Amelie Marleen Hübner

A Role of Interoception for Emotion Inference: Behaviour and Neurophysiology in Healthy Individuals and Frontotemporal Dementia

- 2022 -

A Role of Interoception for Emotion Inference: Behaviour and Neurophysiology in Healthy Individuals and Frontotemporal Dementia

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium (Dr. rer. nat.)

im Fach Psychologie

eingereicht im Fachbereich Psychologie und Sportwissenschaft

der Westfälischen Wilhelms-Universität Münster

von

Amelie Marleen Hübner

geboren am 08.12.1992 in Hameln

Dekan:	Prof. Dr. Guido Hertel
Erste Gutachterin:	Prof. Dr. Ricarda I. Schubotz
Zweiter Gutachter:	Prof. Dr. Markus Junghöfer
Datum der mündlichen Prüfung:	_____
Datum der Promotion:	_____

Acknowledgement

Eine Doktorarbeit zu erreichen war immer mein großer Wunsch. Wenn ich auf die letzten Jahre zurückblicke, waren diese sehr lehrreich, voller schöner und wichtiger Erfahrungen, aber auch nicht immer leicht. Ich kann mich sehr glücklich schätzen so viele Menschen in meinem Leben zu haben, die mich immer unterstützen und an mich glauben. All denjenigen möchte ich im Folgenden meine Dankbarkeit ausdrücken.

Zunächst möchte ich meiner Erstbetreuerin, Prof. Dr. Ricarda Schubotz, danken. Insbesondere dafür, dass sie mir die Möglichkeit gegeben und mir das Vertrauen geschenkt hat, die Doktorarbeit gemeinsam mit der Psychotherapieausbildung zu bewältigen. Danke für die Unterstützung im Laufe der Jahre, für die Anleitung, die Motivation und die Inspiration auf wissenschaftlicher und auch auf persönlicher Ebene. Des Weiteren möchte ich meinem Zweitbetreuer Prof. Dr. Markus Junghöfer danken für die Übernahme und das Interesse an meiner Arbeit und seine sehr freundliche, aufgeschlossene Art. Ein besonderer Dank gilt Dr. Ima Trempler für ihre wissenschaftliche Kompetenz und Freundschaft. Ich hatte großes Glück mit dir gemeinsam an diesem Projekt arbeiten zu dürfen. Ich habe viel von dir lernen können, wir konnten aber auch über alles reden, wofür ich sehr dankbar bin. Darüber hinaus möchte ich den ehemaligen und aktuellen Mitgliedern der Arbeitsgruppe Biologische Psychologie an der Universität Münster danken: Lena Schliephake, Benjamin Jainta, Sophie Siestrup, Falko Mecklenbrauck, Nadiya El-Sourani, Marlen Roehe, Nina Heins, Jennifer Pomp und Rosari Naveena Selvan. Danke für die vielen guten Gesprächen und Hilfe, wann immer sie benötigt wurde. Ein ebenso großer Dank geht an Monika Mertens und Jutta Linke, die mir bei der Durchführung der Experimente und organisatorischen Fragen immer zur Seite standen.

Darüber hinaus möchte ich den Menschen ganz besonders danken, die mich in den vergangenen Jahren in meinem Alltag besonders begleitet haben: meine Freunde und Familie. Ich hätte nie gedacht, dass ich in der Psychotherapieausbildung so viele wunderbare Freunde

finden würde. Maike Richter, Sarah Kinkel, Carolin Bosche, Lena Waltemate, Hendrik Hackfeld, Emil Vollmers, Thole Hoppen und meine Wegbegleiterin seit dem Studium Aline Kasek, eure Freundschaft ist etwas ganz Besonderes für mich. Danke für so viele großartige, lustige und unterstützende Momente in den letzten Jahren. Außerdem möchte ich Benjamin Schäfer und Marius Messinger für den Ausgleich an vielen Abenden und die Nachbarschaft danken, aus der eine tiefe Freundschaft gewachsen ist. Danke auch Jan Middell für die Unterstützung in der Promotion und meinen besten Freundinnen aus der Heimat, Jana Haas und Marina Burre, die immer an meiner Seite sind und seien werden. Ein besonderer Dank gilt auch Dr. Michael Heilemann, der mich durch seine bemerkenswerte und inspirierende Art auf meinem Weg maßgeblich geprägt hat.

Zu guter Letzt möchte ich meiner Familie danken – ich wäre heute nicht da wo ich bin, wenn ihr mich nicht in allem unterstützt hättet. Meine Mutter, Kerstin Hübner, von der ich sicherlich die psychologische Ader geerbt habe. Du hast immer Zutrauen in mich und meine Fähigkeiten, hast mich immer ermutigt und bist bedingungslos für mich da. Ich verdanke dir so vieles von dem, wer ich heute bin, dass ich es kaum in Worte fassen kann. Zudem habe ich großes Glück, eine so tolle große Schwester zu haben. Danke, Annabel Hübner, für deine starke, selbstbewusste Art, die ich mir immer zum Vorbild genommen habe. Und natürlich auch ein großer Dank an meinem Schwager, Daniel Hacker, der der beste Zuwachs für die Familie ist, den man sich wünschen kann. Danke an meinen Freund, Constantin Jahnke, für deine Liebe und den Halt, den du mir gibst. Dich an meiner Seite zu haben, ist mein größtes Glück und ich freue mich auf die gemeinsame Zeit, die vor uns liegt.

Letztlich habe ich eine Doktorarbeit insbesondere wegen einer Person angestrebt – meinem Vater. Leider konnte er diesen Weg nicht mehr zu Ende begleiten. Nichts auf der Welt hätte ich mir mehr gewünscht, Papa, als dass du diese Arbeit noch hättest lesen können. Ich widme diese Dissertation meinem Vater, Christoph Hübner.

Table of Contents

Summary **7**

Zusammenfassung **9**

List of Original Publications **11**

1. Theoretical and Empirical Background **12**

1.1 The Concept of Interoception 13

1.2 The Interoceptive Predictive Coding Approach to Emotion 16

1.2.1 Theories of Emotion 16
1.2.2 Predictive Coding Model of Interoception 17
1.2.3 Cortical Hierarchy and Cortical Representations of Interoception 19

1.3 The Role of Interoception in Emotion and Social Cognition 22

1.3.1 Interoception and Emotion 22
1.3.2 Interoception and Social Cognition 24

1.4 Frontotemporal Dementia – An Impairment of Interoceptive Inference? . 26

1.4.1 Cognitive Profile of Frontotemporal Dementia 26
1.4.2 Impaired Interoception in Frontotemporal Dementia 28

2. Research Questions **30**

3. Research Articles **33**

3.1 Study I: Interoceptive Sensibility Predicts the Ability to Infer Others' Emotional States 33

3.2 Study II: Selective Impairment of Emotion Inference in Behavioural Variant Frontotemporal Dementia? A Pilot Study. 58

3.3 Study III: Interindividual Differences in Interoception Modulate Behavior and Brain Responses in Emotional Inference 91

4. General Discussion and Future Directions **131**

4.1 Summary of the Presented Studies 131

4.2 Effects of Interoception on Emotion Recognition 133

4.2.1 Recognizing Emotion with Different Interoceptive Abilities 133
4.2.2 Neural Correlates of Dynamically Unfolding Emotion 136

4.3 The Role of Interoception in Probabilistic Inference of Emotion 139

4.3.1 Interoception and Adaptation to Emotion Probabilities 139
4.3.2 Neural Correlates of Context-Sensitive Emotion Inference 142
4.3.3 Interoceptive Abilities and Processing of Interoceptive Prediction Errors 145

4.4 Domain Specific and Unspecific Effects of Interoception 147
4.5 Emotion Inference in Frontotemporal Dementia 148
4.5.1 Recognition of Emotional Changes in Frontotemporal Dementia 149
4.5.2 Context-Sensitive Emotional Inference in Frontotemporal Dementia 152
4.5.3 Domain-Specific Impairments in Frontotemporal Dementia 154
4.6 Critical Reflection and Practical Applications 156
5. Conclusion 159
References 160
Abbreviations 175
List of Figures 176
Curriculum Vitae 177
Declarations 179

Summary

Emotional sensations and inferring the emotional state of others, for example from facial expressions, are essential aspects of our everyday lives. Thereby, emotions are associated with a broad range of physiological changes. Recent models of emotion suggest that our emotional experiences and the way we perceive the emotions of others largely depend on predictions for the causes of bodily sensations that are based on our past experiences. The sense of the body's internal states is called interoception and people differ in terms of how sensible they are to interoceptive signals and changes, and how accurately they detect internal bodily sensations. A better ability to perceive interoceptive signals might lead to a richer emotional experience and this, in turn, could facilitate a better understanding of others' emotions. In contrast, clinical manifestations of impaired interoception include, among others, difficulties in recognizing emotions, which can be observed in patients suffering from behavioural variant frontotemporal dementia (bvFTD).

The present thesis aimed to study the effects of interoceptive abilities on emotion recognition and emotional context adaptation as well as their underlying neural networks and physiological correlates. Moreover, the present work aimed at investigating potential impairments of interoceptive predictive processing in patients with bvFTD. To this end, three studies were conducted, which involved behavioural measurements, sympathetic autonomic activity assessed by skin conductance response (SCR) as well as functional magnetic resonance imaging (fMRI). In all studies, the participants' task was to discriminate emotional expressions that evolved from a neutral facial expression into a happy or fearful one, with different probabilities implemented for the occurrence of valence.

The first study assessed the link between sensibility of interoceptive changes and adaptation to emotion probabilities in young healthy subjects. Higher interoceptive sensibility (IS) was found to facilitate the recognition of emotional changes in facial expressions of others

and to entail a more precise adaptation to biased probabilities of emotional expressions. Moreover, in individuals with higher IS, sympathetic arousal increased in response to more surprising and unexpected emotions, indicating heightened arousal in response to violations of expectations in the course of learning from increasing probability of a specific valence. Based on these results, the second study focused on potential impaired emotion recognition and deficits in predictive processing in bvFTD patients. Patients with bvFTD compared to healthy controls showed deficient emotion recognition as a function of uncertainty about emotional changes, which can be attributed to impaired interoceptive abilities in bvFTD. Deviant modulation of autonomic sympathetic activity suggests that patients cannot effectively process interoceptive prediction errors. Finally, the third study focused on effects of differences in IS and interoceptive accuracy (IAcc) and on the neural regions involved in the processing of emotion recognition and context-sensitive emotion inference. The results provide evidence for different neural implementations of IS and IAcc and amplified processing of interoceptive prediction errors in the insula and caudate nucleus in individuals with higher IAcc.

The results of this thesis profoundly contribute to a deeper understanding of the role of interoception in emotion recognition and context-sensitive emotion inferences and the underlying neurophysiology in healthy individuals and patients with bvFTD.

Zusammenfassung

Emotionale Empfindungen und Rückschlüsse auf die emotionalen Zustände anderer Personen, beispielsweise über ihre Gesichtsausdrücke, sind ein wesentlicher Aspekt unseres täglichen Lebens. Dabei sind Emotionen mit einer Vielzahl von physiologischen Veränderungen verbunden. Neuere Modelle zu Emotionen gehen dabei davon aus, dass unsere emotionalen Erfahrungen und die Art und Weise, wie wir die Emotionen anderer Personen wahrnehmen, weitgehend von Vorhersagen über die Ursachen unserer Körperempfindungen abhängen, die auf unseren früheren Erfahrungen beruhen. Der Sinn für die inneren Zustände des Körpers wird als Interozeption bezeichnet und Menschen unterscheiden sich darin, wie sensibel sie für interozeptive Signale und Veränderungen sind und wie genau sie innere Körperempfindungen erkennen. Eine bessere Fähigkeit, interozeptive Signale wahrzunehmen, könnte dabei mit einem stärkeren emotionalen Erleben einhergehen, welches wiederum ein besseres Verständnis der Emotionen anderer Personen begünstigen könnte. Im Gegensatz dazu sind klinische Manifestationen einer gestörten Interozeption unter anderem Schwierigkeiten beim Erkennen von Emotionen, wie sie auch bei Patienten mit verhaltensbetonter Variante der Frontotemporalen Demenz (bvFTD) beobachtet werden.

Ziel der vorliegenden Arbeit war es, die Auswirkungen interozeptiver Fähigkeiten auf die Emotionserkennung und die emotionale Kontextanpassung sowie die zugrunde liegenden neuronalen Netzwerke und physiologischen Korrelate zu untersuchen. Darüber hinaus sollte eine mögliche Beeinträchtigung der interozeptiven prädiktiven Verarbeitung bei Patient/innen mit bvFTD untersucht werden. Zu diesem Zweck wurden drei Studien durchgeführt, die Verhaltensmessungen, die Messung der autonomen Sympathikusaktivität durch die Hautleitfähigkeit (SCRs) sowie funktionelle Magnetresonanztomographie (fMRT) umfassten. In allen Studien bestand die Aufgabe der Teilnehmer/innen darin, emotionale Gesichtsausdrücke zu unterscheiden, die sich von einem neutralen Gesichtsausdruck zu einem

glücklichen oder ängstlichen Ausdruck entwickelten, wobei unterschiedliche Wahrscheinlichkeiten für das Auftreten der Valenzen implementiert wurden.

Die erste Studie untersuchte den Zusammenhang zwischen der Sensibilität für interozeptive Veränderungen und der Emotionserkennung sowie Anpassung an emotionale Kontexte bei jungen gesunden Proband/innen. Es zeigte sich, dass eine höhere interozeptive Sensibilität (IS) das Erkennen von emotionalen Veränderungen im Gesichtsausdruck anderer Personen erleichtert und eine präzisere Anpassung an die verzerrten Wahrscheinlichkeiten emotionaler Ausdrücke mit sich bringt. Darüber hinaus stieg bei Personen mit höherer IS die physiologische Erregung als Reaktion überraschendere und unerwartete Emotionen an, was auf eine stärkere physiologische Aktivierung als Reaktion auf Erwartungsverletzungen hindeutet. Auf der Grundlage dieser Ergebnisse konzentrierte sich die zweite Studie auf mögliche Defizite in der Emotionserkennung und prädiktiven Verarbeitung bei Patient/innen mit bvFTD. Die Patient/innen mit bvFTD zeigten eine schlechtere Erkennung von Emotionen in Abhängigkeit von der Ungewissheit über die Veränderung der Emotion, was auf eine Beeinträchtigung der interozeptiven Fähigkeiten bei bvFTD zurückgeführt werden kann. Eine abweichende Modulation der autonomen sympathischen Aktivität deutet darauf hin, dass die Patient/innen interozeptive Vorhersagefehler nicht effektiv verarbeiten können. Die dritte Studie schließlich konzentrierte sich auf die Auswirkungen von Unterschieden in der IS und der interozeptiven Genauigkeit (IAcc) sowie auf die neuronalen Regionen, die an der Verarbeitung der Emotionserkennung und der Anpassung an den emotionalen Kontext beteiligt sind. Die Ergebnisse deuten auf eine unterschiedliche neuronale Implementierung von IS und IAcc und eine verstärkte Verarbeitung von interozeptiven Vorhersagefehlern der Insula und im Nucleus caudatus bei Personen mit höherer IAcc hin. Die Ergebnisse dieser Arbeit ermöglichen ein tieferes Verständnis der Rolle der Interozeption bei der Emotionserkennung und kontextsensitiven emotionalen Schlussfolgerungen sowie der zugrunde liegenden Neurophysiologie bei gesunden Personen und Patient/innen mit bvFTD.

List of Original Publications

This thesis is based on the following original research articles:

Study I.	Hübner, A. M., Trempler, I., Gietmann, C., & Schubotz, R. I. (2021). Interoceptive Sensibility Predicts the Ability to Infer Others' Emotional States. *Plos one*, *16*(10), e0258089.
Study II.	Hübner, A. M., Trempler, I., Kloth, N., Duning, T., Johnen, A., & Schubotz, R. I. (submitted to journal). Selective Impairment of Emotion Inference in Behavioural Variant Frontotemporal Dementia? A Pilot Study. *Manuscript submitted for publication.*
Study III.	Hübner, A. M., Trempler, I., & Schubotz, R. I. (submitted to journal). Interindividual Differences in Interoception Modulate Behavior and Brain Responses in Emotional Inference. *Manuscript submitted for publication.*

1. Theoretical and Empirical Background

Throughout our lives, our bodies are constantly sending information about our internal conditions to the brain. Most of the time the processing of these signals takes place without us being aware of it, for example of signals that help keeping our blood pressure constant. Sometimes we become aware of these sensations, such as the beating of our heart or the feeling of "butterflies in our stomach" when we are excited or falling in love. This sense of the body's internal states is called interoception and interoception plays an essential role in our everyday life. For example, for decision-making it is crucial to be able to sense the internal state of the body, i.e., the "gut feeling" when making decisions. Interoceptive abilities are thought to have important implications for our mental and physical health and are proposed to influence a number of psychological processes, such as emotion perception. However, people differ in the extent to which they are sensible to these inner states and precisely these differences are thought to play a role in how strongly we feel emotions and perceive the emotions of others.

Traditionally, both in cognitive (neuro)science and in the early days of emotion research, the brain was understood as a "stimulus-response" organ consisting of neurons that remain inactive until stimulated (Barrett & Simmons, 2015). For a long time, this was also the prevailing concept regarding emotions and it was assumed that an emotional sensation (e.g., anger) is followed by a physical reaction (e.g., increasing heartbeat). However, the concept of emotion has changed fundamentally over the last decades. For instance, the pioneering work of Antonio Damasio proposed that emotional sensations begin with changes in bodily states, called "somatic markers" (Damasio, 1994, 1996). The assumption that emotional experiences arise from contextualized perception and interpretation of bodily signals has been further developed in the framework of predictive coding, which holds that emotions are determined by predictions about the causes of bodily sensations based on past experiences (Seth, 2013; Seth, Suzuki, & Critchley, 2012). Accordingly, it is suggested that clinical manifestations such as deficits in

emotional awareness, difficulty in identifying emotions or inferring the mental state of others may occur when interoceptive signals can no longer be used to guide behaviour (Brewer, Cook, & Bird, 2016; Quadt, Critchley, & Garfinkel, 2018; Sowden, Brewer, Catmur, & Bird, 2016). These deficits are also observed in patients suffering from behavioural variant frontotemporal dementia (bvFTD), a neurodegenerative disease mainly characterized by progressive changes in personality, such as loss of empathy and socially inappropriate behaviour (Johnen & Bertoux, 2019; Rascovsky et al., 2011). However, although predictive coding models of interoceptive processing and emotion are theoretically well elaborated, only few empirical studies have examined the effects that interindividual differences or impairments in interoceptive abilities have on, for example, emotion recognition in others.

In this thesis, I will focus on effects of differences in interoceptive abilities on emotion recognition and emotional context adaptation, their underlying neural networks and physiological correlates. In the following sections, I will give a detailed description of the concept of interoception. I will then briefly review existing theories of emotion and the role of interoception, including their neural implementations, followed by a summary of current findings on the role of interoceptive processing in emotion and social cognition. Finally, I will discuss frontotemporal dementia (FTD), with a focus on the behavioural variant of frontotemporal dementia (bvFTD), a pathological neurodegenerative disorder with deficits in emotional awareness and disturbances in emotion recognition, possibly due to deficits in interoceptive prediction.

1.1 The Concept of Interoception

The concept of interoception emerged in the early 20th century and has gained increasing popularity in the academic literature since. The first known use of the term interoception refers back to Sherrington (1906) and his book "The Integrative Action of the Nervous System". Sherrington refers to interoception as sensory representations of the interior

of the body (viscera) and differentiates them from exteroceptive sensations that originate outside the body, like touch, to which he also included sensations related to injury (nociception) and temperature (thermosensation) (Ceunen, Vlaeyen, & Van Diest, 2016).[1] More recent definitions view interoception as not restricted to visceroception[2] and include sensory inputs associated with physiological states of all organs (Craig, 2002, 2003). According to this currently prevailing notion, interoception is defined as "the sense of the physiological condition of the body" (Craig, 2002). This definition also includes sensations of joints, tendons and muscle movements and the position of the body (proprioception). The signals about the body's physiological state and motivational needs form the afferent axis to internal (autonomic and hormonal) physiological control, such as low-level blood chemistry monitoring, representations of skin and body temperature and touch-evoked sensations (Craig, 2009; Tsakiris & Critchley, 2016). Through homeostatic control and allostatic adaptation, interoception supports the stability of the organism and guides behaviour through feelings such as hunger, thirst and dyspnoea (Berntson, Cacioppo, & Quigley, 1993; Tsakiris & Critchley, 2016).

People's responses to questions about their bodily sensations often differ from the actual physical signals (e.g., Köteles, 2021). Accordingly, it proves difficult to measure interindividual differences in interoceptive abilities. This has led to inconsistency in the literature to date regarding the methods used for assessing interoceptive abilities and to the emergence of several different facets of interoception based on the various measurements (Garfinkel & Critchley, 2013). In this thesis, I refer to a distinction according to Garfinkel et al. (2015), proposing that interoception is a three-dimensional construct differentiating between interoceptive sensibility (IS), interoceptive accuracy (IAcc) and interoceptive awareness (IAw). IS quantifies the self-

[1] Although Sherrington does not specifically use the word interoception in his book, he speaks of "interoceptors" and "interoceptive segments". He uses the term "interoceptive" to refer to receptors located within the viscera and excludes all other receptors in the body, which he defines as either exteroceptive or proprioceptive (Ceunen et al., 2016).

[2] Visceroception is defined as the part of interoception that perceives and processes signals arising from internal organs such as the heart, lungs or stomach (Jänig, 1996).

reported beliefs about bodily-awareness and the perceived sensibility to detect internal bodily fluctuations (Ernst, Böker, et al., 2013; Terasawa, Shibata, Moriguchi, & Umeda, 2013). IS can be assessed using questionnaires, such as the Body Perception Questionnaire (BPQ; Porges, 1993) or the Multidimensional Assessment of Interoceptive Awareness, Version 2 (MAIA-2; Mehling, Acree, Stewart, Silas, & Jones, 2018). IAcc is defined as performance in accurately detecting and tracking internal bodily sensations (Ceunen, Van Diest, & Vlaeyen, 2013).[3] Since the heartbeat is a distinct and regular internal event that can easily be quantified, the accuracy of the perception of the heartbeat is a widely used measure of IAcc. The perception of heartbeats can be assessed by counting the heartbeats during a specified time periods ("heartbeat detection", HDT; Schandry, 1981) or through reported synchrony of the heartbeats with external stimuli, e.g. a tone ("heartbeat discrimination"; Brener & Kluvitse, 1988).[4] Accurate perception of interoceptive signals can also be provided by other sensations such as gastric or respiratory sensations, although to date there are only few standardized tasks for measuring non-cardiovascular types of IAcc (e.g., Van Den Houte et al., 2021; Van Dyck et al., 2016).

In addition, it is important to emphasize that objective measures of interoception (e.g., quantified by accuracy of heartbeat perception) are not necessarily correlated with subjective measures reported, for example, in questionnaires (e.g., Ceunen et al., 2013; Forkmann et al., 2016; Meessen et al., 2016). The dissociation between objective and subjective measures of interoceptive abilities has been conceptualized as interoceptive awareness (IAw), which is the third and final component in the distinction according to Garfinkel et al., (2015). IAw is defined as a metacognitive component of interoception and measures the extent to which the objective

[3] In a prior distinction of interoception presented by Garfinkel and Critchley (2013), the term "interoceptive sensitivity" was used to describe the accurate detection of body signals. In the distinction of Garfinkel et al. (2015), the term was changed to "interoceptive accuracy" referring to common usage by other authors such as Ceunen, Van Diest, & Vlaeyen (2013).

[4] In the literature, the terms "heartbeat detection task", "heartbeat counting task" or "heartbeat perception task" are used inconsistently. In recent references, the task of reporting heartbeats during specific time periods is often referred to as the "heartbeat counting task" (e.g., Zamariola et al., 2018), to distinguish it from tasks of heartbeat discrimination. In this thesis, with respect to the above explanation and the original use of Schandry (1981), the term "heartbeat detection task" and the abbreviation "HDT" is used when the task is to count the heartbeat during a specific time interval.

performance (IAcc) matches the subjective evaluation, e.g. determined using a receiver-operating-characteristic (ROC) curve (Barrett, Dienes, & Seth, 2013). A high IAw score is proposed to reflect a person's ability to evaluate when good or bad interoceptive decisions are made (Garfinkel et al., 2015). Although the clinical importance of IAw is increasingly emerging, there is still inconsistency regarding the classification and distinct measures of this metacognitive facet of interoception (Murphy, Catmur, & Bird, 2019).

In summary, interoception supports the physical stability of the organism through receiving and appraising internal bodily signals and through physiological control. However, the effects of interoception go beyond this homeostatic and allostatic control: interoception is thought to be crucial for mental and physical health and, most notably, to be a precursor to emotional experience, as will be discussed in the following section.

1.2 The Interoceptive Predictive Coding Approach to Emotion

1.2.1 Theories of Emotion

An interoceptive concept of emotion and the hypothesis that emotions have a physical basis, was first formulated by the psychologist William James as early as 1884.[5] Until then, it was proposed that a mental affection (emotion) occurs in response to an external event and that this emotion gives rise to a bodily expression. James proposed that viscero-afferent feedback and physical changes precede emotions, stating that "bodily changes follow directly the perception of the exciting fact, and that our feeling of the same changes as they occur is the emotion" (James, 1884). Since the publication of James' theory[6], there has been an ongoing debate about the extent to which emotions are embodied (e.g., Barbalet, 1999; Ellsworth, 1994)

[5] The idea that emotions are partly based on bodily responses is also reflected in early philosophical writings (e.g. Descartes, 1649/1989).

[6] The "James' theory" is often referred to as the "James-Lange theory", because in 1885, one year after the publication of James' theory, the Danish physiologist Carl Lange presented a similar theory in his book "Über Gemütsbewegungen" (The emotions), which essentially agreed with James' theory.

and various extensions of the theory have been developed. Schachter and Singer (1962) also considered visceral arousal to be a prerequisite for emotions, but emphasized that emotional experience is determined by the subsequent cognitive appraisal of bodily arousal. In the 1990s, Antonio Damasio proposed the "somatic marker hypothesis", stating that emotional events begin with unconscious changes in bodily states, so called "somatic markers " (Damasio, 1994, 1996). He describes emotions as "action programmes" that trigger a biologically preset bodily response (and a particular facial expression) and are associated with previous situations and their past outcomes (Damasio & Carvalho, 2013). Changes in the state of the body are sensed and mapped by the interoceptive system and sensory information about these changes are sent to the brain, where they are represented as "somatic markers" that guide decision-making and can be consciously experienced as feelings (Barrett, 2017; Damasio & Carvalho, 2013).[7] Damasio initially relied on observations of patients suffering from damage to prefrontal areas, who, given a normal intellect, showed a lack of physiological responses, emotions and feelings in response to arousing stimuli, in addition to difficulties in decision making (Damasio, 1996). Damasio (1996) concluded that decision-making deficits following such an injury result from the incapacity to use emotions to guide future behaviour based on past experience. Overall, these frameworks provide a starting point and form a key component for recent views on emotion, such as the predictive coding model of interoception assuming emotions as "interoceptive inference" (Seth, 2013; Seth et al., 2012).

1.2.2 Predictive Coding Model of Interoception

Before approaching the predictive coding model of interoception, it is first important to introduce predictive processing as a fundamental principle of brain function (Bubic, Cramon,

[7] Damasio classified emotions and feelings as different phenomena. While he referred to emotions as "action programmes largely triggered by external stimuli", he describes that feelings arise only when the bodily changes are perceived, i.e., as the mental experiences that can (but does not necessarily have to) accompany bodily states (Damasio & Carvalho, 2013).

& Schubotz, 2010). The classical perspective on neural processing conceived the brain as a "stimulus-response" organ that reacts to sensory information from the environment through bottom-up pathways. In contrast, according to the framework of predictive coding, brains do not passively await sensory input, but instead continuously predict the most likely sensory stimulation based on internal, so-called generative models (Clark, 2013). Prediction is proposed to be hierarchically organised, with higher cortical areas generating predictive models that are communicating through feedback connections to lower sensory areas (Friston, 2005). Mismatches between predicted models and actual information are called prediction errors (Seth et al., 2012). According to the model of predictive coding, the brain is continuously attempting to minimize these prediction errors, for example by updating predictions until they match the sensory signals (Seth, 2013). Prediction errors are weighted differently in terms of precision, meaning that the reliability of the signals determines their influence on subsequent hierarchical processing (Friston, 2005). The higher the precision of a prediction error, the greater is its impact on a possible revision of the prediction model. Dynamic precision-weighting can accordingly modulate the balance between top-down and bottom-up signal flow: highly precise prediction errors would bias bottom-up processing, whereas highly precise predictions would bias top-down processing (Friston, 2008; Seth, 2013).

The framework of predictive coding has been suggested and experimentally demonstrated for neural and cognitive processes underlying, e.g., perception, cognition, and action (Bubic et al., 2010; Friston, 2010; Rao & Ballard, 1999). In recent years, this notion has been extended to interoception, assuming that emotional experiences are determined by predictions about the expected interoceptive state of the body (Barrett & Simmons, 2015; Seth et al., 2012). Following the principles of predictive coding, emotional states are proposed to be generated by active top-down inference, so-called interoceptive inferences, i.e., actively derived predictive models of the causes of interoceptive signals based on past experiences (Seth, 2013). Similarity of actual interoceptive signals is computed against the predictions, their related

energy costs and potential benefits for the body (Barrett et al., 2016). Differences between predicted and actual interoceptive signals convey information about interoceptive changes and the resulting prediction error drives autonomic reflexes to ensure physiological homeostasis or allostasis, e.g., sympathetic smooth-muscle vasodilatation as a reflexive response to predicted interoceptive consequences of "blushing with embarrassment" (Seth & Friston, 2016). Again, precision plays a key role here as it regulates the relative weights of predictions and prediction errors and can be optimized, for example, by orientating attention to interoceptive changes (Ainley, Apps, Fotopoulou, & Tsakiris, 2016; Seth, 2013).

In contrast to the presented classical approaches to emotions, the interoceptive predictive coding model neither assumes a clear separation between cognitive and noncognitive processing (e.g., Schachter & Singer, 1962), nor does it postulate emotions as physical states separate from feelings (e.g., Damasio & Carvalho, 2013). However, the interoceptive predictive coding approach builds on the idea presented in these models that emotional experiences are the result of a contextualized perception of changes in physical state and extends the assumption including a hierarchy of (precision-weighted) predictions and prediction errors (Seth & Friston, 2016). The notion of hierarchical interoceptive predictive processing is complemented by a hierarchical structure of neuroanatomical brain regions involved in interoceptive processing, as will be described in the following section.

1.2.3 Cortical Hierarchy and Cortical Representations of Interoception

An interoceptive hierarchy has been described that extends from spinal visceral afferents to subcortical structures and projecting to the anterior insula cortex (AIC), the anterior cingulate cortex (ACC) and the orbitofrontal cortex (OFC) situated at the top of the interoceptive hierarchy (Ainley et al., 2016; Critchley & Harrison, 2013; Seth & Friston, 2016) (for an overview of a proposed neural circuit for hierarchical integration of interoceptive information, see Figure 1).

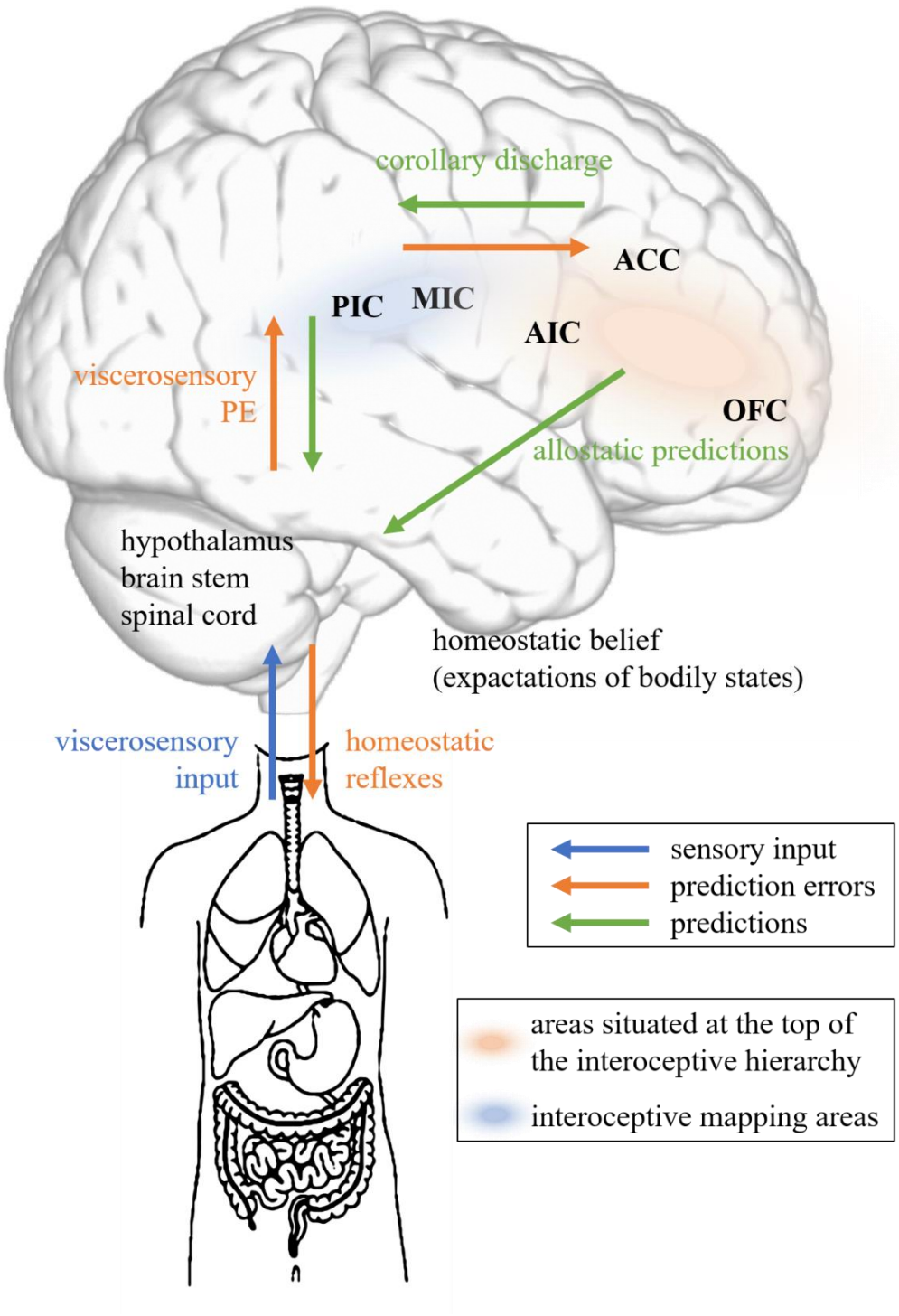

Figure 1. A (simplified) summary of the proposed neuroanatomy for integrating interoceptive information. AIC, ACC and OFC are situated at the top of the interoceptive hierarchy, embodying a generative model of viscerosensory inputs that enables one to infer and predict current and future bodily states (Stephan et al., 2016). These regions receive ascending projections from the posterior and mid-insula, which are thought to signal prediction errors. According to Barrett and Simmons (2015), new interoceptive signals are transmitted to the middle and posterior insula, where differences between the predicted and actual interoceptive signal are computed, with the posterior insula as the primary interoceptive mapping area in the cortex. AIC, ACC and OFC possess connections to hypothalamus, brain stem nuclei and spinal cord (relayed in part by amygdala, periaqueductal gray (PAG), and basal ganglia), that

presumably convey allostatic predictions which modulate the set-points of homeostatic reflexes (Stephan et al., 2016). Descending projections from these areas could send the same prediction to posterior and mid-insula as an efference copy or corollary discharge to which viscerosensory inputs can be compared. The resulting prediction errors are returned via ascending connections to AIC, ACC and OFC, allowing for adjustment of allostatic predictions. ACC: anterior cingulate cortex; AIC: anterior insular cortex; MIC: midinsular cortex; OFC: orbitofrontal cortex; PE: prediction error; PIC: posterior insular cortex. Adapted from Stephan et al. (2016).

These areas are suggested to embody a generative model of interoceptive responses and provide predictions that serve as homeostatic set-points (Seth & Friston, 2016) Among these areas, the insula appears central, activated by all interoceptive and affective stimuli and operating as a comparator that registers top-down predictions against bottom-up prediction errors (Craig, 2009; Gu, Hof, Friston, & Fan, 2013; Seth, 2013). Structurally, the insular cortex is located bilaterally beneath the temporal and frontal lobes and has widespread bidirectional connectivity to parietal, frontal, and limbic regions, making it ideally positioned to both detect and effect changes in physiological state and integrate interoceptive and exteroceptive signals (Deen, Pitskel, & Pelphrey, 2011; Seth, 2013). Within the insula, a posterior-to-anterior gradient is suggested, whereby posterior regions support primary (objective) physical features of interoceptive signals, whereas the (subjective) integration of interoceptive and motivational information is processed in the AIC (Craig, 2009, 2011; Gu et al., 2013). Moreover, the AIC is discussed as a key region for mediating the influence of interoceptive signals on behaviour and functions as a hub between brain networks involved in directing attention to external stimuli and inward attention to one's own body (Barrett & Simmons, 2015; Farb, Segal, & Anderson, 2013; Menon & Uddin, 2010).

Structurally and functionally the AIC is closely connected to the ACC and forms a cortical "salience network" involved in the detection and filtering of salient stimuli (Medford & Critchley, 2010; Menon & Uddin, 2010). The regions can be functionally distinguished, with the AIC more commonly considered a "limbic sensory area" and the ACC a "limbic motor area", as it is associated with autonomic and emotional control (Craig, 2002). Interestingly, both

AIC and ACC contain a high density of large, spindle-shaped neurons, so-called von Economo neurons (VENs) (von Economo, 1926). VENs are projection neurons that are about 4.6 times the size of neighbouring pyramidal neurons (Allman et al., 2011; Allman, Watson, Tetreault, & Hakeem, 2005). They are considered to enable rapid information integration and to play a central role in intuitive decision making (Allman et al., 2005). VENs have been found to express dopamine D3 receptor, which is thought to signal the expectation of reward under uncertainty (Allman et al., 2005; Sokoloff & Schwartz, 2002). Moreover, the serotonin 2b receptor is expressed on the VENs as well as in the human stomach (where it promotes smooth muscle contractions), providing a potential explanation for "gut feelings" (Mayer, 2011). Accordingly, VENs can be associated with the ability to make intuitive decisions that allow adaptation to a rapidly changing environment, which is required, for example, in social and emotional contexts (Allman et al., 2005).

1.3 The Role of Interoception in Emotion and Social Cognition

1.3.1 Interoception and Emotion

A fundamental assumption of the interoceptive predictive coding model is that the interplay of interoceptive predictions and interoceptive prediction errors determines emotional states, or more precisely, that the perception of interoceptive changes and resulting physiological arousal affects emotional experience (Herbert, Pollatos, & Schandry, 2007; Wiens, 2005). Consequently, it can be derived that in individuals with higher interoceptive abilities interoceptive prediction errors are cascaded up the interoceptive hierarchy rather than being suppressed by lower-level predictions (e.g., Ainley et al., 2016). Higher interoceptive abilities might, thus, correspond to increased precision of interoceptive prediction errors (e.g., by focused attention) and, consequently, be associated with experiencing higher emotional arousal (Ainley et al., 2016). A number of studies show strong associations between IAcc and

the intensity of emotional experience (e.g., Barrett et al., 2004; Critchley et al., 2004; Herbert et al., 2007; Pollatos, Herbert, Matthias, & Schandry, 2007) and identify the perception of interoceptive signals as crucial sources of variability in emotion experience and intuition (Dunn et al., 2010; Pollatos, Kirsch, & Schandry, 2005). Consistent with the model, Ainley and colleagues proposed that the affective meaning of "making a mistake" is captured as an interoceptive prediction error and found evidence that individuals with higher IAcc have higher aversion to making mistakes (Ainley et al., 2016; Ainley, Brass, & Tsakiris, 2014).

Moreover, several brain imaging studies support the assumption of strong connections between brain regions situated at the top of the interoceptive hierarchy (especially the AIC), the perception of one's bodily state and the experience of emotions (see Adolfi et al., 2017 for a meta-analysis). Specifically, greater activation in the insula, the ACC, and inferior frontal gyrus (IFG) has been repeatedly found during different variants of the HDT (see Schulz, 2016 for a meta-analysis). Wang et al. (2019) recently showed, that interoceptive attention was associated with increased AIC activation and that AIC activation was predictive of individual differences in IAcc. The interoception-related clusters presented are also thought to be involved in cerebral processing of emotions and the insula in particular is suggested to play an essential role in feeling emotions (Terasawa et al., 2021). Correspondingly, lesions or removal of the insula have been found to affect the recognition of emotions through interoceptive processing and result in reduced accuracy in the perception of highly aroused emotions (Motomura et al., 2019; Terasawa et al., 2015; Terasawa et al., 2021).

In summary, within the framework of interoceptive predictive coding it can be assumed that emotion recognition benefits from adequate access to one's own interoceptive signals. Beyond their relevance for the perception of one's own emotions, interoceptive inferences are suggested to be related to the processing of others' emotions and are thought to play a fundamental role in social cognition.

1.3.2 Interoception and Social Cognition

One proposition that emerged from the finding that a better ability to perceive bodily responses leads to a richer emotional experience, is that a richer emotional experience, in turn, facilitates a better understanding of others' emotions (Arnold, Winkielman, & Dobkins, 2019; Decety, 2011; Zaki, 2014). An underlying assumption is that our knowledge about behaviour or emotions of others is generated from our optimal explanations for perceived (motor and visceral) behaviour of others. These explanations rely on beliefs about what would have been produced by ourselves if we had been in the same intentional and emotional state (Ondobaka, Kilner, & Friston, 2017). Accordingly, inferences about the intentions and emotional states of others are proposed to be based on exteroceptive, proprioceptive and interoceptive information. For example, this can be applied to how we share bodily sensations of others: observers' predictions about their own interoceptive states play an important role in understanding others' feelings, such as the sense of feeling cold when we see someone shivering. This feeling results from interoceptive predictions, which are generated in the observer's deep hierarchical model and simultaneously generate exteroceptive predictions about a person's shivering (Ondobaka et al., 2017). In line with this notion, studies provide evidence that IAcc is positively related to higher cognitive and affective empathy (Grynberg & Pollatos, 2015). Moreover, individuals with higher IAcc were found to be more sensitive to emotional facial expressions of others (Terasawa, Moriguchi, Tochizawa, & Umeda, 2014).

These results provide preliminary evidence that higher interoceptive abilities facilitate the recognition of emotional expressions. However, individuals exhibit considerable variability in their interoceptive abilities. Further evidence is needed to clarify exactly what impact this variability has on emotion recognition, physiological parameters and brain activity. For example, there is little evidence on whether differences in interoceptive abilities are related to differences in physiological parameters that have been shown to reflect responses to prediction

errors, such as skin conductance responses (SCRs). SCR is one of the most commonly used psychophysiological measures of arousal in many scientific fields, including but not limited to psychology and psychiatry (Dawson, Schell, & Filion, 2007; Zaman, Van de Pavert, Van Oudenhove, & Van Diest, 2020). The electrodermal response reflects the activity of the sympathetic branch of the autonomic nervous system, which controls sweat gland activation and in turn influences skin conductance (Dawson et al., 2007; Zaman et al., 2020). SCR has been shown to indicate preparatory or anticipatory responses to upcoming events (Bach et al., 2010; Geuter et al., 2017). Following from this, further questions arise: whether individuals with higher interoception are more sensitive to learn from emotional contexts due to more precise interoceptive predictions, whether emotional context adaptation is accompanied by changes in physiological arousal and which brain regions play a role in this process.

Studies using facial expressions are particularly well suited for this purpose, as facial expressions are one of the primary sources of inferring emotions of others. Faces, particularly those expressing emotions, automatically attract our attention and we are able to infer the emotional meaning in a few hundred milliseconds (Haxby, Hoffman, & Gobbini, 2002; Wood, Rychlowska, Korb, & Niedenthal, 2016). When observing an emotional expression, the sensorimotor system is used to simulate the motor processes that the expresser is likely to use to generate the motor movements seen in the facial expression (Wood et al., 2016). When a perceived facial expression is simulated, we partially activate the corresponding emotional state in ourselves, which provides a basis for inferring the underlying emotion of the expresser. This can be done explicitly or implicitly (i.e., with or without facial expression). The important thing is that the meaning of the emotional expression is derived from our own prior exteroceptive and interoceptive experience of being in the presumed emotional condition (Wood et al., 2016).

