (Weight-Restored from Anorexia Nervosa Group); HC (Healthy Controls Group); IUS-12 (Intolerance of Uncertainty Scale); OCI-R (Obsessive Compulsive Inventory Revised); TAS-20 (20-Item Toronto Alexithymia Scale); EDI-3-ID (Interceptive Deficits Subscale from the Eating Disorders Inventory 3). Depression, Stress and Anxiety refer to the subscale scores from the DASS-21 (21-Item Depression, Anxiety and Stress Scale); WCST_PCT (Percentage of Correct WCST Score); WCST_PRES_PCT (Percentage of WCST Preservative Score). Bolded values denote statistical significance ($p < 0.05$).				

Secondary Measures Analysis Results

Women with and weight restored from AN did not provide significantly poorer Prior Prospective Self-Efficacy Beliefs compared to HCs.

Although it was predicted that women with and weight restored from AN would provide poorer Prior Prospective Self-Efficacy Beliefs on average compared to HCs, a linear regression showed that group was not a statistically significant predictor of participants' Prior Prospective Self-Efficacy Beliefs prior to completing the HCT (Figure B.2; Table B.3).

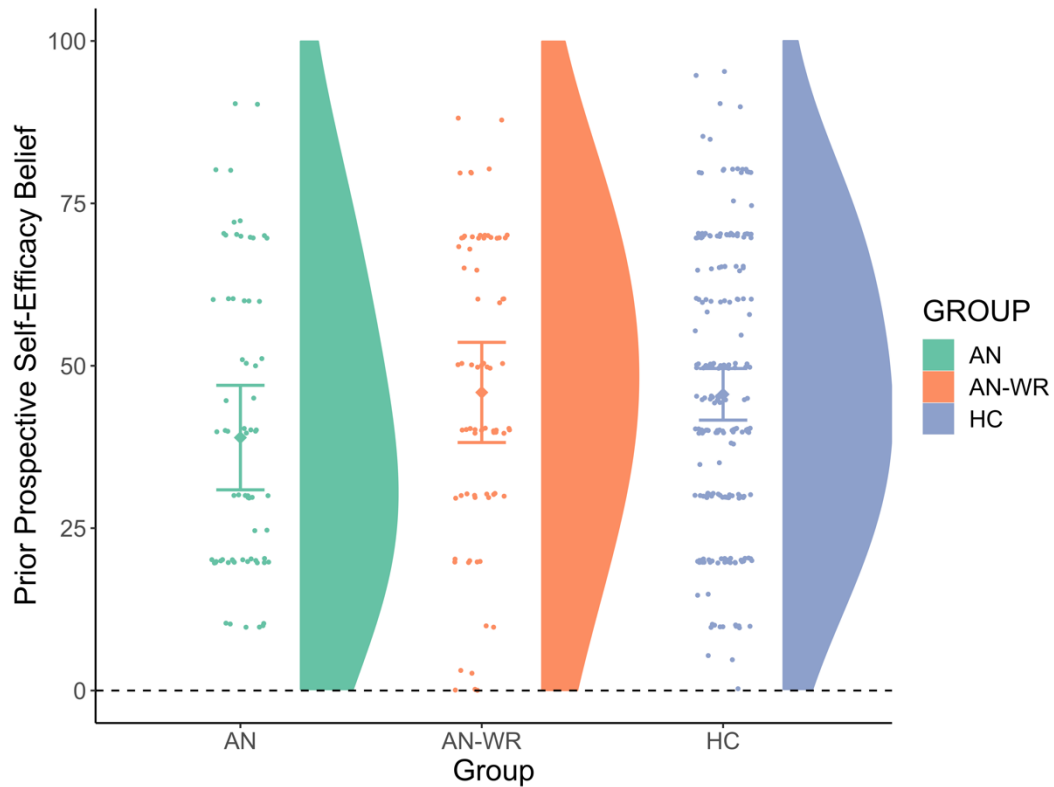


Figure B.3. Average of Prior Prospective Self-Efficacy Beliefs per group (Chapter 3).

Table B.3. Group Differences on Prior Prospective Self-Efficacy Beliefs (Chapter 3).

<i>Predictors</i>	β (SE)	t	95% CI	<i>p</i>
(Intercept)	41.68(6.91)	6.91	28.14 – 55.23	<0.001
AN	-6.24(4.37)	4.37	-14.81 – 2.32	0.153
AN-WR	0.01(4.21)	4.21	-8.23 – 8.26	0.997
AGE	0.16(0.27)	0.27	-0.36 – 0.68	0.552
Random Effects				
σ^2	499.96			
τ_{00} STUDY SITE	0.00			
N STUDY SITE	4			
Observations	189			
Marginal R^2 / Conditional R^2	0.015 / NA			

N.B. Age was used as a control variable and study site was used as a random effect. Abbreviations: AN (Acute Anorexia Nervosa Group); AN-WR (Weight-Restored from Anorexia Nervosa Group); HC (Healthy Controls Group). Bolded values denote statistical significance ($p < 0.05$).

Performance Error Percentage and Difference.

In the preregistration, it was reported that I was going to analyse the performance error percentage differences (misestimation of performance) between the three groups. As seen on

the histogram below (Figure B.4), the data were positively skewed to the right, following a log-normal distribution. Hence, values were logarithmically transformed, and the distribution became more normal as seen on the histogram below (Figure B.5). Although not initially preregistered, a between-group analysis on the difference between participants' retrospective beliefs and actual performance (interoceptive accuracy) was also run, given that the values were more normally distributed (Figure B.6).

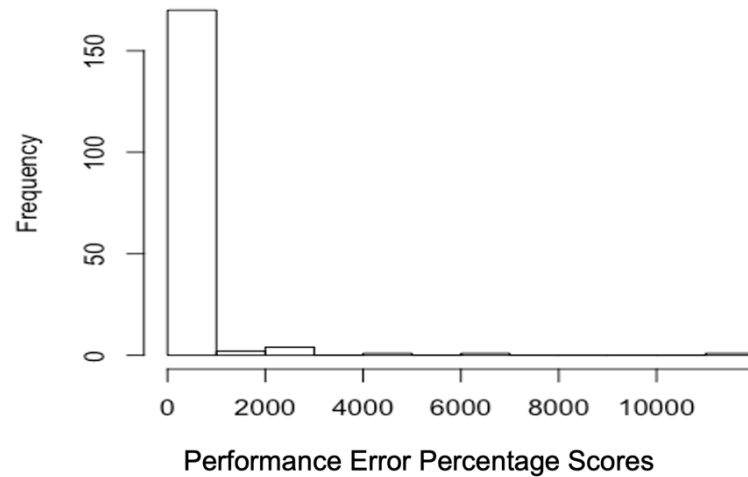


Figure B.4. Histogram of Performance Error Percentage Scores before being logarithmically transformed (Chapter 3).