Correspondingly, difficulties in the recognition of emotional facial expressions of others arise when interoception is impaired, as in alexithymia, a personality construct characterized by an impaired ability to recognize emotions based on difficulties in perceiving internal bodily

signals (Brewer et al., 2016; Grynberg et al., 2012). Alexithymia is prevalent in disorders associated with neuropathological degeneration of VENs and functional deficits of the AIC, such as in bvFTD (Seeley, 2010; Seeley et al., 2006).

1.4 Frontotemporal Dementia – An Impairment of Interoceptive Inference?

1.4.1 Cognitive Profile of Frontotemporal Dementia

Frontotemporal dementia (FTD) is a progressive neurological syndrome with varying clinical presentations and underlying pathologies and is a common cause of dementia at a young age, i.e., dementia occurring in midlife or earlier (Chaudhri, Pathak, Kumar, & Yadav, 2016). FTD is considered the third most frequent form of dementia in persons 65 years and older, and is the second most frequent form of dementia in persons aged less than 65 years (Brunnström, Gustafson, Passant, & Englund, 2009; Ratnavalli, Brayne, Dawson, & Hodges, 2002). Unlike other forms of dementia, prominent cognitive deficits may not appear until later in the course of the disease, while behavioural symptoms dominate the heterogeneous clinical picture in early stages, which raises unique diagnostic challenges (Silveri, 2007). Because of the diagnostic difficulties associated with FTD, the true prevalence of this disorder is likely to be underestimated (Bott, Radke, Stephens, & Kramer, 2014). Epidemiologic work suggests a range of an estimated point prevalence of 15–22/100.000 (Knopman & Roberts, 2011) and incidence of 2.7– 4.1/100.000 (Chaudhri et al., 2016).

As first described by Arnold Pick in 1892, FTD refers to a heterogeneous group of clinical syndromes characterized by progressive, focal neurodegeneration of the frontal and temporal lobes.[8] The neurodegeneration affects brain regions relevant to motivation,

[8]The neurologist Arnold Pick was the first to describe cases of dementia with circumscribed atrophy in frontotemporal regions based on autopsies in 1892 (Pick, 1892). Alois Alzheimer further investigated this issue and described the cellular pathology of intraneuronal inclusions, which were later known as "Pick bodies" (Alzheimer, 1911; McKhann et al., 2001). The definition of the clinical entity of this disease later became known as "Pick's disease" or "Morbus Pick". However, these terms were replaced in subsequent years by the term FTD (McKhann et al., 2001).

personality, social cognition, attention, executive functioning and language (Bott et al., 2014; Pasquier & Petit, 1997). Three clinical subtypes of FTD are distinguished, characterized by impairments in language and/or executive functions: bvFTD, semantic primary progressive aphasia (svPPA) and primary progressive non-fluent/agrammatic aphasia (nfvPPA) (Neary et al., 1998; Piguet, Hornberger, Mioshi, & Hodges, 2011). While both subtypes of primary progressive aphasia (PPA) are characterized by initial disturbance and decline of language functioning, bvFTD is associated with initial and progressive deterioration in social functioning and changes in personality (Bott et al., 2014). bvFTD accounts for half of all FTD diagnoses and is reported to be more common in men (Johnson et al., 2005). Characteristic for bvFTD are early onset (within the first three years) personality changes in the form of socially inappropriate behaviour, loss of empathy and/or sympathy, loss of manners and/or impulsivity, while episodic memory and visuospatial functions are largely preserved (Rascovsky et al., 2011). While in patients with PPA interoception is thought to be mostly preserved, recent literature suggests that in bvFTD, weak emotional signals lead to overreliance on cognitive appraisal rather than internal sensations (Mendez, 2021). Accordingly, especially in bvFTD characteristic personality changes might be based on impaired interoception, which is why in the following I will focus on this variant of FTD.

As noted above, the symptom pattern is a consequence of early-stage ACC and fronto-insular degeneration, which are regions known to contain a high density of VENs (Binney et al., 2017; Kim et al., 2012; Mandelli et al., 2016; Rosen et al., 2002). Consistently, studies show that the number of VENs is reduced by 74% in bvFTD patients, which is not attributable to a general neuronal loss. In contrast, patients with Alzheimer's disease (AD) exhibit normal VEN numbers and morphology, indicating a selective and disease-specific VEN loss in bvFTD (Seeley et al., 2006). Moreover, Zhou et al. (2010) found reduced connectivity of the fronto-insular salience network comprising AIC and ACC in bvFTD. They suggested that reduced salience network connectivity may disintegrate homoeostatic and interoceptive processing

streams, leading to degraded self-representations and impaired social behaviour. A question that follows is whether characteristics of this disease, such as personality changes, loss of empathy and socially inappropriate behaviour result from impaired interoceptive predictive processing, i.e., that patients do not gain awareness of interoceptive changes and therefore are unable to update predictive models in response to interoceptive prediction errors.

1.4.2 Impaired Interoception in Frontotemporal Dementia

In recent years, there has been a growing body of literature examining the link between impairments in interoceptive inferences and physical or mental disorders, contributing to a better understanding of various health conditions such as fatigue, depression, eating disorders, autism and anxiety (Quadt et al., 2018). Clinical manifestations of impaired interoceptive predictive processing include, for example, alexithymia (i.e. deficits in emotional awareness), difficulty in recognizing emotions from facial expressions, or inferring mental states of others (Arnold et al., 2019; Brewer et al., 2016; Sowden et al., 2016). These impairments have been consistently found in patients suffering from bvFTD (Bora, Walterfang, & Velakoulis, 2015; Elamin et al., 2012). In addition, patients with bvFTD have also been found to have an impaired ability to recognize emotions based on facial expressions (see Bora, Velakoulis, & Walterfang, 2016 for a meta-analysis).

However, there is mixed empirical evidence regarding impaired interoception in bvFTD. Balconi et al., (2015) found that bvFTD patients are able to explicitly rate emotions in terms of valence and arousal, but cannot implicitly "feel" the emotion, as indicated by lower implicit autonomic reactivity assessed by SCR. García-Cordero et al., (2016) found impaired interoceptive accuracy and awareness along with an overestimation of their own performance in bvFTD patients. In contrast, Marshall et al. (2017) found that interoceptive accuracy was not impaired in patients with bvFTD. However, they more recently showed that automatic imitation of dynamic facial expressions was reduced in bvFTD patients and was related to impaired

emotion identification (Marshall et al., 2018). Furthermore, difficulties in probabilistic learning and updating in response to exteroceptive prediction errors have also been described in patients suffering from bvFTD. Clinical studies revealed that bvFTD patients showed impairment in a probabilistic learning task, which requires participants to learn which outcomes are predicted by prior information (Dalton, Weickert, Hodges, Piguet, & Hornberger, 2013; Weickert et al., 2013). In addition, automatic detection of changes (or unexpected events) (Hughes & Rowe, 2013) and integration of social contextual information have been found to be impaired in bvFTD (O'Callaghan et al., 2016).

Although some previous research points to reduced interoceptive processing in bvFTD patients which might lead to deficits in emotion recognition, no empirical studies so far have directly linked impaired emotion recognition to deficits in interoceptive prediction. As mentioned previously, patients may be unaware of interoceptive changes and hence may not be able to update prediction models in response to interoceptive prediction errors. Therefore, to gain further insights into a possible impairment of interoceptive prediction in patients with bvFTD, it is of interest to investigate the recognition of emotions in patients with bvFTD in a predictive setting, i.e., in the probabilistic inference of emotional states.

2. Research Questions

As reviewed in the above sections, recent theories on emotions suggest that emotional experiences result of contextualized perceptions of interoceptive changes. Interoceptive inference and its underlying brain regions are theoretically well-founded and have become the topic of a growing body of work in the last years. However, some compelling questions regarding the considerable variability that individuals display in their interoceptive abilities and the influence that this variability or impairments in interoceptive abilities have on emotion recognition so far remain unanswered. The studies composing the present thesis aimed to clarify the effects of differences in interoceptive abilities and possible impairments in bvFTD patients on emotion recognition, physiological arousal and brain activity. Specifically, behavioural experiments along with physiological measures assessed by SCR and an fMRI experiment were conducted to investigate these relations in a probabilistic emotion classification task. To enable the prediction of unfolding emotional expressions as they are perceived in our everyday life, in contrast to emotion recognition paradigms commonly used in clinical contexts, dynamic (videotaped) facial stimuli were used. To this end, videotaped facial stimuli were recorded with actors/actresses posing from a neutral to an emotional expression. The task in all three experiments required participants to indicate whether these neutral facial expressions evolved into a positive or negative valence, i.e., a happy or fearful expression. The facial expressions varied in emotional intensity at the end of the video to convey different degrees of uncertainty. In addition, different probabilities for the occurrence of happy or fearful faces per block were implemented to assess participants' propensity to efficiently update their predictive model. Parallel to this task, a non-emotional speeded classification control task was employed to assess domain specificity of the effects. In addition to the probabilistic classification tasks, the interoceptive measures IS and IAcc were collected in the respective studies using the MAIA-2 questionnaire and the HDT. Within the three studies, I posed the following questions:

1. Do individuals with higher interoceptive abilities recognize others' emotions more easily, are they more sensitive to learn from biased probabilities of emotional expressions and is this reflected in a modulation of physiological arousal?
2. Is bvFTD associated with impaired discrimination of changes in emotional facial expressions and do patients exhibit difficulties in learning from biased probabilities of these expressions?
3. Are brain regions associated with interoceptive processing involved in emotion recognition and adaptation to emotional context, and does the degree of activation in brain regions related to interoceptive processing increase with higher interoceptive ability?

In **Study I**, I examined the link between differences in IS and emotion recognition using behavioural measures (reaction time) and SCR. I hypothesised that higher sensibility for interoceptive changes reflects higher precision of interoceptive signals, leading to a facilitation of emotional change detection and a more precise adaptation to emotion probabilities. In addition, I expected individuals with higher IS to show increased sympathetic arousal (as assessed by SCR) when discriminating emotions and that higher IS is associated with increased sympathetic arousal in response to more surprising and less predictable emotions. Based on the findings of Study I, I conducted **Study II** on potential deficits in interoceptive predictive mechanisms in bvFTD patients. I examined whether patients with bvFTD showed similar response time and SCR patterns to individuals with low IS which were observed in Study I. Specifically, I expected bvFTD patients compared to healthy controls to show increased reaction times and decreased SCRs when discriminating emotional changes and to be unable to learn the block alternating probabilistic imbalance, as reflected by an absence of differences in reaction time between expectable and surprising emotions. Finally, in **Study III** I focused on differential effects of differences in IAcc and IS on emotion recognition and probabilistic

context adaptation. Using fMRI, I investigated regions involved in the processing of dynamically unfolding emotional expressions and signalling of interoceptive prediction errors. I expected that higher interoceptive abilities would positively correlate with the degree of activation increase in brain regions related to interoceptive predictive processing.

3. Research Articles

3.1 Study I: Interoceptive Sensibility Predicts the Ability to Infer Others' Emotional States

Running title: Interoception and emotion inference

Hübner, A. M., Trempler, I., Gietmann, C., & Schubotz, R. I. (2021).

Plos one, *16*(10), e0258089

PLOS ONE

RESEARCH ARTICLE

Interoceptive sensibility predicts the ability to infer others' emotional states

Amelie M. Hübner[1☯*], **Ima Trempler**[1,2☯], **Corinna Gietmann**[1], **Ricarda I. Schubotz**[1,2]

1 Department of Psychology, University of Muenster, Muenster, Germany, **2** Otto-Creutzfeldt-Center for Cognitive and Behavioural Neuroscience, University of Muenster, Muenster, Germany

☯ These authors contributed equally to this work.
* huebnera@uni-muenster.de

Abstract

Emotional sensations and inferring another's emotional states have been suggested to depend on predictive models of the causes of bodily sensations, so-called interoceptive inferences. In this framework, higher sensibility for interoceptive changes (IS) reflects higher precision of interoceptive signals. The present study examined the link between IS and emotion recognition, testing whether individuals with higher IS recognize others' emotions more easily and are more sensitive to learn from biased probabilities of emotional expressions. We recorded skin conductance responses (SCRs) from forty-six healthy volunteers performing a speeded-response task, which required them to indicate whether a neutral facial expression dynamically turned into a happy or fearful expression. Moreover, varying probabilities of emotional expressions by their block-wise base rate aimed to generate a bias for the more frequently encountered emotion. As a result, we found that individuals with higher IS showed lower thresholds for emotion recognition, reflected in decreased reaction times for emotional expressions especially of high intensity. Moreover, individuals with increased IS benefited more from a biased probability of an emotion, reflected in decreased reaction times for expected emotions. Lastly, weak evidence supporting a differential modulation of SCR by IS as a function of varying probabilities was found. Our results indicate that higher interoceptive sensibility facilitates the recognition of emotional changes and is accompanied by a more precise adaptation to emotion probabilities.

OPEN ACCESS

Citation: Hübner AM, Trempler I, Gietmann C, Schubotz RI (2021) Interoceptive sensibility predicts the ability to infer others' emotional states. PLoS ONE 16(10): e0258089. https://doi.org/10.1371/journal.pone.0258089

Editor: Delphine Grynberg, Universite de Lille, FRANCE

Received: December 16, 2020

Accepted: September 20, 2021

Published: October 6, 2021

Copyright: © 2021 Hübner et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting information files. For reasons of transparency additionally all files and our raw/primary data are available under the following link: https://osf.io/sy8gr/.

Funding: The author(s) received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

1 Introduction

Interoception, defined as the sense of the internal physiological state of the body [1], has gained growing interest in recent years because of its impact on physical and mental health, as well as on the processing of emotion. Classical appraisal theories of emotion processing postulate that emotional experiences to arise from the contextualised perception and interpretation of bodily responses to external stimuli [2–4]. The somatic marker hypothesis [5, 6] incorporated this view emphasizing viscerosensory origins of emotions [7]. This notion has been developed further under the framework of predictive coding, which presumes that emotional experiences are determined by inferences of the causes of bodily sensations based on past

experiences. The interoceptive predictive coding model holds that emotional states are determined by the interplay of interoceptive predictions and interoceptive prediction errors [8–10]. Mismatches between descending interoceptive predictions and primary interoceptive afferents convey information about interoceptive changes and activate autonomous responses to restore physiological homeostasis or allostasis [9, 11]. Importantly, this interplay is flexibly tuned to the current reliability of exteroceptive and interoceptive signals by means of *precision*, regulating the relative weight accorded to prediction errors and predictions. Thus, highly precise prediction errors relative to prediction gives bias to bottom-up processing, whereas highly precise predictions relative to prediction errors give bias to top-down processing [12, 13].

Interoceptive precision may be a key to the striking differences individuals show in their interoceptive abilities. High sensitivity to interoceptive changes and, correspondingly, in emotional experience, is suggested to correspond with the ability to raise the precision of interoceptive prediction errors by focused attention [14]. As a result, in individuals with high versus low interoceptive sensitivity, interoceptive predictions are updated more frequently and thus become increasingly precise. The downside of this continual precision optimization can be observed in individuals with anxiety disorders, which promote increased attention to bodily signals [15]. In contrast, low interoceptive sensitivity has been shown to be accompanied by alexithymia, i.e., deficits in identifying and describing one's own and others' emotions [16, 17]. However, it should be noted that individuals could be differentiated on the basis of various measures of interoceptive abilities. Garfinkel and co-workers [18] distinguish between *interoceptive accuracy* (IAcc), operationalized as performance on objective detection of the heartbeat, *interoceptive sensibility* (IS) that quantifies the self-reported belief concerning one's own perception of bodily signals, and *interoceptive awareness* (IAw), defined as a metacognitive measure of the correspondence between objective IAcc and subjective evaluation of one's own interoception. Moreover, differences in interoceptive abilities could be associated with differences in physiological parameters that have been shown to reflect prediction error responses. For example, habituation through repeated exposure to a stimulus (i.e., decreased prediction error) is reflected in a decrease in skin conductance responses (SCRs) [19, 20]. In addition, SCRs have been shown to indicate preparatory or anticipatory reactions to upcoming events [21–23]. Thus, it is conceivable that individuals with a subjectively or objectively high interoceptive ability might also show stronger prediction error responses of the autonomic nervous system.

Individual differences in interoception could contribute to differences in social cognition. Studies provide evidence that the objective sensitivity for interoceptive changes, i.e., IAcc, relates to the ability to infer mental states of others (i.e., theory of mind) [24, 25]. Increased IAcc, as assessed by the heartbeat perception task, is related to an increased perceived arousal elicited by emotional stimuli [26, 27]. Moreover, individuals with higher IAcc are more sensitive to emotional facial expressions of others [28]. Consequently, the concept of interoceptive prediction has recently also augmented models of social cognition, i.e., inferring others' intentions and emotional states based on exteroceptive, interoceptive and proprioceptive information. According to the concept of sensorimotor simulation, the sensorimotor system serves as a route for recognising facial expressions of emotion (see [29] for a review). During the observation of an emotional expression, we use our sensorimotor system to simulate the motor plan that the expresser is likely to use to produce the motor movements seen in the facial expression. This can be done explicitly, i.e. with facial mimicry, or without; crucially, the emotional meaning of the expression is inferred from our own prior exteroceptive and interoceptive experience of being in the presumed emotional state. Finally, the same principle can be applied to understanding others' actions and mental states and how we share their bodily sensations [30]. For example, the predictions of one's own interoceptive states that establish the sense of

feeling of cold when one sees someone shivering. Consequently, emotion recognition benefits from adequate access to one's own interoceptive cues. Thus, if individuals' interoceptive abilities are reflected in their propensity to learn from prediction errors to acquire increasingly precise interoceptive predictive models, this would also make them better at inferring the emotional states of others.

The present study tested these assumptions by investigating the relationship between IS, assessed by the Multidimensional Assessment of Interoceptive Awareness, Version 2 (MAIA-2; [31]), and performance during a probabilistic emotion classification task with videotaped facial stimuli. We measured reaction times (RTs) along with SCRs as a widely used psychophysiological marker of changes in autonomic sympathetic arousal [32, 33] and activation to emotional stimuli [34, 35]. During the experiment, participants were required to indicate whether a neutral facial expression develops into a happy or fearful expression. Facial expressions at the end of the video varied in intensity to introduce different levels of uncertainty. Critically, we implemented different probabilities for the occurrence of either happy or fearful faces per block to assess participants' propensity to efficiently update their predictive model. The varying probability and predictability of stimuli were quantified by information-theoretic measures, i.e., Shannon *surprise* and *entropy*, respectively [36]. The 'surprise' of an event, meaning its improbability, is given by the negative logarithm of the probability, whereas 'entropy' measures the average surprise of all possible events and quantifies the expected information of events regarding their predictability [37].

Regarding the discrimination of emotional change (hypothesis 1, H1), we expected that higher emotional intensity should lead to decreased RTs and increased SCRs reflecting decreased discrimination uncertainty [38, 39]. In addition, this effect should be more pronounced for fearful vs. happy facial expressions considering the consistent evidence for a superior recognition of happy faces compared to other facial expressions [40–42]. Therefore, we hypothesized that individuals with higher compared to lower IS would show decreased RTs and increased SCRs when discriminating emotions, especially when a fearful face of low intensity was presented.

Regarding the probabilistic context adaptation (hypothesis 2, H2), we expected that participants with higher vs. lower IS would learn the block-wise changing probabilistic imbalance more accurately due to stronger attentional precision-weighting [43]. This would be reflected in a positive correlation of surprise and entropy with RTs and SCRs in participants with higher IS but no correlation in participants with lower IS.

To explore whether the effects would be specific to emotion discrimination, we employed a non-emotional speeded classification control task, where participants were asked to discriminate the gender of a presented neutral face while facial stimuli developed from pixelated to high resolution. Paralleling the emotional condition, videos differed in resolution intensity (i.e., degree of pixilation) at the end of the video, with low resolution for low and high resolution for high intensity videos as well as in the probability for the occurrence of male or female faces.

2 Materials and methods

2.1 Participants

Forty-six right-handed healthy young volunteers (36 female, 10 male) with normal or corrected-to-normal vision were included in the present study. The participants' age ranged from 18 to 32 years (22.9 ± 3.5 years). None of them reported a history of neurological or psychiatric disorders. The study protocol was conducted in accordance with ethical standards of the Declaration of Helsinki and approved by the Local Ethics Committee of the University of

Muenster. Each participant submitted a signed informed consent form and received either reimbursement or course credits for their participation afterwards. Individuals provided written informed consent to permit for all potentially identifying information to be published.

2.2 Stimulus material

The stimulus material was created as part of a project on emotion recognition in patients with behavioural variant frontotemporal dementia in cooperation with the University Hospital Muenster, Germany [44]. The stimuli consisted of short videos with a mean duration of 3.00 s (± 0.39), which were displayed on a grey background. Videos depicted male or female faces posing from neutral to either happy or fearful emotional expressions (emotional condition) or male or female faces from pixelated to high resolution (non-emotional condition). Moreover, facial expressions in the emotional condition and image resolution in the non-emotional condition differed in terms of intensity (high/low) to introduce different levels of uncertainty. To control for potential effects of mouth opening on emotion recognition [45], fearful and happy expressions of high and low intensity were each presented in two different versions, i.e., with the mouth open and closed. Note that this factor was not considered in the statistical analyses to reduce the complexity of the statistical model.

To create these stimuli, we recorded several short video sequences (~ 2.08 to 4.40 s) of four actors and four actresses from four age groups each (20–30, 35–45, 50–60, 65–75 years). Actors and actresses were instructed to perform either a happy or a fearful facial expression of either high or low intensity and with mouth open or close. In order to achieve comparable length and development of the enfolding emotions, the subsequent videos were edited and cut using Adobe Premiere Pro CC (Adobe Systems Software, Dublin, Ireland) and Wondershare Filmora Version 8.5.1 (https://filmora.wondershare.com/). The tenth frame before the first (emotional) movement in the face was determined as a start frame which, thus, formed the neutral facial expression. The end frame was determined as the frame after which the emotion had reached its highest intensity and remained constant for another 20 frames. A total of 786 of these first and last frames were extracted and rated in two online-based pilot studies. In the first pilot study, a total of 60 participants (44 females; 27.4 ± 10.9; range 18–65 years) were asked to rate half of the start and end frames with regard to the valence and its intensity of the presented face on a 9-point scale ranging from 1 (*strong fear*), 5 (*neutral*) to 9 (*strong joy*) (see S1 File–containing all the supporting tables and figures). Moreover, if the participants experienced the emotion as neither fearful nor happy, they could enter a different emotion in a provided text field. In the second pilot study with a sample of 48 participants (33 females; 31.6 ± 14.8 years old; range 17–68 years), each participant rated valence, intensity and arousal of their subjective feelings elicited by the pictures on a scale ranging from 1 (*fear high / negative valence / calm*, *relaxed*), 5 (*neutral*) to 9 (*happy high / positive valence / exited*, *activated*). Based on the results of the pilot studies, we selected two videos per actor/actress for each of the eight conditions (happy vs. fearful x high vs. low intensity x mouth open vs. closed) that best met the main task's requirements (i.e., neutral start, high/low intensity of happy/fearful expression in the end) resulting in 128 different videos in total. The final videos had a size of 800x800 pixels, a framerate of 25 frames per second, and an average length of 3.01 s. To create economically valid stimuli, we accepted that the emotional videos would slightly vary in length ($SD = 0.39$) (see S1 File). Therefore, we included video duration as a covariate in our analyses to control for possible confounds on the participants' RT.

For the non-emotional condition, videos of each actor or actress with a constantly neutral expression were recorded. To ensure that all presented faces developed in the same time curve from pixelated to high resolution, they were provided with a Gaussian soft focus using

Premiere Pro CC. The Gaussian soft focus was set to a value of 190 at the first frame, such that the faces were completely pixelated and not recognizable, as ensured by a further pilot study. From the fifth frame on, a linear decay was inducted. To introduce different levels of intensity comparable to the emotion task, the videos either dissolved completely or ran out with a remaining soft focus of 20 at the end frame. We additionally extended the grey background ellipsoid around the faces to ensure that noticeable features like the hairline would not favour ceiling effects. For each actor and actress, we created two videos with high and low intensity with a size of 1080x864 pixels, a framerate of 25 frames per second and an average length of 2.5 s.

2.3 Task

During the experiment, participants were seated in front of a computer screen located at a distance of about one meter. Videos were presented at the centre of the screen, separated by an interstimulus interval of 4 s during which a fixation cross was displayed centrally on the screen (Fig 1). The participants were asked to watch the presented videos attentively and to respond as fast and accurately as possible as soon as they recognized the emotion in the emotional condition or the gender of a presented face in the non-emotional condition. Participants responded by button press on a two-button response box, using their right-hand index and middle fingers. Stimulus-response mappings were counterbalanced across participants.

The task consisted of 32 emotional and non-emotional blocks with 16 consecutive videos each, i.e., a total of 512 trials (256 per condition). Thus, in the emotional condition each of the 128 videos was shown twice. Blocks were combined into four runs consisting of eight blocks each, i.e., after the reoccurrence of eight blocks of the same condition (emotional or non-emotional condition) the condition changed and remained the same for another eight blocks. At the beginning of each block, instructions were presented on the screen indicating whether the

Fig 1. Schematic diagram of the task. a) Trials are presented in a mixed block- and event-related design with emotional (e) and non-emotional (ne) task blocks. **b)** At the beginning of each block, instructions indicate whether the participant has to indicate the emotion or the gender of the upcoming faces. One block includes 16 consecutive video trials. **c)** Videos depict actors/actresses posing from neutral to either happy or fearful facial expressions (blue background) or male or female actors/actresses from pixelated to high resolution (yellow background). The actor and actresses in the figure have given written informed consent (as outlined in PLOS consent form) to publish the photographs.

https://doi.org/10.1371/journal.pone.0258089.g001

participant must respond to either the emotional expression or the gender of the faces. Each experimental block was followed by a break of 8 s, during which participants were given the information that the block had ended, followed by instructions regarding the upcoming block. Accordingly, the overall task for the emotional condition lasted about 44 min (2624 sec), for the control task about 39 min (2368 sec).

Within each block, either a high (75%) or low (25%) probability for the occurrence of fearful or happy faces in the emotional condition and either a high (75%) or low (25%) probability of male or female faces in the non-emotional condition were implemented. Stimuli were presented in a pseudo-randomized order ensuring that videos of one actor/actress were never repeated across consecutive trials. Transitions between the block types were balanced across the experiment.

Prior to the experiment, participants performed a short training session to get accustomed to the task. The training consisted of one block of 16 trials each and an equal probability for the different conditions.

The randomisation was programmed using MATLAB R2019b (The MathWorks Inc., Natick, MA, USA) and stimuli were presented using the Presentation software (Version 19.0, Neurobehavioral Systems, Inc., Berkeley, CA).

2.4 Assessment of interoceptive sensibility

For the self-assessment of interoceptive sensibility, the Multidimensional Assessment of Interoceptive Awareness, Version 2 (MAIA-2, [31]), a test that quantifies the self-reported belief concerning one's own perception of bodily signals, was used. The MAIA-2 is a state-trait questionnaire with 37 items, which consist of eight subscales corresponding to its eight-factor structure: noticing (awareness of uncomfortable, comfortable, and neutral body sensations), not-distracting (tendency to ignore or distract oneself from sensations of pain or discomfort), not-worrying (emotional distress or worry with sensations of pain or discomfort), attention regulation (ability to sustain and control attention to body sensation), emotional awareness (awareness of the connection between body sensations and emotional states), self-regulation (ability to regulate psychological distress by attention to body sensations), body listening (actively listens to the body for insight), and trust (experiences one's body as safe and trustworthy) [46]. Participants are asked to indicate on a 5-point Likert scale from 0 (*never*) to 5 (*always*) how often each statement applies to them in everyday life. Results of prior studies support the validity of the MAIA-2 scales, with Cronbach's alpha for the eight scales ranging from 0.64 to 0.83 [31]. For our analyses, we calculated a total score per participant through reverse coding the corresponding items and summing all items. Together with other questionnaires assessing the participants' emotion processing (see S1 File), the MAIA-2 was completed prior to the main task.

2.5 Skin conductance response acquisition

SCRs were acquired using the BrainVision Recorder Version 1.20.0801 (Brain Products, Munich, Germany). Two Ag/AgCl electrodes were placed on the annular and middle fingers of the participant's left hand and 0.5%-NaCl electrode paste (GEL101; Biopac Systems) was used. Data were recorded at 500 Hz, using a sampling interval of 2000 μS. Preprocessing and data analysis were performed using PSPM [4.1.1], available at pspm.souorceforge.net. Skin conductance data were converted back to a waveform signal with 100 Hz time resolution, filtered with a unidirectional first-order Butterworth high pass filter with a cut-off frequency of 0.05 Hz, according to current recommendations [47]. Data were down-sampled to 10 Hz. The entire SCR time series was then *z*-transformed for each participant to account for interindividual

differences in responsiveness [48]. The data were visually checked for artifacts, but no formal artifact rejection was implemented. The analysis of stimulus-locked (evoked) responses was done following the general linear convolution model (GLM) approach on a single-trial level. To this end, we extracted trial-by-trial estimates and ran the GLM with one regressor per trial. Each trial in the experiment was modelled as a Dirac delta function centred on the event onset, convolved with a canonical skin conductance response function (SCRF) and its first derivative [49]. From the estimated amplitude parameters for the canonical SCRF and its derivative, the response for each condition was reconstructed [47].

2.6 Data analysis

Basic statistical analyses were performed using R, version 3.6.2 [50]. For both behavioural and SCR data, false or missing responses were excluded from the analyses. Behavioural performance was defined by reaction times (RTs).

Comparisons of RTs and SCRs were carried out separately for the emotion and gender task. More specifically, we tested whether valence/gender, intensity and information-theoretic quantities, i.e., Shannon's surprise $I(x_i)$ and entropy $H(X)$ [36], reflecting the inverse probability and predictability of a single stimulus, respectively, could predict (variance in) RTs and SCRs on a single-trial level. While 'surprise' is a measure of the improbability of a particular event, 'entropy' measures the expected or average surprise over all events and thus reflects the predictability of an event within a particular context [37]. Shannon's surprise was based on the frequency of a trial of a specific valence/gender x_i normalized by the sum of all past trials in the block:

$$p(x_i) = \frac{n(x_i) + 1}{\sum x_t + 1}$$

The prior counts before observing the first trial in the block were set to 1/2 for the two factor levels of valence (happy, fearful) and gender (male, female). The surprise $I(x_i)$ of each stimulus event given by the negative logarithm of this probability quantifies the amount of information provided by the current stimulus:

$$I(x_i) = -\ln\ p(x_i)$$

Finally, entropy $H(X)$ measures the average surprise of all possible outcomes and quantifies the expected information of a stimulus regarding its predictability:

$$H(X) = \sum_i -p(x_i)\ln\ p_k(x_i)$$

We conducted generalized linear mixed-effects analyses using R, version 3.6.2 (R Core Team, 2019) via the package *lme4*, version 1.1.21 [51]. As the distribution of single-trial RTs was positively skewed, RTs were transformed to the natural logarithm to more closely approximate a normal distribution. Moreover, Q-Q plots indicated that residuals of RTs and SCR data were normally distributed. For the factors valence/gender and intensity, we used effect coding, with -1 for happy and 1 for fearful expressions, -1 for females and 1 for males, and -1 for high and 1 for low intensity expressions. Surprise and entropy were centered at individual levels, whereas MAIA-2 score was centered at the group level. Each model was fit with valence/gender and intensity (and their interaction), Shannon's surprise and entropy and their respective interaction with the MAIA-2 score as fixed effects, and with a random intercept for each subject. For the emotional condition we added video duration to the models to control for effects of varying video length on RTs. Statistical significance for each fixed effect was calculated via

lmerTest, version 3.1.1 [52], using the Satterthwaite's approximation to denominator degrees of freedom. The significance level was set to $\alpha = .05$. For posteriori pairwise comparisons we used *lsmeans* [53] with the Tukey adjustment for multiple tests and a high (25) and low (-25) level of the centred MAIA-2 score to assess differences between high and low IS participants.

In addition, we calculated Bayesian linear multilevel models in R [50] via the *brms* package and Stan using default priors [54, 55]. Regression coefficients and 95% credible intervals (CIs; i.e., Bayesian confidence intervals) are reported, meaning that the respective parameter falls within this interval with a 95% probability and indicating statistical significance on a 5% level if the interval does not contain zero.

3 Results

3.1 Emotional change recognition (H1)

3.1.1 Behavioural data. Because the percentage of false alarms (2.27%) and missing data (0.14%) was low in the emotional condition, we restricted the behavioural analyses to RTs. The linear mixed-effect model predicting RTs revealed a significant main effect of IS, $b = -0.01$, $\beta = -0.28$, $t = -3.16$, $p = 0.003$. In line with our hypothesis, the negative gradient shows decreased RTs with increasing IS. We found a significant main effect of valence, $b = 0.03$, $\beta = 0.10$, $t = 5.75$, $p < 0.001$, with increased RTs for fearful facial expressions, and a main effect of intensity, $b = 0.02$, $\beta = 0.07$, $t = 3.49$, $p < 0.001$, driven by increased RTs for low intensity of an expression. Moreover, we found a significant interaction between valence and intensity, $b = 0.01$, $\beta = 0.04$, $t = 5.54$, $p < 0.001$, a significant interaction between intensity and IS, $b = -0.001$, $\beta = -0.02$, $t = -2.37$, $p = 0.02$, as well as a significant three-way interaction between valence, intensity and IS, $b = 0.001$, $\beta = 0.02$, $t = 2.37$, $p = 0.02$ (Fig 2a). Post-hoc tests comparing a low (-25) and a high (25) level of the centred MAIA-score revealed that, participants with lower IS showed increased RTs for high compared to low intensity of happy expressions, $b = -0.06$, $t = -4.72$, $p < 0.001$, and of fearful ones, $b = -0.08$, $t = -6.59$, $p < 0.001$. In contrast, participants with higher IS showed shorter RTs for high compared to low intensity of fearful expressions, $b = -0.09$, $t = -6.42$, $p < 0.001$, but no difference for happy expressions, $b = 0.009$, $t = -0.74$ $p = 0.996$. With decreasing IS, we found shorter RTs for happy compared to fearful

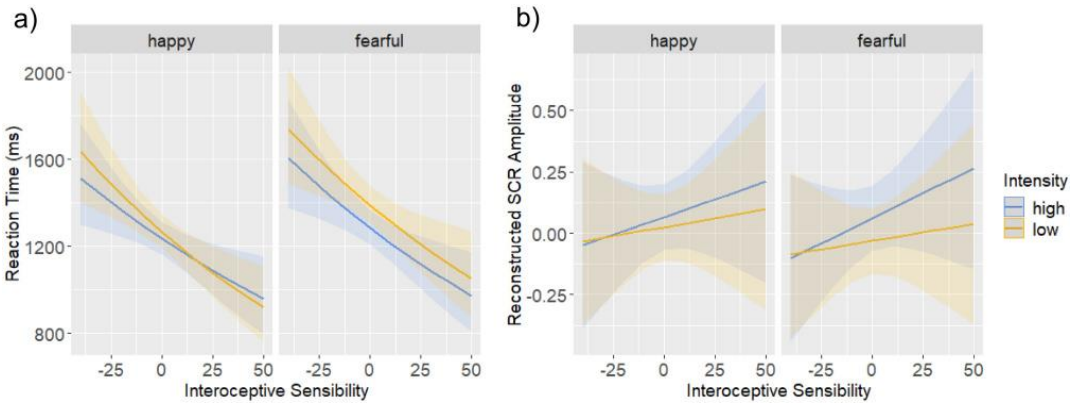

Fig 2. Effects of valence and intensity. Marginal effects of valence (happy, fearful) and intensity (high, low) as a function of interoceptive sensibility assessed by the MAIA-2 questionnaire (Mehling et al., 2018) on **a)** reaction time and **b)** skin conductance response (SCR). The solid lines depict the regression fit, and the shaded areas show the 95% confidence intervals.

https://doi.org/10.1371/journal.pone.0258089.g002

expressions of both low, b = -0.07, t = -5.94, $p < 0.001$, and high intensity, b = -0.05, t = -4.27, $p < 0.001$. With increasing IS, RTs became shorter for happy compared to fearful expressions of low intensity, b = -0.11, t = -9.47, $p < 0.001$, but no difference when intensity was high, b = -0.03, t = -2.27, p = 0.313. In line with our hypothesis, participants with higher IS were thus faster in detecting emotional changes and slowed down their responses only to low intensity fearful expressions. Decreasing IS was accompanied by a gradual increase in RTs with increasing difficulty of the condition, i.e., from happy high to fearful low intensity expressions. In accordance with our assumptions, video length did not predict RTs, b = -0.007, β = -0.008, t = -1.11, p = 0.267.

The Bayesian logistic multilevel model on valence, intensity, surprise and entropy estimates, IS and their respective interactions, revealed significant main effects of valence, intensity and IS, as well as interaction effects of valence and intensity, intensity and IS, as well as a significant three-way interaction between valence, intensity and IS. A table with regression coefficients and corresponding 95% CIs for each variable predicting RTs in the emotional condition is given in the S1 File.

3.1.2 SCR data. The linear mixed-effect model predicting SCRs revealed a trend for an effect of intensity corresponding to our hypothesis, b = -0.03, β = -0.02, t = -1.79, p = 0.07, driven by increased reconstructed SCR for high emotional expressions (Fig 2b). No main effect of valence (p = 0.41) or IS (p = 0.49) on SCR amplitudes, as well as no interaction effects were observed (all $p > 0.34$). The Bayesian logistic multilevel model predicting SCRs revealed no significant main or interaction effects.

3.2 Probabilistic context adaptation (H2)

3.2.1 Behavioural data. We tested whether Shannon's surprise and entropy of a single emotional expression in interaction with individual IS (i.e., continuously varying MAIA-2 scores) were predictive of the participants' performance. The linear mixed-effect model predicting RTs revealed a significant main effect of surprise, b = -0.014, β = -0.02, t = -3.34, $p < 0.001$, and entropy, b = 0.059, β = 0.02, t = 2.10, p = 0.03, as well as significant interactions between IS and surprise, b = 0.001, β = 0.03, t = 4.58, $p < 0.001$, and between IS and entropy, b = 0.003, β = 0.02, t = 2.62, p = 0.01. Specifically, as expected in participants with higher IS surprise and entropy were positively correlated with increased RTs reflecting, in turn, a behavioral advantage in the course of learning from increasing probability of a specific valence. Contrary to our expectation, there also was a negative correlation in participants with lower IS (Fig 3a). The Bayesian logistic multilevel model revealed corresponding main effects of surprise and entropy, as well as interaction effects of surprise and entropy with IS (see S1 File).

3.2.2 SCR data. In the linear mixed-effect model predicting SCRs, we observed weak evidence for our hypothesis suggesting an interaction effect between IS and surprise, b = 0.003, β = 0.02, t = 1.78, p = 0.07, and between IS and entropy, b = 0.019, β = 0.02, t = 1.79, p = 0.07, but no main effects of surprise (p = 0.572) and entropy (p = 0.316) (Fig 3b). Higher IS, thus, appeared to be accompanied by increased SCRs to unpredicted and unpredictable valences. However, the Bayesian logistic multilevel model did not capture these effects and did not reveal main or interaction effects.

3.3 Control task

3.3.1 Behavioural data. In the non-emotion control condition, i.e., the gender task, the percentage of false alarms (5.17%) and missing data (0.78%) was significantly higher than in the emotional condition, although still relatively low (M = 0.07, SD = 0.05, $t(45)$ = 6.54, $p < 0.001$). The linear mixed regression model for the prediction of RTs revealed a significant

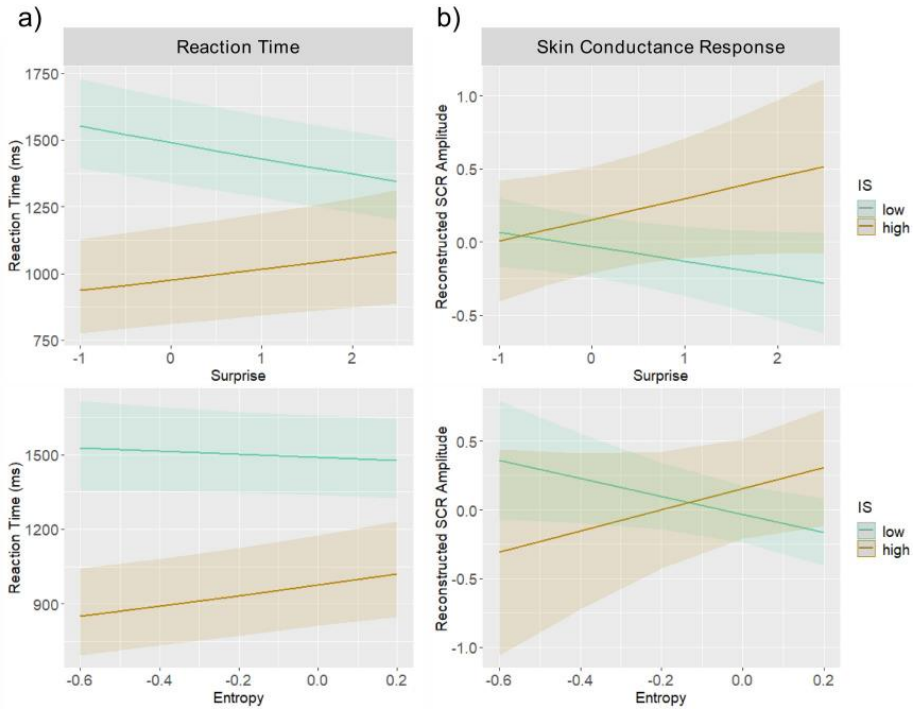

Fig 3. Effects of Shannon's surprise and entropy. Results of the linear mixed effect regression model predicting **a)** reaction time **b)** and skin conductance response (SCR) by mean-centered surprise (top panel) and entropy (lower panel) as a function of interoceptive sensibility (IS) assessed by the MAIA-2 questionnaire [31]. For visualization purposes, marginal effects at representative values of the IS score (-25 for low, 25 for high IS) are presented. The solid lines depict the regression fit, and the shaded areas show the 95% confidence intervals.

https://doi.org/10.1371/journal.pone.0258089.g003

main effect of gender, $b = 0.016$, $\beta = 0.05$, $t = 5.53$, $p < 0.001$, driven by increased RTs for female faces compared to male faces. As expected, the main effect of intensity was also significant, $b = -0.031$, $\beta = -0.09$, $t = -10.25$, $p < 0.001$, with decreased RTs for faces presented with high resolution compared to low resolution. In contrast to the emotional condition, the interaction between gender and intensity was not significant ($p = 0.62$) and there was no interaction effect between gender, intensity and IS ($p = 0.76$) (see S1 File). Likewise, the Bayesian logistic multilevel model also revealed a main effect of gender $b = 0.02$, 95%-CI [0.01, 0.02] and intensity $b = -0.03$, 95%-CI [-0.04, 0.03].

However, contrary to our expectation the linear mixed regression model as well as the Bayesian logistic multilevel model for the prediction of RTs revealed a significant main effect of entropy, $b = 0.14$, $\beta = 0.04$, $t = 4.13$, $p < 0.001$, as well as a trend for a significant interaction between surprise and IS, $b = 0.00$, $\beta = -0.02$, $t = 1.90$, $p = 0.057$. No other main or interaction effects were significant (all $p \geq 0.48$).

3.3.2 SCR data. The linear mixed effect model on reconstructed SCR amplitudes did not reveal significant main effects of gender ($p = 0.88$), intensity ($p = 0.53$) or interaction effects with IS ($p \geq 0.29$). Likewise, the linear mixed effect model did not reveal significant main effects with entropy and surprise or interaction effects with entropy (all $p \geq 0.19$). Nevertheless, we found an interaction effect of surprise and IS, $b = 0.004$, $\beta = 0.02$, $t = 2.15$, $p = 0.03$.

Likewise, the Bayesian logistic multilevel model predicting SCRs in the non-emotional control condition revealed no significant main or interaction effects, except the interaction effect of surprise and IS, $b = 0.00$, 95%-CI [0.00, 0.01].

4 Discussion

The present study tested whether an individual's interoceptive sensibility (IS) is positively related to the speed of recognizing emotional changes in others' facial expressions and to the ability to exploit biased probabilities to adapt expectations of facial emotions. We found (1) decreased RTs in individuals with increasing IS. Individuals with decreased IS were slower in recognizing low vs. high emotional intensity and fearful vs. happy valence of facial expressions. Individuals with higher IS slowed down their responses only for the most difficult condition, i.e., when a fearful face of low intensity was presented. Regarding differences in the exploitation of biased probabilities of happy vs. fearful faces, (2) participants with higher IS were faster in recognizing more probable (i.e., less surprising) and more predictable (i.e., less uncertain) facial emotions, whereas participants with lower IS displayed lower RTs for more probable as well as more predictable facial emotions. A trend for corresponding effects of surprise and entropy were also observed for the SCRs. Finally, we found mixed evidence for the hypothesis that (3) interindividual differences in IS particularly impact emotion but not gender expectation and recognition. Although IS had no impact on RTs or SCRs in the gender discrimination task, the IS score tended to interact with gender surprise similar as with emotion surprise.

4.1 Interoceptive sensibility accompany lower thresholds for emotional change recognition

Our behavioral results confirmed the hypothesis that self-reported sensibility to one's own interoceptive states (*interoceptive sensibility*, IS) modulates the recognition of emotional changes in facial expressions of others. Specifically, while individuals with increased IS showed decreased RTs only for high compared to low intensity of fearful expressions, lower IS participants showed a graduation of performance depending on both valence and intensity by becoming slower at recognizing fearful vs. happy and low vs. high intensity expressions.

While our study is the first to show a relationship between IS and emotion recognition, previous studies also suggested that interoception influences the sensitivity to others' emotions. Terasawa et al. [28] previously reported that individuals with high interoceptive accuracy (IAcc), assessed by the heartbeat detection task, show lower thresholds for the detection of various emotions in facial expressions, especially of happy ones. In addition, individuals with higher sensitivity to visceral changes experience emotional stimuli as more arousing [27, 56]. Our results extend these findings by showing that when tested for emotion recognition speed, individuals with lower IS benefit more from high intensities of facial emotions than individuals with higher IS. Considering that in daily life our spontaneous facial expressions are mostly of low to intermediate intensity, whereas high intensity is an exception [57–59], individuals with low IS conceivably display more difficulties in inferring emotional states from others' facial expressions in their daily social interaction.

Regarding the valence of facial expressions, participants with increased IS recognized fearfulness equally as well as happiness when intensity was high, whereas individuals with lower IS showed a general advantage for the recognition of happy faces. Previous studies found that happy faces are recognized faster and more accurately than other types of facial expressions [40–42], even when presented at low intensity [60, 61]. Theoretical explanations for this effect include that the expression of happiness requires simpler physical changes and occurs more frequently in our daily life [42, 62, 63]. On the contrary, although fearful facial expressions are

detected fast e.g. [64, 65], they are less well *recognized* than other facial expressions [10, 62, 65–67]. In the present study, the participants' task was to discriminate between happy and fearful expressions such that RT measures reflected recognition rather than detection of the emotional expression. The ability to perceive one's own bodily signals has been shown to lead to enhanced emotional discernment in social processing by promoting empathic abilities as well as emotional expressions [68–70]. Against this backdrop, we take the discrimination advantage of higher IS in our sample to be particularly evident for those stimuli that are especially difficult to be unequivocally discriminated, that is, fearful in comparison to happy faces and emotional expressions of low vs. high intensity.

As for the SCRs, although we expected an increased amplitude for fearful vs. happy expressions as SCRs are assumed to reflect physiological arousal to highly activating stimuli which are to be detected fast [34, 35], our findings did not provide evidence for this assumption. A possible explanation could be the lateralisation of the SCR profiles of the hands in response to emotional facial stimuli. Banks et al. [71] found stronger SCRs to anger, disgust and fear on the right hand, while higher amplitudes were found for the left hand in responses to sad, happy and neutral faces. Thus, possible differences in SCR amplitudes in response to fearful vs. happy faces could have been influenced by the SCR recordings from the left hand in our study. Alternatively, the difference between emotion recognition and detection could also provide an explanation for why no SCR difference was found between the two conditions: while emotion detection is associated with high arousal, for which SCR is a highly sensitive measure, emotion recognition implies cognitive appraisal, which possibly involves only subtle changes in SCR. Thus, as increases in SCRs are induced by multiple emotional states, it was found not to be a very specific measure of different emotions [72]. In line with this, there was weak evidence to support our hypothesis that high vs. low expressions (i.e., higher arousal) would elicit stronger SCRs, as has been observed in previous studies [73].

However, we neither found general differences between participants with high vs. low IS in SCR amplitude nor that higher IS facilitates sympathetic arousal in certain experimental conditions. Although previous EEG studies reported increased amplitudes for both the P300 component and slow waves in response to emotional pictures in individuals with high vs. low interoceptive sensibility [27, 56], other studies did not find overall differences in SCR varying with interoceptive sensibility [74]. Our findings provide no evidence for a significant correlation between the IS score or emotion processing and autonomous arousal. As discussed in the following, SCR patterns could be rather characterized by increased attention to these signals.

4.2 Interoceptive sensibility facilitates context-sensitive emotional inference

Using a novel task with videos showing neutral faces, developing to emotional expressions and occurring with different probabilities, the present results confirmed our hypothesis that higher IS is accompanied by a more precise adaptation of emotional predictions. RTs of individuals with higher IS increased when the neutral facial expression unfolded into a rather unexpected, i.e., surprising emotion (as measures by the current probability of a specific emotion occurrence), and decreased with increasing predictability, i.e., lower entropy of an emotion occurrence. In contrast, low IS individuals showed the opposite effects, with decreased RTs to surprising and unpredictable emotions. This correlational pattern suggests that higher IS individuals benefited from the implemented probabilistic context more than individuals with lower IS.

According to Ainley et al. [43], an attentional mechanism is at the basis of higher IS individuals' enhanced precision of prediction errors. Jiang and co-workers [75] report that attention

during a face-scene discrimination task, i.e., independent of emotion processing, accelerates prediction error processing as reflected in an increased neural patterns classifier's performance in distinguishing between expected and unexpected signals. Moreover, using EEG, Petzschner et al. [76] find that interoceptive attention modulates the cortical processing of heartbeats, as suggested by increased heartbeat-evoked potentials. Attention might, thus, manifest as increased precision of the heartbeat or other context-specific ascending interoceptive signals, and reduced precision of currently irrelevant interoceptive signals. This precision-weighting would ultimately lead to more precise interoceptive predictions. Accordingly, expected vs. unexpected negative vs. neutral facial expressions lead to decreased heartbeat-evoked potentials [77]. Furthermore, better perception of visceral cues facilitates unaware conditioned responding as well as the prediction of shocks [78, 79].

Against this background our findings further show, probably due to stronger attentional precision-weighting, individuals with higher IS learn better from, and hence build stronger expectations based on biased probabilities of emotional expressions. This in turn leads to a faster recognition of predicted emotions, but also to hesitation when unexpected emotions occur. In contrast, individuals with decreased IS show the opposite pattern by responding slower to predictable but faster to surprising stimuli. Moreover, although the effects on SCRs were generally weak and should be interpreted with caution, the (non-significant) trend for a modulation of SCRs by surprise and entropy in interaction with IS paralleled this RT pattern, with surprise and entropy positively covarying with SCRs in higher IS individuals but negatively in lower IS individuals. It was previously suggested that SCRs reflect an orienting reflex in response to novel or unexpected stimuli [19, 80]. Thus, the increase of SCRs as a function of entropy and surprise in higher IS individuals is in accordance with our hypotheses, reflecting adaptation to biased probabilities and heightened arousal in response to violations. However, the effect found in low IS individuals was rather unexpected. As a suggestion, our findings could be interpreted along the lines of a dual processing mode [81] assuming a supervisory attentional system [82]: In individuals with higher IS, expected emotions may be processed more automatically, whereas the classification of rather unexpected emotions requires increased cognitive control. As for the opposite effects in individuals with low IS, the emotional classification processes could generally be less controlled, which would be reflected in high RTs and SCRs even for frequently occurring emotions. Thus, the present results provide evidence that differences in IS regarding context-sensitive emotional inference are reflected in differences in bodily responses to emotional stimuli that vary in probability of their occurrence, although the concrete processes underlying this relationship must be clarified in future studies.

Finally, it has been suggested that learning of emotion concepts and affective predictions depending on past experience (i.e., memory), affected by various environmental influences (i.e., stability and habitual patterns of selective attention), determine whether sensory input is experienced as emotional or not [10, 83]. Consequently, prediction might even change how we perceive neutral stimuli [84, 85]. Since we used dynamic video-taped stimuli starting with a neutral facial expression evolving into an emotional expression, it is possible that individuals with higher IS already perceived neutral facial expressions as more emotional according to the respective emotional context, which led to a recognition advantage when prediction was fulfilled. To validate this interpretation, future studies could examine whether interindividual differences in IS determine the influence of affective predictions on the perception of neutral stimuli.

4.3 Domain specific and domain unspecific effects of interoceptive sensibility

Since attention is suggested to modulate precision-weighting of prediction errors in general [75] and also represents a key construct in various subscales of the MAIA-2 questionnaire

[31], effects of IS on emotion discrimination could be simply due to an improved ability to focus attention on any type of sensory input rather than being specifically caused by differences in interoceptive abilities. We therefore implemented a non-emotional control task where participants had to indicate the gender of neutral faces in videos developing from highly pixelated to high resolution in order to assess detection of changes independent of observed bodily cues [30]. The recognition of gender did not differ between these individuals, that is, no evidence was found for a general discrimination advantage in high vs. low IS individuals. However, regarding the probabilistic manipulation generating an expectation bias for female or male faces, surprise and IS tended to interact, albeit not significantly, such that RTs increased as a function of surprise more in high than in low IS individuals. This effect suggests a subtle impact of prediction errors on behavior in higher IS individuals which is independent of the prediction domain.

Our finding is in accordance with the postulation of at least three different areas of predictive functioning, i.e., the exteroceptive, interoceptive, and proprioceptive dimension, underlying perception, emotion, and action, respectively [9]. These can operate both amodal and multimodal, depending on the level of the predictive hierarchy at which predictions are violated. It is conceivable that in the present study, prediction errors in the exteroceptive modality (i.e., in the gender task) also elicited bodily signals on which higher IS individuals direct their focus more than lower IS individuals. As to the emotion task, the interaction was much more pronounced and an additional interaction effect with entropy was observed. Although our findings suggest partially domain-spanning interindividual differences in predictive performance, interoceptive sensitivity and information appear to play a greater role for the recognition of emotion than gender.

4.4 Strengths and limitations

Although the present study is the first investigating whether higher IS comes with better recognition and expectation of both emotional and non-emotional stimuli, the evidence for domain-specific and domain-general effects of IS has to be further validated in future studies. For instance, in order to determine interindividual differences in interoceptive abilities, we used a self-report questionnaire, i.e., the MAIA-2, assuming that different dimensions of interception are correctly tapped by subjective assessment. As several studies report an independence of IS and IAcc as an objective measure of interoceptive abilities [18, 86–89], future approaches could additionally use IAcc to examine the influence of interoceptive abilities on emotion discrimination.

Moreover, we acknowledge the relatively small sample size that plays a role especially with regard to the investigation of interindividual variation. Given that only trends could be observed regarding the effects of SCR, these results should be interpreted with caution. However, our study is particularly characterized by the fact that we created dynamically enfolding emotional facial expressions to investigate the recognition of emotional changes as well as the impact of differences in IS in an economically valid way. Only recently, the need to study dynamic and real situational emotions (as opposed to static Emoji-like expressions) was emphasized in order to get a valid picture of how emotional meaning is inferred [90].

Since social factors play a striking role in the development and maintenance of mental illness [91], our findings provide further evidence that interoception is an important feature of diagnosis and treatment of different psychiatric disorders [92]. Consequently, approaches to improve interoceptive abilities are worth discussing. As such, a therapeutic approach called mindful awareness in body-oriented therapy (MABT) could be appointed, specifically designed to teach skills of interoceptive awareness through a combination of psychoeducation

and somatic approaches [93]. Future studies could investigate whether approaches to improve interoception also positively influence social cognition.

5 Conclusion

The present study suggests that interoceptive sensibility facilitates the speed of recognition of emotional changes in facial expressions of others, potentially mediated by an increased attention to bodily signals. Furthermore, higher interoceptive sensibility entails a more precise adaptation to biased probabilities of emotional valences, pointing to a stronger reliance on situationally adjusted prediction. Correspondingly, bodily responses tend to increase for less probable emotions. Future studies can build on these findings by assessing corresponding effects in clinical populations associated with interoceptive dysfunctions such as anxiety disorders or alexithymia.

Supporting information

S1 File.
(DOCX)

Acknowledgments

We are thankful to Maike Brautmeier for her great help in recording the videos and Marlen Roehe for constructive criticism on the final version of the manuscript. We thank Andreas Johnen for his support with the conceptualization of the task design. Also, we sincerely thank Nina Heins for her support in programming the experiment, Falko Mecklenbrauck for his help in data analysis and Monika Mertens for the assistance in data acquisition.

Author Contributions

Conceptualization: Amelie M. Hübner, Ima Trempler, Ricarda I. Schubotz.

Data curation: Amelie M. Hübner, Corinna Gietmann.

Formal analysis: Amelie M. Hübner, Ima Trempler.

Investigation: Amelie M. Hübner, Corinna Gietmann.

Methodology: Amelie M. Hübner, Ima Trempler.

Project administration: Ricarda I. Schubotz.

Resources: Ricarda I. Schubotz.

Software: Amelie M. Hübner, Ima Trempler.

Supervision: Ricarda I. Schubotz.

Visualization: Ima Trempler.

Writing – original draft: Amelie M. Hübner, Ima Trempler.

Writing – review & editing: Corinna Gietmann, Ricarda I. Schubotz.

References

1. Craig A. D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nature reviews neuroscience*, 3(8), 655. https://doi.org/10.1038/nrn894 PMID: 12154366
2. Schachter S., & Singer J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review*, 69(5), 379.

3. James W. (1884). What is an Emotion?(188–205). *Mind*, 9, 34.
4. Lange C. G. (1885). The mechanism of the emotions. *The Classical Psychologists*, 672–684.
5. Bechara A., Damasio H., & Damasio A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, 10(3), 295–307. https://doi.org/10.1093/cercor/10.3.295 PMID: 10731224
6. Damasio A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 351 (1346), 1413–1420. https://doi.org/10.1098/rstb.1996.0125 PMID: 8941953
7. Damasio A. R. (2010). Self Comes to Mind: Constructing the Conscious Brain. London: Vintage Books.
8. 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 PMID: 24126130
9. Seth A. K., & Friston K. J. (2016). Active interoceptive inference and the emotional brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1708). https://doi.org/10.1098/rstb.2016.0002 PMID: 28080961
10. Smith R., Parr T., & Friston K. J. (2019). Simulating Emotions: An Active Inference Model of Emotional State Inference and Emotion Concept Learning. *Frontiers in Psychology*, 10(December), 1–24. https://doi.org/10.3389/fpsyg.2019.02844 PMID: 31920873
11. Seth A. K., Suzuki K., & Critchley H. D. (2012). An interoceptive predictive coding model of conscious presence. *Frontiers in Psychology*, 3(JAN), 1–16. https://doi.org/10.3389/fpsyg.2011.00395 PMID: 22291673
12. Friston K. (2008). Hierarchical models in the brain. *PLoS Computational Biology*, 4(11).
13. Friston K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11 (2), 127–138. https://doi.org/10.1038/nrn2787 PMID: 20068583
14. Ainley V., Apps M. A. J., Fotopoulou A., & Tsakiris M. (2016). "Bodily precision": A predictive coding account of individual differences in interoceptive accuracy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1708).
15. 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 PMID: 19751958
16. Herbert B. M., Herbert C., & Pollatos O. (2011). On the relationship between interoceptive awareness and alexithymia: is interoceptive awareness related to emotional awareness? *Journal of Personality*, 79(5), 1149–1175. https://doi.org/10.1111/j.1467-6494.2011.00717.x PMID: 21241306
17. Brewer R., Cook R., & Bird G. (2016). Alexithymia: A general deficit of interoception. *Royal Society Open Science*, 3(10). https://doi.org/10.1098/rsos.150664 PMID: 27853532
18. 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 PMID: 25451381
19. Geuter S., Boll S., Eippert F., & Büchel C. (2017). Functional dissociation of stimulus intensity encoding and predictive coding of pain in the insula. ELife, 6, 1–22. https://doi.org/10.7554/eLife.24770 PMID: 28524817
20. Barry R. J. (2009). Habituation of the orienting reflex and the development of preliminary process theory. *Neurobiology of learning and memory*, 92(2), 235–242. https://doi.org/10.1016/j.nlm.2008.07.007 PMID: 18675927
21. Bach D. R., Flandin G., Friston K. J., & Dolan R. J. (2010). Modelling event-related skin conductance responses. *International Journal of Psychophysiology*, 75(3), 349–356. https://doi.org/10.1016/j.ijpsycho.2010.01.005 PMID: 20093150
22. Delgado M. R., Nearing K. I., LeDoux J. E., & Phelps E. A. (2008). Neural circuitry underlying the regulation of conditioned fear and its relation to extinction. Neuron, 59(5), 829–838. https://doi.org/10.1016/j.neuron.2008.06.029 PMID: 18786365
23. Marschner A., Kalisch R., Vervliet B., Vansteenwegen D., & Büchel C. (2008). Dissociable roles for the hippocampus and the amygdala in human cued versus context fear conditioning. *Journal of Neuroscience*, 28(36), 9030–9036. https://doi.org/10.1523/JNEUROSCI.1651-08.2008 PMID: 18768697
24. Sowden S., Brewer R., Catmur C., & Bird G. (2016). The specificity of the link between alexithymia, interoception, and imitation. *Journal of Experimental Psychology: Human Perception and Performance*, 42(11), 1687–1692. https://doi.org/10.1037/xhp0000310 PMID: 27786535
25. Shah P., Catmur C., & Bird G. (2017). From heart to mind: Linking interoception, emotion, and theory of mind. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 93, 220.

26. Dunn B. D., Galton H. C., Morgan R., Evans D., Oliver C., Meyer M., et al. (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 PMID: 21106893

27. Pollatos O., Kirsch W., & Schandry R. (2005). On the relationship between interoceptive awareness, emotional experience, and brain processes. *Cognitive Brain Research*, 25(3), 948–962. https://doi.org/10.1016/j.cogbrainres.2005.09.019 PMID: 16298111

28. Terasawa Y., Moriguchi Y., Tochizawa S., & Umeda S. (2014). Interoceptive sensitivity predicts sensitivity to the emotions of others. *Cognition and Emotion*, 28(8), 1435–1448. https://doi.org/10.1080/02699931.2014.888988 PMID: 24559130

29. Wood A., Rychlowska M., Korb S., & Niedenthal P. (2016). Fashioning the face: sensorimotor simulation contributes to facial expression recognition. *Trends in cognitive sciences*, 20(3), 227–240. https://doi.org/10.1016/j.tics.2015.12.010 PMID: 26876363

30. Ondobaka S., Kilner J., & Friston K. (2017). The role of interoceptive inference in theory of mind. Brain and Cognition, 112, 64–68. https://doi.org/10.1016/j.bandc.2015.08.002 PMID: 26275633

31. 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), 1–12. https://doi.org/10.1371/journal.pone.0208034

32. Critchley H. D. (2002). Electrodermal responses: What happens in the brain. *Neuroscientist*, 8(2), 132–142. https://doi.org/10.1177/107385840200800209 PMID: 11954558

33. Kleckner I. R., Zhang J., Touroutoglou A., Chanes L., Xia C., Simmons W. K., et al. (2017). Evidence for a large-scale brain system supporting allostasis and interoception in humans. *Nature Human Behaviour*, 1(5), 69. https://doi.org/10.1038/s41562-017-0069 PMID: 28983518

34. Bradley M. M., Codispoti M., Cuthbert B. N., & Lang P. J. (2001). Emotion and motivation I: defensive and appetitive reactions in picture processing. *Emotion*, 1(3), 276. PMID: 12934687

35. Bernat E., Patrick C. J., Benning S. D., & Tellegen A. (2006). Effects of picture content and intensity on affective physiological response. *Psychophysiology*, 43(1), 93–103. https://doi.org/10.1111/j.1469-8986.2006.00380.x PMID: 16629689

36. Shannon C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3), 379–423.

37. Strange B. A., Duggins A., Penny W., Dolan R. J., & Friston K. J. (2005). Information theory, novelty and hippocampal responses: unpredicted or unpredictable?. *Neural Networks*, 18(3), 225–230. https://doi.org/10.1016/j.neunet.2004.12.004 PMID: 15896570

38. Calder A. J., Young A. W., Rowland D., & Perrett D. I. (1997). Computer-enhanced emotion in facial expressions. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 264(1383), 919–925. https://doi.org/10.1098/rspb.1997.0127 PMID: 9265191

39. Matsumoto D., & Hwang H. S. (2011). Judgments of facial expressions of emotion in profile. *Emotion*, 11(5), 1223. https://doi.org/10.1037/a0024356 PMID: 21942701

40. Ekman P. (1982). Does the face provide accurate information? *Emotion in the Fuman Face*.

41. Kirita T., & Endo M. (1995). Happy face advantage in recognizing facial expressions. *Acta Psychologica*, 89(2), 149–163.

42. Leppänen J. M., & Hietanen J. K. (2003). Affect and Face Perception: Odors Modulate the Recognition Advantage of Happy Faces. *Emotion*, 3(4), 315–326. https://doi.org/10.1037/1528-3542.3.4.315 PMID: 14674826

43. Ainley V., Apps M. A. J., Fotopoulou A., & Tsakiris M. (2016). "Bodily precision": A predictive coding account of individual differences in interoceptive accuracy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1708).

44. Hübner, A. M., Trempler, I., Kloth, N., Duning, T., Johnen, A., Schubotz, R. I. (2021). Selective Impairment of Emotion Inference in Behavioural Variant Frontotemporal Dementia? A Pilot Study. Manuscript in preparation.

45. Langeslag S. J., Gootjes L., & van Strien J. W. (2018). The effect of mouth opening in emotional faces on subjective experience and the early posterior negativity amplitude. *Brain and cognition*, 127, 51–59. https://doi.org/10.1016/j.bandc.2018.10.003 PMID: 30316954

46. 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 PMID: 23133619

47. Bach D. R., Friston K. J., & Dolan R. J. (2013). An improved algorithm for model-based analysis of evoked skin conductance responses. *Biological Psychology*, 94(3), 490–497. https://doi.org/10.1016/j.biopsycho.2013.09.010 PMID: 24063955

48. Bach D. R., Flandin G., Friston K. J., & Dolan R. J. (2009). Time-series analysis for rapid event-related skin conductance responses. *Journal of Neuroscience Methods*, 184(2), 224–234. https://doi.org/10.1016/j.jneumeth.2009.08.005 PMID: 19686778

49. Bach D. R., Daunizeau J., Friston K. J., & Dolan R. J. (2010). Dynamic causal modelling of anticipatory skin conductance responses. *Biological psychology*, 85(1), 163–170. https://doi.org/10.1016/j.biopsycho.2010.06.007 PMID: 20599582

50. R Core Team. (2019). A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

51. Bates D., Maechler M., Bolker B., Walker S., Christensen R. H. B., Singmann H., et al. (2015). Package 'lme4'. *Convergence*, 12(1), 2.

52. Kuznetsova A., Brockhoff P. B., & Christensen R. H. (2017). lmerTest package: tests in linear mixed effects models. Journal of statistical software, 82(1), 1–26.

53. Lenth R. V. (2016). Least-squares means: the R package lsmeans. *Journal of statistical software*, 69 (1), 1–33.

54. Bürkner P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, 80(1), 1–28. https://doi.org/10.18637/jss.v080.i01

55. Carpenter B., Gelman A., Hoffman M. D., Lee D., Goodrich B., Betancourt M., et al. (2017). Stan: A Probabilistic Programming Language. *Journal of Statistical Software*, 76(1). https://doi.org/10.18637/jss.v076.i01

56. 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 PMID: 17543405

57. Matsumoto D., & Hwang H. C. (2014). Judgments of subtle facial expressions of emotion. Emotion, 14 (2), 349. https://doi.org/10.1037/a0035237 PMID: 24708508

58. Hess U., Adams R. B., Grammer K., & Kleck R. E. (2009). Face gender and emotion expression: Are angry women more like men? *Journal of Vision*, 9(12), 19. https://doi.org/10.1167/9.12.19 PMID: 20053110

59. Motley M. T., & Camden C. T. (1988). Facial expression of emotion: A comparison of posed expressions versus spontaneous expressions in an interpersonal communication setting. *Western Journal of Communication (Includes Communication Reports)*, 52(1), 1–22.

60. Hess U., Blairy S., & Kleck R. E. (1997). The intensity of emotional facial expressions and decoding accuracy. *Journal of Nonverbal Behavior*, 21(4), 241–257. https://doi.org/10.1023/A:1024952730333

61. Palermo R. (2004). SpringerLink—Behavior Research Methods, Volume 36, Number 4. *Behavior Research Methods*, 36(4), 634–638. https://doi.org/10.3758/BF03206544

62. Adolphs R. (2002). Recognizing emotion from facial expressions: psychological and neurological mechanisms. *Behavioral and Cognitive Neuroscience Reviews*, 1(1), 21–62. https://doi.org/10.1177/1534582302001001003 PMID: 17715585

63. Montagne B., Kessels R. P. C., De Haan E. H. F., & Perrett D. I. (2007). The emotion recognition task: A paradigm to measure the perception of facial emotional expressions at different intensities. *Perceptual and Motor Skills*, 104(2), 589–598. https://doi.org/10.2466/pms.104.2.589-598 PMID: 17566449

64. Bayet L., Quinn P. C., Laboissière R., Caldara R., Lee K., & Pascalis O. (2017). Fearful but not happy expressions boost face detection in human infants. *Proceedings of the Royal Society B: Biological Sciences*, 284(1862). https://doi.org/10.1098/rspb.2017.1054

65. Smith F. W., & Rossit S. (2018). Identifying and detecting facial expressions of emotion in peripheral vision. *PLoS ONE*, 13(5), 1–15. https://doi.org/10.1371/journal.pone.0197160 PMID: 29847562

66. Palermo R., & Coltheart M. (2004). Photographs of facial expression: Accuracy, response times, and ratings of intensity. *Behavior Research Methods, Instruments, & Computers*, 36(4), 634–638.

67. Rapcsak S. Z., Galper S. R., Comer J. F., Reminger S. L., Nielsen L., Kaszniak A. W., et al. (2000). Fear recognition deficits after focal brain damage: a cautionary note. *Neurology*, 54(3), 575. https://doi.org/10.1212/wnl.54.3.575 PMID: 10680785

68. Arnold A. J., Winkielman P., & Dobkins K. (2019). Interoception and Social Connection. *Frontiers in Psychology*, 10, 1–6.

69. Decety J. (2011). Dissecting the neural mechanisms mediating empathy. *Emotion Review*, 3(1), 92–108.

70. Zaki J. (2014). Empathy: a motivated account. *Psychological Bulletin*, 140(6), 1608. https://doi.org/10.1037/a0037679 PMID: 25347133

71. Banks S. J., Bellerose J., Douglas D., & Jones-Gotman M. (2012). Bilateral skin conductance responses to emotional faces. *Applied psychophysiology and biofeedback*, 37(3), 145–152. https://doi.org/10.1007/s10484-011-9177-7 PMID: 22407530

72. Kreibig S. D. (2010). Autonomic nervous system activity in emotion: A review. *Biological psychology*, 84(3), 394–421. https://doi.org/10.1016/j.biopsycho.2010.03.010 PMID: 20371374

73. Fusar-Poli P., Landi P., & O'Connor C. (2009). Neurophysiological response to emotional faces with increasing intensity of fear: A skin conductance response study. *Journal of Clinical Neuroscience*, 16 (7), 981–982. https://doi.org/10.1016/j.jocn.2008.09.022 PMID: 19362481

74. Werner N. S., Kerschreiter R., Kindermann N. K., & Duschek S. (2013). Interoceptive awareness as a moderator of affective responses to social exclusion. *Journal of Psychophysiology*, 27(1), 39–50. https://doi.org/10.1027/0269-8803/a000086

75. Jiang J., Summerfield C., & Egner T. (2013). Attention sharpens the distinction between expected and unexpected percepts in the visual brain. *Journal of Neuroscience*, 33(47), 18438–18447. https://doi.org/10.1523/JNEUROSCI.3308-13.2013 PMID: 24259568

76. Petzschner F. H., Weber L. A., Wellstein K. V., Paolini G., Do C. T., & Stephan K. E. (2019). Focus of attention modulates the heartbeat evoked potential. *NeuroImage*, 186, 595–606. https://doi.org/10.1016/j.neuroimage.2018.11.037 PMID: 30472370

77. Gentsch A., Sel A., Marshall A. C., & Schütz-Bosbach S. (2019). Affective interoceptive inference: Evidence from heart-beat evoked brain potentials. *Human Brain Mapping*, 40(1), 20–33. https://doi.org/10.1002/hbm.24352 PMID: 30159945

78. 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 PMID: 11554668

79. Raes A. K., & De Raedt R. (2011). Interoceptive awareness and unaware fear conditioning: Are subliminal conditioning effects influenced by the manipulation of visceral self-perception? *Consciousness and Cognition*, 20(4), 1393–1402. https://doi.org/10.1016/j.concog.2011.05.009 PMID: 21684179

80. Bradley M. M. (2009). Natural selective attention: Orienting and emotion. *Psychophysiology*, 46(1), 1–11. https://doi.org/10.1111/j.1469-8986.2008.00702.x PMID: 18778317

81. Schneider W., & Chein J. M. (2003). Controlled & automatic processing: behavior, theory, and biological mechanisms. *Cognitive science*, 27(3), 525–559.

82. Norman, D. A., & Shallice, T. (1980). Attention to Action: Willed and Automatic Control of Behavior. Technical Report No. 8006.

83. Barrett L. F. (2017). The theory of constructed emotion: an active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience*, 12(1), 1–23. https://doi.org/10.1093/scan/nsw154 PMID: 27798257

84. Anderson E., Siegel E., White D., & Barrett L. F. (2012). Out of sight but not out of mind: Unseen affective faces influence evaluations and social impressions. *Emotion*, 12(6), 1210. https://doi.org/10.1037/a0027514 PMID: 22506501

85. Siegel E. H., Wormwood J. B., Quigley K. S., & Barrett L. F. (2018). Seeing what you feel: Affect drives visual perception of structurally neutral faces. *Psychological science*, 29(4), 496–503. https://doi.org/10.1177/0956797617741718 PMID: 29485945

86. Calì G., Ambrosini E., Picconi L., Mehling W. E., & Committeri G. (2015). Investigating the relationship between interoceptive accuracy, interoceptive awareness, and emotional susceptibility. *Frontiers in Psychology*, 6, 1–13.

87. Garfinkel S. N., & Critchley H. D. (2013). Interoception, emotion and brain: new insights link internal physiology to social behaviour. Commentary on:: "Anterior insular cortex mediates bodily sensibility and social anxiety" by Terasawa et al. (2012). *Social Cognitive and Affective Neuroscience*, 8(3), 231–234. https://doi.org/10.1093/scan/nss140 PMID: 23482658

88. Bornemann B., Herbert B. M., Mehling W. E., & Singer T. (2014). Differential changes in self-reported aspects of interoceptive awareness through 3 months of contemplative training. *Frontiers in Psychology*, 5(OCT), 1–13. https://doi.org/10.3389/fpsyg.2014.01504 PMID: 25610410

89. Farb N., Daubenmier J., Price C. J., Gard T., Kerr C., Dunn B. D., et al. (2015). Interoception, contemplative practice, and health. *Frontiers in Psychology*, 6(JUN), 1–26. https://doi.org/10.3389/fpsyg.2015.00763

90. Barrett L. F., Adolphs R., Marsella S., Martinez A. M., & Pollak S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological science in the public interest*, 20(1), 1–68. https://doi.org/10.1177/1529100619832930 PMID: 31313636

91. Silva M., Loureiro A., & Cardoso G. (2016). Social determinants of mental health: a review of the evidence. *The European Journal of Psychiatry*, 30(4), 259–292.

92. Khalsa S. S., & Lapidus R. C. (2016). Can interoception improve the pragmatic search for biomarkers in psychiatry?. *Frontiers in psychiatry*, 7, 121. https://doi.org/10.3389/fpsyt.2016.00121 PMID: 27504098

93. Price C. J., & Hooven C. (2018). Interoceptive awareness skills for emotion regulation: Theory and approach of mindful awareness in body-oriented therapy (MABT). *Frontiers in psychology*, 9, 798. https://doi.org/10.3389/fpsyg.2018.00798 PMID: 29892247

Supplementary Material

1. Stimulus Material

In the first pilot study a total of 60 participants (44 females; 27.4 ± 10.9; range 18-65 years) were asked to rate half of the start and end frames with regard to the depicted valence and its intensity on a 9-point scale ranging from 1 (*strong fear*), 5 (*neutral*) to 9 (*strong joy*). We recoded the variable so that we had one variable for the happy expressions with values ranging from 0 (*neutral*) to 4 (*strong joy*) and one for the fearful expressions with values ranging from 0 (*neutral*) to 4 (*strong fear*). We calculated two-way ANOVAs on the rating of the end and start frames of the final set of selected videos including the factors VALENCE and INTENSITY. For the end frames, there was a main effect of INTENSITY, $F(1, 124) = 164.41$, $p < 0.001$, but no main effect of VALENCE, $p = 0.904$, and no interaction of VALENCE and INTENSITY, $p = 0.795$. No differences between the conditions were found for the start frames (all $p \geq .185$) (Figure S1). Moreover, the corresponding ANOVA was calculated with video duration as dependent variable. We found a small but significant effect of INTENSITY, $F(1, 124) = 5.84$, $p = 0.017$, but no effect of VALENCE and no interaction effect between the two factors ($p \geq .182$) (Figure S2).

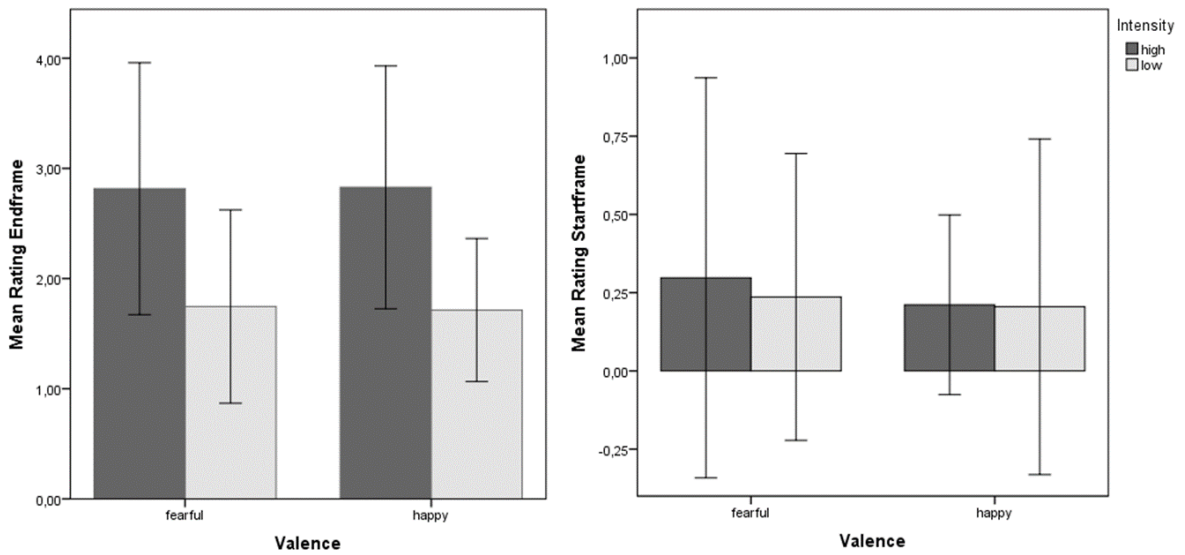

Figure S1. Mean intensity rating for the two conditions of the end and start frame of the final video set. Error bars represent two standard deviations from the mean.

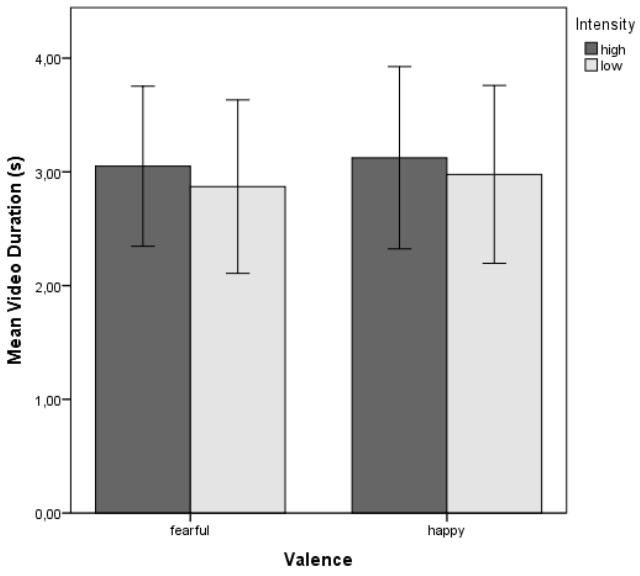

Figure S2. Mean duration of the final video set depending on condition. Error bars represent two standard deviations from the mean.

2. Additional Questionnaires

In addition to the MAIA-2, the questionnaire of self-assessment of emotional competencies (SEK-27) (Berking & Znoj, 2008) was completed, which includes nine scales with 27 items, three for each scale: attention, body awareness, clarity, understanding, regulation, acceptance, tolerance resilience, self-support, willingness to confront. Based on the close link between interoception and empathy and the assumption that a precise prediction of internal states also implies increased empathy for others (Ernst, Northoff, Böker, Seifritz, & Grimm, 2013; Ondobaka et al., 2017), we also added the Saarbruecken personality questionnaire in the self-report version (SPF) (Paulus, 2009). The SPF is a german version of the Interpersonal Reactivity Index (IRI) by Davis (1983) consisting of 16 items covering a total of four scales: empathic concern, fantasy, personal distress, assumption of perspectives. We also added the short form of the Trait Emotional Intelligence Questionnaire (TEIQue-SF). This is a 30-item questionnaire designed to measure global trait emotional intelligence (trait EI) based on the long form of the TEIQue (Furnham & Petrides, 2003).

To explore the relationship between the questionnaires and RTs, we calculated Pearson's correlation coefficients separately for the performance in the emotion and gender task. For the

correlation of RTs with the MAIA-2, we found significant negative correlations for the emotion task (r = -0.429, p = 0.002) as well as for the gender task (r = -0.299, p = 0.044). For the SPF, we found significant negative correlations for the emotion task (r = -0.379, p = 0.01) and the gender task (r = -0.296, p = 0.046). The correlation between the mean RT per participant and the SEK did not reach significance neither in the emotion task (r = -0.250, p = 0.094) nor in the gender task (r = -0.124, p = 0.41). Also, there were no significant correlations of TEI-score with RTs in the emotion task (r = -0.078, p = 0.61) or the gender task (r = 0.066, p = 0.66).

3. Results of the Bayesian logistic multilevel model predicting RTs

Table S1. Bayesian logistic multilevel model predicting RTs in the emotional condition

Coefficient	*b*	l-95% CI	u-95% CI
Video duration	-0.01	-0.02	0.01
Valence	0.03	0.03	0.04
Intensity	0.03	0.02	0.03
IS	-0.01	-0.01	-0.00
Surprise	-0.01	-0.02	-0.01
Entropy	0.05	0.00	0.09
Valence x Intensity	0.01	0.01	0.02
Valence x IS	0.00	-0.00	0.00
Intensity x IS	-0.00	-0.00	-0.00
Surprise x IS	0.00	0.00	0.00
Entropy x IS	0.00	0.00	0.01
Valence x Intensity x IS	0.00	0.00	0.00

4. Control Task

Figure S3. Marginal effects of gender (female, male) and intensity (high, low) as a function of interoceptive sensibility on **a)** reaction time and **b)** skin conductance response (SCR). The solid lines depict the regression fit, and the shaded areas show the 95% confidence intervals.

3.2 Study II: Selective Impairment of Emotion Inference in Behavioural Variant Frontotemporal Dementia? A Pilot Study.

Running title: Selective Impairment of Emotion Inference in bvFTD

Hübner, A. M., Trempler, I., Kloth, N., Duning, T., Johnen, A., & Schubotz, R. I.

Manuscript submitted for publication

Selective Impairment of Emotion Inference in Behavioral Variant Frontotemporal Dementia? A Pilot Study

Amelie M. Hübner[1 *], Ima Trempler[1,2 *], Nadine Kloth[3], Thomas Duning[4], Andreas Johnen[3], Ricarda I. Schubotz[1,2]

[1] Department of Psychology, University of Muenster, Germany

[2] Otto-Creutzfeldt-Center for Cognitive and Behavioural Neuroscience, University of Muenster, Germany

[3] Department of Neurology, University Hospital Muenster, Germany

[4] Department of Neurologie, Neuropsychology and Neurophysiologie, Hospital Bremen-Ost, Gesundheit Nord gGmbH, Bremen, Germany

[*] Shared first authorship

Running title: Selective Impairment of Emotion Inference in bvFTD

Corresponding Author:
Amelie M. Hübner
Institute for Psychology
Fliednerstraße 21
D-48149 Münster
Phone: 0049-251-8334192
E-Mail: huebnera@uni-muenster.de

Abstract

According to predictive coding models of interoception, sensitivity to others' emotions can be linked to actively inferred predictive models of the causes of interoceptive afferents that might underlie a person's behaviour. Consequently, the impairment in emotion recognition often observed in patients with behavioural variant frontotemporal dementia (bvFTD) could be due to deficits in interoceptive predictive mechanisms. The present study explored this question by investigating whether patients with bvFTD (n = 7), compared to healthy controls (n = 7) show difficulties in the discrimination of changes in others' emotional facial expressions and in learning from biased probabilities of these expressions. We recorded reaction times (RTs) and skin conductance responses (SCRs) while participants performed a probabilistic classification task, requiring them to indicate whether a neutral facial expression dynamically turned into a happy or fearful expression which could reach a low or high intensity. Probabilities of emotional expressions were varied between blocks to generate a bias for the more frequently encountered emotion. Healthy controls showed decreased RTs only to low intensity fearful expressions, whereas patients with bvFTD showed a gradual increase in RTs with increasing difficulty of the condition, i.e., from low-intensity fearful to high-intensity happy expressions. While the groups did not differ in their RTs with respect to exploiting the biased probabilities of happy vs. fearful faces, increasing predictability of upcoming facial expressions was associated with decreasing SCRs in healthy controls, but not in patients. Our results suggest that difficulties in recognition of facial emotions in bvFTD increase with uncertainty about the change in the emotional stimulus. They also provide evidence for an impaired prediction of emotional states within particular emotional contexts.

Keywords: bvFTD, interoception, emotion recognition, skin conductance response, emotional inference, predictive coding

1. Introduction

According to the concept of predictive coding, the brain not only responds to incoming sensory information, but constantly predicts information through probabilistic inference from current sensory input [1,2]. While evidence supporting this model mostly comes from exteroceptive perception, recent work suggests that predictive coding may also hold for the interoceptive system [3,4]. Interoception, which involves the prediction of the internal physiological state of one's own body [5], is of growing interest, particularly because it is thought to play a fundamental role in the development and perception of emotion [6-11].

The predictive coding model of interoception suggests that emotions result from actively inferred predictive models of the causes of interoceptive afferents [4]. Interoceptive predictions comprise descending efferent autonomic effects on peripheral physiology, which are compared to ascending afferent interoceptive signals [3,12]. Resulting mismatches, i.e. interoceptive prediction errors, convey information about interoceptive changes. The weight afforded to either predictions or prediction errors (i.e., *precision*) regulates the interplay between top-down and bottom-up interoceptive processing and has been suggested to play a role in intra- and interindividual differences in interoceptive abilities and emotional arousal. In particular, attentional mechanisms have been postulated as the basis for increased precision of prediction errors [13] and it has been shown that interoceptive attention modulates the cortical processing of interoceptive changes [14].

With this in mind, the question arises as to what problems occur when interoceptive predictive processing is dysfunctional. Deficits of predictive mechanisms have been proposed to underlie a wide range of diseases with neuropsychiatric symptoms, such as autism spectrum disorder [15], somatization disorders [16], attention deficit hyperactivity disorder [17], Parkinson's disease [18] and schizophrenia [19,20]. With regard to dysfunctional interoceptive prediction, clinical manifestations may arise when interoceptive signals can no longer be used

to guide behaviour, which in turn can lead to aberrant emotional awareness (i.e., alexithymia), difficulties in identifying emotions through facial expressions (i.e., emotion recognition) or inferring mental states of others (i.e., theory-of-mind) [3,21]. These characteristics are, among others, symptoms of patients suffering from behavioural variant frontotemporal dementia (bvFTD) [22-24]. Behavioural variant FTD is a neurodegenerative disease mainly characterized by early onset and progressive changes in personality and motivation, e.g. apathy, loss of empathy and socially inappropriate behaviour [25]. In contrast to other forms of dementia, noticeable cognitive deficits may only appear in later course of the disease, while behavioural symptoms dominate the heterogeneous clinical presentations in early stages, raising unique diagnostic challenges [26-28].