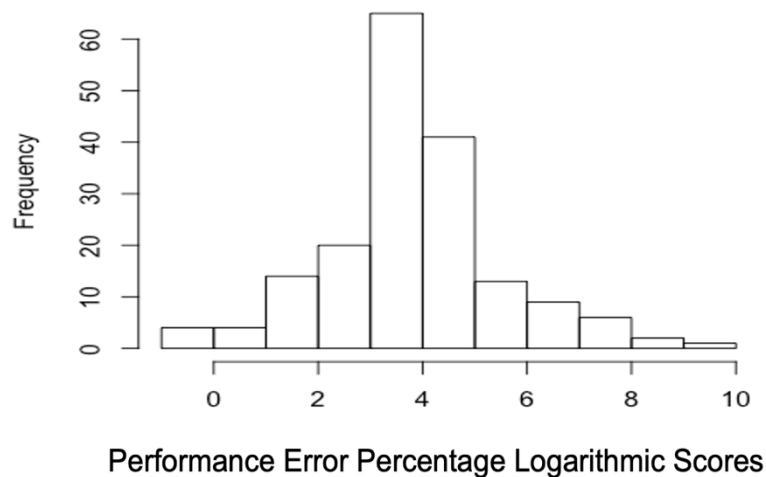


Figure B.5. Histogram of the Logarithmically transformed Performance Error Percentage Scores (Chapter 3).

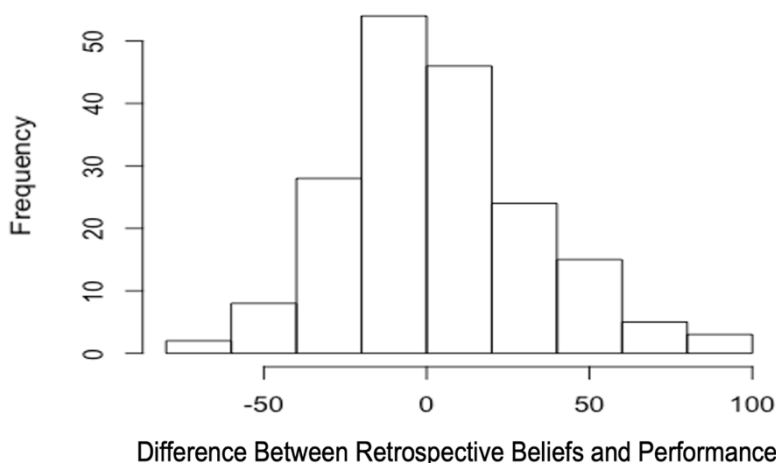


Figure B.6. Histogram of the difference between Posterior Retrospective Self-Efficacy Beliefs and Performance (Chapter 3).

Women with and weight restored from AN underestimated their Performance Retrospectively, while HCs overestimated their Performance

It was expected that the clinical groups would significantly misestimate their performance retrospectively, as indexed by their greater Performance Error Percentage Scores in comparison to the HCs. However, non-significant differences were found between participants' logarithmically transformed Performance Error Percentage Scores (Table B.4a). The difference between Posterior Retrospective Self-Efficacy Beliefs and Performance (non-preregistered) was also examined and this misestimation was found to differ in the two clinical groups relative to the HCs (Table B.4b). Specifically, the difference between beliefs and performance on the task was positive for the HCs and negative for both clinical groups. Further within-group comparisons examining the main effects revealed that the HCs significantly overestimated their Performance (i.e., they predicted they would do better than they did in reality), while the two clinical groups underestimated their Performance (although this effect was not statistically significant; Table B.4c). In summary, even though there is a difference in misestimation between the groups, one cannot claim that the clinical groups misestimate more or less than HCs, due to the opposite directionality of these differences.

Table B.4. Results from the Secondary, Behavioural Analyses (Chapter 3).

<i>Predictors</i>	$\beta(SE)$	<i>t</i>	95% CI	<i>p</i>
Table B.4a: Performance Error Percentage (Log)				
(Intercept)	3.43(0.54)	6.32	2.36 – 4.50	<0.001
AN	0.02(0.34)	0.06	-0.65 – 0.69	0.950

AN-WR	0.07(0.34)	0.21	-0.60 – 0.74	0.837
AGE	0.02(0.02)	0.75	-0.03 – 0.06	0.453

Table B.4b: Difference between Posterior Retrospective Self-Efficacy Beliefs and Performance

Intercept)	-1.82(9.13)	-0.20	-19.82 – 16.19	0.842
AN	-15.44(5.77)	-2.68	-26.82 – -4.07	0.008
AN-WR	-13.99(5.67)	-2.45	-25.18 – -2.81	0.015
AGE	0.43(0.35)	1.23	-0.26 – 1.13	0.221

Table B.4c: Difference between Posterior Retrospective Self-Efficacy Beliefs and Performance within Group

AN	-7.73(4.87)	-1.59	-17.62 - 2.17	-0.122
AN-WR	-4.33(5.19)	-0.83	-14.87 - 6.22	0.410
HC	8.86(2.73)	3.24	3.45 - 14.27	0.002

N.B. Age was used as a control variable in the analyses looking at Group effects on the difference between Posterior Retrospective Self-Efficacy Beliefs and Performance. Abbreviations: AN (Acute Anorexia Nervosa Group); AN-WR (Weight-Restored from Anorexia Nervosa Group); HC (Healthy Controls Group). Bolded values denote statistical significance ($p < 0.05$).

Women with AN showed a correlation between their Performance and Performance Confidence Estimates.

To evaluate whether the two clinical groups underestimated Performance in the cardiac interoception task, and specifically whether they would show poorer interoceptive awareness (correspondence between Performance and Performance Confidence Estimates), Pearson's correlations were run, both across groups and also within each group separately. There was a non-significant ($r = -0.066$, $p = .375$) correlation between Performance and Performance Confidence across all groups. When examining correlations between Performance and Performance confidence within each group separately a significant relationship in women with AN was found, such that higher Performance Confidence Estimates were related to higher Performance ($r_{AN} = .439$, $p_{AN} = .008$). The correlation between these two variables was not significant in the AN-WR and HC groups ($r_{AN-WR} = -.240$, $p_{AN-WR} = .158$; $r_{HC} = -.144$, $p_{HC} = .127$). Contrary to the prediction, these findings suggest that the AN group may have better Interoceptive Awareness compared to the AN-WR and HC groups.

Women with AN showed a poorer correlation between their Interoceptive Sensibility scores and Interoceptive Accuracy (i.e., Performance) than HCs.

To better understand the extent to which the discrepancy between interoceptive sensibility (subjective measure of one's ability to attend to interoceptive signals; here, measured via the interoceptive deficits subscale from the EDI-3; (Garner, 2004)) and Performance explain interoceptive awareness, an Interoceptive Trait Prediction Error (ITPE) z-score was obtained per participant, expecting the AN and AN-WR groups to show an underestimation of their Performance as indexed by their more negative ITPE z-score values in comparison to HCs.

When looking at the ITPE z-scores at face value (M(SD)), both clinical groups had a negative value (ITPE_{AN}= -0.46(1.22), ITPE_{AN-WR}= -0.25(1.61)) suggesting they underestimated their own interoceptive abilities. The HCs had a positive value (ITPE_{HC}= 0.21(1.56)), suggesting they overestimated their interoceptive abilities. Using a linear regression to examine ITPE z-score differences between groups, significant group differences were found, suggesting that the AN group significantly underestimated their interoceptive abilities in comparison to HCs, as indexed by the significantly more negative ITPE z-score, yet this was not the case for the AN-WR group (Table B.5). These results suggest that in general, the AN group underestimated their Performance, despite the positive correlation that was found above, between Performance Confidence and Performance. It is noteworthy, that these analyses do not allow examination of latent variables that may be driving behaviour and thus the focus here is placed on the belief updating analyses instead.

Table B.5. ITPE z-score Differences Between Groups (Chapter 3).

<i>Predictors</i>	$\beta(SE)$	<i>t</i>	95% CI	<i>p</i>
HC (Intercept)	0.21(0.14)	1.14	-0.07 – 0.49	0.146
AN	-0.67(0.30)	-2.27	-1.25 – -0.09	0.025
AN-WR	-0.46(0.29)	-1.57	-1.04 – 0.12	0.118
Abbreviations: AN (Acute Anorexia Nervosa Group); AN-WR (Weight-Restored from Anorexia Nervosa Group); ITPE (Interoceptive Trait Prediction Error). Bolded values denote statistical significance ($p < 0.05$).				

Women with AN reported significantly lower confidence than HCs following HCT Performance.

It was predicted that both clinical groups would report significantly lower confidence throughout the experiment in comparison to HCs, especially during the HCT (hereafter referred to as Performance Confidence). The difference in the Confidence estimates of the AN-WR group as opposed to the HC group following Prior and Posterior Prospective Self-Efficacy

Beliefs was not significantly different (Table B.6a and B.6c). In line with the prediction, the AN group reported to be significantly less confident after the HCT trials, i.e., when reporting how confident they were in the number of heartbeats they felt, as opposed to the HCs (Table B.6b). A trend towards significance was also found in Posterior Prospective Confidence estimates, in that the AN group were less confident than the HCs (Table B.6c).

Table B.6. Group Effect on Confidence Estimates (Chapter 3).

<i>Predictors</i>	$\beta(SE)$	<i>t</i>	95% <i>CI</i>	<i>p</i>
Table B.6a: Group Differences on Prior Prospective Confidence				
HC (Intercept)	67.09(2.26)	29.72	62.63 - 71.54	<0.001
AN	-3.09(4.69)	-0.66	-12.34 - 6.16	0.511
AN-WR	-6.09(4.54)	-1.34	-15.05 - 2.88	0.182
Table B.6b: Group Differences on Performance Confidence				
HC (Intercept)	59.08(2.30)	25.64	54.54 - 63.63	<0.001
AN	-17.63(4.79)	-3.68	-27.07 - -8.18	<0.001
AN-WR	-6.38(4.64)	-1.37	-15.53 - 2.78	0.171
Table B.6c: Group Differences on Posterior Prospective Confidence Estimate				
HC (Intercept)	72.66(1.86)	39.01	68.98 - 76.33	<0.001
AN	-7.10(3.87)	-1.84	-14.73 - 0.53	0.068
AN-WR	-2.94(3.75)	-0.79	-10.34 - 4.45	0.433
Abbreviations: AN (Acute Anorexia Nervosa Group); AN-WR (Weight-Restored from Anorexia Nervosa Group); HC (Healthy Controls Group). Bolded values denote statistical significance ($p < 0.05$).				

Computational Modelling: Model Presentation and Comparisons, Control and Exploratory Analyses

Computational Modelling Analyses

In the next part of this Results section, I present the baseline model comparison, the analyses ran to examine whether the winning model was better than the competing models, further analyses for participants' precision-weighted Learning Rate, and exploratory analyses on between-group differences on the actual Learning Rates.

1. Baseline Model Comparison

As explained in Section 3.2 (Chapter 3), for validation and control purposes of the main models of interest, two sets of baseline models were created and then compared. The first set of models (one model per evidence measure) assumed a perfect learning rate ($\lambda = 1$, the participant uses

the evidence as their posterior beliefs). The second set of baseline models (actually, one model given its baseline nature) assumed no learning ($\lambda = 0$, the participant uses the prior as their posterior beliefs). In the first step for the modelling analyses, the baseline models were compared. To do this, I compared the models using Performance and Posterior Retrospective Self-Efficacy Beliefs as evidence (separate models), and Performance Confidence and EDI-3-ID as different proxies for precision. Due to the high correlation ($r=0.87$, $p<.001$) between Posterior Retrospective Self-Efficacy Beliefs (when used as evidence), and Posterior Prospective Self-Efficacy Beliefs, the precision-weighted Learning Rate was very close to 1 (i.e., a perfect learning rate, where the participant uses the evidence as their posterior beliefs). Therefore, the baseline model which assumed 100% learning was the best fit, both across all groups and within each group separately (Models S1.3, S2.3, S3.3, S4.3 in Table B.7).

Table B.7. MAE and BIC Values for the Baseline Models (Chapter 3).