To date, disturbances in emotion recognition and theory of mind in bvFTD have not been directly linked to deficits in interoceptive prediction. Clinical studies suggest that updating in response to exteroceptive prediction errors is impaired in bvFTD, manifesting as difficulties in change detection [29], impaired decision making based on integration of social contextual information [30], and deficits in probabilistic learning in cognitive tasks [31]. Empirical evidence for impaired interoception in bvFTD has been inconclusive. Chiu and colleagues [32] report a general underestimation of the emotional intensity of negative facial expressions, whereas Balconi et al. [33] found that patients were able to explicitly appraise these expressions in terms of valence and arousal. Compared with healthy controls and patients with Alzheimer's disease, this appraisal was accompanied by a lower implicit autonomic responsiveness as assessed by skin conductance responses (SCRs). García-Cordero et al. [34] found impaired interoceptive accuracy and awareness in bvFTD patients, whereas Marshall et al. [35] did not find a difference between bvFTD patients and control participants while performing the same task. However, they recently showed that automatic imitation of dynamic facial expressions was reduced in bvFTD and correlated with impaired emotion identification [36].

In the present study, we assumed that the signals of interoceptive changes in bvFTD patients might be insufficient to update models in response to interoceptive prediction errors. We examined differences in reaction times (RTs) and SCRs between patients with bvFTD and healthy controls during a probabilistic emotion classification task using videotaped facial stimuli. During the task, participants were required to indicate whether a neutral facial expression transitioned into a happy or a fearful expression, whereby facial expressions at the end of the video varied in intensity. Thus, different levels of uncertainty were introduced, that is, from low intensity fearful expressions to high intensity happy expressions. Moreover, we implemented different probability levels for the occurrence of happy and fearful faces per block in order to assess participants' propensity to efficiently update their prediction model based on a priori likelihood. The varying probability and predictability of stimuli were quantified by information-theoretic measures, i.e., Shannon surprise (the improbability of an event, given by the negative logarithm of the probability) and entropy (the average surprise of all possible events) [37,38]. The task was previously used in a sample of healthy young participants [39], where we investigated variations in RTs and SCRs as a function of interindividual differences in interoceptive sensibility (IS). Based on the results of this previous study, we expected that patients with bvFTD would show similar RT and SCR patterns to those observed in individuals with low IS. Specifically, we set up the following hypotheses:

(H1) Compared to healthy controls, we expected patients with bvFTD to show increased RTs and decreased SCRs when discriminating emotional changes, especially when a fearful vs. happy face was presented and intensity was low vs. high. (H2) Moreover, we expected that healthy controls, but not patients with bvFTD, would learn the block alternating probabilistic imbalance; this should lead to a positive correlation of surprise and entropy with RTs and SCRs in healthy controls, but not in patients.

To assess the specificity of the effects, we employed a non-emotional speeded classification control task, where patients and control subjects were asked to discriminate

gender of emotionally neutral faces while facial stimuli developed from pixelated to high resolution.

2. Materials and Methods

2.1 Participants

Due to the coronavirus pandemic, we had to complete our data collection earlier than intended in order to protect participants and patients as high-risk groups. Seven right-handed patients diagnosed with bvFTD (1 female, 6 males; 66.43 ± 9.31 years old, range 58-81 years) participated in the study. Patients were recruited from the outpatient Memory Clinic at the Department of Neurology at the University Hospital Münster. All recruited patients had previously received a diagnosis of bvFTD based on a comprehensive diagnostic process including at least structural MRI of the brain, extensive history-taking and caregiver interviews as well as a comprehensive neuropsychological examination. All patients fulfilled the current revised diagnostic criteria for bvFTD [25]. A group of seven healthy participants (4 females, 3 males; 58.50 ± 10.75 years old; range 48-73 years) served as healthy control subjects. No participant had a history of other neurological or psychiatric diseases.

All participants gave written informed consent prior to participation. The study was performed in accordance with the Declaration of Helsinki and had been approved by the ethics committee of the University Hospital Muenster.

2.2 Neuropsychological Testing

Patients completed a neuropsychological test battery assessing core cognitive functions. This included the Montreal Cognitive Assessment (MoCA), a screening tool designed for mild cognitive disorders taking into account different cognitive areas such as attention and concentration, executive functions, memory, language, visuoconstructive skills, conceptual

thinking, arithmetic and orientation [40]; the German version of the Frontal Assessment Battery (FAB-D), constructed to assess behavioral and frontal-executive functions [41,42]; the reduced version of the Social Cognition and Emotional Assessment (mini-SEA), composed of two different subtests including a theory of mind evaluation (Faux-Pas test) and a facial emotion recognition test, to assess impaired emotional and social functions [43,44]; the Dementia Apraxia Test (DATE), designed to differentiate bvFTD from Alzheimer's disease based on apraxia profiles [45]; two subscales of the Regensburg Word Fluency test (RWT) [46], which is a neuropsychological method for the detection of executive dysfunction based on the Controlled Oral Word Association Test (COWAT) by Benton [47]; and the German version of the Hospital Anxiety and Depression Scale (HADS-D), a self-assessment tool used to assess fears and depressive behaviors in the context of physical illness [48,49]. Mean test values of all applied tests along with standard deviations for the two samples are shown in Table 1.

In addition, family caregivers completed various questionnaires about their impression of patients' symptoms and behaviour. These included the Frontal Behavioral Inventory (FBI), an questionnaire assessing behavioural symptoms [50]; the Saarbruecken personality questionnaire (SPF) [51], which is the German equivalent of the Interpersonal Reactivity Scale (IRI; [52]) consisting of 16 items covering a total of four scales: empathic concern, fantasy, personal distress and perspective taking; the informant-rated Apathy Evaluation Scale (AES-I) [53], a tool to characterize and quantify apathy in individuals ages 55 and older; and the Toronto Alexithymia Scale (TAS), a personality trait questionnaire to assess alexithymia [54]. Mean test values along with standard deviations for the patients are shown in the Supporting Files containing all the supporting tables and figures.

Table 1. Demographic and clinical data

	Mean (± SD)		*p*-value[a]
	Patients (*n = 7*)	Controls (*n= 7*)	
Age (years)	66.43 (9.31)	58.50 (10.75)	0.181
Education (years)	14.07 (3.14)	13.92 (2.63)	0.926
MoCA	21.00 (5.92)	25.50 (1.87)	0.097
FAB-D	13.43 (2.99)	16.67 (1.51)	0.036
mini-SEA	25.29 (4.61)	28.17 (2.40)	0.197
DATE	42.14 (8.11)	52.67 (2.50)	0.011
RWT			
Subscale S-words	6.57 (3.87)	16.00 (5.55)	0.004
Wordlist animals	16.00 (8.49)	24.33 (5.68)	0.066
HADS			
Depression	3.86 (1.95)	3.00 (1.79)	0.430
Anxiety	5.57 (2.64)	7.00 (3.90)	0.449

MoCa, Montreal Cognitive Assessment; Fab-D, Frontal Assessment Battery; mini-SEA, Social Cognition and Emotional Assessment; DATE, Dementia Apraxia Test; RWT, Regensburg Word Fluency test; HADS, Hospital Anxiety and Depression Scale

[a] *p*-value of independent *t*-tests

2.3 Stimulus Material

The stimuli used were the same as reported in Hübner et al. [39]. Stimuli consisted of short videos with a mean duration of 3.00 s shown on a grey background. In the emotional condition, videos depicted different actresses and actors posing from neutral to either happy or fearful emotional facial expressions. In addition, the facial expressions were shown with different intensities, i.e., either with high or low intensity. To control for possible effects of mouth opening on the recognition of emotions [55], each condition (fearful/happy expressions of high/low intensity) was presented in two different versions, i.e., with the mouth open and closed. To reduce the complexity of the statistical model this factor was not considered in the statistical analyses. Two videos of one of eight actors/actresses were selected for each of the eight conditions (happy vs. fearful x high vs. low intensity x mouth open vs. closed) resulting

in 128 different videos in total. The final videos had a size of 800x800 pixels and a framerate of 25 frames per second. We accepted that the emotional videos vary slightly in length (*SD* = 0.39 s) in order to create ecologically valid stimuli and included the duration of the videos as a covariate in the analyses to control for possible confounds on RTs. For details on the creation and piloting of the stimulus material, see Hübner et al. [39].

For the non-emotional condition, we recorded videos of each actor or actress with a continual neutral expression. The videos were edited so that faces slowly emerged from a pixelated version to a high-resolution image using a Gaussian soft focus in Premiere Pro CC. At the first frame, the faces were entirely pixelated and not recognizable (as ensured by a further pilot study) and a linear decay was inducted from the fifth frame on. In order to generate different levels of intensity comparable to the emotion task, the created videos either dissolved completely or ran out with a remaining soft focus of 20 at the end frame. To ensure that we did not favour ceiling effects due to noticeable features like the hairline, the grey background was extended ellipsoid around the faces. Overall, we created two videos per actor/actress with high and low intensity with a size of 1080 by 864 pixels, a framerate of 25 frames per second and a length of 2.5 s.

2.4 Task

Patients and controls were seated during the experiment in front of a computer screen located at a distance of about one meter. Videos were presented at the centre of the screen and were separated by an interstimulus interval of 4 s during which a fixation cross was displayed in the middle of the screen (Figure 1). Participants were asked to closely attend to the presented videos and to respond as fast and accurately as possible as soon as they recognized the emotion or the gender of the presented face in the emotional or the non-emotional condition, respectively. Patients and controls responded by button press on a two-button response box,

using the index and middle fingers of the right-hand. We counterbalanced stimulus-response mappings across participants.

Prior to the experiment, patients and controls were instructed and then performed a short training session to get accustomed to the task. The training consisted of one block of 16 trials each with equal probability for the different conditions. The task consisted of eight blocks, four emotional and four non-emotional blocks with 32 consecutive videos each, i.e., a total of 256 trials (128 per condition). After repeating four blocks of the same condition (emotional or non-emotional condition), participants could take a break before the condition changed and four blocks of the other condition were shown. At the beginning of each block, instructions were presented on the screen indicating whether the task was to respond to the emotional expression or the gender of the faces. Each block was followed by a pause of 8 seconds during which the information was given that the block had ended, followed by instructions for the next block. Accordingly, the emotion recognition task took about 22 minutes, the control task took about 19 minutes.

Within each block, we implemented either a high (.75) or low (.25) probability for the occurrence of fearful or happy faces in the emotional condition and either a high (.75) or low (.25) probability of male or female faces in the non-emotional condition. To ensure that videos of one actor or actress were not repeated across consecutive trials, the stimuli were presented in a pseudo-randomized order. Moreover, we balanced transitions between the block types across the experiment.

The randomisation was programmed using MATLAB R2019b (The MathWorks Inc., Natick, MA, USA) and stimuli were presented using Presentation software (Version 19.0, Neurobehavioral Systems, Inc., Berkeley, CA).

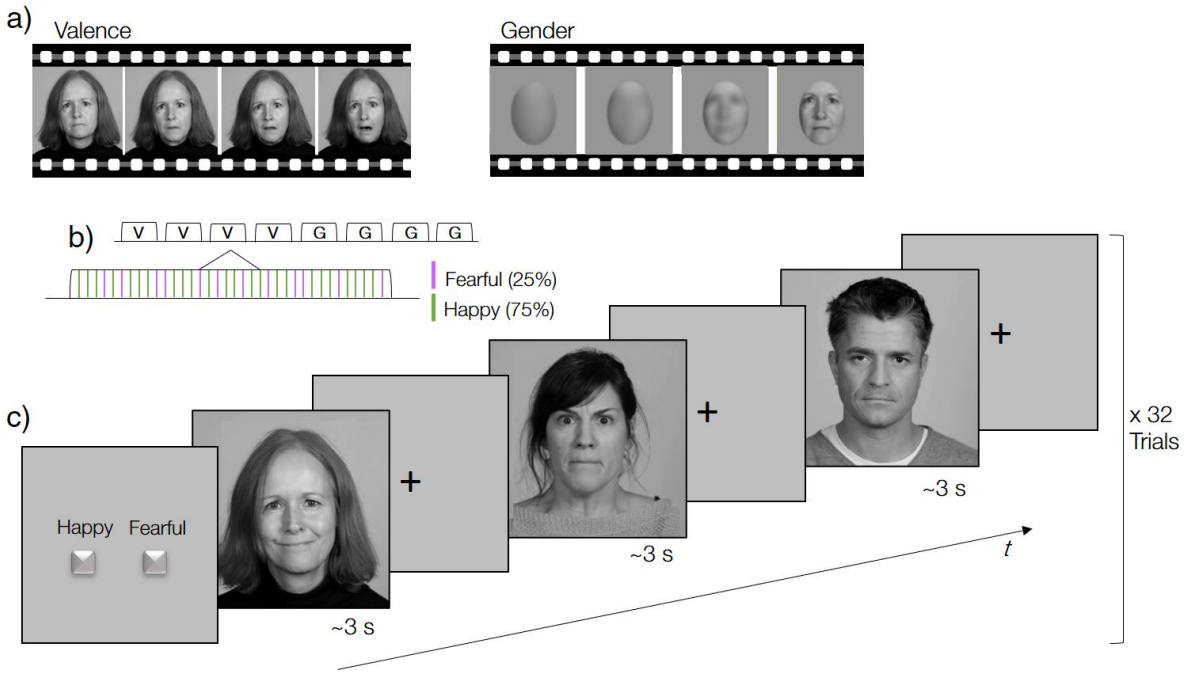

Fig 1. Schematic representation of the classification task. a) Short video clips were presented during which participants were asked to indicate the valence (V) of a neutral face developing to either a happy or fearful expression in the emotional condition, or the gender (G) of the face developing from highly pixelated to high resolution in the non-emotional control condition. **b)** Trials were presented in a mixed block and event-based design. Within blocks, trials per category occurred with different probabilities. **c)** At the beginning of each block, it was specified whether the participant had to indicate the emotion or gender of the upcoming faces. A block comprises 32 consecutive video trials. The actors and actresses in the picture have given their written consent to the publication of the photos.

2.5 Skin conductance response (SCR) acquisition

We acquired SCRs using the BrainVision Recorder Version 1.20.0801 (Brain Products, Munich, Germany) and placed two Ag/AgCl electrodes on the annular and middle fingers of the participant's left hand, using 0.5%-NaCl electrode paste (GEL101; Biopac Systems). Data were recorded at 500 Hz, using a sampling interval of 2000 µS. Preprocessing and data analysis were performed with PSPM [4.1.1], available at pspm.souorceforge.net. SCR data were converted back to a waveform signal with 100 Hz time resolution. The data were filtered with a unidirectional first-order Butterworth high pass filter with a cut-off frequency of 0.05 Hz and down sampled to 10 Hz, according to established recommendations [56]. To account for interindividual differences in responsiveness, we then *z*-transformed the entire SCR time series

for each participant [57]. No formal artefact rejection was implemented, but the data were visually checked for artefacts. The analysis of stimulus-locked (evoked) responses was done following the general linear convolution model (GLM) approach on a single-trial level. We then extracted trial-by-trial estimates and ran the GLM with one regressor per trial, modelling each trial as a Dirac delta function centred on the event onset, convolved with a canonical SCR function and its first derivative [58]. Finally, we reconstructed the response for each condition from the estimated amplitude parameters for the canonical SCR function and its derivative [56].

2.6 Data Analysis

Basic statistical analyses were performed using R, version 3.6.2 [59]. For both RTs and SCRs, false or missing responses were excluded from the analyses. Behavioural performance was defined by RTs.

Both, comparisons of RTs and SCRs were carried out separately for the emotion and the control task. More precisely, it was tested whether valence/gender, intensity and information-theoretic quantities, i.e., Shannon's surprise $I(x_i)$ and entropy $H(X)$ [37], could predict (variance in) RTs and SCRs on a single-trial level. While 'Surprise' refers to the improbability of a particular event, 'entropy' measures the expected or average surprise over all trials and thus reflects the predictability of an event within a given context [38]. In our paradigm, Shannon's surprise was based on the frequency of a trial of a specific valence/gender x_i normalized by the sum of the previous trials in the block:

$$p(x_i) = \frac{n(x_i) + 1}{\sum x_t + 1}$$

We set the prior counts before observing the first trial in the block to 1/2 for the two factor levels of valence (happy, fearful) and gender (male, female). In accordance with the following formula, the surprise $I(x_i)$ of each trial given by the negative logarithm of this probability quantifies the amount of information provided by the current stimulus:

$$I(x_i) = -\ln p(x_i)$$

Entropy $H(X)$ measures the average surprise of all possible outcomes and quantifies the expected information of an event in terms of its predictability:

$$H(X) = \sum_i -p(x_i) \ln p_k(x_i)$$

The generalized linear mixed-effects analyses were conducted using R, version 3.6.2 [59] via the package *lme4*, version 1.1.21 [60]. We transformed RTs to the natural logarithm to more closely approximate a normal distribution and used Q-Q plots to check if residuals of RTs and SCR data were normally distributed. For the factors valence/gender and intensity, effect coding with -1 for happy and 1 for fearful expressions, -1 for female and 1 for male faces, and -1 for high and 1 for low intensity expressions, was used. Each model was fit with valence/gender and intensity (and their interaction), Shannon's surprise and entropy and their respective interaction with the factor group (patients vs. controls) as fixed effects, and with a random intercept for each subject. Moreover, we added the participants' age as well as video duration as fixed effects. Statistical significance for each fixed effect was calculated via *lmerTest*, version 3.1.1 [61], using the Satterthwaite's approximation to denominator degrees of freedom. The significance level was set to $\alpha = .05$. For posteriori pairwise comparisons between patients and healthy controls, we used *lsmeans* [62] with the Tukey adjustment for multiple tests.

Additionally, we calculated Bayesian linear multilevel models in R [59] via the *brms* package and Stan using default priors [63,64]. Note that regression coefficients and 95% credible intervals (CIs; i.e., Bayesian confidence intervals) are reported, which means that the respective parameter falls within this interval with a 95% probability and indicates statistical significance on a 5% level if the interval does not contain zero.

3. Results

3.1 Emotional Change Detection

Behavioural data. We restricted all analyses to RTs because the percentage of false alarms (FAs) and misses (Ms) of healthy controls (FAs: M = 2.01%, SD = 0.99; Ms: M = 0.33%, SD = 0.88) as well as of patients with bvFTD (FAs: M = 5.13%, SD = 4.23; Ms: M = 2.46%, SD = 2.40) was generally low in the emotional condition.

The linear multilevel model predicting RTs revealed a significant main effect of valence, b = -0.14, t =-5.58, p < 0.001, characterized by increased RTs for fearful compared to happy facial expressions, a significant main effect of intensity, b = -0.09, t = -3.37, p < 0.001, with increased RTs for low vs. high intensity, but no significant main effect of group, b = -0.13, t = -0.987, p = 0.336. While age had a significant effect, b = 0.02, t = 3.38, p = 0.006, video length did not predict RTs, b = 0.003, t = 0.02, p = 0.988.

In line with our hypothesis H1, there was a significant three-way interaction of valence, intensity and group, b = 0.10, t = 2.10, p = 0.036, which was also supported by the results of the Bayesian logistic multilevel model, b = 0.10, 90%-CI = [0.01, 0.20]. Patients with bvFTD showed increased RTs for low compared to high intensities of expressions, especially when they were fearful as opposed to happy. In contrast, healthy controls also showed increased RTs for low compared to high intensity of fearful expressions, but no difference for happy expressions (Figure 2a). Thus, as expected healthy controls slowed down their responses only to low intensity fearful expressions, whereas patients with bvFTD showed a gradual increase in RTs with increasing difficulty of the condition, i.e., from happy high to fearful low intensity expressions.

SCR data. Contrary to our expectation, the linear mixed-effect model predicting SCRs revealed no significant main effects of valence, b = 0.08, t = 0.37, p = 0.710, intensity, b = 0.23, t = 1.00, p = 0.320, or group, b = -1.00, t = 1.73, p = 0.084, on SCR amplitudes, as well as no

interaction effects (all $p > 0.202$) (Figure 2b). Likewise, the Bayesian logistic multilevel model predicting SCRs did not reveal any significant main effects of valence, intensity and group and no interactions of these factors.

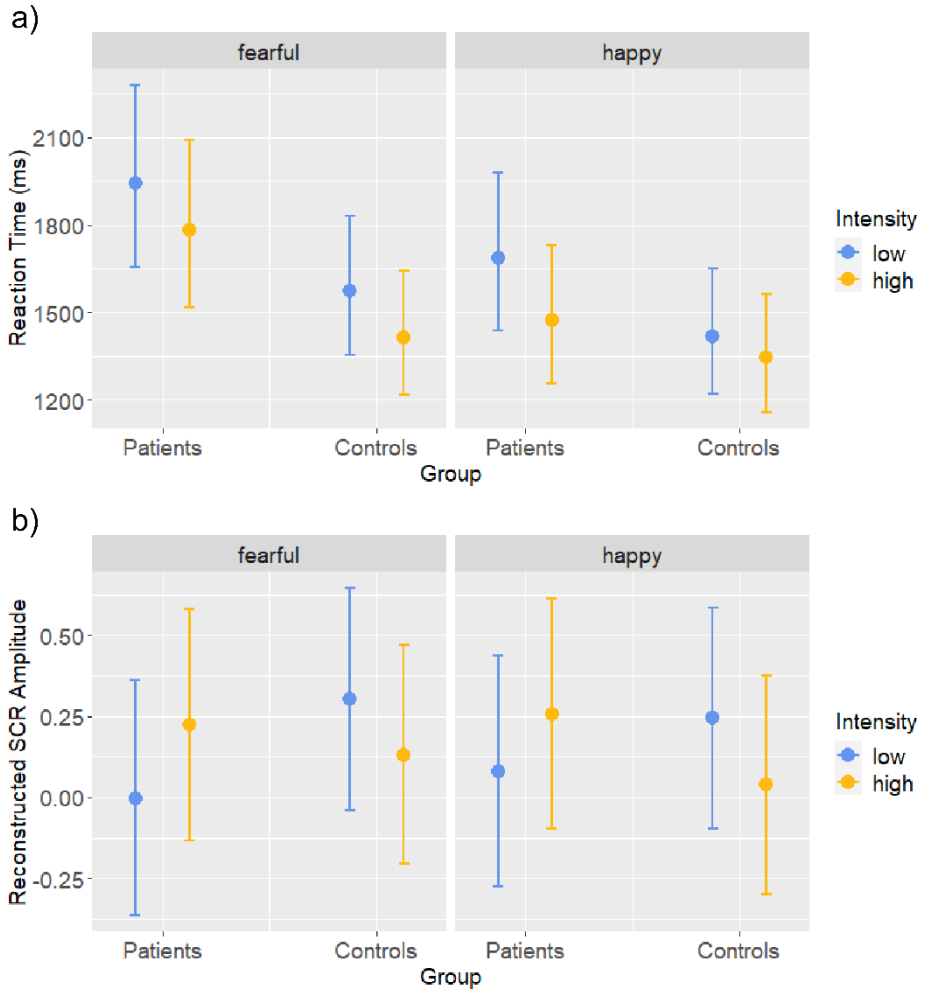

Fig 2. Effects of valence (happy, fearful) and intensity (high, low) on a) reaction time and b) skin conductance response (SCR) in the emotion condition separated for patients with bvFTD and healthy controls.

3.2 Probabilistic Context Adaptation

Behavioural data. To capture group differences in probabilistic context adaptation, we tested whether Shannon's surprise and entropy of a single emotional expression in interaction with group were predictive of RTs. The linear mixed-effect model revealed a significant main effect of surprise, $b = -0.06$, $t = -4.08$, $p < 0.001$, with shorter RTs to more surprising events, but no significant main effect of entropy, $b = -0.01$, $t = -1.40$, $p = 0.163$. Contrary to our

hypothesis (H2), there were no significant interactions of group with surprise, b = 0.001, t = 0.77, p = 0.441, and entropy, b = -0.18, t = -1.66, p = 0.097 (Figure 3a). The Bayesian logistic multilevel model supported these results revealing a corresponding main effect of surprise, b = -0.06, 90%-CI = [-0.08, -0.03], but no main effect of entropy, b = -0.11, 90%-CI = [-0.26, 0.05], and no interaction effects of group with surprise, b = 0.01, 90%-CI = [-0.02, 0.05], and entropy, b = -0.18, 90%-CI = [-0.39, 0.03].

SCR data. No main effects of surprise, b = 0.06, t = 0.47, p = 0.640, and entropy, b = -0.26, t = 0.38, p = 0.707, on SCRs and no interaction effect of group and surprise, b = 0.28, t = 1.64, p = 0.101, were observed. Supporting our hypothesis (H2), an interaction effect between group and entropy, b = 2.18, t = 2.31, p = 0.021, was observed (Figure 3b). As expected, healthy controls showed increased SCRs to unpredictable valences, whereas patients with bvFTD did not. The Bayesian logistic multilevel supported the interaction of entropy and group, b = 2.18, 90%-CI = [0.37, 4.05].

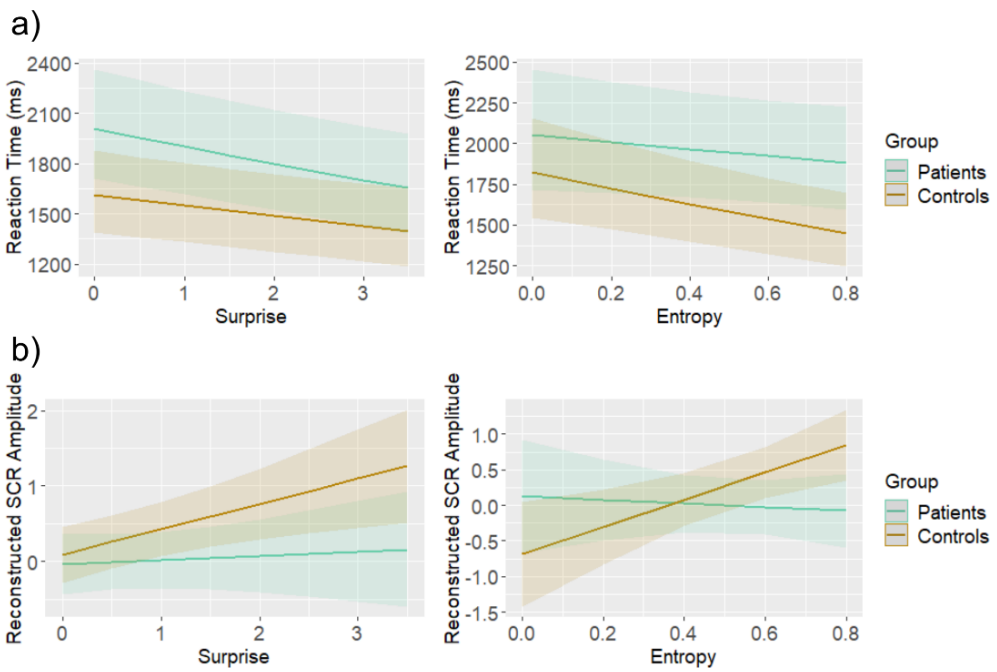

Fig 3. Effects of surprise and entropy on a) reaction time and b) skin conductance response (SCR) in the emotion condition separated for patients with bvFTD and healthy controls.

3.3 Control task

Behavioural data. In the non-emotion control condition for the recognition of gender, the percentage of false alarms and missing data of healthy controls (FAs: M = 3.35%, SD = 2.54; Ms: M = 0.33%, SD = 3.16) as well as of patients (FAs: M = 7.70%, SD = 4.22; Ms: M = 4.69%, SD = 5.94) was also relatively low.

The linear mixed regression model for the prediction of RTs revealed significant main effects of age, $b = 0.01$, $t = 3.24$, $p = 0.008$, gender, $b = 0.06$, $t = 2.5$, $p = 0.012$, and intensity, $b = 0.12$, $t = -4.95$, $p < 0.001$. Main effects of gender and intensity were driven by increased RTs to male vs. female faces and low vs. high intensity facial expressions (i.e. higher vs. lower degree of pixelation at the end of the video), respectively (see Supporting Files). No significant main effect of group ($p > 0.07$) and no interaction effects of group with intensity or gender were observed ($p > 0.840$). Moreover, no significant main or interaction effects were observed for surprise and entropy (all $p > 0.521$). Likewise, the Bayesian logistic multilevel model revealed main effects of gender, $b = 0.06$, 95%-CI [0.01, 0.11], and intensity, $b = -0.12$, 95%-CI [-0.17, -0.07], but no further significant effects.

SCR data. The linear mixed effect model on reconstructed SCR amplitudes revealed a significant main of entropy, $b = -1.75$, $t = -2.31$, $p = 0.021$, and a significant interaction effect of entropy and group $b = 2.14$, $t = 2.07$, $p = 0.040$ (see Supporting Files). This was also supported by the Bayesian model, with $b = -1.73$, 95%-CI [-3.21, -0.24] for entropy, and $b = 2.11$, 95%-CI [0.15, 4.07] for the interaction between entropy and group. Contrary to the effect found in the emotional condition, this effect was driven by decreased SCRs to unpredictable vs. predictable genders in patients, whereas in controls no modulation by entropy was observed.

4. Discussion

In the present work, we investigated whether behavioural variant frontotemporal dementia (bvFTD) is associated with impairment in the recognition of unfolding emotional expressions and adaptation to emotional contexts. Using a probabilistic emotion classification task, we found that healthy control participants showed prolonged reaction times (RTs) only to low intensity fearful expressions, while patients with bvFTD showed a gradual increase in RTs with increasing difficulty of the condition, i.e., from high-intensity happy expressions to low-intensity fearful expressions (H1). While no RT differences between the groups were found regarding the exploitation of biased probabilities of happy vs. fearful faces, a higher predictability of upcoming facial expressions came with decreasing SCRs in healthy controls but not in patients with bvFTD (H2). Finally, the results of the control task suggest that while bvFTD does not impair discrimination performance per se, differences to healthy participants regarding probabilistic context adaptation can be observed independent of the domain.

4.1. Emotional change detection in bvFTD

While in healthy controls a reduced speed of emotional change detection could only be observed for low intensity of fearful expressions compared to all other conditions (i.e., for the most difficult condition), bvFTD patients showed a gradual increase of response times from happy expressions of high intensity over happy expressions of low intensity and fearful expressions of high intensity to fearful expressions of low intensity. Deficits in emotion recognition and detection in bvFTD have been consistently observed in previous studies during the verbal categorization of facial expressions (for review see [65]), as well as for musical, vocal, or physical expression of emotions [22,66-68]. Moreover, there is consistent evidence that the processing of negative emotions, such as fear or disgust is more impaired than the processing of happy expressions [70-71]. The present results complement these few existing

studies in bvFTD patients, providing evidence for an effect of both intensity and valence on the detection of dynamically unfolding facial expressions of emotions, indicative of increasing difficulties in emotion recognition when uncertainty about the change in the emotional stimulus increases.

We interpret the observed deficits to suggest impaired interoceptive abilities in bvFTD. In a previous study using the same task in young healthy subjects, we assessed participants' interoceptive sensibility (IS) on the basis of their self-reported belief regarding their perception of bodily signals [39,72]. There, we observed that lower IS was associated with a slower detection of emotional changes, as also occurred in the study reported here [39]. Equivalent to the pattern shown by our patients, young healthy participants with reduced IS were slower to recognize low compared to high emotional intensity and fearful compared to happy facial expressions, whereas people with higher IS exhibited slower responses only in the most difficult condition, i.e. when presented with a fearful low intensity face. Thus, just as healthy individuals with low IS, patients with bvFTD seem to be less challenged by positive valence and high intensities of emotional expressions and may be able to partially compensate for their interoceptive deficits. Difficulties in understanding social signals, inappropriate behaviour in social interactions, as well as reduced empathy, which are all characteristic of bvFTD, might therefore be linked to the inability to recognize the relatively weak expressions of emotions that are common in everyday life [73,74], including moderate negative reactions of others to one's own behaviour.

Consistent with our interpretation of impaired interoception in bvFTD, Garcia-Cordero et al. [34] found impaired interoceptive accuracy in patients with bvFTD relative to healthy controls using a heartbeat detection task and an association of impaired task performance with reduced grey matter volume in the anterior insula and fronto-temporal regions. It has been suggested that the anterior insular complex and anterior cingulate cortex are critical for our interoceptive abilities [5,75]. This is where long-range spindle neurons, also called Von

Economo neurons (VENs), are located, allowing rapid integration of information [76]. The VENs are stained with antibodies of the D3 receptor for dopamine, which plays a role in modulating expectancy under uncertainty [77]. In addition, the serotonin 2b receptor, which is also found in the human stomach, is expressed on the VENs, providing a possible explanation for the "gut feeling" [78]. Early-stage fronto-insular degeneration has been described in patients with bvFTD, along with a 74% reduction in VENs that is not due to overall neuronal loss [79-82]. Functionally, reduced connectivity in the salience network comprising the anterior insula and anterior cingulate cortex has been found [83]. Impaired integrity of this network could therefore lead to difficulties in recognising changes in bodily states, perceiving emotions and social behaviour, and evaluating internal sensations that convey a "gut feeling" [69,84].

Although previous studies found that patients with bvFTD show lower baseline tonic levels and decreased phasic psychophysiological arousal to emotional stimuli in the sympathetic and parasympathetic nervous system [33,85-87], we did not observe general differences or valence- or intensity-dependent differences in SCR amplitude between patients with bvFTD and healthy controls. These findings are consistent with our previous study [39], which also did not provide evidence for a significant relationship between the IS score and SCR during emotion processing. As will be discussed below, we therefore suggest that the SCR pattern may reflect increased attention to (i.e., increased precision of) interoceptive prediction errors rather than reflecting interoceptive signals themselves.

4.2. Context-sensitive emotional inference in bvFTD

Healthy controls, as well as patients with bvFTD, detected an emotion faster when its valence was more surprising due to our probabilistic block manipulation. Modulation of RTs by surprise or entropy did not differ between groups. Accordingly, our assumption of an impaired adaptation of emotional predictions in bvFTD was not confirmed by behavioural

findings. In our previous study [39], individuals with higher IS detected expected and predictable emotions faster, whereas individuals with lower IS responded faster to unexpected and unpredictable emotions. In the present study, both healthy controls and bvFTD patients showed the pattern that was found for lower IS individuals in the previous study, i.e., shorter RTs to unexpected emotions. It is possible that the performance advantage resulting from the increasing likelihood of a particular valence was masked in control subjects by a hasty response to unexpected events that arguably attracted greater attention. Alternatively, the results could be due to a generally low IS in the healthy control group, which we did not assess in the present study. Moreover, the older age of the subjects might have played a role since studies show that difficulties in probabilistic learning increase with age [88,89].

However, although no differences in performance were found between healthy subjects and patients, in line with our hypothesis a modulation of SCR by entropy was observed. Emotional entropy (reflecting the predictability of an emotion within a particular context) positively covaried with SCRs in healthy controls but not in patients with bvFTD. Previous studies showed that SCR amplitudes can reflect an orienting reflex in response to novel or unexpected stimuli [90,91]. Moreover, the reported effects are in line with Hübner et al. [39] where young individuals with high (vs. low) IS showed the same SCR modulation by entropy as the healthy control subjects in the present study. We suggest that the increase of SCRs as a function of entropy in healthy individuals reflects heightened precision weighting of incoming input when predictability is low and, in turn, adaptation to particular emotional contexts when predictability is high [13,92]. The fact that no corresponding modulation of SCR by entropy was observed in patients with bvFTD provides first evidence that an impaired attentional gain control of interoceptive prediction error might be fundamental to the disease [3]. As interoceptive signals can no longer be used to guide a context-sensitive behavioural adjustment, patients are unable to adapt appropriately to specific social and/or emotional situations.

Inappropriate behaviours and reactions, which are common in these patients, might well reflect a consequence of this impairment.

4.3 Domain-specific and domain-general differences between bvFTD and healthy controls

To test whether effects in bvFTD patients were specific to emotion processing, and not merely reflected general deficits in face discrimination or implicit probabilistic adaptation, we used a non-emotional control task. While no differences between the groups were observed for the gender discrimination, we found an interaction effect between group and entropy on SCRs that was, however, opposite to the effect observed in the emotion condition.

The relatively few studies investigating face identity recognition in bvFTD patients revealed inconsistent findings. Van den Stock et al. [67] found no impairment in an identity matching task compared with an emotion matching task in bvFTD. In contrast, Kumfor et al. [93] using emotional stimuli mixed with neutral ones suggested the recognition of facial identity to be impaired. In the present study, we found no differences between groups regarding the discrimination of gender, even at higher levels of difficulty (i.e., higher degree of pixilation). Our findings thus indicate that an impairment in the discrimination of facial expressions in bvFTD is indeed specific to the recognition of different emotions.

Interestingly, regarding the probabilistic manipulation generating an expectation bias for female or male faces, we found an opposite effect to that observed in the emotional condition: though there were no differences with regard to a modulation of RTs by surprise or entropy, SCRs increased as a function of predictability in patients but not in controls. Speculatively, higher SCRs may reflect increasing influence of top-down modulation when predictability increases in an attempt to compensate for degraded visual input, as it has been reported for patients with bvFTD [94]. However, this finding suggests that patients are indeed influenced by the probabilistic structure of the task, but not if the categories for which the probabilities are implemented are emotional ones.

4.4 Limitations and strengths

Studies on the precise neuropsychological profile of patients with bvFTD are essential to improve early diagnosis, symptomatic treatment and disease management [95-97]. The present pilot study is the first to investigate whether patients with bvFTD show impairments in the recognition of emotional and non-emotional stimuli in probabilistic contexts. Moreover, our study is one of the first using dynamically unfolding emotional facial expressions to investigate the recognition of emotional changes in an ecologically valid way.

However, we acknowledge the small sample size and point out that the results need to be validated in future studies. Nevertheless, most of the effects we observed in patients with bvFTD in this study are consistent with those of Hübner et al. [39] with participants with low IS showing similar SCR patterns than patients with bvFTD. Accordingly, our findings provide first evidence for impaired interoceptive processing in patients suffering from bvFTD.

4.5 Conclusion

The present pilot study suggests that patients with bvFTD are slower to detect emotional changes in the facial expressions of others than healthy individuals, possibly due to an impairment in their interoceptive abilities. Deviant skin conductance modulation in this process indicates that these patients cannot effectively process interoceptive prediction errors. These results give new insights into the neuropsychological profile of patients with bvFTD, providing a starting point for further studies on impaired interoceptive processing in this disease.

References

1. Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138. https://doi.org/10.1038/nrn2787

2. Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87.

3. 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

4. Seth, A. K., & Friston, K. J. (2016). Active interoceptive inference and the emotional brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1708). https://doi.org/10.1098/rstb.2016.0007

5. Craig, A. D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nature reviews neuroscience*, *3*(8), 655.

6. Adolphs, R. (2002). Recognizing emotion from facial expressions: psychological and neurological mechanisms. *Behavioral and Cognitive Neuroscience Reviews*, *1*(1), 21–62.

7. 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

8. 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

9. Tsakiris, M., & Critchley, H. (2016). Interoception beyond homeostasis: affect, cognition and mental health.

10. Damasio, A. R., Tranel, D., & Damasio, H. C. (1991). Behavior: theory and preliminary testing. *Frontal lobe function and dysfunction*, 217.

11. Damasio, A. R. (2000). The Feeling of What Happens: Body and Emotion in the Making of Consciousness. New York: Harcourt Brace.

12. Seth, A. K., Suzuki, K., & Critchley, H. D. (2012). An interoceptive predictive coding model of conscious presence. *Frontiers in Psychology*, *3*(JAN), 1–16. https://doi.org/10.3389/fpsyg.2011.00395

13. Ainley, V., Apps, M. A. J., Fotopoulou, A., & Tsakiris, M. (2016). “Bodily precision”: A predictive coding account of individual differences in interoceptive accuracy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1708). https://doi.org/10.1098/rstb.2016.0003

14. Petzschner, F. H., Weber, L. A., Wellstein, K. V., Paolini, G., Do, C. T., & Stephan, K. E. (2019). Focus of attention modulates the heartbeat evoked potential. *NeuroImage*, *186*(November 2018), 595–606. https://doi.org/10.1016/j.neuroimage.2018.11.037

15. Van de Cruys, S., Evers, K., Van der Hallen, R., Van Eylen, L., Boets, B., de-Wit, L., & Wagemans, J. (2014). Precise minds in uncertain worlds: predictive coding in autism. *Psychological review*, 121(4), 649.

16. Owens, A. P., Allen, M., Ondobaka, S., & Friston, K. J. (2018). Interoceptive inference: From computational neuroscience to clinic. *Neuroscience & Biobehavioral Reviews*, 90, 174–183. https://doi.org/10.1016/j.neubiorev.2018.04.017

17. Gonzalez-Gadea, M. L., Chennu, S., Bekinschtein, T. A., Rattazzi, A., Beraudi, A., Tripicchio, P., ... & Ibanez, A. (2015). Predictive coding in autism spectrum disorder and attention deficit hyperactivity disorder. *Journal of neurophysiology*.

18. Trempler, I., Bürkner, P. C., El-Sourani, N., Binder, E., Reker, P., Fink, G. R., & Schubotz, R. I. (2020). Impaired context-sensitive adjustment of behaviour in Parkinson's disease patients tested on and off medication: An fMRI study. *NeuroImage*, 212, 116674.

19. 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), 48-58.

20. Brown, H., Adams, R. A., Parees, I., Edwards, M., & Friston, K. (2013). Active inference, sensory attenuation and illusions. *Cognitive processing*, 14(4), 411-427.

21. Brewer, R., Cook, R., & Bird, G. (2016). Alexithymia: A general deficit of interoception. *Royal Society Open Science*, *3*(10). https://doi.org/10.1098/rsos.150664

22. Snowden, J. S., Austin, N. A., Sembi, S., Thompson, J. C., Craufurd, D., & Neary, D. (2008). Emotion recognition in Huntington's disease and frontotemporal dementia. *Neuropsychologia*, 46(11), 2638–2649. https://doi.org/10.1016/j.neuropsychologia.2008.04.018

23. Johnen, A., & Bertoux, M. (2019). Psychological and cognitive markers of behavioral variant frontotemporal dementia–A clinical neuropsychologist's view on diagnostic criteria and beyond. *Frontiers in neurology*, *10*, 594.

24. Elamin, M., Pender, N., Hardiman, O., & Abrahams, S. (2012). Social cognition in neurodegenerative disorders: a systematic review. *Journal of Neurology, Neurosurgery & Psychiatry*, 83(11), 1071-1079.

25. Rascovsky, K., Hodges, J. R., Knopman, D., Mendez, M. F., Kramer, J. H., Neuhaus, J., van Swieten, J. C., Seelaar, H., Dopper, E. G. P., Onyike, C. U., Hillis, A. E., Josephs, K. A., Boeve, B. F., Kertesz, A., Seeley, W. W., Rankin, K. P., Johnson, J. K., Gorno-Tempini, M.-L., Rosen, H., … Miller, B. L. (2011). Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain*, 134(9), 2456–2477. https://doi.org/10.1093/brain/awr179

26. Bott, N. T., Radke, A., Stephens, M. L., & Kramer, J. H. (2014). Frontotemporal dementia: diagnosis, deficits and management. *Neurodegenerative disease management*, 4(6), 439-454.

27. Mendez, M. F., & Perryman, K. M. (2002). Neuropsychiatric Features of Frontotemporal Dementia: Evaluation of Consensus Criteria and Review. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 14(4), 424–429. https://doi.org/10.1176/jnp.14.4.424

28. Silveri, M. C., Reali, G., Jenner, C., & Puopolo, M. (2007). Attention and memory in the preclinical stage of dementia. *Journal of Geriatric Psychiatry and Neurology*, 20(2), 67-75.

29. Hughes, L. E., & Rowe, J. B. (2013). The Impact of Neurodegeneration on Network Connectivity: A Study of Change Detection in Frontotemporal Dementia. *Journal of Cognitive Neuroscience*, 25(5), 802–813. https://doi.org/10.1162/jocn_a_00356

30. O'Callaghan, C., Bertoux, M., Irish, M., Shine, J. M., Wong, S., Spiliopoulos, L., Hodges, J. R., & Hornberger, M. (2016). Fair play: Social norm compliance failures in behavioural variant frontotemporal dementia. *Brain*, 139(1), 204–216. https://doi.org/10.1093/brain/awv315

31. Dalton, M. A., Weickert, T. W., Hodges, J. R., Piguet, O., & Hornberger, M. (2013). Impaired acquisition rates of probabilistic associative learning in frontotemporal dementia is associated with fronto-striatal atrophy. *NeuroImage:* Clinical, 2, 56–62. https://doi.org/10.1016/j.nicl.2012.11.001

32. Chiu, I., Gfrörer, R. I., Piguet, O., Berres, M., Monsch, A. U., & Sollberger, M. (2015). "Now I see it, now I don't": Determining Threshold Levels of Facial Emotion Recognition for Use in Patient Populations. *Journal of the International Neuropsychological Society*, 21(7), 568–572. https://doi.org/10.1017/S1355617715000557

33. Balconi, M., Cotelli, M., Brambilla, M., Manenti, R., Cosseddu, M., Premi, E., ... & Borroni, B. (2015). Understanding emotions in frontotemporal dementia: the explicit and implicit emotional cue mismatch. *Journal of Alzheimer's Disease*, 46(1), 211-225.