Model Comparison across all Participants	N	MAE	BIC
S1.1 Baseline: 0 Learning	183	14.4	1586
S1.2 Baseline: 100% Learning: Evidence was Performance	183	22.86	1650
S1.3 Baseline: 100% Learning: Evidence was Posterior Retrospective Self-Efficacy	183	7.46	1417
Model Comparison in the HCs only			
S2.1 Baseline: 0 Learning	114	13.97	978.8
S2.2 Baseline: 100% Learning: Evidence was Performance	114	23.2	1017
S2.3 Baseline: 100% Learning: Evidence was Posterior Retrospective Self-Efficacy	114	7.19	854
Model Comparison in the AN group only			
S3.1 Baseline: 0 Learning	34	16.3	302.2
S3.2 Baseline: 100% Learning: Evidence was Performance	34	19.53	300.7
S3.3 Baseline: 100% Learning: Evidence was Posterior Retrospective Self-Efficacy	34	8.08	268.7
Model Comparison in the AN-WR group only			
S4.1 Baseline: 0 Learning	35	13.97	312
S4.2 Baseline: 100% Learning: Evidence was Performance	35	25	335.3
S4.3 Baseline: 100% Learning: Evidence was Posterior Retrospective Self-Efficacy	35	7.74	299.4
N.B. MAE refers to Mean Absolute Error and BIC refers to Bayesian Information Criterion. These two measures are used to examine model fit, with smaller values suggesting better model fit.			

2. Linear regressions ran to complement the results presented in Table 3.6

Table B.8a reports the results from the examination of whether the difference between the Predicted Posterior Self-Efficacy Beliefs using Performance and Posterior Retrospective Self-Efficacy Beliefs as Evidence in the two separate models, and Performance Confidence as the Precision Proxy was statistically significant, i.e., whether there was a significantly better model. In the first instance the analysis was ran across all groups, and in the second it was only run in the HC group. In Table B8b I examined whether the difference between the Precision Proxies was statistically significant when using Posterior Retrospective Self-Efficacy Beliefs as the evidence. These analyses were run to explore the statistical significance of the results presented in Table 3.6

Table B.8. Additional Linear Regressions for the Belief Updating Predictions (Chapter 3).

<i>Predictors</i>	$\beta(SE)$	<i>t</i>	95% CI	<i>p</i>
Table B.8a: Exploratory analyses for Prediction A				
Predicted Posteriors; Difference between Evidences, Performance Confidence as Precision Proxy. Across Groups				
Intercept	5.38(0.86)	6.27	3.69 - 7.07	<0.001
Predicted Posteriors; Difference between Evidences, Performance Confidence as Precision Proxy. Within the Healthy Control Group (HC)				
Intercept	5.64(1.14)	4.94	3.38 - 7.90	<0.001
Predicted Posteriors; Difference between Evidences, EDI-3-ID as Precision Proxy. Across Groups				
Intercept	6.52(0.97)	6.72	4.61 – 8.44	<0.001
Predicted Posteriors; Difference between Evidences, EDI-3-ID as Precision Proxy. Within the Healthy Control Group (HC)				
Intercept	6.58(1.22)	5.41	4.17 – 8.99	<0.001
Table B.8b: Exploratory analyses for Prediction A				
Predicted Posteriors; Difference between Precision Proxies, Posterior Retrospective Self-Efficacy Beliefs as Evidence; Across Groups				
Intercept	0.48(0.29)	1.65	-0.09 – 1.06	0.100
Predicted Posteriors; Difference between Precision Proxies, Posterior Retrospective Self-Efficacy Beliefs as Evidence; Within HC				
Intercept	0.26(0.25)	1.01	-0.25 – 0.76	0.315
Bolded values denote statistical significance ($p < 0.05$).				

- Control analyses on precision-weighted Learning Rates with Wisconsin Card Sorting Task (WCST) scores

Non-preregistered exploratory analyses using the percentage of correct WCST scores and the percentage of WCST preservative scores were carried out to examine whether differences in Learning Rates between the groups were explained by set-shifting difficulties. However, neither analysis was significant (Table B.9).

Table B.9. Control analyses for Learning Rate using WCST scores (Chapter 3).

<i>Predictors</i>	<i>β(SE)</i>	<i>t</i>	<i>95% CI</i>	<i>p</i>
Analysis with Group and WCST Percentage Score as Predictor Variables				
(Intercept)	0.42(0.08)	5.13	0.26 – 0.58	<0.001
AN	-0.07(0.05)	-1.29	-0.17 – 0.04	0.199
AN-WR	0.01(0.04)	0.27	-0.07 – 0.09	0.789
WCST_PCT	0.001(0.001)	0.95	-0.001 – 0.003	0.334
Analysis with Group and WCST Preservative Percentage Score as Predictor Variables				
(Intercept)	0.54(0.05)	10.49	0.44 – 0.64	<0.001
AN	-0.06(0.05)	-1.21	-0.17 – 0.04	0.229
AN-WR	0.02(0.04)	0.40	-0.07 – 0.10	0.688
WCST_PRES_PCT	-0.002(0.002)	-1.01	-0.005 – 0.002	0.316
Abbreviations: AN (Acute Anorexia Nervosa Group); AN-WR (Weight-Restored from Anorexia Nervosa Group); WCST_PCT (Percentage of Correct WCST Score); WCST_PRES_PCT (Percentage of WCST Preservative Score). Bolded values denote statistical significance.				

4. Analyses on Actual Learning Rates

The actual Learning Rates of the three groups following the assumptions of a Bayesian Learning Framework were explored and the data were analysed in the form of a linear regression based on the structural equivalence between the two (see Section 4.2, Chapter 4 and (Saramandi et al., 2022) for detailed methodology). Specifically, I performed one linear regression, using in both as the outcome variable participants' Belief Update (i.e., the difference between the Posterior and Prior Prospective Self-Efficacy Beliefs) and as predictor variables the interaction between Group and Prediction Error (calculated as the difference between Posterior Retrospective and Prior Prospective Self-Efficacy Beliefs). In this linear regression, the coefficient of the interaction term corresponds (due to the aforementioned structural equivalence) to the difference between the Learning Rates of the groups.

Interestingly, when comparing participants' actual Learning Rates (I calculated one Learning Rate per group), the AN-WR group showed a significantly lower actual Learning Rate vs. HCs, as predicted, but this difference was not significant when looking at the AN group vs. the HCs (Table B.10).

Table B.10. Analysis on Actual Learning Rates (Chapter 3).

<i>Predictors</i>	$\beta(SE)$	<i>t</i>	95% CI	<i>p</i>
Intercept	0.08(0.83)	0.10	-1.56 – 1.72	0.924
Prediction Error	0.79(0.05)	16.08	0.69 – 0.89	<0.001
AN x Prediction Error	0.11(0.11)	1.03	-0.10 – 0.33	0.305
AN-WR x Prediction Error	-0.21(0.10)	-2.15	-0.40 – -0.02	0.033

N.B. The first linear regression model uses participants' actual Performance as evidence, and the second linear regression model uses participants' Posterior Retrospective Self-Efficacy Beliefs as evidence. Prior refers to Prior Prospective Self-Efficacy Beliefs. Abbreviations: AN (Acute Anorexia Nervosa Group); AN-WR (Weight-Restored from Anorexia Nervosa Group). Bolded values denote statistical significance.

Exploratory Analyses in the acute and weight restored AN groups

In this part the results from the exploratory analyses on counterfactual beliefs, and correlations between the main measures and aspects of insight into illness are presented.

Exploratory Analyses with Everyday, Counterfactual Beliefs

Next, I explored how everyday, counterfactual beliefs on hopes and fears as measured via items developed by the research team (see Table 3.3, Chapter 3) predicted the Prior and Posterior Prospective Self-Efficacy Beliefs of the AN and AN-WR groups, separately. A significant effect of the Fear subscale on the Posterior Prospective Self-Efficacy Beliefs in both the AN and AN-WR groups was found (Table B.11d). Specifically, patients with less overall fear on gaining weight and losing control over their eating behaviour had higher Posterior Prospective Self-Efficacy beliefs. However, the effects of the fear subscale on patients' Prior Prospective Self-Efficacy Beliefs (Table B.11c) were non-significant and this was also the case for the effects of the hope scale on both their Prior and Prospective Self-Efficacy Beliefs (Table B.11a, B.11b). In both the AN and AN-WR groups a statistical trend effect of overall Counterfactual Belief Scores (i.e., Hope and Fear subscales combined) on patients' Posterior Prospective Self-Efficacy Beliefs (Table B.11f) was observed, but not on their priors (Table B.11e). That is, a higher total Counterfactual Belief Score was indicative of more optimistic Posterior

Prospective Self-Efficacy Beliefs. These results yield further validity to the experimental task presented in Chapter 3 revealing that the key measure, namely posterior prospective beliefs about interoception following a HCT relates to everyday explicit beliefs patients have about the future.

Table B.11. Results from Exploratory Analysis with Counterfactual Beliefs (Chapter 3).

<i>Predictors</i>	$\beta(SE)$	<i>t</i>	95% CI	<i>p</i>
B.11a: “Hopes” Predicting Prior Prospective Self-Efficacy Beliefs				
AN	28.25(16.03)	1.76	-4.40 – 60.90	0.088
Hopes	1.51(2.26)	0.70	-3.09 – 6.11	0.508
AN-WR	46.43(8.36)	5.55	29.37 – 63.49	<0.001
Hopes	0.06(1.31)	0.05	-2.60 – 2.73	0.962
B.11b: “Fears” Predicting Prior Prospective Self-Efficacy Beliefs				
AN	32.43(5.80)	5.94	20.62 – 44.24	<0.001
Fears	2.16(1.47)	1.47	-0.83 – 5.14	0.150
AN-WR	41.66(6.56)	6.34	28.27 – 55.04	<0.001
Hopes	1.21(1.23)	0.99	-1.28 – 3.71	0.329
B.11c: “Hopes” Predicting Posterior Prospective Self-Efficacy Beliefs				
AN	29.09(14.08)	2.07	0.41 – 57.77	0.047
Hopes	1.38(1.98)	0.70	-2.66 – 5.41	0.492
AN-WR	43.54(10.16)	4.28	22.81 – 64.27	<0.001
Hopes	-00.13(1.59)	-0.80	-3.37 – 3.12	0.937
B.11d: “Fears” Predicting Posterior Prospective Self-Efficacy Beliefs				
AN	31.34(4.96)	6.32	21.23 – 41.44	<0.001
Fears	2.52(1.25)	2.01	-0.04 – 5.07	0.053
AN-WR	30.71(7.62)	4.03	15.17 – 46.24	<0.001
Fears	2.87(1.42)	2.02	-0.03 – 5.77	0.052
B.11e: Overall Counterfactual Beliefs (Hopes and Fears Combined) Predicting Prior Prospective Self-Efficacy Beliefs				
AN	22.99(11.47)	2.00	-0.38 – 46.36	0.054
Hopes and Fears Combined	3.23(2.21)	1.45	-1.29 – 7.71	0.156
AN-WR	35.07(12.31)	2.85	9.96 – 60.19	0.008
Hopes and Fears Combined	2.39(2.37)	1.01	-2.45 – 7.23	0.322
B.11f: Overall Counterfactual Beliefs (Hopes and Fears Combined) Predicting Posterior Prospective Self-Efficacy Beliefs				
AN	21.27(9.89)	2.15	1.13 – 41.41	0.039
Hopes and Fears Combined	3.55(1.90)	1.86	-0.33 – 7.43	0.071
AN-WR	17.46(14.42)	1.21	-11.96 – 46.87	0.235

Hopes and Fears Combined	5.18(2.78)	1.86	-0.49 – 10.85	0.072

Abbreviations: AN (Acute Anorexia Nervosa Group); AN-WR (Weight-Restored from Anorexia Nervosa Group). Bolded values denote statistical significance ($p < 0.05$). Please note that the p values here are marginally significant and should be interpreted with caution.

Additional analyses with items from the EDI-3 questionnaire

Given the similarity between some of the counterfactual belief items from the questionnaire developed by the research team and items from the EDI-3 (Garner, 2004), a validated questionnaire widely used in eating disorders questionnaires, correlational analyses between items of the EDI-3 and Posterior Prospective Self-Efficacy Beliefs were run in each clinical group separately. Items DT_16, DT_49, B_64, and ID_44 (from the EDI-3; Table B.12) are similar to items (9), (11) from the insight questionnaire developed by the research team (Table 3.3, Chapter 3). Thus, I wanted to explore whether the pattern of the results would be the same to the results reported in Table B.11d when running correlations between the self-efficacy estimates and selected items from the validated, EDI-3 questionnaire. Specifically, the following items were used: item 16 (“*I am terrified of gaining weight*”); item 49 (“*If I gain a pound, I worry that I will keep gaining*”); item 44 (“*I worry that my feelings will get out of control*”); and item 64 (“*When I am upset, I worry that I will start eating*”). However, none of the correlations were statistically significant.

Table B.12. Correlations between Self-Efficacy Beliefs and EDI-3 Items (Chapter 3).

	POST PROSP AN	POST PROSP ANWR
DT_16: “I am terrified of gaining weight”	-0.04	0.02
DT_49: “If I gain a pound, I worry that I will keep gaining”	0	-0.03
B_64: “When I am upset, I worry that I will start eating”	0.16	-0.07
ID_44: “I worry that my feelings will get out of control”	-0.11	0.1

N.B. Heatmap with results from correlational analyses. The table presents the r values from the correlations. * = $p < 0.05$, ** = $p < 0.001$. POST PROSP refers to Posterior Prospective Self-Efficacy Beliefs; AN refers to the Anorexia Nervosa group; AN-WR refers to the weight-restored Anorexia Nervosa group.