34. 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

35. Marshall, C. R., Hardy, C. J., Russell, L. L., Clark, C. N., Dick, K. M., Brotherhood, E. V., ... & Warren, J. D. (2017). Impaired interoceptive accuracy in semantic variant primary progressive aphasia. *Frontiers in neurology*, 8, 610.

36. Marshall, C. R., Hardy, C. J., Russell, L. L., Clark, C. N., Bond, R. L., Dick, K. M., ... & Warren, J. D. (2018). Motor signatures of emotional reactivity in frontotemporal dementia. Scientific reports, 8(1), 1-13.

37. Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*(3), 379–423.

38. Strange, B. A., Duggins, A., Penny, W., Dolan, R. J., & Friston, K. J. (2005). Information theory, novelty and hippocampal responses: unpredicted or unpredictable?. *Neural Networks*, *18*(3), 225-230.

39. Hübner, A. M., Trempler, I., Gietmann, C., & Schubotz, R. I. (2021). Interoceptive sensibility predicts the ability to infer others' emotional states. *Plos one*, *16*(10), e0258089.

40. Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., ... & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*(4), 695-699.

41. Benke, T., Karner, E., & Delazer, M. (2013). FAB-D: German version of the frontal

assessment battery. *Journal of neurology, 260*(8), 2066-2072.

42. Dubois, B., Slachevsky, A., Litvan, I., & Pillon, B. F. A. B. (2000). The FAB: a frontal assessment battery at bedside. *Neurology, 55*(11), 1621-1626.

43. Bertoux, M., Volle, E., Funkiewiez, A., de Souza, L. C., Leclercq, D., & Dubois, B. (2012). Social Cognition and Emotional Assessment (SEA) is a marker of medial and orbital frontal functions: a voxel-based morphometry study in behavioral variant of frontotemporal degeneration. *Journal of the International Neuropsychological Society, 18*(6), 972-985.

44. Funkiewiez, A., Bertoux, M., de Souza, L. C., Lévy, R., & Dubois, B. (2012). The SEA (Social Cognition and Emotional Assessment): A clinical neuropsychological tool for early diagnosis of frontal variant of frontotemporal lobar degeneration. *Neuropsychology, 26*(1), 81.

45. Johnen, A., Frommeyer, J., Modes, F., Wiendl, H., Duning, T., & Lohmann, H. (2016). Dementia Apraxia Test (DATE): a brief tool to differentiate behavioral variant frontotemporal dementia from Alzheimer's dementia based on apraxia profiles. *Journal of Alzheimer's Disease, 49*(3), 593-605.

46. Aschenbrenner, S., Tucha, O., & Lange, K. W. (2000). Regensburger Wortflüssigkeits-Test: RWT. Hogrefe, Verlag für Psychologie.

47. Benton, A. L. (1968). Differential behavioral effects in frontal lobe disease. *Neuropsychologia, 6*,53–60

48. Petermann, F. (2015). Hospital anxiety and depression scale, deutsche version (HADS-D). *Zeitschrift für Psychiatrie, Psychologie und Psychotherapie.*

49. Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta psychiatrica scandinavica, 67*(6), 361-370.

50. Kertesz, A., Davidson, W., & Fox, H. (1997). Frontal behavioral inventory: diagnostic criteria for frontal lobe dementi. *Canadian Journal of Neurological Sciences,* 24(1), 29-36.

51. Paulus, C. (2009). Der Saarbrücker Persönlichkeitsfragebogen SPF (IRI) zur Messung von Empathie: psychometrische Evaluation der deutschen Version des interpersonal reactivity index.

52. Davis, M. (1983): Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44*, 113-126.

53. Marin, R. S. (1991). Apathy: a neuropsychiatric syndrome. *The Journal of neuropsychiatry and clinical neurosciences.*

54. Bagby, R. M., Parker, J. D., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia Scale—I. Item selection and cross-validation of the factor structure. *Journal of psychosomatic research, 38*(1), 23-32.

55. Langeslag, S. J., Gootjes, L., & van Strien, J. W. (2018). The effect of mouth opening in emotional faces on subjective experience and the early posterior negativity amplitude. *Brain and cognition, 127*, 51-59.

56. Bach, D. R., Friston, K. J., & Dolan, R. J. (2013). An improved algorithm for model-based analysis of evoked skin conductance responses. *Biological Psychology, 94*(3),

490–497. https://doi.org/10.1016/j.biopsycho.2013.09.010

57. Bach, D. R., Flandin, G., Friston, K. J., & Dolan, R. J. (2009). Time-series analysis for rapid event-related skin conductance responses. *Journal of Neuroscience Methods, 184*(2), 224–234.

58. Bach, D. R., Flandin, G., Friston, K. J., & Dolan, R. J. (2010). Modelling event-related skin conductance responses. *International Journal of Psychophysiology, 75*(3), 349–356.

59. R Core Team. (2019). A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.*

60. Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., … Bolker, M. Ben. (2015). Package 'lme4.' *Convergence, 12*(1), 2.

61. Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). *lmerTest: tests in linear mixed effects models. R package v. 2.0–30*.

62. Lenth, R. V. (2016). Least-squares means: the R package lsmeans. *Journal of statistical software, 69*(1), 1-33.

63. Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software, 80*(1), 1–28. https://doi.org/doi:10.18637/jss.v080.i01

64. Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A Probabilistic Programming Language. *Journal of Statistical Software, 76*(1). https://doi.org/10.18637/jss.v076.i01

65. Kumfor, F., Hazelton, J. L., De Winter, F.-L., de Langavant, L. C., & Van den Stock, J. (2017). Clinical Studies of Social Neuroscience: A Lesion Model Approach. In A. Ibáñez, L. Sedeño, & A. M. García (Eds.), Neuroscience and Social Science (pp. 255–296). Springer International Publishing. https://doi.org/10.1007/978-3-319-68421-5_12

66. Keane, J., Calder, A. J., Hodges, J. R., & Young, A. W. (2002). Face and emotion processing in frontal variant frontotemporal dementia. *Neuropsychologia, 40*(6), 655-665.

67. Van den Stock, J., De Winter, F.-L., de Gelder, B., Rangarajan, J. R., Cypers, G., Maes, F., Sunaert, S., Goffin, K., Vandenberghe, R., & Vandenbulcke, M. (2015). Impaired recognition of body expressions in the behavioral variant of frontotemporal dementia. *Neuropsychologia, 75*, 496–504. https://doi.org/10.1016/j.neuropsychologia.2015.06.035

68. Omar, R., Henley, S. M., Bartlett, J. W., Hailstone, J. C., Gordon, E., Sauter, D. A., ... & Warren, J. D. (2011). The structural neuroanatomy of music emotion recognition: evidence from frontotemporal lobar degeneration. *Neuroimage, 56*(3), 1814-1821.

69. Eckart, J. A., Sturm, V. E., Miller, B. L., & Levenson, R. W. (2012). Diminished disgust reactivity in behavioral variant frontotemporal dementia. *Neuropsychologia, 50*(5), 786-790.

70. Lavenu, I., Pasquier, F., Lebert, F., Petit, H., & Van derLinden, M. (1999). Perception of emotion in frontotemporal dementia and Alzheimer disease.AlzheimerDisease & Associated Disorders,13, 96–101.

71. Sturm, V., Hua, A., Lwi, S., Rankin, K., Rosen, H., Miller, B., ... & Seeley, W. (2015). Baseline autonomic dysfunction relates to cognitive empathy impairment in behavioral variant frontotemporal dementia (P1. 230).

72. 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

73. Matsumoto, D., & Hwang, H. C. (2014). Judgments of subtle facial expressions of emotion. *Emotion*, *14*(2), 349.

74. Motley, M. T., & Camden, C. T. (1988). Facial expression of emotion: A comparison of posed expressions versus spontaneous expressions in an interpersonal communication setting. *Western Journal of Communication (Includes Communication Reports)*, *52*(1), 1–22.

75. Critchley, H. D., Wiens, S., Rotshtein, P., Öhman, A., & Dolan, R. J. (2004). Neural systems supporting interoceptive awareness. *Nature neuroscience*, *7*(2), 189.

76. Allman, J., Hakeem, A., & Watson, K. (2002). Book review: Two phylogenetic specializations in the human brain. *The Neuroscientist*, *8*(4), 335-346.

77. Rushworth, M. F., & Behrens, T. E. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature neuroscience*, *11*(4), 389.

78. Mayer, E. A. (2011). Gut feelings: the emerging biology of gut–brain communication. *Nature Reviews Neuroscience*, *12*(8), 453.

79. Seeley, W. W., Crawford, R., Rascovsky, K., Kramer, J. H., Weiner, M., Miller, B. L., & Gorno-Tempini, M. L. (2008). Frontal paralimbic network atrophy in very mild behavioral variant frontotemporal dementia. Archives of neurology, 65(2), 249-255.

80. Kim, E. J., Sidhu, M., Gaus, S. E., Huang, E. J., Hof, P. R., Miller, B. L., ... & Seeley, W. W. (2012). Selective frontoinsular von Economo neuron and fork cell loss in early behavioral variant frontotemporal dementia. *Cerebral cortex*, 22(2), 251-259.

81. Mandelli, M. L., Vitali, P., Santos, M., Henry, M., Gola, K., Rosenberg, L., ... & Gorno-Tempini, M. L. (2016). Two insular regions are differentially involved in behavioral variant FTD and nonfluent/agrammatic variant PPA. *Cortex*, *74*, 149-157.

82. Binney, R. J., Pankov, A., Marx, G., He, X., McKenna, F., Staffaroni, A. M., ... & Rosen, H. J. (2017). Data-driven regions of interest for longitudinal change in three variants of frontotemporal lobar degeneration. *Brain and behavior*, *7*(4), e00675.

83. Zhou, J., Greicius, M. D., Gennatas, E. D., Growdon, M. E., Jang, J. Y., Rabinovici, G. D., ... & Seeley, W. W. (2010). Divergent network connectivity changes in behavioural variant frontotemporal dementia and Alzheimer's disease. *Brain*, *133*(5), 1352-1367.

84. Adolfi, F., Couto, B., Richter, F., Decety, J., Lopez, J., Sigman, M., ... & Ibáñez, A. (2017). Convergence of interoception, emotion, and social cognition: a twofold fMRI meta-analysis and lesion approach. *Cortex*, *88*, 124-142.

85. Joshi, A., Mendez, M. F., Kaiser, N., Jimenez, E., Mather, M., & Shapira, J. S. (2014). Skin conductance levels may reflect emotional blunting in behavioral variant frontotemporal dementia. *The Journal of neuropsychiatry and clinical neurosciences*, *26*(3), 227-232.

86. Fletcher, P. D., Downey, L. E., Golden, H. L., Clark, C. N., Slattery, C. F., Paterson, R. W., ... & Warren, J. D. (2015). Auditory hedonic phenotypes in dementia: A behavioural and neuroanatomical analysis. *Cortex*, *67*, 95-105.

87. Sturm, V. E., Sible, I. J., Datta, S., Hua, A. Y., Perry, D. C., Kramer, J. H., ... & Rosen, H. J. (2018). Resting parasympathetic dysfunction predicts prosocial helping deficits in behavioral variant frontotemporal dementia. *Cortex, 109*, 141-155.

88. Pietschmann, M., Endrass, T., Czerwon, B., & Kathmann, N. (2011). Aging, probabilistic learning and performance monitoring. *Biological psychology, 86*(1), 74-82.

89. Simon, J. R., Howard Jr, J. H., & Howard, D. V. (2011). Age differences in implicit learning of probabilistic unstructured sequences. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 66*(1), 32-38.

90. Bradley, M. M. (2009). Natural selective attention: Orienting and emotion. *Psychophysiology*, 46(1), 1–11. https://doi.org/10.1111/j.1469-8986.2008.00702.x

91. Geuter, S., Boll, S., Eippert, F., & Büchel, C. (2017). Functional dissociation of stimulus intensity encoding and predictive coding of pain in the insula. *ELife, 6*, 1–22. https://doi.org/10.7554/eLife.24770

92. Feldman, H., & Friston, K. (2010). Attention, uncertainty, and free-energy. Frontiers in human neuroscience, 4, 215.

93. Kumfor, F., Hutchings, R., Irish, M., Hodges, J. R., Rhodes, G., Palermo, R., & Piguet, O. (2015). Do I know you? Examining face and object memory in frontotemporal dementia. *Neuropsychologia, 71*, 101-111.

94. Virani, K., Jesso, S., Kertesz, A., Mitchell, D., & Finger, E. (2013). Functional neural correlates of emotional expression processing deficits in behavioural variant frontotemporal dementia. *Journal of psychiatry & neuroscience: JPN, 38*(3), 174.

95. Neary, D., Snowden, J. S., Gustafson, L., Passant, U., Stuss, D., Black, S. A., ... & Benson, D. F. (1998). Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria. *Neurology, 51*(6), 1546-1554.

96. Chow, T. W., Pio, F. J., & Rockwood, K. (2011). An international needs assessment of caregivers for frontotemporal dementia. *Canadian Journal of Neurological Sciences, 38*(5), 753-757.

97. Manoochehri, M., & Huey, E. D. (2012). Diagnosis and management of behavioral issues in frontotemporal dementia. *Current neurology and neuroscience reports, 12*(5), 528-536.

Supplementary Material

Patients (questionnaires were completed by the family caregivers, who were asked about their subjective impression of the patients) (*n* = *7*)	
FBI	31.29 (18.41)
SPF	35.29 (8.67)
empathic concern	10.57 (3.82)
fantasy	6.14 (1.46)
personal distress	12.28 (3.68)
assumption of perspectives	6.28 (2.87)
AES-I	33.00 (14.09)
TAS	2.91 (0.70)

Table S1. Additional clinical data for patients

FBI, Frontal Behavioral Inventory; SPF, Saarbruecken personality questionnaire; AES-I, informant-rated Apathy Evaluation Scale; TAS, Toronto Alexithymia Scale.

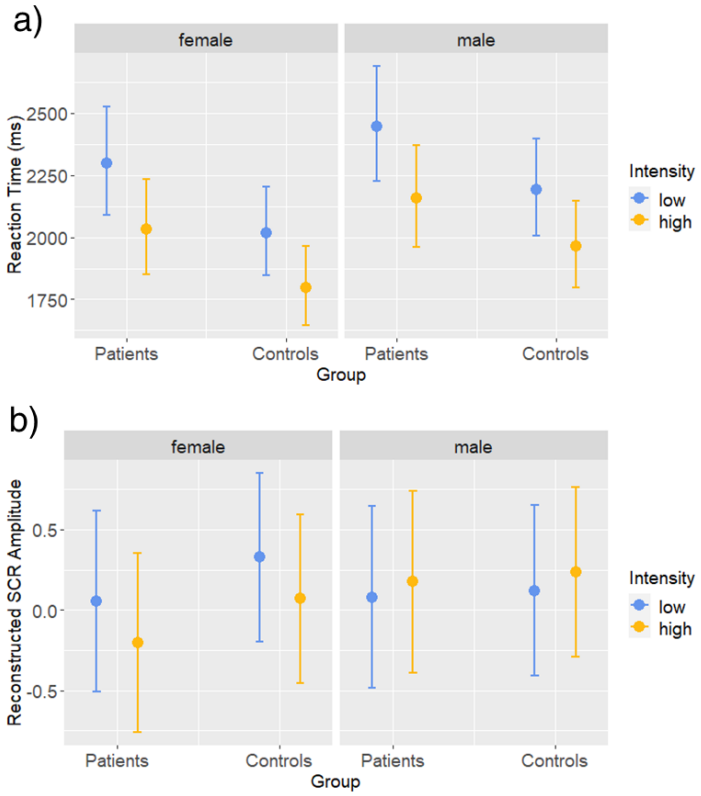

Figure S1. Effects of gender (female, male) and intensity (high, low) on a) reaction time and b) skin conductance response (SCR) in the non-emotion condition separated for patients with bvFTD and healthy controls.

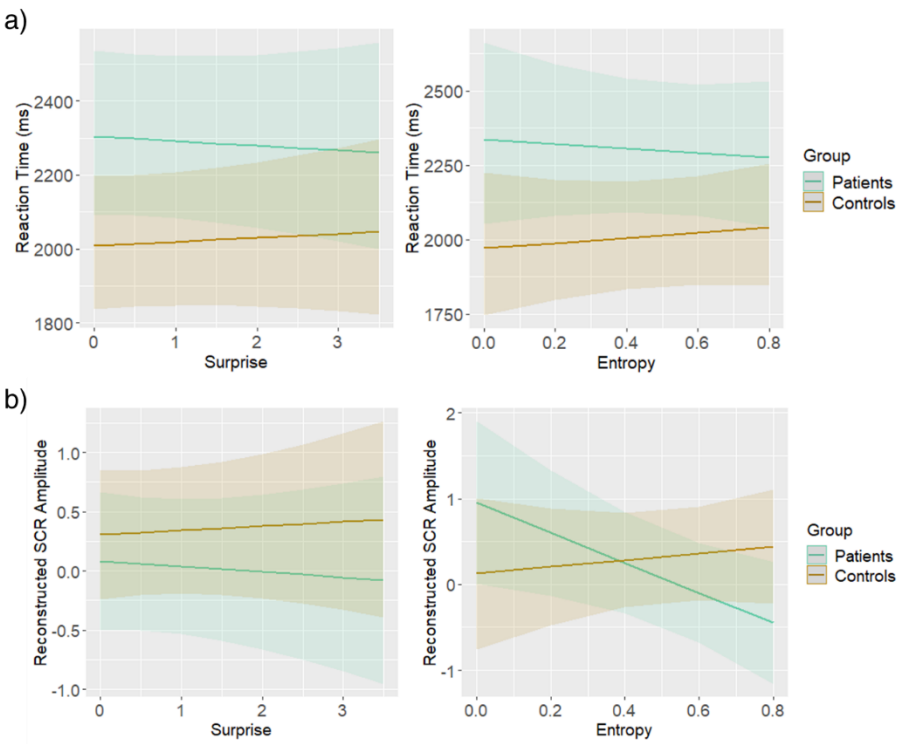

Figure S2. Effects of surprise and entropy on a) reaction time and b) skin conductance response (SCR) in the non-emotion condition separated for patients with bvFTD and healthy controls.

3.3 Study III: Interindividual Differences in Interoception Modulate Behavior and Brain Responses in Emotional Inference

Running title: Neural correlates of interoception

Hübner, A. M., Trempler, I., & Schubotz, R. I.

Manuscript submitted for publication

Interindividual differences in interoception modulate behavior and brain responses in emotional inference

Amelie M. Hübner[1], Ima Trempler[1,2], Ricarda I. Schubotz[1,2]

[1] Department of Psychology, University of Muenster, Germany

[2] Otto-Creutzfeldt-Center for Cognitive and Behavioral Neuroscience, University of Muenster, Germany

Running title: Neural correlates of interoception

Corresponding Author:
Amelie M. Hübner
Institute for Psychology
Fliednerstraße 21
D-48149 Münster
Phone: 0049-251-8334192

E-Mail: huebnera@uni-muenster.de

Abstract

Emotional experiences are proposed to arise from contextualized perception of bodily responses, also referred to as interoceptive inferences. The recognition of emotions benefits from adequate access to one's own interoceptive information. However, direct empirical evidence of interoceptive inferences and their neural basis is still lacking. The present fMRI study used a probabilistic emotion classification task with videotaped dynamically unfolding facial expressions. In a first step, we aimed to determine functional areas involved in the processing of dynamically unfolding emotional expressions. We then tested whether individuals with higher interoceptive accuracy (IAcc) or higher interoceptive sensitivity (IS) benefit more from the contextually given likelihood of emotional valence and whether brain regions reflecting individual IAcc and/or IS play a role in this. Individuals with higher IS benefitted more from the biased probability of emotional valence. Brain responses to more predictable emotions elicited a bilateral activity pattern comprising the inferior frontal gyrus and the posterior insula. Importantly, individual IAcc scores positively covaried with brain responses to more surprising and less predictable emotional expressions in the insula and caudate nucleus. The current results provide empirical evidence for the notion that IAcc and IS have different neural bases. We show for the first time that IAcc score is associated with enhanced processing of interoceptive prediction errors, particularly in the anterior insula. A higher IS score, in contrast, seems more likely to be associated with a stronger weighting of attention to interoceptive changes processed by the posterior insula and ventral prefrontal cortex.

Keywords: interoception, emotion recognition, emotional inference, predictive coding

1. Introduction

Interoception is defined as the sense of internal physiological states of the body (Craig, 2002) and plays an important role in better understanding of emotional processes, and related to this, the emergence of mental disorders (Farb et al., 2015; Paulus & Stein, 2006; Seth, 2013). As precursors to current theories of emotion processing, classical appraisal theories of emotions (James, 1884; Lange, 1885; Schachter & Singer, 1962) as well as the somatic marker hypothesis (Bechara, Damasio, & Damasio, 2000; Damasio, 1996) have postulated early on that emotional experiences arise from the contextualized perception and interpretation of bodily responses to external stimuli. Building on and extending these theories on emotion processing, the interoceptive predictive coding model assumes that emotional content is determined by interoceptive predictions, meaning actively inferred predictive models of the causes of bodily sensations based on past experiences (Seth, 2013; Seth & Friston, 2016). Interoceptive prediction errors, i.e., mismatches between descending interoceptive predictions and ascending primary interoceptive afferents, convey information about interoceptive changes and activate corresponding (sympathetic or parasympathetic) systems to restore physiological homeostasis or allostasis (Seth & Friston, 2016; Seth et al., 2012). The relative weight accorded to predictions and prediction errors at multiple levels of the predictive hierarchy is described in terms of *precision*, i.e. the reliability of (exteroceptive or interoceptive) signals (Friston, 2008, 2010). It has been proposed that individual differences in interoceptive precision also determine differences in the awareness of interoceptive sensations and, correspondingly, in emotional experience (Ainley et al., 2016).

Notably, interoception is of central importance not only for the perception of one's own emotional states, but also for those of other individuals (Ondobaka et al., 2017; Shah, Catmur, & Bird, 2017). Increased interoceptive abilities are linked to perceived arousal elicited by emotional stimuli (Dunn et al., 2010; Pollatos et al., 2005), higher sensitivity for emotional

facial expressions of others (Cook, Brewer, Shah, & Bird, 2013; Terasawa et al., 2014) and a facilitation of the recognition of emotional changes, accompanied by a more precise adaptation to emotion probabilities (Hübner, Trempler, Gietmann, & Schubotz, 2021). Low interoceptive abilities on the other hand have been related to alexithymia, a subclinical phenomenon characterized by difficulties in identifying and describing feelings and associated with impaired expression recognition (Brewer et al., 2016; Cook et al., 2013; Herbert, Herbert, & Pollatos, 2011; Sowden et al., 2016). Nevertheless, there are different restrictive definitions and still no consensus on the exact conception of interoception. In this paper, we refer to a distinction based on various measures of interoceptive abilities according to Garfinkel, Seth, Barrett, Suzuki, & Critchley, (2015), differentiating between interoceptive accuracy (IAcc), operationalized as accuracy in detection of the heartbeat, interoceptive sensibility (IS) as self-reported measure regarding the individual evaluation of the ability to perceive body signals, and interoceptive awareness (IAw), defined as a metacognitive component in the sense of the degree of concordance between IS and IAcc.

The neuroanatomical brain regions involved in interoceptive processing extend from viscero-motor areas (VMAs), through the insula cortex and the anterior cingulate cortex (ACC), to the orbitofrontal cortex (OFC), which is assumed to be situated at the top of the interoceptive hierarchy (Ainley et al., 2016; Barrett & Simmons, 2015; Garfinkel & Critchley, 2013; Seth & Friston, 2016). Regarding these interoceptive pathways, structurally and functionally closely connected insula and ACC are thought to be the most important cortical region activated by most interoceptive and affective stimuli (Ainley et al., 2016; Barrett & Simmons, 2015; Craig, 2009; Medford & Critchley, 2010; Menon & Uddin, 2010). Imaging studies provide evidence that these clusters are also involved in cerebral processing of emotions. For example, recent studies show overlaps of interoceptive signalling, emotion perception and social cognition, especially in the right anterior insular cortex (AIC) and frontotemporal regions (Pollatos et al., 2005; Zaki, Davis, & Ochsner, 2012; see Adolfi et al., 2017 for a meta-analysis). Consequently,

it can be assumed that increased cortical processing in these areas implies an advantage not only for the more accurate appraisal of one's own bodily signals, but that this better access to one's own interoceptive signals is also beneficial for recognizing the emotions of others.

In the present fMRI study, we aimed to investigate the impact of individuals' scoring in IAcc and IS on their brain activity and performance (reaction times) in a probabilistic emotion classification task. The experimental paradigm required participants to indicate whether a videotaped neutral facial expression develops into a happy or fearful expression. To introduce different levels of uncertainty, facial expressions varied in intensity at the end of the video. Moreover, we implemented different probabilities for the occurrence of happy or fearful faces per block to assess participants' propensity to efficiently update their predictive model, quantified by information-theoretic measures, i.e., Shannon *surprise* (i.e., the improbability of an event) and *entropy* (i.e., the inverse predictability of a single stimulus), respectively (Shannon, 1948; Strange, Duggins, Penny, Dolan, & Friston, 2005; see Methods for further details).

We tested whether individuals with higher IS or IAcc recognize others' emotions more easily and are more sensitive to learn from biased probabilities of emotional expressions. Here, it is important to emphasize that objective measures of IAcc are not necessarily correlated with subjective measures of IS (e.g., Ceunen, Van Diest, & Vlaeyen, 2013; Forkmann et al., 2016; Meessen et al., 2016). However, as little is yet known about the differential effects of interindividual differences in IS and IAcc on emotion recognition, we initially expected that both higher IAcc and higher IS would have a beneficial effect on emotion recognition and context adaptation. In particular, we expected individuals with higher IS and IAcc to show lower thresholds for emotion recognition, reflected in decreased reaction times for emotional expressions (Hypothesis 1, H1). In addition, they should benefit more from a biased probability of an emotion, which is reflected in decreased reaction times for expected emotions (Hypothesis 2, H2) (see also Hübner et al., 2021).

On the neural level, we firstly investigated functional areas specialized in the processing of dynamically unfolding emotional facial expressions depending on intensity and valence. We expected high vs. low intensity of emotional expressions (Hypothesis 3, H3) to engage brain regions involved in the perception of changeable facial attributes and recognition of facial expressions, i.e., the inferior frontal gyrus (IFG), fusiform gyrus, insula and amygdala (Calder et al., 2003; Furl et al., 2010; Gallese, Keysers, & Rizzolatti, 2004; Haxby et al., 2002; Jung et al., 2021). For fearful vs. happy unfolding facial expressions, we expected increased neural activation especially in the amygdala (Adolphs, Tranel, Damasio, & Damasio, 1994; Breiter et al., 1996; Calder, 1996; Fusar-Poli et al., 2009). Secondly, regarding the probabilistic context adaptation, we expected stimulus-bound *surprise* to be positively correlated with activity in regions involved in the signalling of (interoceptive) prediction errors, i.e., AIC, ACC and striatum (Schiffer & Schubotz, 2011; Seth, 2013) (Hypothesis 4, H4), and *entropy* to be positively correlated with activity in the hippocampus, caudate nucleus, ACC, AIC and orbitofrontal cortex (OFC) (Ahlheim, Stadler, & Schubotz, 2014; Bechara et al., 2000; Herry et al., 2007; Schiffer, Ahlheim, Wurm, & Schubotz, 2012; Strange et al., 2005) (Hypothesis 5, H5). Finally, we expected the degree of activation increase in brain regions related to interoceptive processing (i.e., insula and ACC) to be positively correlated with IS and IAcc (Hypothesis 6, H6).

2. Materials and Methods

2.1 Participants

Forty-two right-handed healthy young volunteers (M_{age} = 21.00 years, *SD* = 2.51, age range = 18-29 years, 34 women, 8 men) with normal or corrected-to-normal vision were included in the present study. None of them reported a history of neurological or psychiatric disorders and all participants met the criteria for MRI scanning. The study protocol was conducted in accordance with ethical standards of the Declaration of Helsinki and approved by the Local Ethics Committee of the University of Muenster. Each participant submitted a signed informed consent form and received either reimbursement or course credits for their participation afterwards. Individuals provided written informed consent to permit for all potentially identifying information to be published.

2.2 Stimulus Material

The emotional stimuli used were the same as reported in Hübner et al. (2021) and consisted of short videos with a mean duration of 3.00 s, displayed on a grey background. In order to avoid ceiling effects due to noticeable features like the hairline, in this task we additionally extended the grey background ellipsoid around the emotional faces. To create valid stimuli which are comparable to facial expressions that one encounters in daily life, we accepted that the emotional videos would slightly vary in length (*SD* = 0.39). However, to control for possible confounds of these variations on the participants' reaction time, we included the video durations as a covariate in our analyses. We created videos with female or male faces posing from neutral expressions to happy or fearful emotional ones. To introduce different levels of uncertainty, emotional facial expressions differed in terms of INTENSITY (high/low). Moreover, we controlled for potential effects of mouth opening on emotion recognition (Langeslag, Gootjes, & van Strien, 2018). To this end, each condition (fearful/happy expressions of

high/low intensity) was presented with the mouth open and closed, i.e., in two different versions. However, to reduce the complexity of the statistical model, we did not include these conditions in the statistical analyses. A detailed explanation of the generation of the stimuli used is described in Hübner et al. (2021).

2.3 Task

During the main task in the MR scanner, the videos were centrally back-projected onto a screen. Participants viewed the screen through a mirror above the head coil. Videos were presented separated by jittered inter-stimulus-intervals (ITI) of 3500, 4000, 4500 or 5000 msec, during which a fixation cross was displayed centrally on the screen. The participants were asked to watch the presented videos attentively and to respond as fast and accurately as possible as soon as they recognized the emotion in the emotional condition of a presented face. Participants responded by pressing one of two response keys with their right index and middle finger. Stimulus-response mappings were counterbalanced across participants. The task consisted of 16 emotional blocks with 16 consecutive videos each, i.e., a total of 256 trials. Thus, each of the 128 videos was shown twice. After each block, participants had a short break of 8 s, during which participants were given the information that the block has ended, followed by information regarding the start of a new block and instructions on key assignment. Accordingly, the overall task lasted about 33 min (2624 sec).

Within each block, either a high (75%) or low (25%) probability for the occurrence of fearful or happy faces was implemented. Stimuli were presented in a pseudo-randomized order ensuring that every each of the eight conditions (happy vs. fearful x high vs. low intensity x mouth open vs. closed) was shown twice in each block and videos of one actor/actress were never repeated across consecutive trials. Transitions between the block types were balanced throughout the experiment. The randomization was programmed using MATLAB R2019b (The MathWorks Inc., Natick, MA, USA) and stimuli were presented using the Presentation software

(Version 19.0, Neurobehavioral Systems, Inc., Berkeley, CA). A schematic diagram of the task is shown in Figure 1.

Prior to the experiment, participants performed a short training session to get accustomed to the task. The training consisted of one block of 16 trials each and an equal probability for the different conditions.

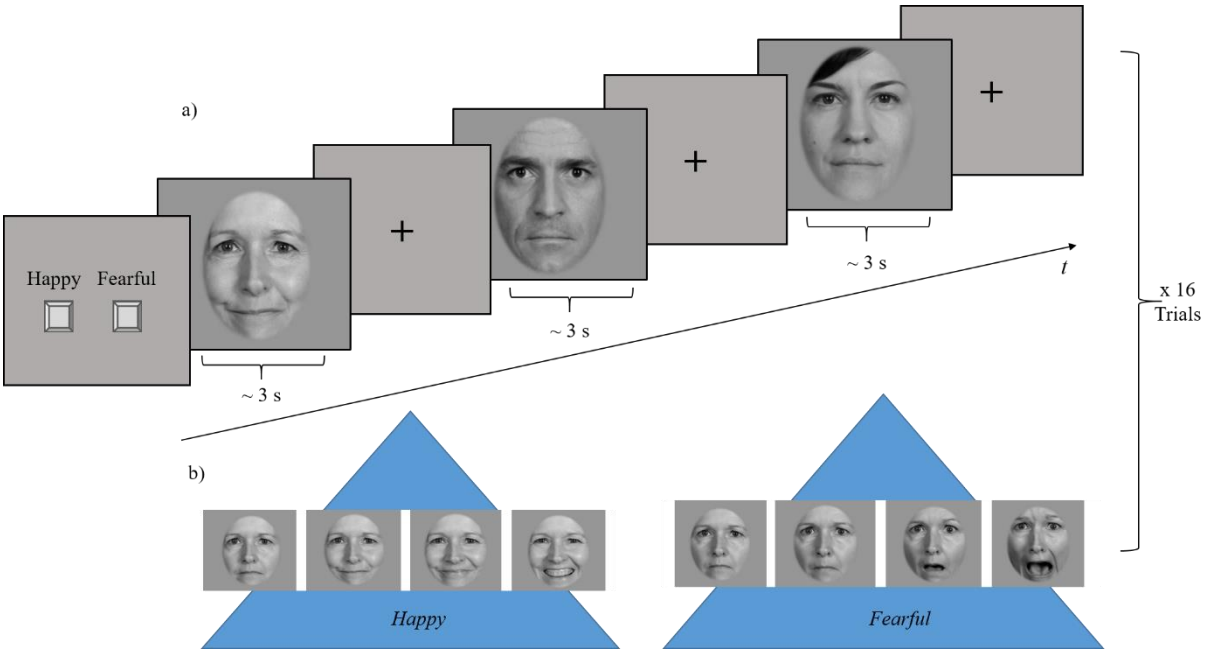

Figure 1. Schematic diagram of the task. **a)** At the beginning of each block, instructions indicate that the participant has to indicate the emotion of the upcoming faces (happy vs. fearful). One block includes 16 consecutive video trials. **c)** Videos depict actors/actresses posing from neutral to either happy or fearful facial expressions. The actor and actresses in the figure have given written informed consent to publish the photographs

2.4 Assessment of interoceptive sensibility (IS) and interoceptive accuracy (IAcc)

For the self-assessment of interoceptive sensibility (IS) we used the Multidimensional Assessment of Interoceptive Awareness, Version 2 (MAIA-2, Mehling et al., 2018). This questionnaire quantifies the self-reported belief of the perception of one's own bodily signals. The MAIA-2 is a state-trait questionnaire with 37 items consisting of the following eight subscales: noticing, i.e., the awareness of uncomfortable, neutral or comfortable body sensations; not-distracting, i.e., the tendency to ignore or distract oneself from sensations such as pain or discomfort; not-worrying, i.e., emotional distress or worry regarding sensations of

pain or discomfort; attention regulation, i.e., the ability to control attention to the sensations of the body; emotional awareness, i.e., the awareness of the connection between body sensations and emotional states; self-regulation, i.e., the perceived ability to regulate distress by attention to body sensations; body listening, i.e., how one's actively listens to the body for insight, and trust, i.e., the experiences of the body as safe and trustworthy (Mehling et al., 2012). Participants have to indicate how often each statement applies to them in everyday life on a 5-point Likert scale from 0 (*never*) to 5 (*always*). The validity of the MAIA-2 scales is indicated with Cronbach's alpha ranging from 0.64 to 0.83 for the eight scales (Mehling et al., 2018). For our statistical analyses, we calculated a total score per participant by summing up the results of the subscales. Together with other questionnaires assessing the participants' emotion processing, the MAIA-2 was completed online by the participants in the using the academic Unipark program (Questback GmbH, 2017).

For the assessment of interoceptive accuracy (IAcc) we used the Heartbeat detection task (HDT) as reported in Terasawa et al. (2014) based on a task developed by Schandry (1981) and Ehlers & Breuer (1992). The HDT task was assessed in a second session on average five days after the main MR session. During the HDT, the participants were asked to count their own heartbeat at different time intervals. Instructions were given on a screen using the presentation software (Version 19.0, Neurobehavioural Systems, Inc., Berkely, CA). The intervals were marked by a short start and end tone. The real heartbeats were recorded using a pulse oximeter and extracted using the Brain Vision Recorder software (Version 1.20.0801; Brain Products GmbH, 2018). The present study included a total of six intervals with a duration of 2 x 35sec, 2 x 25sec and 2 x 45sec (see Terasawa et al., 2014). For each trial, an accuracy score was derived using the following formula presented in Garfinkel et al. (2015):

$$1 - \left(\left|n\text{beats}_{real} - n\text{beats}_{reported}\right|\right) / \left(\left(n\text{beats}_{real} + n\text{beats}_{reported}\right)/2\right)$$

The resulting accuracy scores were averaged over the six trials, yielding an average value for each participant varying between zero and one, with higher values indicating higher interoceptive accuracy (Hartet al., 2013).

2.5 fMRI data acquisition

Whole-brain imaging data were collected with a 3-T Siemens Magnetom Prisma MR tomograph (Siemens, Erlangen, Germany) using a 20-channel head coil. To acquire structural data for each participant a 3D-multiplanar T1 weighted gradient-echo (MPRAGE) sequence was used. Scanning parameters were set to 192 slices, a 256 x 256 mm^2 field of view (FoV), a repetition time (TR) of 2.28 msec, an echo time (TE) of 2.28 msec and slice thickness of 1 mm.

Functional blood oxygenation level-dependent (BOLD) images were acquired using a T2*-weighted single-shot echo planar imaging (EPI) sequence (64 ×64 data acquisition matrix, 192 mm field of view (FoV), 90° flip angle, time of repetition (TR) = 2000 ms, echo time (TE) = 30 msec). Each volume consisted of 33 adjacent axial slices with a slice thickness of 3 mm and a gap of 1 mm, resulting in a voxel size of 3 ×3 × 4 mm. To provide a whole brain coverage the images were acquired in ascending order with scanning axis tilting 20 degrees up the AC-PC plane in order to lessen signal dropouts in orbitofrontal regions. In order to minimize head motion, the head was tightly fixated with cushions. Stimuli were projected on a screen positioned behind the scanner bore and presented in the center of the field of vision by a video projector. The subjects viewed the screen by a 45°mirror, which was fixated on the top of the head coil and adjusted for each participant to provide a good view of the entire screen.

2.6 Behavioral data analysis

Basic statistical analyses were performed using R, version 3.6.2 (R Core Team, 2019). Behavioral performance was measured in terms of reaction times (RTs). Incorrect or missing responses were excluded from analyses. For the comparisons of RTs, we tested whether

VALENCE, INTENSITY and information-theoretic quantities, i.e., Shannon's surprise $I(x_i)$ and entropy $H(X)$ (Shannon, 1948), could predict (variance in) RTs on a single-trial level. While 'surprise' measures the improbability of a particular event, 'entropy' measures the expected or average surprise over all events and thus reflects the predictability of a stimulus in a given context (Strange et al., 2005). For our analysis, we based Shannon's surprise on the frequency of a trial of a certain valence x_i normalized by the sum of all past trials in the block:

$$p(x_i) = \frac{n(x_i) + 1}{\sum x_t + 1}$$

For the two factor levels of valence (happy, fearful), prior counts before observing the first trial in a block were set to ½. The surprise $I(x_i)$ of each stimulus measures the amount of information provided by the current stimulus. This is given by the negative logarithm of the probability:

$$I(x_i) = -\ln p(x_i)$$

Entropy $H(X)$ is given by the average surprise of all possible outcomes. As illustrated in the following formula, it quantifies the expected information of an event regarding its predictability:

$$H(X) = \sum_i -p(x_i) \ln p_k(x_i)$$

We conducted generalized linear mixed-effects analyses using R, version 3.6.2 (R Core Team, 2019) via the package *lme4*, version 1.1.21 (Bates et al., 2015). We transformed RTs to the natural logarithm in order to achieve a more closely approximate to normal distribution. For the factors VALENCE and INTENSITY effect coding was used, with -1 for happy and 1 for fearful expressions, and 1 for high and -1 for low intensity expressions. While surprise and entropy were centered at individual levels, the MAIA-2 score and HDT score was centered at group level. Each model was fit with VALENCE and INTENSITY (and their interaction), Shannon's surprise and entropy and their respective interaction with the MAIA-2 score and the HDT score separately as fixed effects, and with a random intercept for each subject. As discussed above,

we added video duration to the statistic models to control for potentially confounding effects of varying video length on RTs. In accordance with our assumptions and our previous work, video length did not predict RTs, $b = 0.000$, $\beta = 0.000$, $t = 1.004$, $p = 0.315$. Statistical significance for each fixed effect was calculated via *lmerTest*, version 3.1.1 (Kuznetsova et al., 2016), using the Satterthwaite's approximation to denominator degrees of freedom. The significance level was set to $\alpha = .05$.

2.7 fMRI data preprocessing

All brain image preprocessing and basic statistical analyses were performed with SPM12 (http://www.fil.ion.ucl.ac.uk/spm/software/spm12/) and custom Matlab scripts (Version R2020b; The MathWorks Inc., Natick, MA, USA). To correct for differences in slice acquisition time, all functional images were slice-timed to the middle slice. All individual functional MR (EPI) images were realigned to correct for three-dimensional motion. Structural images were co-registered (rigid body transformation) to the mean functional image and then segmented into native space tissue components. Functional images were then normalised to the standard MNI space (Montreal Neurological Institute, Montreal, QC, Canada) and spatially smoothed with a Gaussian kernel of full width at half-maximum of $6mm^3$. Motion correction estimates were inspected visually as those participants who exceeded a maximum of 4 mm head movements between two scans in the x, y, and z dimensions would have been excluded from further analyses.

2.8 fMRI design specification

The statistical analysis of functional images was conducted using SPM12 implementing a general linear model (GLM) for serially autocorrelated observations (Friston et al., 1995; Worsley and Friston, 1995). To investigate functional areas specialized in the processing of emotion recognition depending on INTENSITY and VALENCE, a GLM was constructed including

four regressors coding for onsets and durations of the specific event types, i.e., happy high (hh), happy low (hl), fear high (fh) and fear low (fl), which were then convolved with the canonical hemodynamic response function (HRF) and regressed against the observed fMRI data. To model variability in the BOLD amplitude as a function of surprise and entropy, additional to each of the four regressors two parametric modulators each were added i.e., one for surprise and one entropy. Each of these modulators were mean-centered before entering the GLM. The subject-specific six rigid-body transformations obtained from realignment were included as regressors of no interest.

On the first level, one-sample *t*-tests for the following contrasts were calculated and submitted to a second-level *t*-test to detect functional areas specialized in the processing of emotion recognition on group level: intensity high (hh ∩ fh) > intensity low (hl ∩ fl), fearful (fh ∩ fl) > happy (hh ∩ hl), as well as the respective opposing contrast, i.e., intensity low (hl ∩ fl) > intensity high (hh ∩ fh), happy (hh ∩ hl) > fearful (fh ∩ fl), and parametric effects of surprise and entropy for all trials (fearful ∩ happy). For second-level whole-brain analyses, statistical differences were defined as $q < .05$ at the cluster level following False Discovery Rate (FDR) correction. Moreover, we defined regions of interest (ROIs) to correlate BOLD activation of the selected contrasts with differences in interoceptive abilities, i.e., interoceptive sensibility (IS, assessed by MAIA-2) and interoceptive accuracy (IAcc, assessed by HDT). We defined hypothesised ROIs of the anterior cingulate cortex (ACC) and insula, as well as caudate nucleus. All ROIs were derived from the automated anatomical labelling (AAL) atlas and created using the SPM Wake Forest University (WFU) Pickatlas toolbox (http://www.fmri.wfubmc.edu/cms/software, version 2.3) (Maldjian et al., 2003). We then used these ROIs to extract each subject's beta values of the activity associated with the defined contrasts during the task, i.e., VALENCE, INTENSITY, SURPRISE and ENTROPY. We performed correlation analyses using SPSS 26 (IBM, New York, USA) to examine the relationship

between BOLD activation of the selected contrasts and interoceptive abilities, with an α-level of $p < .05$, two-tailed.

3 Results

3.1 Behavioural results

Emotion recognition. As consistently reported in other studies, IS (i.e., continuously varying MAIA-2 total scores) and IAcc (i.e., individual HDT score) did not significantly correlate, $r = 0.001$, $p = 0.995$. In line with our hypothesis, the linear mixed-effect model predicting RTs revealed a significant main effect of valence, $b = 0.14$, $\beta = 0.007$, $t = 17.67$, $p < 0.001$, with increased RTs for fearful facial expressions compared to happy ones, and a main effect of INTENSITY, $b = -0.05$, $\beta = 0.007$, $t = -6.51$, $p < 0.001$, driven by increased RTs for low vs. high intensity of an expression. In contrast to Hübner et al. (2021), no significant interaction between VALENCE and INTENSITY was found, $b = -0.01$, $\beta = 0.01$, $t = -1.16$, $p = 0.247$.

Contrary to our hypothesis, the linear mixed-effect model predicting RTs revealed no significant main effects of individuals' IS, $b = 0.00$, $\beta = 0.001$, $t = 0.02$, $p = 0.986$, or IAcc, $b = 0.09$, $\beta = 0.065$, $t = 1.44$, $p = 0.15$. However, we found a significant interaction between VALENCE and IS, $b = 0.001$, $\beta = 0.004$, $t = 2.106$, $p = 0.03$, as well as between VALENCE and IAcc, $b = -0.085$, $\beta = 0.027$, $t = -3.085$, $p = 0.002$. In accordance with our hypothesis, with increasing IS, participants showed significantly longer RTs for fearful faces, but not for happy ones. In contrast to what we expected, with increasing IAcc participants showed significantly longer RTs for the recognition of happy faces but no difference for fearful ones (Figure 2a, 2b).

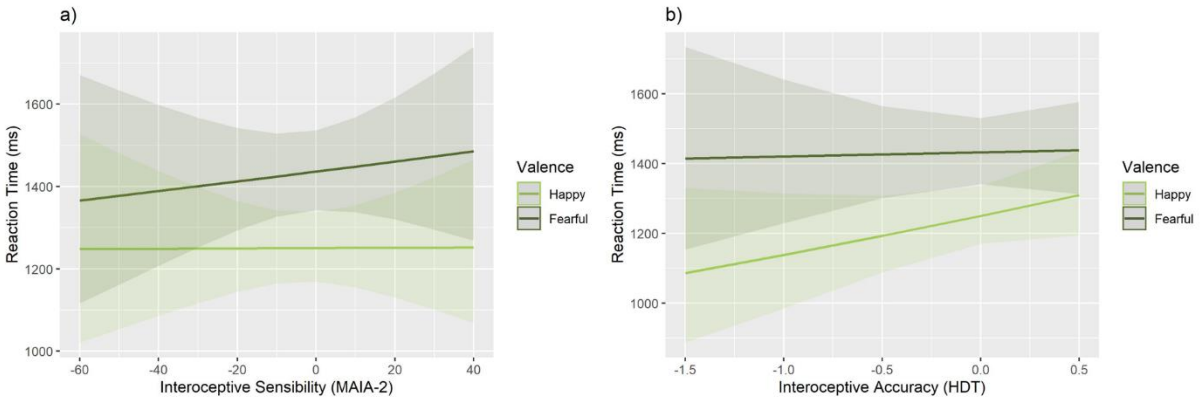

Figure 2. Significant effects of a) interoceptive sensibility assessed by the MAIA-2 questionnaire (Mehling et al., 2018) and b) interoceptive accuracy assessed by HDT-score, on reaction time divided for VALENCE (happy, fearful)

Probabilistic context adaptation. We tested whether Shannon's surprise and entropy of a single emotional expression in interaction with individual IS and IAcc were predictive of the participants' performance. The linear mixed-effect model predicting RTs revealed a significant main effect of surprise, b = -0.014, β = 0.004, t = -3.258, p = 0.001, and entropy, b = 0.048, β = 0.024, t = 2.003, p = 0.04. This effect was driven by increasing RTs with increasing entropy and decreasing RTs with increasing surprise. Regardless of differences in IS or IAcc, participants showed longer RTs to less predictable events, but shorter RTs to more surprising events, corresponding an expected positive correlation of entropy and RTs, as well as an unexpected negative correlation of surprise and RTs, respectively. Furthermore, we found a significant interaction between IS and entropy, b = 0.002, β = 0.001, t = 2.138, p = 0.032. As expected, only for participants with high IS, but not for those with low IS, entropy was positively correlated with RTs (Figure 3). Finally, there were no significant interactions for IS and surprise or IAcc and surprise or entropy (all p > 0.31).

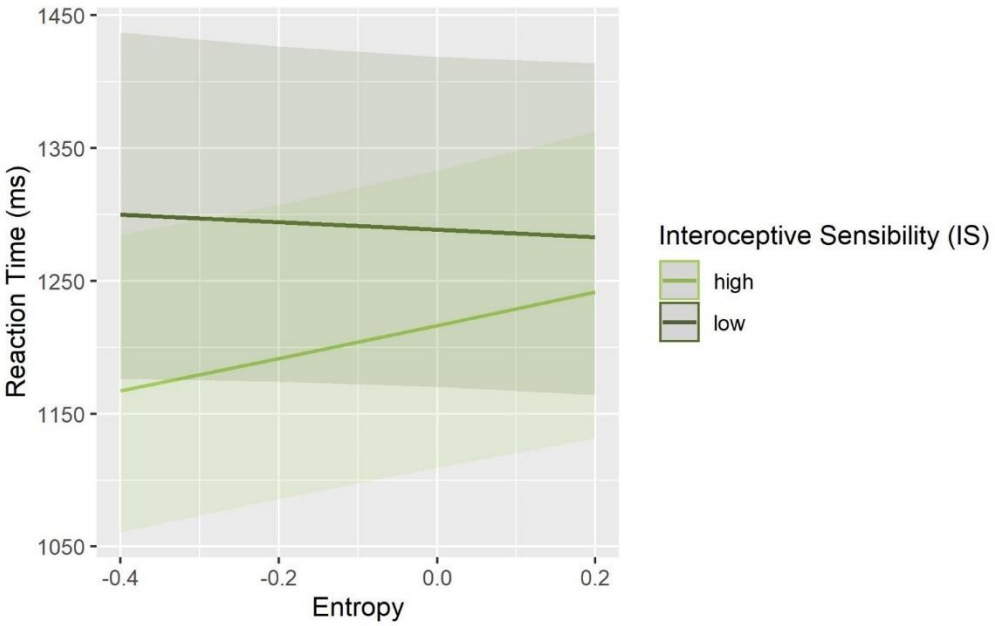

Figure 3. Results of the linear mixed effect regression model predicting reaction time (ms) by mean-centered entropy as a function of interoceptive sensibility (IS) assessed by the MAIA-2 questionnaire (Mehling et al., 2018). For illustrative purposes, we performed a median split to divide participants to two discrete groups with high and low IS. The solid lines depict the regression fit, and the shaded areas show the 95% confidence intervals.

3.2 fMRI results

Main effects of emotional valence and intensity. The contrast between high and low emotional intensity [(hh ∩ fh) > (hl ∩ fl)] revealed significant bilateral activation in the inferior occipital gyrus, as well as unilateral activation in the fusiform gyrus (Figure 4a, Table 1). For the reverse contrast [(hl ∩ fl) > (hh ∩ fh)] (Figure 4b), the analysis revealed higher frontal activations exclusively in the right hemisphere, including middle frontal gyrus, supplementary motor area, and anterior dorsal insula (Table 1). To identify the network associated with emotional valence, we contrasted fearful faces (fh ∩ fl) with happy faces (hh ∩ hl), which revealed significant bilateral activation in the temporal lobe, as well as the inferior frontal gyrus (IFG) and the anterior insula (Table 1, Figure 4c). No significant activation for the reverse contrast [(hh ∩ hl) > (fh ∩ fl)] was found.

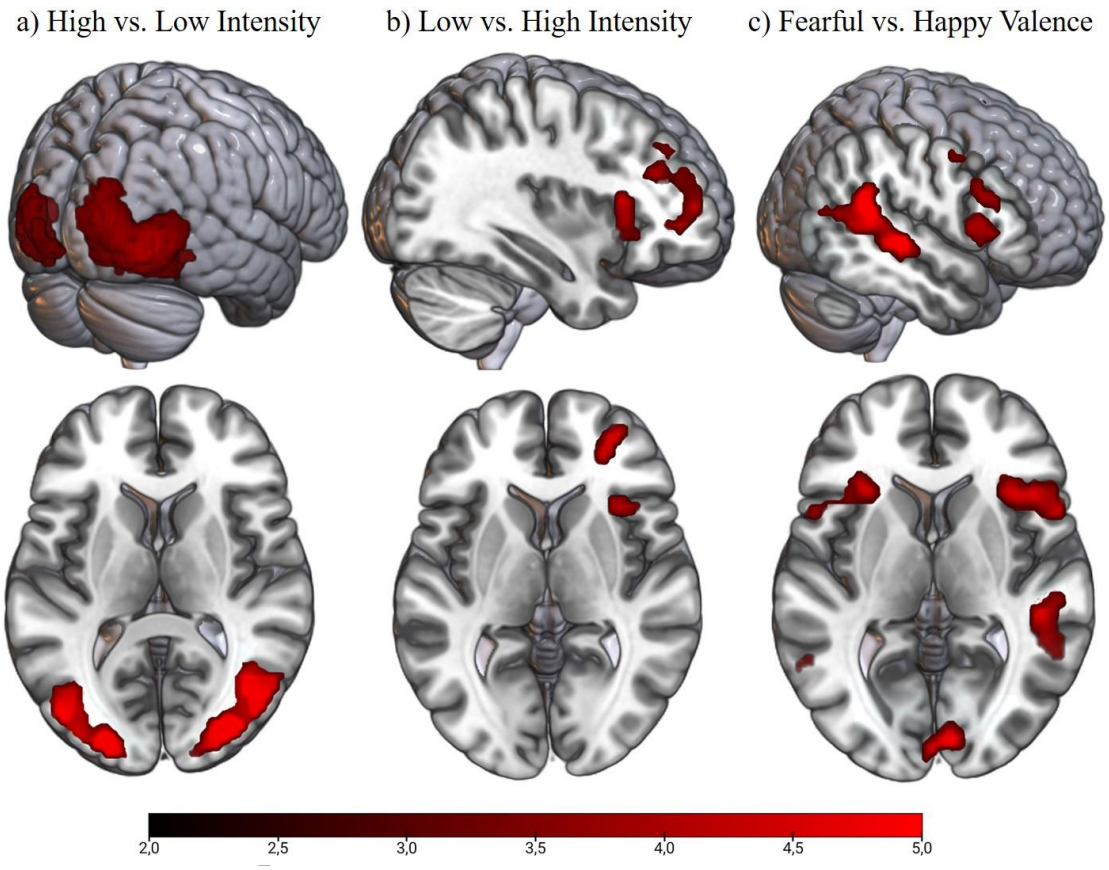

Figure 4. fMRI main effects at $p < .05$, whole-brain FDR cluster-corrected. (a) There was statistically significantly increased activation during high vs. low emotional intensity in the inferior occipital gyrus, as well as unilateral activation in the fusiform gyrus. (b) Low vs. high intensity of emotions elicited significant activation in frontal regions, comprising right-hemispheric activation of the right frontal gyrus, ACC and anterior insula. (c) Fearful vs. happy faces increased activation (amongst others) in the temporal cortex as well as the inferior frontal gyrus and anterior insula. Note: Since no significant activation was found, the reverse contrast happy > fearful faces is not displayed in the Figure.

Table 1. Maxima of activation from the conjunction analysis of the contrast images of Intensity high > Intensity low, Intensity low > Intensity high and Fearful > Happy at $p < 0.05$ cluster-level FDR-corrected. Labels are reported according to the AAL atlas.

Region Label	Cluster Extent	t-value	MNI Coordinates		
			x	y	z
Intensity high > Intensity low					
Right Inferior Occipital Gyrus	827	9.98	45	-70	2
Right Fusiform Gyrus		6.59	39	-58	-13
Left Inferior Occipital Gyrus	579	8.82	-45	-79	5
Left Middle Occipital Gyrus		7.36	-27	-94	14
Intensity low > Intensity high					
Right Middle Frontal Gyrus	295	4.93	30	50	8
Right Superior Frontal Gyrus	83	4.93	15	11	50
Right Anterior Cingulate Cortex		3.80	9	23	29
Right Supplementary Motor Area		3.67	9	23	53
Right Anterior Insula	84	4.63	36	17	11
Fearful > Happy					
Right Superior Temporal Gyrus	515	6.79	48	-43	17
Left Cerebelum	802	6.35	-6	-76	-31
Left Middle Temporal Gyrus	141	5.83	-57	-52	11
Left Anterior Insula	284	5.81	-27	23	2
Left Inferior Frontal Gyrus (p. Triangularis)		4.69	-45	23	20
Left Inferior Frontal Gyrus (p. Orbitalis)		4.59	-48	11	20
Right Anterior insula	508	5.49	36	23	-4
Right Inferior Frontal Gyrus (p. Orbitalis)		5.49	45	20	5
Right Precentral Gyrus		5.18	48	5	47
Left Precentral Gyrus	141	5.29	-33	-25	53
Right Supplementary Motor Area	225	5.01	6	17	53
Left Superior Parietal Lobule	73	4.32	-30	-52	41

Parametric effects of surprise and entropy. Corresponding with our hypothesis, we ran parametric analyses capturing regions more activated during more surprising events *(positive surprise)* as well as increased activation for events with decreasing predictability *(positive entropy)*. We further modelled the respective reverse contrasts, i.e., *negative surprise* and *negative entropy*. For negative surprise and positive entropy, however, no significant activation

could be observed. For more surprising events, we found significant unilateral activation in the left postcentral gyrus (Table 2, Figure 5b). The brain response to more predictable events (i.e., *negative entropy*) showed a bilateral activity pattern comprising - amongst others - the inferior frontal gyrus (IFG), posterior insula and middle temporal gyrus (Table 2, Figure 5a).

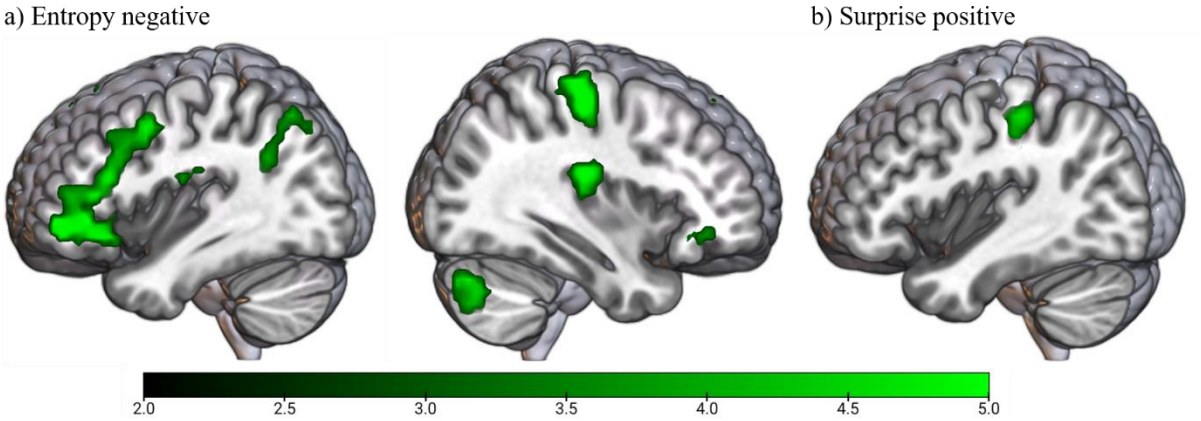

Figure 5. fMRI results at $p < 0.05$, whole-brain FDR cluster-corrected for the parametric effects of surprise and entropy. (A) Decreasing entropy (i.e., *negative entropy*) elicited significant bilateral activity pattern comprising - amongst others - the IFG, the left posterior insula as well as the middle temporal gyrus. (B) Shown is the statistically significant increased activation with increasing surprise of an emotion (i.e., *positive surprise*), with unilateral activation in the postcentral gyrus.

Table 2. Maxima of activation from the conjunction analysis of the parametric effect images of *surprise positive* and *entropy negative* at $p < 0.05$ cluster-level FDR-corrected. Labels are reported according to the AAL atlas.

Region Label	**Cluster Extent**	**t-value**	**MNI Coordinates**		
			x	**y**	**z**
Surprise positive					
Left Postcentral Gyrus	173	7.17	-33	-28	44
Left Superior Parietal Lobule	57	6.47	-30	-46	56
Entropy negative					
Left Middle Frontal Gyrus	1797	7.17	-42	50	-7
Left IFG (p. Orbitalis)		6.84	-45	26	-10
Left IFG (p. Triangularis)		6.47	-42	38	-4
Right Precentral Gyrus	207	6.13	36	-22	53
Right Posterior insula	313	5.77	36	-19	20
Right Middle Temporal Gyrus		5.61	66	-31	-1
Left Middle Temporal Gyrus	131	5.36	-63	-37	-7
Left Medial Frontal Cortex	48	5.14	-6	47	-13
Left Posterior insula	36	5.09	-33	-28	23
Left Angular Gyrus	175	5.05	-48	-64	38
Right Cerebelum Crus 1	208	4.95	33	-76	-31
Right Postcentral Gyrus	68	4.57	57	-7	41
Right IFG (p. Orbitalis)	49	4.57	42	32	-10
Right IFG (p. Triangularis)		4.13	51	41	-7
Right Precuneus	42	3.89	6	-58	35

*ROI activation relating to interoceptive sensibility (IS) and interoceptive accuracy (IAcc).*We correlated each participant's MAIA-2 total score and HDT score with the extracted beta values of the defined contrasts for the following regions of interest (ROIs): ACC and insula for all contrasts; caudate nucleus in addition for the parametric effects of surprise and entropy, as this region has repeatably found to be activated by perceptual prediction errors (Schiffer & Schubotz, 2011; Schiffer et al., 2012). For the MAIA-2 score as a measurement of IS, no significant correlations could be found (all $p > 0.294$). For an explorative analysis and to offer a basis for further research in the area, a table with significant correlations of the eight MAIA-2 subscales with the extracted beta weights is provided in the Supplementary material (Table S1).

Regarding differences in interoceptive accuracy (IAcc), parameter estimates of the right ACC activity for the contrast VALENCE (Fearful vs. Happy) marginal significantly correlated with the HDT score of the participants (r = 0.282, p = 0.071, two-tailed). Activity of ACC or insula for the contrast of INTENSITY were not related to differences in the HDT (all p > 0.81). For the parametric effects of entropy, parameter estimates of left (r =.330, p =.033, two-tailed) and right caudate activity (r = 0.362, p = 0.018, two-tailed) as well as (marginal) left insula activity (r = 0.299, p = 0.054, two-tailed) significantly correlated with the HDT scores, i.e., individuals with greater IAcc showed significantly higher activation in caudate and tendentially also in left insula for the occurrence of less predictable events (Figure 6a, 6b). Consistent with these findings, beta weights of ROI activity in left caudate for the surprise contrast were also significantly correlated to interoceptive accuracy (r = 0.356, p = 0.021; right caudate r = 0.285, p = 0.067, two-tailed) (Figure 6c).

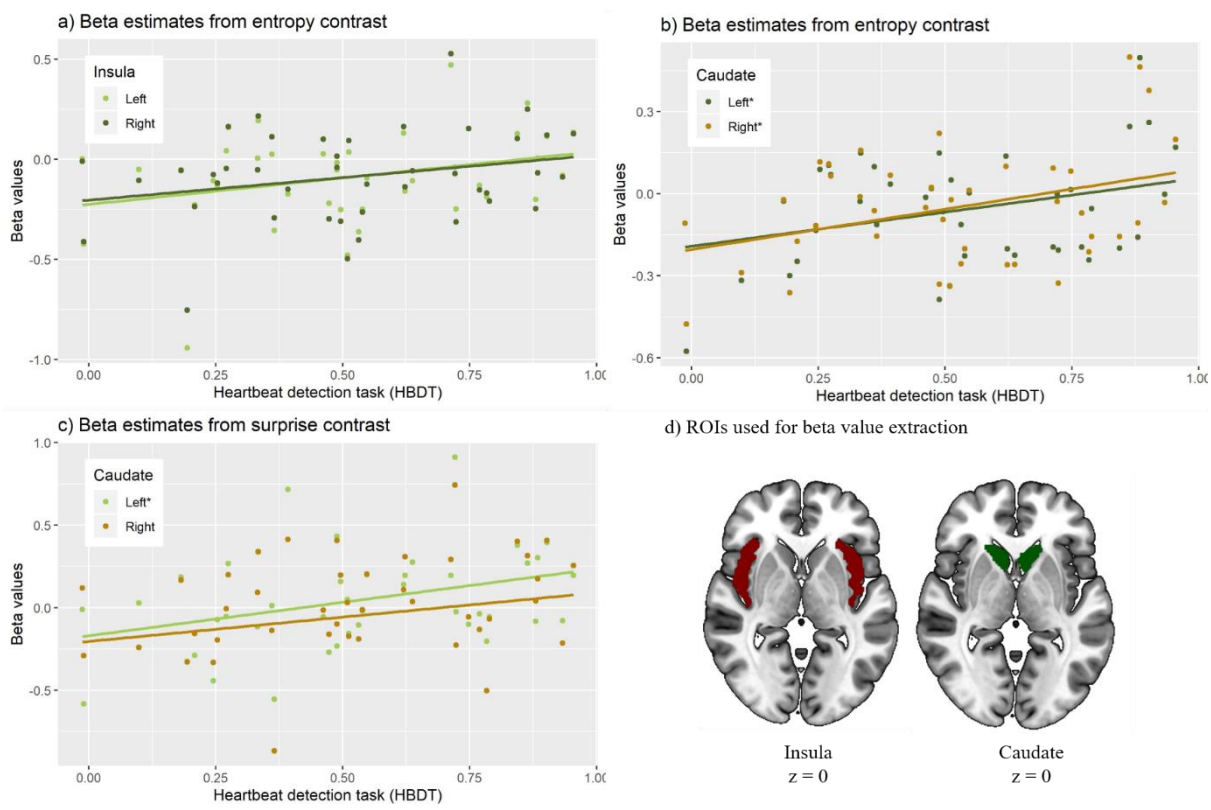

Figure 6. Region of interest (ROI) fMRI data and beta value extraction for entropy and surprise contrast. a) Activation for differences of the predictability of emotions (entropy) scaled with

participants' interoceptive accuracy (IAcc) as indicated by HDT score. Left (r = 0.330, p = 0.033, two-tailed) and right caudate activity (r = 0.362, p = 0.018, two-tailed) significantly correlated with the HDT scores. b) Marginal significant correlation of HDT and parameter estimates of the left insula activity (r = 0.299, p = 0.054, two-tailed) for beta estimates of the entropy contrast. c) Significant correlation of beta estimates of the left caudate in the surprise contrast with interoceptive accuracy (r = 0.356, p = 0.021; right caudate r = 0.285, p = 0.067, two-tailed). d) Insula and caudate ROI used for beta value extraction.

4. Discussion

In the current fMRI study, we investigated whether an individual's interoceptive sensibility (IS) and interoceptive accuracy (IAcc) are related to the speed of recognizing emotional changes in facial expressions of others, and if so, whether this relation depends on the valence, intensity and current probability of emotions. Furthermore, we investigated the neural basis of dynamically unfolding emotional expressions and interoceptive abilities.

Individuals' IS and IAcc had no effect on the speed of detecting the emergence of emotional expression, but interestingly, these parameters differed with regard to their impact on recognizing emotional valence. On the neural level, both valence and intensity of emotion had distinct effects on different brain sites including, among others, the anterior insula cortex (AIC) and the frontal lobes. Participants with higher IS recognized facial emotions faster when they were statistically more likely, whereas no such effects emerged for IAcc. Independent of IS and IAcc, more predictable emotions elicited bilateral brain activity in the IFG and the posterior insula. Finally, as expected, participants' IAcc correlated positively with activation in the caudate nucleus and insula when confronted with less predicted or less predictable emotions.

4.1. Recognizing emotional changes with different interoceptive abilities: the effect of valence and intensity

Our behavioral results partly confirmed the hypothesis that IS and IAcc modulate the recognition of emotional changes in facial expressions of others. Contrary to Hübner et al., (2021), higher IS scores were not associated with lower thresholds for emotional change

recognition, and individuals with lower IS did not benefit more from high intensities of facial emotions. In the present study, we added IAcc as an objective measure of interoceptive abilities but could not find any corresponding evidence for a generally lower detection threshold or interaction with intensity of emotions either.

Regarding the recognition of emotions with different valences, happy faces were recognized faster than fearful faces. This is in line with consistent evidence for superior recognition of happy faces compared to other facial expressions (Ekman, 1982; Hess, Blairy, & Kleck, 1997; Kirita & Endo, 1995; Leppänen & Hietanen, 2003; Palermo, 2004), and may be due to the fact that happiness requires simpler physical changes and occurs more frequently in daily life (Adolphs, 2002; Leppänen & Hietanen, 2003; Montagne, Kessels, De Haan, & Perrett, 2007).

Interestingly, differences in IS and IAcc interacted with emotional valence, indicating that individuals with high vs. low IS showed longer RTs for fearful faces, but not for happy ones, whereas with high vs. low IAcc individuals showed increased RTs for happy ones. Previous studies suggested that interoceptive abilities facilitate the sensitivity to recognize others' emotions (Hübner et al., 2021; Terasawa et al., 2014). Yet, there is evidence suggesting that individuals with higher IAcc are less accurate in recognizing negative emotions, which has been explained by the fact that with increasing IAcc, emotions are more regulated (Georgiou, Mai, Fernandez, & Pollatos, 2018). Accordingly, the ability to regulate interoceptive signals by weighting of attention might lead to interindividual differences in emotion recognition. This could also serve as an explanation for the observed contrasting interactions between IS and IAcc and emotional valence, as will be discussed further below.

As hypothesised, high vs. low intensity of emotional expressions revealed significant activation in the inferior occipital gyri and in the right fusiform gyrus. The processing of emotional face features is proposed to involve pathways from inferior occipital gyrus to fusiform gyrus, comprising the fusiform face area (FFA), which is known as a central region in

the face-selective cortical network (Goffaux, Schiltz, Mur, & Goebel, 2013; Grill-Spector, Knouf, & Kanwisher, 2004; Richler & Gauthier, 2014). Thereby, the inferior occipital gyrus is proposed to be involved in early visual processing and basic perception of facial features, whereas the fusiform gyrus processes the identity of the faces (Haxby, Hoffman, & Gobbini, 2000). Recently, the fusiform gyrus was found to show enhanced activation during emotionally arousing stimuli (Frank, Costa, Averbeck, & Sabatinelli, 2019) and fusiform volume was found to be positively correlated with performance in emotional intensity recognition (Jung et al., 2021). Moreover, Sato et al. (2004) found increased activation in the inferior occipital gyrus and fusiform gyrus for dynamical compared to static emotional facial expressions. In line with these studies, we suggest that the observed increased activity for high vs. low intensity of emotional expressions is related to enhanced perceptual processing for dynamic characteristics of faces.

In contrast, low vs. high emotional intensity elicited higher activation in the right middle frontal gyrus, right ACC, right supplementary motor area and anterior dorsal insula. It can be assumed that the detection of low- compared to high-intensity emotions is more difficult and requires a more detailed evaluation. The dorsolateral prefrontal cortex (dlPFC) is proposed to play a key role in attentional processing of emotional information (Mondino, Thiffault, & Fecteau, 2015) and to contribute to the distinction of the specific emotion that is expressed (Tsuchida & Fellows, 2012). Impairment of the dlPFC and ACC have been associated with deficits in interpersonal perception (Mah, Arnold, & Grafman, 2004), for example becoming evident in patients suffering from behavioral variant frontotemporal dementia (bvFTD) (Hübner et al., 2022; Keane, Calder, Hodges, & Young, 2002). The observed increased activation in the anterior dorsal insula also supports the assumption of more profound evaluative processing of emotions with low intensity, as this area plays a crucial role in the integration of affective and evaluative processes (Berntson et al., 2011). More specifically, the AIC has been

linked to emotion recognition by modulating arousal potentially mediated through interoceptive inferences (Motomura et al., 2019).

Correspondingly, we also observed increased AIC and IFG activation for fearful vs. happy faces. As fearful compared to happy faces are comparatively difficult to detect, as also suggested by our behavioral findings, this activation increase might reflect higher cognitive effort for the recognition of fearful expressions. Neuroimaging studies show IFG involvement in response inhibition in go/no-go tasks (Aron, Robbins, & Poldrack, 2004; Logan & Cowan, 1984). Especially right IFG activation is therefore believed to indicate the suppression of response tendencies (Menon, Adleman, White, Glover, & Reiss, 2001; Rubia, Smith, Brammer, & Taylor, 2003) and attentional control for the detection of important cues (Hampshire, Chamberlain, Monti, Duncan, & Owen, 2010). Moreover, evidence suggests that the right IFG processes emotional communicative signals in facial expression of others (Nakamura et al., 1999) and that impairment of the IFG is associated with alexithymia (Hobson et al., 2018). Accordingly, these findings indicate that the observed neuronal activity for fearful vs. happy faces primarily reflects an activation increase due to the more difficult detection and evaluation of physical changes in the development of fearful expressions, rather than a stronger fear response as observed in fear conditioning mediated by the amygdala (LeDoux, 2003).

4.2. Context-sensitive emotional inference and connections to interoceptive sensibility and accuracy

Another part of our results confirmed our hypothesis that the subjective individual evaluation of the ability to perceive body signals (*interoceptive sensibility*, IS) is associated with a more precise adaptation of emotional predictions. In line with the results observed in Hübner et al. (2021), RTs of individuals with higher IS decreased with increasing predictability of emotions, i.e., lower entropy of an emotion occurrence, whereas low IS individuals showed the opposite effects, namely decreased RTs to unpredictable emotion. These findings suggest

that higher IS individuals benefited from the implemented probabilistic context more than individuals with lower IS. We attribute these results to stronger attentional precision-weighting of individuals with higher IS, which might lead them to learn better from biased probabilities of emotional expressions and thus build stronger expectations (Ainley et al., 2016; Hübner et al., 2021). Correspondingly, using EEG recordings, studies show that interoceptive attention increases cortical processing of heartbeats (Petzschner et al., 2019) and attention independent of emotion processing accelerates the processing of prediction errors (Jiang, Summerfield, & Egner, 2013). Valence-specific affective predictions of expected or unexpected facial expressions have been shown to reduce or enhance cortical responses respectively to interoceptive signals as a function of the probability of repetition (Gentsch, Sel, Marshall, & Schütz-Bosbach, 2019). Moreover, when subjects' attention was tuned to their bodily signals by using a HDT, their following conditioned responding was found to be facilitated (Raes & De Raedt, 2011), and individuals with higher IAcc are better in predicting shocks (Katkin, Wiens, & Öhman, 2001).

However, the question remains why interindividual differences in IS and not IAcc were associated with a faster recognition of predicted emotions and to hesitation when unexpected emotions occurred. The MAIA-2 dimensions are constructed as elements of conscious processes of interoception, accessible to self-report (Mehling et al., 2012). Thereby, the MAIA-2 focuses on regulatory aspects of interoception, i.e., the tendency to ignore or distract oneself from perceived sensations, the ability to sustain and control attention to body sensation or the ability to regulate psychological distress by attention to body sensations (Mehling, Acree, Stewart, Silas, & Jones, 2018). Accordingly, high values in IS represent the perceived awareness of emotional states, but also how easily attention to these sensations can be weighted, sustained or regulated. The assessment of IAcc through the HDT provides a quantification for the objective ability of accurately detecting interoceptive signals (i.e., the heartbeat) which may or may not go hand in hand with better attention regulation to these. In our paradigm, to sustain

and control attention towards interoceptive changes might be the key ability to raise the precision of interoceptive predictions. Thus, particularly higher IS might result in more frequent and more precisely updated interoceptive predictions.

Regarding the probabilistic context adaptation on the neural level, we rather expected more surprising and less predictable emotions to activate regions involved in the signalling of interoceptive prediction errors i.e., AIC, ACC, hippocampus or caudate nucleus (Ahlheim et al., 2014; Bechara et al., 2000; Herry et al., 2007; Schiffer et al., 2012; Seth, 2013; Strange et al., 2005). However, for more surprising expressions only a single activation in the left postcentral gyrus increased, presumably reflecting simply a more abrupt lifting of the response finger in reaction to more surprising stimuli.

Instead, we found significantly increased activation for more predictable events (*negative entropy*) with a bilateral activity pattern comprising the IFG, posterior insula and middle temporal gyrus. In general, such prediction enhancement effects are proposed to reflect increased activation of predicted elements (Press, Kok, & Yon, 2020; Press & Yon, 2019). Interoceptive processing in the insula is proposed to involve posterior-to-anterior regions, with posterior regions supporting primary objective physical features of interoceptive information, whereas the AIC serves the subjective integration of interoceptive and motivational signals (Craig, 2009; Gu et al., 2013; Seth, 2013b; Seth et al., 2012). For example, the processing of cool temperatures was associated with activation in the posterior insula, while subjective ratings of these stimuli were processed in the AIC (Craig, Chen, Bandy, & Reiman, 2000). Moreover, the latter region has been suggested to respond to interoceptive prediction errors (Paulus & Stein, 2006; Singer, Critchley, & Preuschoff, 2009), which is why we expected AIC rather for *positive entropy*, i.e., decreased predictability.

Our results support the presented theoretical assumptions, revealing an activation increase for more predictable emotional stimuli exclusively in the posterior part of the insula. As this area is proposed to provide primary interoceptive representation of the physiological

condition of the body, it can be derived that if an emotional valence in facial expression of others is perceived over several trials (i.e., accompanied by increased predictability), physiological representations of the body are created and retained. Thus, these physiological representations could lead to higher activity in the corresponding areas (i.e., the posterior insula) and, moreover, influence the perception of subsequent information by interoceptive predictions. In accordance, we observed increased IFG activation for more predictable facial expressions. Previous research showed increased stimulus-evoked activity in the right IFG and right temporal-parietal junction to be associated with better performance on the next trial, suggesting these regions to be involved in stimulus-triggered reorienting of attention toward behaviorally relevant stimuli (Craig, 2009; Weissman, Roberts, Visscher, & Woldorff, 2006). These findings are consistent with our behavioral results and our assumption that the ability to sustain and control attention towards interoceptive changes might be crucial to raise the precision of interoceptive predictions.

4.3. Interoceptive abilities and their neuronal representations

Finally, we aimed to test more specifically whether the extent of activation increase in brain regions subserving interoceptive processing, i.e., insula and ACC, positively correlates with IS or IAcc. For the parametric effects of entropy and surprise we further included the caudate nucleus as a ROI, as this area was often reported for perceptual prediction errors (Schiffer & Schubotz, 2011; Schiffer et al., 2012). We expected individuals with higher IS or IAcc to build stronger interoceptive predictions and hence show increased activation for more surprising and less predictable events in regions which process perceptual or interoceptive prediction errors. Indeed, the individual IAcc score was positively correlated with the caudate response to entropy and surprise, and with the insula response to entropy. Hence, there was a higher BOLD activation in these regions for the occurrence of less predictable or more

surprising events, or both, in individuals with higher vs. lower IAcc scores. No such correlations were observed for the IS.

Imaging studies found significant activation of the insula, ACC and IFG during different variations of the HDT (see Schulz, 2016 for a meta-analysis). Interoceptive attention has been associated with increased AIC activation (Chong, Ng, Lee, & Zhou, 2017; Wang et al., 2019) and interoceptive awareness was shown to change intra-insula signal flows (Kuehn, Mueller, Lohmann, & Schuetz-Bosbach, 2016). Moreover, IAcc has been found to positively covary with activation in the insula cortex, cingulate gyrus and somatosensory cortex (Critchley et al., 2004; Pollatos et al., 2007; Wang et al., 2019).

In contrast, individual IS scores did not correlate with beta estimates of ACC, insula or caudate nucleus. Indeed, little is known about brain areas that may reflect individual IS traits. Recently, Smith et al. (2021) could not find any correlation between the functional connectivity of the insula with individuals' MAIA scores, and hence suggested that brain areas reflecting the individual IS score show only minor overlap with regions reflecting the individual IAcc score. Moreover, Wang et al. (2020) even found a higher IS to be associated with decreased connectivity of AIC and ACC to temporal and frontal regions. These findings are also consistent with our results indicating that higher IAcc, but not IS, positively correlates with activation in brain regions processing perceptual and interoceptive prediction errors, i.e., insula and caudate nucleus. Since the behavioral effects related to the recognition of different emotional valences and the adaptation of emotional predictions differed between IS and IAcc, it is not surprising that different brain regions are also associated with the individual IS and IAcc values. A stronger relationship between IAcc and AIC and striatal activation might indicate stronger processing of interoceptive prediction errors in highly IAcc individuals. Higher IS might rather be associated with the ability to sustain and control attention to interoceptive changes, which might be processed, inter alia, in the posterior insula and the IFG (as the posterior insula provides the primary interoceptive representation of the body’s physiological state and the IFG

is involved in reorienting of attention toward relevant stimuli, see section 4.2). These assumptions still need to be investigated in further studies. Nevertheless, our study is the first to provide new insights and direct empirical investigations of the different interoceptive facets, their impact on interoceptive predictions and neuronal representations.

4.4 Conclusion

Our results provide new insights into how interoceptive accuracy and sensibility shape on emotion recognition and interoceptive predictions at the behavioral and the brain level. In particular, higher IS was shown to involve more precise adaptation to biased probabilities of emotion. Moreover, BOLD effects indicate a different neural basis of interoceptive accuracy and sensibility. Future studies can build on these results and expand knowledge about different interoceptive abilities and their neural implementations, e.g., to provide a basis for the diagnosis and treatment of psychiatric disorders based on interoception deficits.

Acknowledgement

We sincerely thank Monika Mertens for the assistance in data acquisition. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Adolfi, F., Couto, B., Richter, F., Decety, J., Lopez, J., Sigman, M., … Ibáñez, A. (2017). Convergence of interoception, emotion, and social cognition: A twofold fMRI meta-analysis and lesion approach. *Cortex*, *88*, 124–142. https://doi.org/10.1016/j.cortex.2016.12.019

Adolphs, R. (2002). Recognizing emotion from facial expressions: psychological and neurological mechanisms. *Behavioral and Cognitive Neuroscience Reviews*, *1*(1), 21–62.

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature*, *372*(6507), 669–672.

Ahlheim, C., Stadler, W., & Schubotz, R. I. (2014). Dissociating dynamic probability and predictability in observed actions-an fMRI study. *Frontiers in Human Neuroscience*, *8*(MAY), 1–13. https://doi.org/10.3389/fnhum.2014.00273

Ainley, V., Apps, M. A. J., Fotopoulou, A., & Tsakiris, M. (2016). "Bodily precision": A predictive coding account of individual differences in interoceptive accuracy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1708). https://doi.org/10.1098/rstb.2016.0003

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences*, *8*(4), 170–177.

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

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, *10*(3), 295–307.

Berntson, G. G., Norman, G. J., Bechara, A., Bruss, J., Tranel, D., & Cacioppo, J. T. (2011). The insula and evaluative processes. *Psychological Science*, *22*(1), 80–86. https://doi.org/10.1177/0956797610391097

Breiter, H. C., Etcoff, N. L., Whalen, P. J., Kennedy, W. A., Rauch, S. L., Buckner, R. L., … Rosen, B. R. (1996). Response and habituation of the human amygdala during visual processing of facial expression. *Neuron*, *17*(5), 875–887.

Brewer, R., Cook, R., & Bird, G. (2016). Alexithymia: A general deficit of interoception. *Royal Society Open Science*, *3*(10). https://doi.org/10.1098/rsos.150664

Calder, A. J. (1996). Facial emotion recognition after bilateral amygdala damage: Differentially severe impairment of fear. *Cognitive Neuropsychology*, *13*(5), 699–745.

Calder, A. J., Keane, J., Manly, T., Sprengelmeyer, R., Scott, S., Nimmo-Smith, I., & Young, A. W. (2003). Facial expression recognition across the adult life span. *Neuropsychologia*, *41*(2), 195–202. https://doi.org/10.1016/S0028-3932(02)00149-5

Ceunen, E., Van Diest, I., & Vlaeyen, J. (2013). Accuracy and awareness of perception: related, yet distinct (commentary on Herbert et al., 2012). *Biological Psychology*, *92*(2), 423–427.

Chong, J. S. X., Ng, G. J. P., Lee, S. C., & Zhou, J. (2017). Salience network connectivity in

the insula is associated with individual differences in interoceptive accuracy. *Brain Structure and Function*, *222*(4), 1635–1644.

Cook, R., Brewer, R., Shah, P., & Bird, G. (2013). Alexithymia, Not Autism, Predicts Poor Recognition of Emotional Facial Expressions. *Psychological Science*, *24*(5), 723–732. https://doi.org/10.1177/0956797612463582

Craig. (2009). How do you feel--now? The anterior insula and human awareness. *Nature Reviews Neuroscience*, *10*(1).

Craig, A D. (2002). *how do you feel Craig 2002*. *3*(August).

Craig, Arthur D, Chen, K., Bandy, D., & Reiman, E. M. (2000). Thermosensory activation of insular cortex. *Nature Neuroscience*, *3*(2), 184–190.

Critchley, H. D., Wiens, S., Rotshtein, P., Öhman, A., & Dolan, R. J. (2004). Neural systems supporting interoceptive awareness. *Nature Neuroscience*, *7*(2), 189–195. https://doi.org/10.1038/nn1176

Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *351*(1346), 1413–1420.

Dunn, B. D., Galton, H. C., Morgan, R., Evans, D., Oliver, C., Meyer, M., … 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

Ekman, P. (1982). Does the face provide accurate information? *Emotion in the Fuman Face*.

Ernst, J., Northoff, G., Böker, H., Seifritz, E., & Grimm, S. (2013). Interoceptive awareness enhances neural activity during empathy. *Human Brain Mapping*, *34*(7), 1615–1624.

Farb, N., Daubenmier, J., Price, C. J., Gard, T., Kerr, C., Dunn, B. D., … Mehling, W. E. (2015). Interoception, contemplative practice, and health. *Frontiers in Psychology*, *6*(JUN), 1–26. https://doi.org/10.3389/fpsyg.2015.00763

Forkmann, T., Scherer, A., Meessen, J., Michal, M., Schächinger, H., Vögele, C., & Schulz, A. (2016). Making sense of what you sense: Disentangling interoceptive awareness, sensibility and accuracy. *International Journal of Psychophysiology*, *109*, 71–80.

Frank, D. W., Costa, V. D., Averbeck, B. B., & Sabatinelli, D. (2019). Directional interconnectivity of the human amygdala, fusiform gyrus, and orbitofrontal cortex in emotional scene perception. *Journal of Neurophysiology*, *122*(4), 1530–1537.

Friston, K. (2008). Hierarchical models in the brain. *PLoS Computational Biology*, *4*(11).

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138. https://doi.org/10.1038/nrn2787

Furl, N., Van Rijsbergen, N. J., Kiebel, S. J., Friston, K. J., Treves, A., & Dolan, R. J. (2010). Modulation of perception and brain activity by predictable trajectories of facial expressions. *Cerebral Cortex*, *20*(3), 694–703. https://doi.org/10.1093/cercor/bhp140

Fusar-Poli, P., Placentino, A., Carletti, F., Landi, P., Allen, P., Surguladze, S., … Politi, P. (2009). Functional atlas of emotional faces processing: A voxel-based meta-analysis of 105 functional magnetic resonance imaging studies. *Journal of Psychiatry and Neuroscience*, *34*(6), 418–432.

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences*, *8*(9), 396–403. https://doi.org/10.1016/j.tics.2004.07.002

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

Gentsch, A., Sel, A., Marshall, A. C., & Schütz-Bosbach, S. (2019). Affective interoceptive inference: Evidence from heart-beat evoked brain potentials. *Human Brain Mapping*, *40*(1), 20–33. https://doi.org/10.1002/hbm.24352

Georgiou, E., Mai, S., Fernandez, K. C., & Pollatos, O. (2018). I see neither your fear, nor your sadness–Interoception in adolescents. *Consciousness and Cognition*, *60*, 52–61.

Goffaux, V., Schiltz, C., Mur, M., & Goebel, R. (2013). Local discriminability determines the strength of holistic processing for faces in the fusiform face area. *Frontiers in Psychology*, *3*, 604.

Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience*, *7*(5), 555–562.

Hampshire, A., Chamberlain, S. R., Monti, M. M., Duncan, J., & Owen, A. M. (2010). The role of the right inferior frontal gyrus: inhibition and attentional control. *NeuroImage*, *50*(3), 1313–1319. https://doi.org/10.1016/j.neuroimage.2009.12.109

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biological Psychiatry*, *51*(1), 59–67. https://doi.org/10.1016/S0006-3223(01)01330-0

Haxby, J. V, Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223–233.