Correlational Analyses with Explicit Beliefs

Insight deficits have been noted in individuals with AN, and I wanted to explore whether there would be a correlation between scores of a questionnaire assessing insight into illness and beliefs about interoception and related measures (here, Prior and Posterior Prospective Self-Efficacy Beliefs, Performance and Learning Rate). None of the correlations between self-

efficacy beliefs, performance and learning rates and standalone item and subscale error scores from the insight into illness questionnaire were significant (Table B.13).

Table B.13. Correlations with Insight into Illness Error Scores (Chapter 3).

	POST			LEARNING	LEARNING
	PRIORS	PROSP	PERFORMANCE	RATE 1	RATE 2
Item 1	-0.06	0.21	-0.04	0.16	-0.08
Item 2	-0.17	-0.2	0.15	-0.11	0.17
Item 3	-0.14	0.25	0.07	0.17	-0.04
Item 4	-0.15	-0.08	0.1	0.01	0.1
Item 5	0.06	0.05	0.13	0.04	-0.07
Item 6	-0.12	0.01	0.09	-0.08	0.04
Item 7	-0.15	0.19	-0.11	0.14	-0.01
Item 8	-0.08	0.04	-0.07	0.01	-0.06
General, Allocentric					
Perspective of Self	-0.16	-0.08	0.11	0	0.1
Health Consequences	-0.04	0.03	0.11	-0.03	-0.01
Perspective Taking	-0.12	0.12	-0.1	0.08	-0.03

N.B. Heatmap with results from correlational analyses. The table presents the r values from the Pearson correlations. * = $p < 0.05$, ** = $p < 0.001$. Items 1-8 refer to the error scores calculated from the respective items of the Insight into Illness Questionnaire (Table 3.3). The last three rows represent the correlations with each subscale. PRIORS refers to the Prior Prospective Self-Efficacy Beliefs; POST PROSP refers to the Posterior Prospective Self-Efficacy Beliefs; Learning Rate 1 refers to the precision-weighted learning rate with Performance Confidence as a proxy for precision of the evidence; Learning Rate 2 refers to the precision-weighted learning rate with EDI-3-ID as a proxy for precision of the evidence.

Appendix C

Appendix C includes the script participants were told before providing Pre-Instructions Baseline Fullness Ratings (Section 1) and supplementary material from control and exploratory analyses (Section 2) for the experiment presented in Chapter 4.

Section 1: Supplemental Methods

Baseline Ratings Instructions

Experimenter: “Now, we will do a task during which I will ask you questions about fullness, before and after drinking water. I will give you detailed instructions for this in a moment. First, we will take two ratings. The first is what we will call ‘Objective’ – in other words, this fullness rating should be based on how full you should logically be now, when accounting for the time of your most recent meal, the amount and type of food you consumed. For example, you may have had a small lunch 2 hours ago, but this is the typical portion and meal you have for lunch and therefore be fuller than someone who may need a larger portion. The second rating is what we will call “Subjective” – in other words, this fullness rating should be based on how full you feel, irrespective of the time of your most recent meal, the amount and type of food you consumed. This rating should be solely based on how you feel right now. For example, you may have had a large portion of salad just an hour ago and not feel completely full, or you might be feeling rather full because the type of food you recently consumed makes you feel bloated. It may be the case that how full you are and how full you feel are similar, or you might feel fuller than you are, or feel less full than you are. Please let me know if you have any questions.” If the participant does not have any questions, take the two ratings and proceed with the main task instructions.

Section 2: Supplemental Results

1. Control analyses for the Prior and Posterior Prospective Fullness Beliefs and Prediction Error.

Tables C.1a, C.1b, and C.1c present the results from the control analyses, using the effects of bladder and stomach fullness as well as knowledge about the bottle’s total volume (in ml) as control variables. The main effects from the primary analyses (see Results) however remained the same.

Table C.1. Controlling for Effects in Stomach Fullness (Chapter 4).

<i>Predictors</i>	<i>β(SE)</i>	<i>t</i>	<i>95% CI</i>	<i>p</i>
Table C.1a: Prior Prospective Fullness Belief (Control Analysis)				
(Intercept)	2.04(1.21)	1.68	-0.33 - 4.41	0.092
AN	0.27(0.73)	0.38	-1.16 - 1.71	0.706
AN-WR	2.26(0.78)	2.9	0.73 - 3.78	0.004
Condition (High Quantity)	-0.06(0.24)	-0.26	-0.54 - 0.41	0.798
Bladder Fullness	-0.00(0.01)	-0.28	-0.02 - 0.02	0.781
Stomach Fullness	-0.00(0.10)	-0.07	-0.02 - 0.02	0.943
Bottle Quantity Knowledge	0.00(0.00)	-0.46	-0.00 - 0.00	0.646
Condition (High Quantity)	-0.70(0.50)	-1.39	-1.69 - 0.29	0.164
AN-WR x Condition (High Quantity)	0.26(0.60)	0.43	-0.92 - 1.44	0.669
Marginal R ² / Conditional R ²	0.138 / 0.669			
Table C.1b: Posterior Prospective Fullness Belief (Control Analysis)				
(Intercept)	2.59(1.27)	2.05	0.11 - 5.08	0.041
AN	0.35(0.74)	0.48	-1.09 - 1.80	0.631
AN-WR	2.76(0.85)	3.23	1.09 - 4.44	0.001
Condition (High Quantity)	-0.24(0.26)	-0.92	-0.74 - 0.27	0.358
Bladder Fullness	-0.00(0.01)	-0.15	-0.02 - 0.02	0.884
Stomach Fullness	-0.00(0.01)	-0.39	-0.03 - 0.02	0.696
Bottle Quantity Knowledge	-0.00(0.00)	-0.80	-0.00 - 0.00	0.422
Condition (High Quantity)	-0.77(0.51)	-1.52	-1.76 - 0.22	0.130
AN-WR x Condition (High Quantity)	0.35(0.62)	0.57	-0.86 - 1.55	0.572
Marginal R ² / Conditional R ²	0.218 / 0.731			
Table C.1c: Prediction Error (Control Analysis)				
(Intercept)	-0.92(1.28)	-0.73	-3.43 - 1.58	0.469
AN	-0.96(0.76)	-1.26	-2.46 - 0.53	0.207
AN-WR	-2.74(0.83)	-3.30	-4.37 - -1.12	0.001
Condition (High Quantity)	-0.28(0.27)	-1.03	-0.82 - 0.26	0.303
Bladder Fullness	0.01(0.01)	0.98	-0.01 - 0.03	0.326
Stomach Fullness	-0.00(0.01)	-0.30	-0.03 - 0.02	0.763
Bottle Quantity Knowledge	0.00(0.00)	1.09	-0.00 - 0.00	0.275
Condition (High Quantity)	0.72(0.57)	1.25	-0.41 - 1.84	0.212
AN-WR x Condition (High Quantity)	1.82(0.65)	2.8	0.55 - 3.10	0.005
Marginal R ² / Conditional R ²	0.090 / 0.566			
Abbreviations: AN (Acute Anorexia Nervosa Group); AN-WR (Weight-Restored from Anorexia Nervosa Group); HC (Healthy Controls Group). Bolded values denote statistical significance ($p < 0.05$)				