Herbert, B. M., Herbert, C., & Pollatos, O. (2011). On the relationship between interoceptive awareness and alexithymia: is interoceptive awareness related to emotional awareness? *Journal of Personality*, *79*(5), 1149–1175.

Herry, C., Bach, D. R., Esposito, F., Di Salle, F., Perrig, W. J., Scheffler, K., … Seifritz, E. (2007). Processing of temporal unpredictability in human and animal amygdala. *Journal of Neuroscience*, *27*(22), 5958–5966. https://doi.org/10.1523/JNEUROSCI.5218-06.2007

Hess, U., Blairy, S., & Kleck, R. E. (1997). The intensity of emotional facial expressions and decoding accuracy. *Journal of Nonverbal Behavior*, *21*(4), 241–257. https://doi.org/10.1023/A:1024952730333

Hobson, H., Hogeveen, J., Brewer, R., Catmur, C., Gordon, B., Krueger, F., … Grafman, J. (2018). Language and alexithymia: Evidence for the role of the inferior frontal gyrus in acquired alexithymia. *Neuropsychologia*, *111*, 229–240.

Hübner, A. M., Trempler, I., Gietmann, C., & Schubotz, R. I. (2021). Interoceptive sensibility predicts the ability to infer others' emotional states. *Plos One*, *16*(10), e0258089.

Hübner, A. M., Trempler, I., Kloth, N., Duning, T., Johnen, A., & Schubotz, R. I. (2022). Selective Impairment of Emotion Inference in Behavioural Variant Frontotemporal Dementia? A Pilot Study. *In Press*.

James, W. (1884). What is an Emotion?(188-205). *Mind*, *9*, 34.

Jiang, J., Summerfield, C., & Egner, T. (2013). Attention sharpens the distinction between expected and unexpected percepts in the visual brain. *Journal of Neuroscience*, *33*(47), 18438–18447. https://doi.org/10.1523/JNEUROSCI.3308-13.2013

Jung, S., Kim, J.-H., Kang, N.-O., Sung, G., Ko, Y.-G., Bang, M., … Lee, S.-H. (2021). Fusiform gyrus volume reduction associated with impaired facial expressed emotion recognition and emotional intensity recognition in patients with schizophrenia spectrum psychosis. *Psychiatry Research: Neuroimaging*, *307*, 111226.

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

Keane, J., Calder, A. J., Hodges, J. R., & Young, A. W. (2002). Face and emotion processing in frontal variant frontotemporal dementia. *Neuropsychologia*, *40*(6), 655–665.

Kirita, T., & Endo, M. (1995). Happy face advantage in recognizing facial expressions. *Acta Psychologica*, *89*(2), 149–163. https://doi.org/10.1016/0001-6918(94)00021-8

Kuehn, E., Mueller, K., Lohmann, G., & Schuetz-Bosbach, S. (2016). Interoceptive awareness changes the posterior insula functional connectivity profile. *Brain Structure and Function*, *221*(3), 1555–1571. https://doi.org/10.1007/s00429-015-0989-8

Lange, C. G. (1885). The mechanism of the emotions. *The Classical Psychologists*, 672–684.

LeDoux, J. (2003). The emotional brain, fear, and the amygdala. *Cellular and Molecular Neurobiology*, *23*(4), 727–738.

Leppänen, J. M., & Hietanen, J. K. (2003). Affect and Face Perception: Odors Modulate the Recognition Advantage of Happy Faces. *Emotion*, *3*(4), 315–326. https://doi.org/10.1037/1528-3542.3.4.315

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review*, *91*(3), 295.

Mah, L., Arnold, M. C., & Grafman, J. (2004). Impairment of social perception associated with lesions of the prefrontal cortex. *American Journal of Psychiatry*, *161*(7), 1247–1255. https://doi.org/10.1176/appi.ajp.161.7.1247

Meessen, J., Mainz, V., Gauggel, S., Volz-Sidiropoulou, E., Sütterlin, S., & Forkmann, T. (2016). The relationship between interoception and metacognition. *Journal of Psychophysiology*.

Mehling, W. (2016). Differentiating attention styles and regulatory aspects of self-reported interoceptive sensibility. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1708), 20160013.

Mehling, W., Acree, M., Stewart, A., Silas, J., & Jones, A. (2018). The multidimensional assessment of interoceptive awareness, version 2 (MAIA-2). *PLoS ONE*, *13*(12), 1–12. 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.

Menon, V., Adleman, N. E., White, C. D., Glover, G. H., & Reiss, A. L. (2001). Error-related

brain activation during a Go/NoGo response inhibition task. *Human Brain Mapping*, *12*(3), 131–143.

Mondino, M., Thiffault, F., & Fecteau, S. (2015). Does non-invasive brain stimulation applied over the dorsolateral prefrontal cortex non-specifically influence mood and emotional processing in healthy individuals? *Frontiers in Cellular Neuroscience*, *9*, 399.

Montagne, B., Kessels, R. P. C., De Haan, E. H. F., & Perrett, D. I. (2007). The emotion recognition task: A paradigm to measure the perception of facial emotional expressions at different intensities. *Perceptual and Motor Skills*, *104*(2), 589–598.

Motomura, K., Terasawa, Y., Natsume, A., Iijima, K., Chalise, L., Sugiura, J., … Umeda, S. (2019). Anterior insular cortex stimulation and its effects on emotion recognition. *Brain Structure and Function*, *224*(6), 2167–2181. https://doi.org/10.1007/s00429-019-01895-9

Nakamura, K., Kawashima, R., Ito, K., Sugiura, M., Kato, T., Nakamura, A., … Fukuda, H. (1999). Activation of the right inferior frontal cortex during assessment of facial emotion. *Journal of Neurophysiology*, *82*(3), 1610–1614.

Ondobaka, S., Kilner, J., & Friston, K. (2017). The role of interoceptive inference in theory of mind. *Brain and Cognition*, *112*, 64–68. https://doi.org/10.1016/j.bandc.2015.08.002

Palermo, R. (2004). SpringerLink - Behavior Research Methods, Volume 36, Number 4. *Behavior Research Methods*, *36*(4), 634–638. https://doi.org/10.3758/BF03206544

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

Petzschner, F. H., Weber, L. A., Wellstein, K. V., Paolini, G., Do, C. T., & Stephan, K. E. (2019). Focus of attention modulates the heartbeat evoked potential. *NeuroImage*, *186*(November 2018), 595–606. https://doi.org/10.1016/j.neuroimage.2018.11.037

Pollatos, O., Kirsch, W., & Schandry, R. (2005). On the relationship between interoceptive awareness, emotional experience, and brain processes. *Cognitive Brain Research*, *25*(3), 948–962. https://doi.org/10.1016/j.cogbrainres.2005.09.019

Pollatos, O., Schandry, R., Auer, D. P., & Kaufmann, C. (2007). Brain structures mediating cardiovascular arousal and interoceptive awareness. *Brain Research*, *1141*, 178–187.

Press, C., Kok, P., & Yon, D. (2020). The perceptual prediction paradox. *Trends in Cognitive Sciences*, *24*(1), 13–24.

Press, C., & Yon, D. (2019). Perceptual prediction: Rapidly making sense of a noisy world. *Current Biology*, *29*(15), R751–R753.

Raes, A. K., & De Raedt, R. (2011). Interoceptive awareness and unaware fear conditioning: Are subliminal conditioning effects influenced by the manipulation of visceral self-perception? *Consciousness and Cognition*, *20*(4), 1393–1402.

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87.

Richler, J. J., & Gauthier, I. (2014). A meta-analysis and review of holistic face processing. *Psychological Bulletin*, *140*(5), 1281.

Rubia, K., Smith, A. B., Brammer, M. J., & Taylor, E. (2003). Right inferior prefrontal cortex

mediates response inhibition while mesial prefrontal cortex is responsible for error detection. *Neuroimage*, *20*(1), 351–358.

Sato, W., Kochiyama, T., Yoshikawa, S., Naito, E., & Matsumura, M. (2004). Enhanced neural activity in response to dynamic facial expressions of emotion: an fMRI study. *Cognitive Brain Research*, *20*(1), 81–91.

Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review*, *69*(5), 379.

Schiffer, A.-M., & Schubotz, R. I. (2011). Caudate nucleus signals for breaches of expectation in a movement observation paradigm. *Frontiers in Human Neuroscience*, *5*, 38.

Schiffer, A. M., Ahlheim, C., Wurm, M. F., & Schubotz, R. I. (2012). Surprised at all the entropy: Hippocampal, caudate and midbrain contributions to learning from prediction errors. *PLoS ONE*, *7*(5). https://doi.org/10.1371/journal.pone.0036445

Schulz, S. M. (2016). Neural correlates of heart-focused interoception: a functional magnetic resonance imaging meta-analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1708), 20160018.

Seth, A. K. (2013a). 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. (2013b). 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. J. (2016). Active interoceptive inference and the emotional brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1708). 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*, *3*(JAN), 1–16. https://doi.org/10.3389/fpsyg.2011.00395

Shah, P., Catmur, C., & Bird, G. (2017). From heart to mind: Linking interoception, emotion, and theory of mind. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *93*, 220.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*(3), 379–423.

Simmons, W. K., Avery, J. A., Barcalow, J. C., Bodurka, J., Drevets, W. C., & Bellgowan, P. (2013). Keeping the body in mind: insula functional organization and functional connectivity integrate interoceptive, exteroceptive, and emotional awareness. *Human Brain Mapping*, *34*(11), 2944–2958.

Singer, T., Critchley, H. D., & Preuschoff, K. (2009). A common role of insula in feelings, empathy and uncertainty. *Trends in Cognitive Sciences*, *13*(8), 334–340.

Smith, S. D., Nadeau, C., Sorokopud-Jones, M., & Kornelsen, J. (2021). The Relationship Between Functional Connectivity and Interoceptive Sensibility. *Brain Connectivity*.

Sowden, S., Brewer, R., Catmur, C., & Bird, G. (2016). The specificity of the link between alexithymia, interoception, and imitation. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(11), 1687–1692. https://doi.org/10.1037/xhp0000310

Strange, B. A., Duggins, A., Penny, W., Dolan, R. J., & Friston, K. J. (2005). Information

theory, novelty and hippocampal responses: Unpredicted or unpredictable? *Neural Networks*, *18*(3), 225–230. https://doi.org/10.1016/j.neunet.2004.12.004

Terasawa, Y., Fukushima, H., & Umeda, S. (2013). How does interoceptive awareness interact with the subjective experience of emotion? An fMRI Study. *Human Brain Mapping*, *34*(3), 598–612. https://doi.org/10.1002/hbm.21458

Terasawa, Y., Moriguchi, Y., Tochizawa, S., & Umeda, S. (2014). Interoceptive sensitivity predicts sensitivity to the emotions of others. *Cognition and Emotion*, *28*(8), 1435–1448. https://doi.org/10.1080/02699931.2014.888988

Tsuchida, A., & Fellows, L. K. (2012). Are you upset? Distinct roles for orbitofrontal and lateral prefrontal cortex in detecting and distinguishing facial expressions of emotion. *Cerebral Cortex*, *22*(12), 2904–2912.

Wang, Xiaoqin, Tan, Y., Van den Bergh, O., von Leupoldt, A., & Qiu, J. (2020). Intrinsic functional brain connectivity patterns underlying enhanced interoceptive sensibility. *Journal of Affective Disorders*, *276*, 804–814.

Wang, Xingchao, Wu, Q., Egan, L., Gu, X., Liu, P., Gu, H., … Gao, Z. (2019). Anterior insular cortex plays a critical role in interoceptive attention. *Elife*, *8*, e42265.

Weissman, D. H., Roberts, K. C., Visscher, K. M., & Woldorff, M. G. (2006). The neural bases of momentary lapses in attention. *Nature Neuroscience*, *9*(7), 971–978. https://doi.org/10.1038/nn1727

Wiebking, C., Duncan, N. W., Tiret, B., Hayes, D. J., Marjaǹska, M., Doyon, J., … Northoff, G. (2014). GABA in the insula—a predictor of the neural response to interoceptive awareness. *Neuroimage*, *86*, 10–18.

Zaki, J., Davis, J. I., & Ochsner, K. N. (2012). Overlapping activity in anterior insula during interoception and emotional experience. *Neuroimage*, *62*(1), 493–499.

Supplementary Material

Table S1. Since the conceptualization of the MAIA-2 questionnaire does not provide a conduction of a total score, we assumed that it might not capture the effects sufficiently. For an explorative analysis and to provide a basis for further research in the area, we therefore correlated each of the eight MAIA-2 subscales with the beta weights, respectively. The table contains all (marginal) significant correlations of the different contrasts and MAIA-2 subscales, *p* =0.5, two-tailed.

MAIA-2 subscales	**Region of interest (ROI)**	***r***	***p value***
Intensity contrast			
MAIA-2 Not-Worrying	L ACC	.281	.071
	R ACC	.327	.034*
MAIA-2 Self-Regulation	L Insula	-.282	.070
	R Insula	-.321	.038*
MAIA-2 Emotional Awareness	L Insula	-.257	.100
	R Insula	-.311	.045*
Valence contrast			
n.s.			
Entropy contrast			
MAIA-2 Not-Worrying	L ACC	.377	.014*
	R ACC	.380	.013*
Surprise contrast			
MAIA-2 Self-Regulation	L ACC	-.371	.016*
	R ACC	-.341	.027*

4. General Discussion and Future Directions

4.1 Summary of the Presented Studies

In the presented work I investigated the role of interoceptive abilities in emotion recognition and emotional context adaptation as well as their physiological correlates and neural implementations. More specifically, I examined whether higher interoceptive abilities reflect higher precision of interoceptive signals, leading to a facilitation of emotional change detection and a more precise adaptation to emotion probability and whether brain regions reflecting interoceptive processing play a role in this. Based on the initial findings of differences in emotion recognition and context adaptation between individuals with low and high IS, I further investigated a possible impairment of interoceptive predictive processing in patients with bvFTD.

In **Study I**, I assessed the link between sensibility of interoceptive changes and adaptation to emotion probabilities in young healthy subjects. The results provide evidence that self-reported sensibility to one's own interoceptive states modulates the recognition of emotional changes in facial expressions of others. Specifically, individuals with higher IS have been found to show lower thresholds for emotion recognition, as reflected in decreased reaction times for emotional expressions, especially of high intensity. Furthermore, higher IS entails a more precise adaptation to biased probabilities of emotional valences, reflected in decreased reaction times for expected emotions. Indications for a corresponding modulation of SCR by IS as a function of varying probabilities were found. Sympathetic arousal increased for more surprising and unexpected emotions in individuals with higher IS, indicating heightened arousal in response to violations of expectations in the course of learning from increasing probability of a specific valence. In summary, Study I provides evidence that IS facilitates the recognition speed of emotional changes in facial expressions of others and entails a more precise adaptation to biased emotional probabilities.

Based on these results, in **Study II** I focused on potential deficits in interoceptive processing and deficient emotion recognition in patients suffering from bvFTD. Patients showed increasing difficulty in emotion recognition compared to healthy controls when uncertainty about the change in the emotional stimulus increased. The observed deficits can be attributed to an impairment of interoception in bvFTD patients. The hypothesis of an impaired adaptation of emotional predictions in bvFTD was partly confirmed. While healthy controls and patients did not differ in their reaction times with respect to the exploiting biased probabilities, increasing predictability of upcoming facial expressions was associated with a decrease in SCRs in healthy controls, but not in patients. The deviant modulation of skin conductance suggests that patients cannot effectively process interoceptive prediction errors. Lastly, these effects were found to be specific for emotion recognition as bvFTD patients did not show impaired discrimination of facial expressions in the control task. To sum up, the results provide indications for deficient emotion recognition in bvFTD that might be based on impaired interoceptive predictions.

Finally, in **Study III** I investigated the effects of individuals' IS and IAcc on performance and brain activity to provide new insights into how IAcc and IS affect emotion recognition and adaptation to biased emotional probabilities on a behavioural level and in the brain. Here, individuals' IS and IAcc had no effect on recognition speed of emotional expression, but IS and IAcc differed with regard to their impact on recognizing emotional valence. Moreover, as observed in Study I, individuals with higher IS benefitted more from the biased probability of emotional valence, whereas there was no interaction between differences in IAcc and adaptation to emotion probabilities. On the neural level, brain responses to more predictable emotions elicited a bilateral activity pattern comprising the IFG and the posterior insula. Importantly, individual IAcc, but not IS scores, positively correlated with brain responses to more surprising and less predictable emotional expressions in the insula and caudate nucleus, indicating an enhanced processing of interoceptive prediction errors in higher

IAcc individuals. In conclusion, the results provide evidence for the notion that different measures of interoceptive abilities (IAcc and IS) have differential effects on emotion recognition and context-sensitive emotion inference. Indications for different neural implementations of IAcc and IS also emerge from the results.

In the following, I will discuss the findings of these studies in regard to how interoceptive abilities affect emotion recognition and adaptation to biased probabilities of emotion, and incorporate the results of this thesis with the current state of research. Moreover, I will discuss findings on functional areas specialized for processing of unfolding emotional expressions as well as regions involved in the signalling of emotion inference and review the results on brain regions related to IAcc and IS. Subsequently, I will discuss how the effects of interoception are specific to emotion processing and to what extent interoceptive processing may be impaired in bvFTD. I will conclude this section with a critical reflection and by presenting practical applications of the results.

4.2 Effects of Interoception on Emotion Recognition

4.2.1 Recognizing Emotion with Different Interoceptive Abilities

In the present studies, a primary focus was set on the modulation of emotional change recognition in facial expressions of others by differences in interoceptive abilities.[9] Most current studies on emotion recognition have been conducted using static facial images. However, static pictures do not reflect the true form and liveliness of dynamic facial expressions as they occur in everyday life (Sato et al., 2004). Dynamic facial expressions of emotions are a natural and powerful way of emotional communication between individuals (Sato et al., 2004). The three studies presented here intentionally used dynamically unfolding faces to obtain a

[9] Note that in the studies presented, participants' task was to discriminate between positive and negative valence (i.e., between happy and fearful expressions). Thus, as discussed in the studies, the task reflects *recognition* rather than *detection* of the emotional expressions.

greater emotive relevance and to more closely emulate emotion perception in daily life. Additionally, according to the concept of sensorimotor simulation, it can be assumed that during the observation of dynamic compared to static facial expressions the sensorimotor system is more engaged and that, thus, effects of differences in interoceptive abilities on emotion recognition are more pronounced (Wood et al., 2016; cf. section 1.3.2).

Regarding the modulation of emotion recognition by differences in interoceptive abilities, I hypothesised that higher sensibility (IS) and accuracy (IAcc) for interoceptive changes are reflected in higher precision of interoceptive signals, resulting in facilitated recognition of others' emotions. The rationale behind this assumption, which e.g. Arnold (2019) refers to as the "enhanced emotional discernment hypothesis", is that a better ability to sense bodily reactions leads to a richer emotional experience, which in turn facilitates a better understanding of others' emotions. Yet, little is known about the effects of individuals' IS on emotion recognition, whereas there are some studies investigating the effects of IAcc on empathy and emotions. IAcc has been reported to be positively related to higher cognitive and affective empathy (Grynberg & Pollatos, 2015). Moreover, individuals with higher IAcc were found to evaluate the experience of emotions more intensely (Barrett, Quigley, Bliss-Moreau, & Aronson, 2004) and to be more sensitive to emotional facial expressions of others (Terasawa et al., 2014).

Extending the finding of these previous studies, the results of Study I suggest that self-reported sensibility to one's own interoceptive states facilities the recognition of emotional changes in facial expressions of others, as reflected in decreased reaction times for recognition of emotions in higher IS individuals. As thoroughly discussed in Study I, happy faces are recognized faster and more accurately, whereas fearful expressions are recognized less well than other types of facial expressions (e.g., Leppänen & Hietanen, 2003, 2004). Accordingly, additional evidence was found that the discrimination advantage of individuals with higher IS occurs primarily for those stimuli that are particularly difficult to recognize unambiguously,

i.e., fearful versus happy faces and low versus high intensity emotional expressions. An individual's IS may therefore have significant implications for everyday social interactions: as spontaneous facial expressions of low to moderate intensity are most common in everyday life (e.g., Matsumoto & Hwang, 2014; Motley & Camden, 1988), individuals with low IS may have greater difficulty in inferring the emotional state of others from their facial expressions and therefore face more difficulty in their daily social interactions.

Contrary to Study I, behavioural results of Study III showed no discrimination advantage for individuals with higher IS or IAcc. Discrepancies in the effects of IS on emotion recognition in Study III compared to Study I could be due to differences in the two samples with regard to the distribution of IS values. Future studies could preliminarily examine interoceptive abilities of the participating subjects and use the interoceptive scores to classify individuals as either high or low interoceptive, or include specifically high and low interoceptive subjects in the sample (e.g., Kuehn, Mueller, Lohmann, & Schuetz-Bosbach, 2016). Nevertheless, the results of Study III indicate that both IS and IAcc modulate the recognition of emotional changes. These parameters differed with regard to their impact on recognizing emotional valence, i.e., individuals with higher IS showed longer response times for fearful faces, whereas individuals with higher IAcc showed increased response times for happy ones. As discussed in Study III, based on previous research, one could speculate that with increasing IAcc emotions are more regulated (e.g., Georgiou, Mai, Fernandez, & Pollatos, 2018), which might also affect the recognition of positive emotions in individuals with higher IAcc. However, since differential modulation of response times by IAcc and IS had not been part of the initial hypotheses, further studies are needed to investigate whether the effects of IAcc and IS on the recognition of emotional valence indeed result from differences in emotion regulation.

Regarding the modulation of sympathetic arousal, in Study I no differences between participants with high compared to low IS in SCR amplitude or facilitation of sympathetic

arousal by higher IS in specific experimental conditions were observed. Previous studies using the electroencephalogram (EEG) have found associations between higher IAcc and amplitudes of both the P300 component and slow waves in response to emotional pictures (e.g., Herbert et al., 2007; Pollatos et al., 2005). Consistent with the findings of Study I, studies using SCR have not detected differences in SCR with higher interoceptive abilities (Werner, Kerschreiter, Kindermann, & Duschek, 2013). Thus, since SCR is thought to reflect an orientation reflex in response to new or unexpected stimuli (e.g., Bradley, 2009; Geuter et al., 2017), it is more likely that a modulation of SCR is observed as a function of the predictability of emotions, as also suggested by the further results, which I will discuss in more detail in section 4.3.

To sum up, the findings provide evidence that higher IS is associated with a lower threshold for emotion recognition, especially for emotions of high intensity. No corresponding effects for differences in IAcc were found, although the results of Study III indicate a differential modulation of the recognition of emotional valence through IS and IAcc respectively. Consistent with previous studies, there is no evidence for an association of IS scores with sympathetic arousal measured by SCR as a function of emotional valence or intensity. Future studies are needed to further clarify the differential effects of individuals' IS and IAcc on emotion recognition. For example, future research could address effects of differences in interoceptive abilities on the recognition of emotions that are comparatively difficult to discriminate, such as disgust versus anger (e.g., Jack, Garrod, & Schyns, 2014).

4.2.2 Neural Correlates of Dynamically Unfolding Emotion

The perception of emotions in facial expressions of others is a fundamental ability that enables us to interact with each other and to infer affective states of others (Dols & Russell, 2017). The perception of facial expressions leads to the recruitment of dedicated and distributed brain networks (Duchaine & Yovel, 2015; Skiba & Vuilleumier, 2020). Since most fMRI studies of facial emotion processing use static facial pictures, there are only few studies

investigating how dynamic emotional faces are processed. Thus, in the following I will discuss findings on functional areas specialized in the processing of unfolding emotional expressions and how these areas are connected to interoception.

In Study III, I conducted a whole-brain analysis to test the hypothesis of respective cortical networks specialized in the processing of dynamically unfolding emotional facial expressions depending on emotional intensity and valence. For the processing of high versus low intensity of emotional expressions a network comprising the inferior occipital gyrus and the right fusiform gyrus was found. The involvement of the areas is consistent with previous findings, revealing increased activation in response to dynamic happy and fearful expressions compared to static expressions in regions including the inferior occipital gyrus, middle occipital gyrus and fusiform gyrus (Sato et al., 2004). As already discussed in Study III, distinct cortical areas for analysing invariant characteristics (e.g., identity) and changeable features (e.g., facial expression and gaze) are engaged when emotional faces are processed, involving pathways from inferior occipital gyrus to fusiform gyrus (Haxby, Hoffman, & Gobbini, 2000; Skiba & Vuilleumier, 2020). As indicated by behavioural results of Study I and Study III, high emotional intensities were comparatively easily recognized (cf. section 4.2.1). Moreover, in the paradigm used in the studies comprising this thesis, the faces of the actors/actresses were repeated several times throughout the task. Accordingly, inferior occipital gyrus and fusiform gyrus activation in response to high compared to low emotional intensity might reflect enhanced perceptual and cognitive processing of variant and invariant facial features of the repetitive emotional faces.

Interestingly, less intense compared to high intense emotions were accompanied by activation in the right middle frontal gyrus, right ACC and anterior dorsal insula. Each of these cortical areas, the dorsolateral prefrontal cortex (dlPFC), ACC and insula are reported to contain a high density of VENs (Cauda et al., 2013; Fajardo et al., 2008). VENs are proposed to be involved in rapid information integration by monitoring changes in the physiological network that regulates bodily signals and social behaviour (Cauda et al., 2013; cf. section 1.2.3).

Correspondingly, studies indicate an involvement of the anterior cingulate and fronto-insular cortices in rapid processing of facial emotional information (Fan et al., 2011). Moreover, recent evidence from brain tumour rejection associated with the insula lobe suggests that lesions of the insula impair interoception, which in turn inhibits arousal recognition and the ability to discriminate emotions (Motomura et al., 2019).

The results of the behavioural data in Study I and Study III indicate that the detection of low compared to high intensity emotions is much more challenging (cf. section 4.2.1). Thus, it seems reasonable that lower intensity of emotions requires deeper evaluative processing and higher cognitive effort to recognize the emotional content. On the neural level, this appears to be reflected in increased activation in areas associated with interoceptive processing and a high density of VENs (i.e., AIC and ACC). This speaks in favour of an involvement of AIC and ACC in the recognition of emotion that are difficult to recognize and supports the assumption that interoceptive processing plays a role in emotion recognition (see also Motomura et al., 2019). These findings also fit the behavioural results of Study I, which suggest that individuals with decreased IS exhibit more difficulties in comparatively challenging conditions (cf. section 4.2.1). It appears that exactly these challenging conditions reveal activation patterns comprising areas associated with interoceptive processing, i.e., AIC and ACC. Indeed, regarding emotion recognition depending on valence, fearful compared to happy expressions were accompanied by increased activation of the AIC along with the inferior frontal gyrus (IFG). This activation pattern matches the above explanation, since fearful compared to happy faces are comparatively difficult to recognize. Both of these areas, AIC and IFG, are functionally associated with access to interoceptive information and underpinning the subjective experience of the emotional state (Terasawa, Fukushima, & Umeda, 2013). Accordingly, it seems plausible that reliance on interoceptive information mediated by areas at the top of the interoceptive hierarchy, i.e., AIC and ACC, might be essential especially for emotions that are difficult to recognize.

To more directly assess the contribution of brain regions involved in interoceptive processing and emotion recognition as a function of IS and IAcc, in Study III I further examined the extent of activation increase in the insula and ACC in response to the emotional stimuli relative to individual IS or IAcc scores (see also section 4.3.3). However, no evidence was found that higher individual IS or IAcc is related to increased activation in the insula or ACC in response to more challenging conditions (i.e., high versus low intensity or fearful versus happy emotional valence). This is also consistent with the behavioural results of Study III, which found no discrimination advantage for individuals with higher IS or IAcc on emotion recognition as a function of valence and intensity (contrary to Study I, see section 4.2.1). Based on the results presented regarding an activation increase in AIC, ACC or IFG in response to more challenging emotional intensities and valences, it nevertheless seems promising for future research to further investigate these associations with respect to interoceptive abilities and the aforementioned regions.

In summary, the results on the neuronal representations of unfolding emotional facial expressions depending on emotional intensity and valence suggest that when uncertainty about the change in the emotional stimuli increases, several brain regions comprising the AIC, ACC or IFG are engaged. These results support the assumption that interoceptive predictive processing plays a role in emotion recognition, as these areas are also associated with processing of interoceptive prediction and interoceptive prediction errors, respectively.

4.3 The Role of Interoception in Probabilistic Inference of Emotion

4.3.1 Interoception and Adaptation to Emotion Probabilities

A starting point of this thesis was the postulation that in individuals with higher interoceptive abilities interoceptive prediction errors are cascaded up the interoceptive hierarchy rather than being suppressed by low-level predictions. Therefore individuals with

higher interoceptive abilities are proposed to exhibit enhanced precision of interoceptive prediction errors and are better able to adaptively weight interoceptive signals relative to others competing for attention (e.g., Ainley et al., 2016; Arnold et al., 2019). Empirical evidence supporting this proposition relates, for example, to increased error aversion in individuals with higher IAcc (Sueyoshi, Sugimoto, Katayama, & Fukushima, 2014). However, this theoretical assumption has never been evaluated in the context of emotional prediction, i.e., probabilistic inference of emotional states. To this end, in the three studies composing this thesis, different probabilities for the occurrence of either happy or fearful faces per block were implemented to assess participants' propensity to efficiently update their predictive model. The presented results clearly indicate that higher IS is associated with a more precise adaptation of emotion predictions. In Study I, reaction times of individuals with higher IS increased when an expression was rather unexpected (i.e., a surprising emotion occurred) and decreased with increasing predictability (i.e., lower entropy of emotion occurrence). In line with this, Study III revealed corresponding effects with respect to entropy, which positively covaried with reaction times in higher IS individuals. Recent EEG studies have shown that interoceptive attention increases cortical processing of heartbeats (e.g., Petzschner et al., 2019) and that attention facilitates faster processing of prediction errors (e.g., Jiang, Summerfield, & Egner, 2013). In line with these previous results, the findings can be attributed to greater precision weighting of attention by individuals with higher IS. Accordingly, the results provide evidence that individuals with higher IS are more likely to learn from biased probabilities of emotional expressions and thus build stronger expectations (Ainley et al., 2016; cf. section 1.3.1). This in turn leads to faster recognition of predicted emotions but also to hesitation when unexpected emotions occur, whereas in individuals with low IS the results indeed suggested an opposite pattern (Study I & Study III).

In accordance with the hypothesis regarding the modulation of physiological parameters, the findings of Study I revealed a paralleled SCR pattern to behavioural measures

with increased SCRs amplitude with greater surprise and lower predictability in individuals with higher IS. As discussed in Study I, given that SCR is thought to reflect an orienting reflex in response to unexpected stimuli (e.g., Bradley, 2009; Geuter et al., 2017), the results suggest that in higher IS individuals prediction errors elicited by surprising emotions are cascaded up the interoceptive hierarchy leading to increased physiological arousal. Previous studies have further shown that regardless of differences in interoceptive abilities, SCR change was heightened after novel pictures and decreased with repetition of stimuli containing neutral, pleasant and unpleasant pictures (Bradley, 2009). Taken together, extending previous studies, the behavioural results of Study I and III and results of modulation of sympathetic arousal in Study I clearly and consistently indicate that in individuals with higher IS, prediction errors cascade up the interoceptive hierarchy, leading to hesitation and increased physiological arousal when surprising and less predictable emotions occur.

Interestingly, results of Study III indicate that differences in IS, but not IAcc, were associated with faster recognition of predicted emotions and hesitation when unexpected emotions occurred. There are several ways to capture the facets of IS and IAcc, with two respective examples being the MAIA-2 and the HDT (e.g., Garfinkel et al., 2015; cf. section 1.1). As discussed in Study III, the MAIA-2 dimensions are constructed as elements of conscious processes of interoception that are accessible to self-report (Mehling et al., 2012). The questionnaire focuses on regulatory aspects as well as aspects of successfully directing attention towards and away from interoceptive signals (Mehling et al., 2018). Thus, higher IS scores represent perceived awareness of emotional states and how effectively attention to these sensations can be weighted, maintained, or regulated. The HDT maps the objective ability to accurately monitor interoceptive signals in a short time period. This ability may be associated with better attentional focus on these signals, but is not necessarily associated with the ability to maintain or control attention towards interoceptive signals. Empirical evidence for this assumption can be seen in the findings that IAcc scores assessed by the HDT are found to be

reduced for longer time intervals (e.g., Zamariola, Maurage, Luminet, & Corneille, 2018). Maintaining and controlling attention towards interoceptive changes may be the key feature to increase the precision of interoceptive predictions (Arnold et al., 2019). Conceivably, in accordance with the above explanations, especially higher IS values could lead to more frequent and accurate interoceptive predictions. However, further studies are required to explicitly investigate this assumption.

In summary, the findings suggest that higher IS individuals benefit more from the implemented probabilistic context than individuals with lower IS and show a more precise adaptation of emotion predictions. Moreover, higher IS individuals exhibit decreased physiological arousal in response to more predictable emotions (reflecting adaptation to biased probabilities) and increased arousal in response to violations. The finding that individuals' IS scores, but not IAcc scores, are related to emotional probability adaptation may be attributable to the proposition that higher IS, rather than higher IAcc, is associated with the ability to sustain and control attention to interoceptive changes. This assumption is also supported by the results on functional areas involved in the signalling of emotion inference (Study III), which I will discuss in the following.

4.3.2 Neural Correlates of Context-Sensitive Emotion Inference

Regarding the probabilistic context adaptation I hypothesised that surprising and less predictable emotions activate regions involved in the signalling of interoceptive prediction errors, e.g., AIC, ACC, hippocampus or caudate nucleus (Ahlheim, Stadler, & Schubotz, 2014; Bechara, Damasio, & Damasio, 2000; Herry et al., 2007; Schiffer, Ahlheim, Wurm, & Schubotz, 2012; Strange et al., 2005). However, contrary to the expectation no such modulation was observed. Instead, the processing of emotions with increasing predictability was found to be accompanied by activation of a network comprising the IFG and posterior insula. Since intuitively one would rather expect a repetition-related reduction in neural responses following

stimulus repetition, a brief explanation will be given as to why stimulus repetition in general can lead to increased brain activity. A stimulus repetition-induced attenuation of activity in the respective cortex has been shown in numerous studies and is referred to as repetition suppression (Grill-Spector, Henson, & Martin, 2006; see Schubotz, 2015, for a discussion and review). However, there is also evidence reporting repetition enhancement effects (Segaert, Weber, de Lange, Petersson, & Hagoort, 2013; Todorovic, van Ede, Maris, & de Lange, 2011). As effects of repetition on neural activity are argued to be based on top-down predictive processing rather than bottom-up neuronal adaptation to unexpected stimuli, a suppression of surprise responses elicited by unexpected events is assumed (Gentsch, Sel, Marshall, & Schütz-Bosbach, 2019; Summerfield et al., 2008).

The results of Study III revealed increased activation for more predictable emotion with a bilateral activity pattern comprising the IFG and posterior insula. IFG activation in response to more predicable emotion might reflect attentional control when facing relevant stimuli (e.g., Hampshire, Chamberlain, Monti, Duncan, & Owen, 2010) and stimulus-triggered reorienting of attention on behaviourally relevant stimuli (e.g., Craig, 2009; Weissman, Roberts, Visscher, & Woldorff, 2006). Moreover, since impairment of the IFG was found to be associated with alexithymia, recent evidence suggests IFG involvement in interoceptive functioning (Hobson et al., 2018). Thus, the results of Study III, revealing an increase in IFG activity as a function of emotion predictability, suggest a role of the IFG in attentional control towards the emotional stimuli and stabilisation of interoceptive predictions. Though, the exact role of an IFG involvement in attentional control and interoceptive predictions ought to be tested in future studies.

Interestingly, with respect to the neural mechanisms underlying the probabilistic context adaptation of emotions, Study III further revealed an activation increase exclusively in the posterior part of the insula with increasing predictability of emotion. A posterior-to-anterior pathway for interoceptive processing within the insula is assumed (Craig, 2009; Jabbi,

Bastiaansen, & Keysers, 2008; cf. section 1.2.3). The posterior insula is considered the primary interoceptive mapping area in the cortex, receiving direct interoceptive input from the ventromedial nuclei of the thalamus (Craig, 2002; Critchley et al., 2004). Barrett and Simmons (2015) propose that subsequent to the transmission of new interoceptive signals to the mid- and posterior insula, here differences between the predicted and the actual interoceptive signal are computed. Thereupon, it was proposed that AIC and ACC are more likely to be involved in processing and reducing prediction errors through their contribution in the intrinsic executive control and attention networks. The present results of Study III suggest that the mapping of interoceptive signals associated with the posterior insula may be crucial for a precise adaptation of emotion predictions. In particular, since regions that integrate interoceptive information, including the posterior insula, were recruited during processing of emotion predictions, the results suggest that adaptation to the emotional context involves the creation and maintenance of physiological representations of the body.

These findings are consistent with the behavioural results of Study I and Study III, which suggest that the ability to maintain and control attention towards interoceptive changes may be key to increase the precision of interoceptive predictions (cf. section 4.3.1). Since individuals with higher IS showed a more precise adaptation of emotion predictions in both Study I and III, higher IS may thus be associated with increased activation in the posterior insula and IFG. However, these relations were not directly examined in Study III, which rather focused on a preliminary investigation of the links between interoceptive abilities and brain regions subserving interoceptive processing and the processing of prediction errors in general (i.e., insula, ACC and caudate nucleus; cf. section 4.3.3). However, it seems promising to assess whether posterior insula and IFG activation during valence-specific predictions of expected and unexpected facial expressions are associated with differences in IS scores. Future studies are warranted to clarify the contribution of the posterior insula and IFG in probabilistic context adaptation of emotion and their relation to interoceptive abilities.

Finally, future studies could also investigate functional connectivity between the posterior and anterior insula during interoceptive predictions. Initial empirical evidence in this regard revealed reduced functional connectivity between the posterior insula and AIC during awareness to interoceptive signals, i.e., the heartbeat (Kuehn et al., 2016). The authors suggested that the selective inhibition along the posterior-to-anterior insula may allow for less noisy and more pronounced perception of interoceptive signals, but only in individuals with high IAcc. Accordingly, it would be of interest to investigate whether, in a paradigm such as the one used in the present studies, connectivity decreases with increasing repetition or predictability of emotions and increases with prediction errors (as the posterior insula is proposed to relay prediction errors to the AIC for reduction, see e.g., Barrett & Simmons, 2015).

4.3.3 Interoceptive Abilities and Processing of Interoceptive Prediction Errors

In Study III, I specifically investigated whether the extent of activation increase in brain regions involved in interoceptive predictive processing, i.e., insula and ACC, is associated with individual IS or IAcc scores. Since the caudate nucleus is commonly associated with perceptual prediction errors, this region was additionally included in the analyses (Schiffer & Schubotz, 2011; Schiffer et al., 2012). Intriguingly, the individual IS score was not related to an activation increase in insula, ACC or caudate nucleus. Indeed, in individuals with higher IAcc a significantly higher BOLD signal increase in the caudate nucleus and left insula in response to less predictable events and higher activation in the caudate nucleus in response to more surprising emotions was observed. These findings are in line with previous studies, demonstrating that IAcc positively covaries with activation in the AIC, cingulate gyrus and somatosensory cortex (Critchley et al., 2004; Pollatos, Schandry, Auer, & Kaufmann, 2007) and that interoceptive attention is related to increased AIC activation (Chong, Ng, Lee, & Zhou, 2017; Xingchao Wang et al., 2019). The assumption that individuals' IAcc is positively related to neural activity in the AIC can be extended by the present results, which suggest a positive

correlation between individuals' IAcc scores and insula as well as striatal activation in response to unpredictable emotions. Since the anterior portion of the insula in particular is thought to be related to the processing and reduction of interoceptive prediction errors (e.g., Barrett & Simmons, 2015), the results provide empirical indications for a contribution of the insula and striatal cortex in processing interoceptive prediction errors that is amplified in individuals with higher IAcc.

Given that the behavioural effects related to emotional context adaptation differed between IS and IAcc in Study III, it is not surprising that individual IS and IAcc scores are associated with different brain regions. Recently, Smith et al. (2021) found functional connectivity of the insula not to be correlated with individuals' IS scores and Wang et al. (2020) even found that higher IS was associated with reduced connectivity of AIC and ACC to temporal and frontal regions. In accordance with previous studies, the results of Study III also suggest that brain areas reflecting individual IS have little overlap with regions reflecting IAcc of individuals. Based on the results of Study III and behavioural results of Study I, one might speculate that the IAcc score might be associated with enhanced processing of interoceptive prediction errors (e.g., AIC and caudate nucleus), whereas a higher IS score seems more likely to be associated with maintenance of attention to interoceptive changes (e.g., posterior insula) and reorienting of attention (e.g., IFG; cf. section 4.3.2). The presented findings provide a first step towards a deeper understanding of the role of interoceptive abilities in emotion prediction and their neural representations. However, further research is warranted to more directly investigate the neural mechanism associated with interoceptive abilities and interoceptive prediction errors.

To sum up, the findings regarding regions involved in the signalling of interoceptive predictions and prediction errors revealed activation increases for more predictable emotion in IFG and posterior insula, suggesting that both areas might be crucial for a precise adaptation of emotion predictions. Moreover, the findings provide indications that adaptation to the

emotional contexts involves the creation and maintenance of physiological representations of the body. Lastly, the results provide evidence for amplified processing of interoceptive prediction errors during the perception of unexpected emotions in individuals with higher IAcc. Overall the findings point to differential effects of the interoceptive measures (IAcc and IS) on emotion recognition and adaptation to biased emotional probabilities, as well as different neural implementations.

4.4 Domain Specific and Unspecific Effects of Interoception

As described in the previous sections, the key to interoception facilitating emotion recognition and learning of emotional contexts may lie in a more precise weighting of attention towards interoceptive signals. However, attention is thought to increase precision-weighting of prediction errors in general (e.g., Jiang et al., 2013). Thus, a further question of this thesis was whether the effects of IS on emotion recognition and emotional context adaptation are specific to the emotional domain. To this end, a non-emotion classification control task was implemented paralleled to the emotion condition. More specifically, the control task was implemented to investigate whether the effects observed in the emotion condition were caused by differences in interoceptive abilities or whether they are attributable to a generally enhanced ability of individuals with higher IS to direct attention to any kind of sensory input. The results of Study I provide no evidence for a general discrimination advantage among individuals with higher IS, nor for a modulation of physiological parameters. However, reaction times tended to increase more for higher IS individuals as a function of surprise (albeit not significantly).

Predictive processing occurs across three distinct modalities: exteroceptive predictions (underlying perception of the external world), proprioceptive predictions (leading to action through active inference), and interoceptive predictions (underlying emotional processing and autonomic regulation) (Seth, 2014; Seth & Friston, 2016). Integrated experiences of the embodied self result from hierarchical predictions across all three dimensions, including

multimodal and amodal predictions depending on the level of the prediction hierarchy at which the predictions are violated. Recent studies suggest an interaction of interoceptive and exteroceptive perception and support the notion that inferences from one domain affect the other (e.g., Allen, Levy, Parr, & Friston, 2019). As examples of domain-spanning interferences, studies indicate that precision of interoceptive states can directly influence exteroceptive uncertainty (Allen et al., 2016, 2019; Boldt, De Gardelle, & Yeung, 2017; Spence, Dux, & Arnold, 2016). It is plausible that in Study I, prediction errors in the control task also elicited bodily signals to which individuals with higher IS directed their attention more. Nevertheless, the results of the control task only revealed tendencies for possible correlations. In the emotion task the interactions were much more pronounced and accompanied by stronger physiological arousal in response to interoceptive prediction errors in individuals with higher IS (Study I). Accordingly, the results might be attributable to stronger bodily signals elicited by emotional compared to non-emotional stimuli, on which individuals with higher IS may rely more. These individuals may therefore also experience greater benefits in emotion recognition and adaption to the emotional context.

In conclusion, the results of Study I suggest partially and rather weak domain-spanning effects in predictive performance as a function of IS. IS appears to play a greater role in the recognition and prediction of emotion than in gender recognition. Given the considerations of recent studies suggesting that all three modalities affect inferences in other domains (e.g., Allen et al., 2019), it might nevertheless prove interesting for future studies to investigate if interoceptive abilities affect performance in non-emotional tasks that require, for example, stability of predictions and flexible responding to prediction errors (e.g. Trempler et al., 2017).