Tables C.2a-C.2e present the results from the between-group comparisons of the Post-Instructions Pre-Sipping Baseline Fullness State Ratings.

Table C.2. Control Analyses on Post-Instructions Pre-Sipping Baseline Fullness State Rating (Chapter 4).

Predictors	β(SE)	t	95% CI	p
Table C.2a: Average of the Post-Instructions Pre-Sipping Baseline Fullness State Rating across the 4 Trials				
(Intercept)	59.24(3.27)	18.11	52.74 - 65.74	<0.001
AN	3.23(5.42)	0.6	-7.53 - 13.99	0.552
AN-WR	-7.20(5.99)	-1.20	-19.10 - 4.69	0.232
R ² / Adjusted R ²	0.027 / 0.006			
Table C.2b: Post-Instructions Pre-Sipping Baseline Fullness State Rating at Condition: Low Quantity (2%)				
(Intercept)	58.62(3.54)	16.56	51.59 - 65.65	<0.001
AN	4.35(5.86)	0.74	-7.30 - 15.99	0.46
AN-WR	-6.62(6.48)	-1.02	-19.49 - 6.25	0.31
R ² / Adjusted R ²	0.025 / 0.004			
Table C.2c: Post-Instructions Pre-Sipping Baseline Fullness State Rating at Condition: Low Quantity (5%)				
(Intercept)	60.14(3.57)	16.86	53.05 - 67.22	<0.001
AN	3.90(5.90)	0.66	-7.83 - 15.63	0.511
AN-WR	-7.04(6.53)	-1.08	-20.01 - 5.93	0.284
R ² / Adjusted R ²	0.025 / 0.003			
Table C.2d: Post-Instructions Pre-Sipping Baseline Fullness State Rating at Condition: High Quantity (12%)				
(Intercept)	56.94(3.31)	17.22	50.37 - 63.50	<0.001
AN	2.06(5.47)	0.38	-8.81 - 12.93	0.707
AN-WR	-8.39(6.05)	-1.39	-20.40 - 3.63	0.169
R ² / Adjusted R ²	0.029 / 0.007			
Table C.2e: Post-Instructions Pre-Sipping Baseline Fullness State Rating at Condition: High Quantity (15%)				
(Intercept)	61.28(3.21)	19.06	54.89 - 67.66	<0.001
AN	2.61(5.32)	0.49	-7.96 - 13.18	0.625
AN-WR	-6.78(5.88)	-1.15	-18.46 - 4.91	0.252
R ² / Adjusted R ²	0.023 / 0.002			
Abbreviations: AN (Acute Anorexia Nervosa Group); AN-WR (Weight-Restored from Anorexia Nervosa Group). Bolded values denote statistical significance ($p < 0.05$)				

- Mediation Analyses for mediating role of traits on Prior Prospective, Experienced Retrospective and Posterior Prospective Fullness Beliefs and Prediction Error.

The results in Tables C.3a., C.3b, C.3c, and C.3d are from the step 2 of the mediating analysis (see Figure 4.2), where one examines the effect of the mediator variable (here, the different traits) on the DV of interest. None of the traits of interest had a significant effect on the DVs of interest and thus the next steps of the mediating analysis were not conducted.

Table C.3. Mediation Analysis Step 2 (Chapter 4).

Predictors	β(SE)	t	95% CI	p
Table C.3a: Effect of Traits on Prior Prospective Fullness Belief				
Depression	0.02(0.03)	-0.56	-0.07 - 0.04	0.579
Marginal R ² / Conditional R ²	0.003 / 0.778			
Stress	0.01(0.03)	0.45	-0.04 - 0.07	0.652
Marginal R ² / Conditional R ²	0.002 / 0.777			
Anxiety	-0.04(0.04)	-0.99	-0.11 - 0.04	0.323
Marginal R ² / Conditional R ²	0.009 / 0.776			
TAS-20	-0.04(0.02)	-1.66	-0.09 - 0.01	0.907
Marginal R ² / Conditional R ²	0.025 / 0.776			
IUS-12	0.01(0.03)	0.36	-0.04 - 0.06	0.723
Marginal R ² / Conditional R ²	0.001 / 0.778			
Table C.3b: Effect of Traits on Experienced Retrospective Fullness Belief				
Depression	-0.00(0.01)	-0.16	-0.02 - 0.02	0.871
Marginal R ² / Conditional R ²	0.000 / 0.232			
Stress	0.00(0.01)	0.06	-0.02 - 0.02	0.956
Marginal R ² / Conditional R ²	0.000 / 0.232			
Anxiety	-0.00(0.01)	-0.12	-0.03 - 0.03	0.902
Marginal R ² / Conditional R ²	0.000 / 0.232			
TAS-20	0.01(0.01)	0.6	-0.01 - 0.02	0.547
Marginal R ² / Conditional R ²	0.002 / 0.232			
IUS-12	0.01(0.01)	0.67	-0.01 - 0.03	0.502
Marginal R ² / Conditional R ²	0.002 / 0.233			
Table C.3c: Effect of Traits on Posterior Prospective Fullness Belief				
Depression	-0.01(0.03)	-0.46	-0.07 - 0.04	0.645
Marginal R ² / Conditional R ²	0.002 / 0.736			
Stress	-0.00(0.03)	-0.02	-0.06 - 0.06	0.987
Marginal R ² / Conditional R ²	0.000 / 0.735			
Anxiety	-0.05(0.04)	-1.40	-0.13 - 0.02	0.162
Marginal R ² / Conditional R ²	0.016 / 0.736			
TAS-20	-0.03(0.02)	-1.37	-0.08 - 0.01	0.172

Marginal R ² / Conditional R ²	0.015 / 0.737			
IUS-12	0.01(0.03)	0.45	-0.04 - 0.07	0.650
Marginal R ² / Conditional R ²	0.002 / 0.741			
Table C3.d: Effect of Traits on Prediction Error				
Depression	0.01(0.03)	0.37	-0.04 - 0.07	0.711
Marginal R ² / Conditional R ²	0.001 / 0.674			
Stress	-0.01(0.03)	-0.37	-0.07 - 0.05	0.710
Marginal R ² / Conditional R ²	0.001 / 0.674			
Anxiety	0.04(0.04)	1.18	-0.03 - 0.12	0.238
Marginal R ² / Conditional R ²	0.011 / 0.674			
TAS-20	0.04(0.02)	1.52	-0.01 - 0.08	0.128
Marginal R ² / Conditional R ²	0.018 / 0.674			
IUS-12	-0.04(0.02)	-1.66	-0.09 - 0.01	0.097
Marginal R ² / Conditional R ²	0.025 / 0.776			
N.B. Depression, Anxiety, and Stress were measured using the DASS-21 (Lovibond & Lovibond, 1995). Abbreviations: TAS-20 (20-Item Toronto Alexithymia Scale; Bagby et al., 1994); IUS-12 (12-Item Intolerance of Uncertainty Scale; Carleton et al., 2007). Bolded values denote statistical significance ($p < 0.05$)				

3. Results of Group differences on Confidence Estimates and Mediation Analyses with Traits as Mediator Variables

The results in Tables C.4a and C.4b are from the between-group differences and the interaction with condition on confidence estimates however none of the results were statistically significant.

Table C.4. Results from Analyses on Confidence Estimates (Chapter 4).

Predictors	β(SE)	t	95% CI	p
Table C.4a: Confidence for Post-Instructions Pre-Sipping Baseline Fullness State Estimate				
Main Effect of Group				
(Intercept)	71.58(2.80)	25.6	66.09- 77.07	<0.001
AN	4.70(4.64)	1.01	-4.39 - 13.79	0.311
AN-WR	-1.53(5.13)	-0.30	-11.58 - 8.52	0.765
Marginal R ² / Conditional R ²	0.013/0.748			
Interaction between Group x Condition				
(Intercept)	72.23(2.91)	24.8	66.53 - 77.93	<0.001
AN	2.77(4.81)	0.57	-6.67 - 12.20	0.566
AN-WR	-0.38(5.32)	-0.1	-10.82 - 10.05	0.942

Condition (High Quantity)	-1.31(1.57)	-0.8	-4.38 - 1.77	0.404
AN x Condition (High Quantity)	3.86(2.60)	1.49	-1.23 - 8.96	0.137
AN-WR x Condition (High Quantity)	-2.29(2.87)	-0.8	-7.92 - 3.34	0.425
Marginal R ² / Conditional R ²	0.016 / 0.750			

Table C.4b: Confidence for Second Fullness State Guess

Group Main Effect				
(Intercept)	72.97(3.00)	24.3	67.09 - 78.85	<0.001
AN	7.14(4.40)	1.62	-1.48 - 15.77	0.105
AN-WR	-3.81(4.87)	-0.78	-13.34 - 5.73	0.434
Marginal R ² / Conditional R ²	0.034 / 0.663			
Interaction between Group x Condition				
(Intercept)	74.31(2.81)	26.5	68.79 - 79.82	<0.001
AN	7.58(4.66)	1.63	-1.55 - 16.71	0.104
AN-WR	-3.48(5.15)	-0.68	-13.58 - 6.61	0.499
Condition (High Quantity)	-2.68(1.85)	-1.45	-6.30 - 0.94	0.146
AN x Condition (High Quantity)	-0.87(3.06)	-0.29	-6.86 - 5.11	0.775
AN-WR x Condition (High Quantity)	-0.64(3.38)	-0.19	-7.26 - 5.98	0.849
Marginal R ² / Conditional R ²	0.040 / 0.660			
N.B. Bolded values denote statistical significance ($p < 0.05$)				

Given that significant differences were found on the effect of group on the Confidence estimates for the First Fullness State Guess and Post-Sipping Fullness State Rating, the steps for mediation analysis were applied. The significant results are presented in the main Results section (Table 4.7) and Tables C5.1 and C5.2 present the non-significant results from the step 2 of the mediation analysis (i.e., effect of mediator variable on DV of interest).

Table C.5. Results from Step 2 of the Mediation Analysis (Chapter 4).

Predictors	β (SE)	t	95% CI	p
Table C.5a: Effect of Traits on Confidence for First Fullness State Guess				
Depression	-0.15(0.17)	-0.91	-0.49 - 0.18	0.363
Marginal R ² / Conditional R ²	0.138 / 0.669			
Anxiety	-0.40(0.22)	-1.80	-0.84 - 0.03	0.071
Marginal R ² / Conditional R ²	0.024 / 0.626			
TAS-20	-0.01(0.14)	-0.8	-0.29 - 0.27	0.937
Marginal R ² / Conditional R ²	0.000 / 0.625			
IUS	0.13(0.16)	0.79	-0.19 - 0.45	0.428
Marginal R ² / Conditional R ²	0.005 / 0.621			
Table C.5b: Effect of Traits on Confidence for Post-Sipping Fullness State Rating				
Depression	-0.03(0.18)	-0.20	-0.38 - 0.31	0.844

Marginal R ² / Conditional R ²	0.000 / 0.723			
Stress	-0.28(0.18)	-1.56	-0.64 - 0.07	0.12
Marginal R ² / Conditional R ²	0.020 / 0.723			
Anxiety	-0.34(0.23)	-1.46	-0.80 - 0.12	0.145
Marginal R ² / Conditional R ²	0.018 / 0.723			
TAS-20	0.13(0.15)	0.85	-0.17 - 0.42	0.397
Marginal R ² / Conditional R ²	0.006 / 0.723			
IUS-12	0.01(0.17)	0.06	-0.33 - 0.35	0.953
Marginal R ² / Conditional R ²	0.000 / 0.719			
N.B. Depression, Anxiety, and Stress were measured using the DASS-21 (Lovibond & Lovibond, 1995). Abbreviations: TAS-20 (20-Item Toronto Alexithymia Scale; Bagby et al., 1994); IUS-12 (12-Item Intolerance of Uncertainty Scale; Carleton et al., 2007). Bolded values denote statistical significance ($p < 0.05$)				