4.5 Emotion Inference in Frontotemporal Dementia

The symptoms of bvFTD in early stages are primarily characterized by altered personality and behaviour in the form of socially inappropriate behaviour, loss of empathy and

manners, while prominent cognitive deficits occur later in the course of the disease (e.g., Silveri, 2007; cf. section 1.4.1). It is conceivable that the characteristic personality and behavioural changes in early stages of bvFTD might be based on impaired interoception, as also suggested by recent neuropsychiatric commentaries (e.g., Mendez, 2021). The investigation of a possible impairment in the recognition of emotional and non-emotional stimuli in probabilistic contexts could provide new insights that are essential for a better understanding of the neuropsychological profile of bvFTD patients (see e.g., Manoochehri & Huey, 2012). Accordingly, Study II investigated disease-specific difficulties in emotional change detection and context-sensitive emotional inference. In the following section, I will discuss the results of Study II with regard to a possible impairment of interoception in bvFTD and in relation to the results of Study I and Study III.

4.5.1 Recognition of Emotional Changes in Frontotemporal Dementia

The results of Study II revealed that bvFTD patients were slower in recognizing emotions and further had greater difficulty in recognizing emotions in more difficult conditions, i.e., patients showed a gradual increase in response times from happy expressions of high intensity to fearful expressions of low intensity. In contrast, healthy controls showed prolonged reaction times only for the most difficult condition, i.e., low intensity of fearful expressions compared to all other conditions. The findings of deficient emotion recognition in bvFTD are in line with a large body of studies reporting deficits in verbal categorization of facial expressions (for reviews, see Kumfor, Hazelton, Winter, Langavant, & Stock, 2017; Kumfor & Piguet, 2012). Difficulties in emotion recognition have also been reported under different stimuli modalities and task conditions, comprising vocal (Snowden et al., 2008), physical (Van den Stock et al., 2015) and musical (Omar et al., 2011) emotional expressions. Moreover, by using short film clips to identify the main characters' emotions, Goodkind and colleagues (2015) found that patients with bvFTD compared to healthy controls have deficits in emotion

recognition, which were particularly pronounced for negative emotions. Greater difficulties in the processing of negative emotions compared to happy expressions were repeatedly observed in bvFTD patients in previous studies (see Bora, Velakoulis, & Walterfang, 2016 for a meta-analysis).

Beyond impaired emotion recognition, studies indicate that bvFTD patients cannot accurately detect interoceptive signals or assess internal sensations (e.g., Balconi et al., 2015; Eckart, Sturm, Miller, & Levenson, 2012; García-Cordero et al., 2016). Thus, observed deficits in Study II may be attributable to an impairment of interoceptive abilities in bvFTD. Some recent studies have already interpreted results on deficits in emotion recognition in bvFTD against the background of impaired interoceptive abilities (e.g., Adolfi et al., 2017; Mendez, 2021; Van den Stock & Kumfor, 2019). The comparison of the results of Study II with the findings of Study I also supports the notion that deficient emotion recognition in bvFTD is based on impaired interoceptive processing. The results of Study I showed an equivalent reaction time pattern in subjects with low compared to high IS, as did the results of Study II for patients with bvFTD compared to healthy controls: individuals with reduced IS were slower in recognizing low compared to high emotional intensity and fearful compared to happy expressions, whereas individuals with higher IS exhibited slower responses only in the most difficult condition, i.e. when a fearful face with low intensity was presented. Hence, like individuals with low IS, patients with bvFTD appear to be less challenged by positive valence and high intensity emotional expressions and to be able to partially compensate for their interoceptive deficits. Potential interoceptive deficits in bvFTD patients seem to be most evident in emotions in which uncertainty about the change in the emotional stimulus is higher. As discussed in Study II, characteristics of bvFTD like difficulties in understanding social cues, inappropriate behaviour in social interactions and diminished empathy, may therefore be related to an inability to recognize relatively weak expressions that are common in everyday life,

encompassing moderate negative reactions of others to one's own behaviour (e.g., Matsumoto & Hwang, 2011; Motley & Camden, 1988).

Recent research that has linked deficits in emotion recognition in bvFTD to impaired interoceptive abilities (e.g., Adolfi et al., 2017; Mendez, 2021; Van den Stock & Kumfor, 2019) mainly referred to dysfunction of the insula in bvFTD. It is assumed that a weak emotional signal favours insufficient interoceptive accuracy, leading to overreliance on cognitive evaluation rather than internal sensations (Mendez, 2021). Early-stage fronto-insular degeneration has been repeatedly described in bvFTD patients and is associated with a 74% decrease in VENs (Kim et al., 2012; Mandelli et al., 2016; Seeley et al., 2008; cf. section 1.4.1). Previous findings indicate that emotion recognition and interoceptive abilities are impaired in bvFTD and each of these deficits has been independently linked to the structural integrity of the insula (see Van den Stock & Kumfor, 2019 for a discussion and review). Consistently, studies found impaired interoceptive accuracy in bvFTD patients that was neuroanatomically related to grey matter volume in the anterior insula and fronto-temporal regions (García-Cordero et al., 2016). The results of Study III indicated that the AIC plays a major role especially in the recognition of less intense emotions (see section 4.2.2). It is important to emphasize that patients suffering from bvFTD exhibit early-stage degeneration in precisely those neural regions, which in Study III were found to be involved in the recognition of more challenging emotional conditions, i.e., ACC and fronto-insular degeneration and selective destruction of VENs (Rosen et al., 2002; cf. section 4.2.2). The synopsis of these findings points to deficient emotion recognition of particularly difficult-to-recognize emotions in bvFTD, which might be related to ACC and fronto-insular degeneration and impaired interoceptive processing. However, this assumption ought to be tested in future studies.

It is worth noticing that a number of factors might potentially contribute to patient-specific differences in interoceptive abilities, including the exact stage of the disease and the extent of atrophy or hypo-metabolism involving the insula (Kumfor, Hazelton, Rushby,

Hodges, & Piguet, 2019). To better account for these aspects, future studies should collect structural and functional data or include an assessment of patients' current interoceptive abilities (e.g., assessed by a simple version of the HDT) and relate it directly to performance in emotion recognition.

Regarding the modulation of physiological parameters, Study II revealed neither general nor valence or intensity specific differences in SCR amplitude between bvFTD patients and healthy controls. Contrary to these results, previous studies found decreased psychophysiological arousal in response to emotional stimuli in addition to impaired emotion recognition in patients with bvFTD (e.g., Joshi et al., 2014; Sturm et al., 2018). However, the results of Study II are in line with the assumption presented in section 4.3.1, that SCR amplitude is more likely to be modulated as a function of surprise and predictability. Correspondingly, a modulation of SCR amplitude by entropy was observed in bvFTD patients compared to healthy controls, as will be discussed in the next section.

4.5.2 Context-Sensitive Emotional Inference in Frontotemporal Dementia

Contrary to expectations, the results of Study II revealed that both healthy controls and patients with bvFTD recognized emotions faster when its valence was more surprising. Since the modulation of response times by predictability did not differ between the groups, the assumption of impaired adaptation of emotional predictions in patients with bvFTD cannot be initially confirmed by the behavioural findings. Comparing the results of Study II with Study I and Study III, it can be seen that healthy controls as well as bvFTD patients showed a similar response pattern as found for individuals with lower IS, i.e., shorter reaction times for unexpected emotions. Relative to Study I and III, healthy controls in Study II would have been expected to show a more precise adaptation of emotional predictions, resulting in decreased reaction times with increasing predictability rather than with more surprising emotions. As discussed in Study I, it is possible that with lower interoceptive abilities emotional classification

processes are less controlled and unexpected emotions attract greater attention and trigger a faster, hasty response. Especially differences in IS might manifest in differential adaptation to emotion probabilities (cf. section 4.3.1). Accordingly, the findings could also be related to generally low IS in the healthy control group in Study II. This would also be supported by recent studies indicating that interoception decreases with age (Murphy, Geary, Millgate, Catmur, & Bird, 2018). However, interoceptive abilities of healthy controls were not assessed in Study II. Future studies should include an assessment of IS or IAcc for a healthy control group to control for possible effects of interoceptive traits.

As hypothesised and in contrast to the behavioural results, Study II revealed a modulation of physiological parameters by entropy. Healthy controls, but not patients with bvFTD, showed greater SCR amplitude in response to less predictable emotions. Since it is assumed that patients have lower interoceptive abilities compared to healthy individuals, these effects fit the results of Study I, which showed an increase of SCRs as a function of entropy for individuals with high, but not low IS. As already discussed in Study I, the increase of SCR as a function of entropy is likely to reflect an adaptation to biased probabilities and a corresponding decrease of physiological arousal in response to more predictable emotions, whereas heightened arousal occurs in response to violations (see also e.g., Bradley, 2009). The fact that no corresponding modulation of SCR was observed in patients with bvFTD might thus be attributable to impaired attentional control of interoceptive signals and interoceptive prediction errors. This assumption is also supported by the fact that the early focus of neuropathology in bvFTD is in the insula, ACC, and the salience network, which are proposed to play a key role in interoceptive predictions and prediction errors (e.g., Mendez, 2021; cf. section 1.2.3). Studies report impaired interoceptive accuracy in bvFTD patients associated with alterations in grey matter volume in the anterior insula (García-Cordero et al., 2016; cf. section 4.5.1), as well as altered pain and temperature perception associated with grey matter volume in the right middle and posterior insula (e.g., Adolfi et al., 2017; Fletcher et al., 2015; see Van den Stock & Kumfor,

2019 for a discussion and review). Consistently, the results of Study III suggest that in the paradigm used in the studies composing this thesis, the posterior insula as a mapping area of interoceptive signals may be crucial for a precise adaptation of emotion predictions (see section 4.3.2). This, in turn, would be consistent with the fact that bvFTD patients, who have early-stage neuropathology in the insula, have difficulty in probabilistic learning of emotions, as insula-mediated interoceptive signals can no longer be used to guide a context-sensitive behavioural adaptation (Mendez, 2021). However, further research is warranted, since the modulation of response times by predictability at the behavioural level did not differ between groups and the assumption of an impaired adaptation of emotional predictions in bvFTD could only partly be confirmed. Here, it is important to consider that only a very small number of patients and controls (seven participants per group) could be included in Study II, which makes further studies essential to validate the findings more thoroughly.

4.5.3 Domain-Specific Impairments in Frontotemporal Dementia

Regarding impaired interoception in bvFTD it is furthermore important to investigate whether the effects can be transferred to other domains and rather reflect a general deficit in face discrimination or implicit probabilistic adaptation. The results for gender discrimination in Study II showed no differences between healthy controls and bvFTD patients in response time depending on gender or the degree of resolution. While gender recognition in bvFTD patients has not yet been investigated in previous studies, the results are in line with previous research that included a control task for recognition of identity to assess emotion specificity for recognition deficits in bvFTD (e.g., Van den Stock et al., 2015). The findings of these studies revealed impaired recognition of static and dynamic body and facial expressions in bvFTD, whereas identity processing was intact. There are some studies that equally point to an impairment in emotion recognition in bvFTD patients while facial identity processing is preserved (e.g., Keane, Calder, Hodges, & Young, 2002), but there is also evidence suggesting

that face perception deficits are underestimated in bvFTD patients (e.g., Kumfor et al., 2015). Indeed, the behavioural results of Study II suggest that the impairment in discriminating facial expressions in bvFTD is specific to emotion recognition.

Regarding the probabilistic context adaptation, unlike the behavioural data, the results for a modulation of SCR revealed an interaction effect between group and entropy. The interaction was rather unexpected and opposite to the effect observed in the emotion condition, i.e., SCRs increased as a function of predictability in patients but not in healthy controls. This finding may indicate that patients are influenced by the probabilistic structure of the task, but not if the probability implementation involves emotional categories. Since the interaction was evident only in physiological parameters and was rather weak, future studies are needed to further investigate implicit probabilistic context adaptation in bvFTD in other domains.

In summary, the results suggest difficulties in emotion recognition in bvFTD especially for low-intensity emotions of negative valence. The results provide preliminary indications that deficient emotion recognition in bvFTD is based on impaired interoceptive abilities. It is plausible that impaired interoceptive abilities and resulting difficulties in emotion recognition contribute to the abnormal social behaviour that is characteristic for this disease (see also Keane et al., 2002). Finally, physiological parameters provide first evidence of difficulty in probabilistic emotion learning in bvFTD.

Since symptoms of bvFTD in early stages are primarily characterized by altered personality and behaviour, the correct diagnosis of bvFTD is often delayed and misdiagnoses as psychiatric disorders are common (Galimberti, Dell'Osso, Altamura, & Scarpini, 2015; Katisko et al., 2019). Of importance in this regard is the lack of specificity of current diagnostic criteria, making studies on the neuropsychological profile of bvFTD patients essential for improving early diagnosis, symptomatic intervention and disease management (Chow, Pio, & Rockwood, 2011; Manoochehri & Huey, 2012; Rascovsky et al., 2011). Even though Study II included only a small sample size and the results should be interpreted with caution, the results

provide crucial indications for aberrant interoceptive processing, emotion recognition and probabilistic emotion context adaptation in bvFTD patients, offering a starting point for future studies.

4.6 Critical Reflection and Practical Applications

The studies presented in this thesis were conducted to gain new empirical insights into effects of interoceptive predictive processing on emotion recognition and adaptation to emotional contexts. However, when studying these relations, a few challenges arise that deserve a further critical reflection. Throughout the previous sections, a number of ideas for further studies that would strengthen the findings and interpretations of the current studies have already been presented. The critical concerns addressed in this section will provide additional directions for future research and point to practical applications of the studies' findings.

One initial issue to consider is the difficulty to quantify interoceptive abilities. A fundamental limitation of the MAIA-2, and of course other questionnaires assessing IS, is that they are self-reported rather than objective. As a result, respondents inevitably report deviations from their own baseline, which means that intraindividual rather than interindividual variability is captured (Mehling et al., 2012). However, a major strength and the reason why the MAIA-2 was used for the studies is its multidimensionality, which allows for a more differentiated assessment of crucial psychological aspects of the perception and evaluation of body sensations (Mehling et al., 2012). The heartbeat detection task (HDT; Schandry, 1981) has emerged as a widely used approach to objectively assess IAcc. The HDT was chosen in Study III to ensure comparability with other studies on IAcc and also because convincing evidence of its validity exists, particularly for individuals with good interoceptive abilities (e.g., Khalsa, Rudrauf, Feinstein, & Tranel, 2009; Pollatos & Schandry, 2004). Nevertheless, the method has been increasingly criticized in recent studies (see Coll, Hobson, Bird, & Murphy, 2021 for a meta-analysis). Several factors have been shown to substantially influence the results obtained with

the HDT (Murphy, Brewer, Hobson, Catmur, & Bird, 2018). For example, HDT scores are thought to be largely contaminated by the influence of non-interoceptive processes (Desmedt, Luminet, & Corneille, 2018), as well as strongly tied to heart rates (e.g., Zamariola et al., 2018) and beliefs about one's own resting heart rate (e.g., Murphy et al., 2018; Ring & Brener, 2018). More precisely, studies show that the manipulation of participants' beliefs about their own resting heart rate alters HDT scores (Ring, Brener, Knapp, & Mailloux, 2015) and accurate knowledge of average heart rate correlates with improved performance (Murphy, Geary, et al., 2018). Accordingly, it may be beneficial for future studies to include alternative interoceptive tasks. Tasks that might be considered for use in future studies could be other variants of the HDT that better account of the critical aspects mentioned (e.g., Brener & Ring, 2016), or tasks that focus on the detection of other bodily signals outside of the cardiac domain, such as interoceptive accuracy of respiration (e.g., Van Den Houte et al., 2021).

The studies included in this thesis are particularly characterized by the fact that dynamically unfolding videos of emotion development were produced to enable the assessment of emotion recognition as closely as possible to the real world. A more realistic implementation of emotion recognition and the studying of dynamic emotions is increasingly demanded (e.g., Barrett, Adolphs, Marsella, Martinez, & Pollak, 2019). It should be noted, however, that this approach likewise addresses only one modality of emotional information, that is, the visual perception of emotional facial expressions without considering the social context (see also Goodkind et al., 2015). Emotional experience is thought to rely on interoceptive input, but the resulting emotion perception (or experience) is also thought to be significantly shaped by contextual emotion-related interpretations (e.g., Burleson & Quigley, 2021). In line, especially in subtypes of frontotemporal dementia, it was shown that the context significantly modulates emotion processing (e.g., Kumfor et al., 2018). The need for research examining the derivation of emotional meanings from facial movements and other social information in contexts of everyday life has only recently been emphasized (Barrett et al., 2019). To incorporate social

contexts more fully in the examination of effects on differences in interoceptive abilities on emotion recognition, for future studies it might be worth considering to include the presentation of short film clips to identify the main characters' emotions (e.g., Goodkind et al., 2015).

Lastly, a caveat of the cross-sectional nature of the presented research is uncertain causality. Due to the complex interplay of cognitive and physiological factors involved in emotional processes, it remains a great challenge to elucidate the causal directionality of emotional processes through correlations between interoceptive abilities and affective states (see also Nord & Garfinkel, 2022; Schoeller, Haar, Jain, & Maes, 2019). In particular, when drawing conclusions regarding impaired interoception in bvFTD patients as compared to other findings, this aspect should be considered in the present work. Recent studies have attempted to address the problem of causality in emotional processes by introducing technologies that allow interference with underlying sensory pathways through controlled somatosensory stimulation (i.e., interoceptive illusions) (Schoeller et al., 2019).

Studies of interoceptive processing are particularly important since disruptions to the interoceptive system may contribute to a variety of clinical conditions, including depression, anxiety and psychosis (Khalsa et al., 2018). Beyond this, interoception may be an important target of existing and novel mental health treatments (Nord & Garfinkel, 2022), such as improving interoceptive awareness in the context of body-oriented therapy (Price & Hooven, 2018). To approach these practical implications, it is essential to gain a better understanding of the interoceptive pathways and the underlying neurophysiology especially in the recognition and inference of emotion, to which the studies in this thesis, despite their limitations, make an important contribution.

5. Conclusion

The studies composing this dissertation are the first to examine emotion recognition in a predictive setting as a function of differences in interoceptive abilities. The studies provide evidence that IS facilitates emotion recognition and entails a more precise adaptation to biased probabilities of emotions, potentially mediated by an increased attention to bodily signals. The results provide indications that, in higher IS individuals, adaptation to biased probabilities of emotion occurrence is also reflected in physiological arousal. Moreover, the findings revealed deficient emotion recognition in bvFTD, which may be attributable to impaired interoception. Furthermore, deviant modulation of sympathetic arousal points to impaired processing of interoceptive prediction errors in bvFTD patients. Importantly, more challenging emotional conditions engage brain regions associated with interoceptive processing (i.e., AIC, ACC). Regions that integrate interoceptive information, including the posterior insula, were recruited during processing of emotion predictions, suggesting a role of the mapping of interoceptive signals in emotional context adaptation. Lastly, the results provide evidence for a contribution of the insula and striatal cortex in processing interoceptive prediction errors that is amplified in individuals with higher IAcc.

The presented studies contribute to the understanding of the role of interoception in emotion recognition and emotional context adaptation and further elucidate the underlying neural networks and physiological correlates in healthy individuals and patients with bvFTD. The results raise intriguing questions for future research, with the systematic disentanglement of the neural mechanisms of interoceptive predictive processing in emotion recognition being a particularly noteworthy aspect that warrants further investigation.

References

Adolfi, F., Couto, B., Richter, F., Decety, J., Lopez, J., Sigman, M., … Ibáñez, A. (2017). Convergence of interoception, emotion, and social cognition: A twofold fMRI meta-analysis and lesion approach. *Cortex*, *88*, 124–142. https://doi.org/10.1016/j.cortex.2016.12.019

Ahlheim, C., Stadler, W., & Schubotz, R. I. (2014). Dissociating dynamic probability and predictability in observed actions-an fMRI study. *Frontiers in Human Neuroscience*, *8*(MAY), 1–13. https://doi.org/10.3389/fnhum.2014.00273

Ainley, V., Apps, M. A. J., Fotopoulou, A., & Tsakiris, M. (2016). "Bodily precision": A predictive coding account of individual differences in interoceptive accuracy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1708). https://doi.org/10.1098/rstb.2016.0003

Ainley, V., Brass, M., & Tsakiris, M. (2014). Heartfelt imitation: high interoceptive awareness is linked to greater automatic imitation. *Neuropsychologia*, *60*, 21–28.

Allen, M., Frank, D., Schwarzkopf, D. S., Fardo, F., Winston, J. S., Hauser, T. U., & Rees, G. (2016). Unexpected arousal modulates the influence of sensory noise on confidence. *Elife*, *5*, e18103.

Allen, M., Levy, A., Parr, T., & Friston, K. J. (2019). In the Body's Eye: The Computational Anatomy of Interoceptive Inference. *BioRxiv*, 1–37. https://doi.org/10.1101/603928

Allman, J. M., Tetreault, N. A., Hakeem, A. Y., Manaye, K. F., Semendeferi, K., Erwin, J. M., … Hof, P. R. (2011). The von Economo neurons in the frontoinsular and anterior cingulate cortex. *Annals of the New York Academy of Sciences*, *1225*(1), 59–71.

Allman, J. M., Watson, K. K., Tetreault, N. A., & Hakeem, A. Y. (2005). Intuition and autism: A possible role for Von Economo neurons. *Trends in Cognitive Sciences*, *9*(8), 367–373. https://doi.org/10.1016/j.tics.2005.06.008

Alzheimer, A. (1911). Uber eigenartige Krankheitsfalle des spateren Alters. *Psychiatr Nervenkr Z Gesamte Neurol Psychiatr*, *4*, 356–385.

Arnold, A. J., Winkielman, P., & Dobkins, K. (2019). Interoception and Social Connection. *Frontiers in Psychology*, *10*(November), 1–6. https://doi.org/10.3389/fpsyg.2019.02589

Bach, D. R., Flandin, G., Friston, K. J., & Dolan, R. J. (2010). Modelling event-related skin conductance responses. *International Journal of Psychophysiology*, *75*(3), 349–356.

Balconi, M., Cotelli, M., Brambilla, M., Manenti, R., Cosseddu, M., Premi, E., … Borroni, B. (2015). Understanding emotions in frontotemporal dementia: The explicit and implicit emotional cue mismatch. *Journal of Alzheimer's Disease*, *46*(1), 211–225. https://doi.org/10.3233/JAD-142826

Barbalet, J. M. (1999). William James' theory of emotions: Filling in the picture. *Journal for the Theory of Social Behaviour*, *29*(3), 251–266.

Barrett, A. B., Dienes, Z., & Seth, A. K. (2013). Measures of metacognition on signal-detection theoretic models. *Psychological Methods*, *18*(4), 535.

Barrett, L. F. (2017a). *How emotions are made: The secret life of the brain*. Pan Macmillan.

Barrett, L. F. (2017b). The theory of constructed emotion: an active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience*, *12*(1), 1–23.

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest*, *20*(1), 1–68.

Barrett, L. F., Quigley, K. S., Bliss-Moreau, E., & Aronson, K. R. (2004). Interoceptive sensitivity and self-reports of emotional experience. *Journal of Personality and Social Psychology*, *87*(5), 684.

Barrett, L. F., Quigley, K. S., & Hamilton, P. (2016). An active inference theory of allostasis and interoception in depression. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1708). https://doi.org/10.1098/rstb.2016.0011

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

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, *10*(3), 295–307.

Berking, M., & Znoj, H. (2008). Entwicklung und Validierung eines Fragebogens zur standardisierten Selbsteinschätzung emotionaler Kompetenzen (SEK-27). *Zeitschrift Für Psychiatrie, Psychologie Und Psychotherapie*, *56*(2), 141–153.

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

Binney, R. J., Pankov, A., Marx, G., He, X., McKenna, F., Staffaroni, A. M., … Rosen, H. J. (2017). Data-driven regions of interest for longitudinal change in three variants of frontotemporal lobar degeneration. *Brain and Behavior*, *7*(4), 1–11. https://doi.org/10.1002/brb3.675

Boldt, A., De Gardelle, V., & Yeung, N. (2017). The impact of evidence reliability on sensitivity and bias in decision confidence. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(8), 1520.

Bora, E., Velakoulis, D., & Walterfang, M. (2016). Meta-Analysis of Facial Emotion Recognition in Behavioral Variant Frontotemporal Dementia. *Journal of Geriatric Psychiatry and Neurology*, *29*(4), 205–211. https://doi.org/10.1177/0891988716640375

Bora, E., Walterfang, M., & Velakoulis, D. (2015). Theory of mind in behavioural-variant frontotemporal dementia and Alzheimer's disease: A meta-analysis. *Journal of Neurology, Neurosurgery and Psychiatry*, *86*(7), 714–719. https://doi.org/10.1136/jnnp-2014-309445

Bott, N. T., Radke, A., Stephens, M. L., & Kramer, J. H. (2014). Frontotemporal dementia: diagnosis, deficits and management. *Neurodegenerative Disease Management*, *4*(6), 439–454. https://doi.org/10.2217/nmt.14.34.Frontotemporal

Bradley, M. M. (2009). Natural selective attention: Orienting and emotion. *Psychophysiology*, *46*(1), 1–11. https://doi.org/10.1111/j.1469-8986.2008.00702.x

Brener, J., & Kluvitse, C. (1988). Heartbeat detection: judgments of the simultaneity of external stimuli and heartbeats. *Psychophysiology*, *25*(5), 554–561.

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.

Brewer, R., Cook, R., & Bird, G. (2016). Alexithymia: A general deficit of interoception. *Royal Society Open Science*, *3*(10). https://doi.org/10.1098/rsos.150664

Brunnström, H., Gustafson, L., Passant, U., & Englund, E. (2009). Prevalence of dementia subtypes: a 30-year retrospective survey of neuropathological reports. *Archives of Gerontology and Geriatrics*, *49*(1), 146–149.

Bubic, A., Yves von Cramon, D., & Schubotz, R. I. (2010). Prediction, cognition and the brain. *Frontiers in Human Neuroscience*, *4*(March), 1–15. https://doi.org/10.3389/fnhum.2010.00025

Burleson, M. H., & Quigley, K. S. (2021). Social interoception and social allostasis through touch: legacy of the somatovisceral afference model of emotion. *Social Neuroscience*, *16*(1), 92–102.

Cauda, F., Torta, D. M. E., Sacco, K., D'Agata, F., Geda, E., Duca, S., … Vercelli, A. (2013). Functional anatomy of cortical areas characterized by von Economo neurons. *Brain Structure and Function*, *218*(1), 1–20. https://doi.org/10.1007/s00429-012-0382-9

Ceunen, E., Van Diest, I., & Vlaeyen, J. (2013). Accuracy and awareness of perception: related, yet distinct (commentary on Herbert et al., 2012). *Biological Psychology*, *92*(2), 423–427.

Ceunen, E., Vlaeyen, J. W. S., & Van Diest, I. (2016). On the origin of interoception. *Frontiers in Psychology*, *7*(MAY), 1–17. https://doi.org/10.3389/fpsyg.2016.00743

Chaudhri, V. K., Pathak, D., Kumar, P., & Yadav, V. (2016). Synthesis, characterization and in vivo anthelmintic activity of some newer benzimidazole derivatives. *Der Pharmacia Lettre*, *8*(3), 127–134. https://doi.org/10.3109/09540261.2013.776523.The

Chong, J. S. X., Ng, G. J. P., Lee, S. C., & Zhou, J. (2017). Salience network connectivity in the insula is associated with individual differences in interoceptive accuracy. *Brain Structure and Function*, *222*(4), 1635–1644.

Chow, T. W., Pio, F. J., & Rockwood, K. (2011). An international needs assessment of caregivers for frontotemporal dementia. *The Canadian Journal of Neurological Sciences. Le Journal Canadien Des Sciences Neurologiques*, *38*(5), 753–757. https://doi.org/10.1017/S0317167100054147

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

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

Craig. (2009). How do you feel--now? The anterior insula and human awareness. *Nature Reviews Neuroscience*, *10*(1).

Craig, A. D. (2002). *how do you feel Craig 2002*. *3*(August).

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

Craig, A. D. (2011). Significance of the insula for the evolution of human awareness of feelings from the body. *Annals of the New York Academy of Sciences*, *1225*(1), 72–82.

Critchley, H. D., & Harrison, N. A. (2013). Visceral influences on brain and behavior. *Neuron*, *77*(4), 624–638.

Critchley, H. D., Wiens, S., Rotshtein, P., Öhman, A., & Dolan, R. J. (2004). Neural systems supporting interoceptive awareness. *Nature Neuroscience*, *7*(2), 189–195. https://doi.org/10.1038/nn1176

Dalton, M. A., Weickert, T. W., Hodges, J. R., Piguet, O., & Hornberger, M. (2013). Impaired acquisition rates of probabilistic associative learning in frontotemporal dementia is associated with fronto-striatal atrophy. *NeuroImage: Clinical*, *2*(1), 56–62. https://doi.org/10.1016/j.nicl.2012.11.001

Damasio, A., & Carvalho, G. B. (2013). The nature of feelings: evolutionary and neurobiological origins. *Nature Reviews Neuroscience*, *14*(2), 143–152.

Damasio, A. R. (1994). Descartes' error and the future of human life. *Scientific American*, *271*(4), 144.

Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *351*(1346), 1413–1420.

Davis, M. H. (1983). Measuring individual differences in empathy: evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, *44*(1), 113.

Dawson, M. E., Schell, A. M., & Filion, D. L. (2007). The electrodermal system In: Cacioppo JT, Tassinary LG, Berntson GG, editors. *Handbook of Psychophysiology*, *2*, 200–223.

Decety, J. (2011). Dissecting the neural mechanisms mediating empathy. *Emotion Review*, *3*(1), 92–108.

Deen, B., Pitskel, N. B., & Pelphrey, K. A. (2011). Three systems of insular functional connectivity identified with cluster analysis. *Cerebral Cortex*, *21*(7), 1498–1506. https://doi.org/10.1093/cercor/bhq186

Descartes, R. (1989). On the passions of the soul (S. Voss, Trans.). *Indianapolis, IN: Hackett. (Original Work Published 1649)*.

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.

Dols, J. M. F., & Russell, J. A. (2017). *The science of facial expression*. Oxford University Press.

Duchaine, B., & Yovel, G. (2015). A Revised Neural Framework for Face Processing. *Annual Review of Vision Science*, *1*(1), 393–416. https://doi.org/10.1146/annurev-vision-082114-035518

Dunn, B. D., Galton, H. C., Morgan, R., Evans, D., Oliver, C., Meyer, M., … 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

Eckart, J. A., Sturm, V. E., Miller, B. L., & Levenson, R. W. (2012). Diminished disgust reactivity in behavioral variant frontotemporal dementia. *Neuropsychologia*, *50*(5), 786–790.

Economo, V. (1926). Eine neue Art Spezialzellen des Lobus cinguli und Lobus insulae. *Zeitschrift Für Die Gesamte Neurologie Und Psychiatrie*, *100*(1), 706–712.

Elamin, M., Pender, N., Hardiman, O., & Abrahams, S. (2012). Social cognition in neurodegenerative disorders: A systematic review. *Journal of Neurology, Neurosurgery and Psychiatry*, *83*(11), 1071–1079. https://doi.org/10.1136/jnnp-2012-302817

Ellsworth, P. C. (1994). William James and emotion: is a century of fame worth a century of misunderstanding? *Psychological Review*, *101*(2), 222.

Ernst, J., Böker, H., Hättenschwiler, J., Schüpbach, D., Northoff, G., Seifritz, E., & Grimm, S. (2013). The association of interoceptive awareness and alexithymia with neurotransmitter concentrations in insula and anterior cingulate. *Social Cognitive and Affective Neuroscience*, *9*(6), 857–863. https://doi.org/10.1093/scan/nst058

Ernst, J., Northoff, G., Böker, H., Seifritz, E., & Grimm, S. (2013). Interoceptive awareness enhances neural activity during empathy. *Human Brain Mapping*, *34*(7), 1615–1624.

Fajardo, C., Escobar, M. I., Buriticá, E., Arteaga, G., Umbarila, J., Casanova, M. F., & Pimienta, H. (2008). Von Economo neurons are present in the dorsolateral (dysgranular) prefrontal cortex of humans. *Neuroscience Letters*, *435*(3), 215–218. https://doi.org/10.1016/j.neulet.2008.02.048

Fan, J., Gu, X., Liu, X., Guise, K. G., Park, Y., Martin, L., … Hof, P. R. (2011). Involvement of the anterior cingulate and frontoinsular cortices in rapid processing of salient facial emotional information. *NeuroImage*, *54*(3), 2539–2546. https://doi.org/10.1016/j.neuroimage.2010.10.007

Farb, N. A. S., Segal, Z. V., & Anderson, A. K. (2013). Attentional modulation of primary interoceptive and exteroceptive cortices. *Cerebral Cortex*, *23*(1), 114–126. https://doi.org/10.1093/cercor/bhr385

Forkmann, T., Scherer, A., Meessen, J., Michal, M., Schächinger, H., Vögele, C., & Schulz, A. (2016). Making sense of what you sense: Disentangling interoceptive awareness, sensibility and accuracy. *International Journal of Psychophysiology*, *109*, 71–80.

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).

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138. https://doi.org/10.1038/nrn2787

Furnham, A., & Petrides, K. V. (2003). Trait emotional intelligence and happiness. *Social Behavior and Personality: An International Journal*, *31*(8), 815–823.

Galimberti, D., Dell'Osso, B., Altamura, A. C., & Scarpini, E. (2015). Psychiatric symptoms in frontotemporal dementia: epidemiology, phenotypes, and differential diagnosis.

Biological Psychiatry, *78*(10), 684–692.

García-Cordero, I., Sedeño, L., de la Fuente, L., Slachevsky, A., Forno, G., Klein, F., … 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). https://doi.org/10.1098/rstb.2016.0006

Garfinkel, S. N., & Critchley, H. D. (2013). Interoception, emotion and brain: new insights link internal physiology to social behaviour. Commentary on:: "Anterior insular cortex mediates bodily sensibility and social anxiety" by Terasawa et al. (2012). *Social Cognitive and Affective Neuroscience*, *8*(3), 231–234. https://doi.org/10.1093/scan/nss140

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

Gentsch, A., Sel, A., Marshall, A. C., & Schütz-Bosbach, S. (2019). Affective interoceptive inference: Evidence from heart-beat evoked brain potentials. *Human Brain Mapping*, *40*(1), 20–33. https://doi.org/10.1002/hbm.24352

Georgiou, E., Mai, S., Fernandez, K. C., & Pollatos, O. (2018). I see neither your fear, nor your sadness–Interoception in adolescents. *Consciousness and Cognition*, *60*, 52–61.

Geuter, S., Boll, S., Eippert, F., & Büchel, C. (2017). Functional dissociation of stimulus intensity encoding and predictive coding of pain in the insula. *ELife*, *6*, 1–22. https://doi.org/10.7554/eLife.24770

Goodkind, M. S., Sturm, V. E., Ascher, E. A., Shdo, S. M., Miller, B. L., Rankin, K. P., & Levenson, R. W. (2015). Emotion recognition in frontotemporal dementia and Alzheimer's disease: A new film-based assessment. *Emotion*, *15*(4), 416.

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, *10*(1), 14–23. https://doi.org/10.1016/j.tics.2005.11.006

Grynberg, D., Chang, B., Corneille, O., Maurage, P., Vermeulen, N., Berthoz, S., & Luminet, O. (2012). Alexithymia and the processing of emotional facial expressions (EFEs): Systematic review, unanswered questions and further perspectives. *PLoS ONE*, *7*(8). https://doi.org/10.1371/journal.pone.0042429

Grynberg, D., & Pollatos, O. (2015). Perceiving one's body shapes empathy. *Physiology and Behavior*, *140*, 54–60. https://doi.org/10.1016/j.physbeh.2014.12.026

Gu, X., Hof, P. R., Friston, K. J., & Fan, J. (2013). Anterior insular cortex and emotional awareness. *Journal of Comparative Neurology*, *521*(15), 3371–3388.

Hampshire, A., Chamberlain, S. R., Monti, M. M., Duncan, J., & Owen, A. M. (2010). The role of the right inferior frontal gyrus: inhibition and attentional control. *NeuroImage*, *50*(3), 1313–1319. https://doi.org/10.1016/j.neuroimage.2009.12.109

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biological Psychiatry*, *51*(1), 59–67. https://doi.org/10.1016/S0006-3223(01)01330-0

Haxby, J. V, Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223–233.

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

Herry, C., Bach, D. R., Esposito, F., Di Salle, F., Perrig, W. J., Scheffler, K., … Seifritz, E. (2007). Processing of temporal unpredictability in human and animal amygdala. *Journal of Neuroscience*, *27*(22), 5958–5966. https://doi.org/10.1523/JNEUROSCI.5218-06.2007

Hughes, L. E., & Rowe, J. B. (2013). The impact of neurodegeneration on network connectivity: a study of change detection in frontotemporal dementia. *Journal of Cognitive Neuroscience*, *25*(5), 802–813.

Jabbi, M., Bastiaansen, J., & Keysers, C. (2008). A common anterior insula representation of disgust observation, experience and imagination shows divergent functional connectivity pathways. *PloS One*, *3*(8), e2939.

Jack, R. E., Garrod, O. G. B., & Schyns, P. G. (2014). Dynamic facial expressions of emotion transmit an evolving hierarchy of signals over time. *Current Biology*, *24*(2), 187–192.

James, W. (1884). What is an Emotion?(188-205). *Mind*, *9*, 34.

Jänig, W. (1996). Neurobiology of visceral afferent neurons: neuroanatomy, functions, organ regulations and sensations. *Biological Psychology*, *42*(1–2), 29–51.

Jiang, J., Summerfield, C., & Egner, T. (2013). Attention sharpens the distinction between expected and unexpected percepts in the visual brain. *Journal of Neuroscience*, *33*(47), 18438–18447. https://doi.org/10.1523/JNEUROSCI.3308-13.2013

Johnen, A., & Bertoux, M. (2019). Psychological and Cognitive Markers of Behavioral Variant Frontotemporal Dementia–A Clinical Neuropsychologist's View on Diagnostic Criteria and Beyond. *Frontiers in Neurology*, *10*(June). https://doi.org/10.3389/fneur.2019.00594

Johnson, J. K., Diehl, J., Mendez, M. F., Neuhaus, J., Shapira, J. S., Forman, M., … Neumann, M. (2005). Frontotemporal lobar degeneration: demographic characteristics of 353 patients. *Archives of Neurology*, *62*(6), 925–930.

Joshi, A., Mendez, M. F., Kaiser, N., Jimenez, E., Mather, M., & Shapira, J. S. (2014). Skin conductance levels may reflect emotional blunting in behavioral variant frontotemporal dementia. *Journal of Neuropsychiatry and Clinical Neurosciences*, *26*(3), 227–232. https://doi.org/10.1176/appi.neuropsych.12110332

Katisko, K., Cajanus, A., Korhonen, T., Remes, A. M., Haapasalo, A., & Solje, E. (2019). Prodromal and early bvFTD: evaluating clinical features and current biomarkers. *Frontiers in Neuroscience*, *13*, 658.

Keane, J., Calder, A. J., Hodges, J. R., & Young, A. W. (2002). Face and emotion processing in frontal variant frontotemporal dementia. *Neuropsychologia*, *40*(6), 655–665.

Khalsa, S. S., Adolphs, R., Cameron, O. G., Critchley, H. D., Davenport, P. W., Feinstein, J. S., … 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., Rudrauf, D., Feinstein, J. S., & Tranel, D. (2009). The pathways of interoceptive awareness. *Nature Neuroscience*, *12*(12), 1494–1496.

Kim, E. J., Sidhu, M., Gaus, S. E., Huang, E. J., Hof, P. R., Miller, B. L., … Seeley, W. W. (2012). Selective frontoinsular von economo neuron and fork cell loss in early behavioral variant frontotemporal dementia. *Cerebral Cortex*, *22*(2), 251–259. https://doi.org/10.1093/cercor/bhr004

Knopman, D. S., & Roberts, R. O. (2011). Estimating the number of persons with frontotemporal lobar degeneration in the US population. *Journal of Molecular Neuroscience*, *45*(3), 330–335.

Köteles, F. (2021). Can We Trust Our Body Feelings? The Discrepancy Between Actual and Perceived Events. In *Body Sensations* (pp. 213–243). Springer.

Kuehn, E., Mueller, K., Lohmann, G., & Schuetz-Bosbach, S. (2016). Interoceptive awareness changes the posterior insula functional connectivity profile. *Brain Structure and Function*, *221*(3), 1555–1571. https://doi.org/10.1007/s00429-015-0989-8

Kumfor, F., Hazelton, J. L., Rushby, J. A., Hodges, J. R., & Piguet, O. (2019). Facial expressiveness and physiological arousal in frontotemporal dementia: phenotypic clinical profiles and neural correlates. *Cognitive, Affective, & Behavioral Neuroscience*, *19*(1), 197–210.

Kumfor, F., Hazelton, J. L., Winter, F.-L. De, Langavant, L. C. de, & Stock, J. Van den. (2017). Clinical studies of social neuroscience: a lesion model approach. *Neuroscience and Social Science*, 255–296.

Kumfor, F., Hutchings, R., Irish, M., Hodges, J. R., Rhodes, G., Palermo, R., & Piguet, O. (2015). Do I know you? Examining face and object memory in frontotemporal dementia. *Neuropsychologia*, *71*, 101–111.

Kumfor, F., Ibañez, A., Hutchings, R., Hazelton, J. L., Hodges, J. R., & Piguet, O. (2018). Beyond the face: How context modulates emotion processing in frontotemporal dementia subtypes. *Brain*, *141*(4), 1172–1185. https://doi.org/10.1093/brain/awy002

Kumfor, F., & Piguet, O. (2012). Disturbance of emotion processing in frontotemporal dementia: a synthesis of cognitive and neuroimaging findings. *Neuropsychology Review*, *22*(3), 280–297.

Lange, C. (1922). The emotions. *In K. Dunlap (Ed.), The Emotions (I. A. Haupt, Trans.; Pp. 33-90). Baltimore: Williams & Wilkins. (Original Work Published 1885).*

Leppänen, J. M., & Hietanen, J. K. (2003). Affect and Face Perception: Odors Modulate the Recognition Advantage of Happy Faces. *Emotion*, *3*(4), 315–326. https://doi.org/10.1037/1528-3542.3.4.315

Leppänen, J. M., & Hietanen, J. K. (2004). Positive facial expressions are recognized faster than negative facial expressions, but why? *Psychological Research*, *69*(1–2), 22–29. https://doi.org/10.1007/s00426-003-0157-2

Mandelli, M. L., Vitali, P., Santos, M., Henry, M., Gola, K., Rosenberg, L., … Gorno-Tempini, M. L. (2016). Two insular regions are differentially involved in behavioral variant FTD and nonfluent/agrammatic variant PPA. *Cortex*, *74*, 149–157. https://doi.org/10.1016/j.cortex.2015.10.012

Manoochehri, M., & Huey, E. D. (2012). Diagnosis and management of behavioral issues in frontotemporal dementia. *Current Neurology and Neuroscience Reports*, *12*(5), 528–536. https://doi.org/10.1007/s11910-012-0302-7

Marshall, C. R., Hardy, C. J. D., Russell, L. L., Clark, C. N., Bond, R. L., Dick, K. M., … Warren, J. D. (2018). Motor signatures of emotional reactivity in frontotemporal dementia. *Scientific Reports*, *8*(1), 1–13. https://doi.org/10.1038/s41598-018-19528-2

Marshall, C. R., Hardy, C. J. D., Russell, L. L., Clark, C. N., Dick, K. M., Brotherhood, E. V., … Warren, J. D. (2017). Impaired interoceptive accuracy in semantic variant primary progressive aphasia. *Frontiers in Neurology*, *8*(NOV), 1–6. https://doi.org/10.3389/fneur.2017.00610

Matsumoto, D., & Hwang, H. C. (2014). Judgments of subtle facial expressions of emotion. *Emotion*, *14*(2), 349.

Matsumoto, D., & Hwang, H. S. (2011). Judgments of facial expressions of emotion in profile. *Emotion*, *11*(5), 1223.

Mayer, E. A. (2011). Gut feelings: The emerging biology of gut-"brain communication. *Nature Reviews Neuroscience*, *12*(8), 453–466. https://doi.org/10.1038/nrn3071

McKhann, G. M., Albert, M. S., Grossman, M., Miller, B., Dickson, D., & Trojanowski, J. Q. (2001). Clinical and pathological diagnosis of frontotemporal dementia: report of the Work Group on Frontotemporal Dementia and Pick's Disease. *Archives of Neurology*, *58*(11), 1803–1809.

Medford, N., & Critchley, H. D. (2010). Conjoint activity of anterior insular and anterior cingulate cortex: awareness and response. *Brain Structure and Function*, *214*(5–6), 535–549.

Meessen, J., Mainz, V., Gauggel, S., Volz-Sidiropoulou, E., Sütterlin, S., & Forkmann, T. (2016). The relationship between interoception and metacognition. *Journal of Psychophysiology*.

Mehling, W., Acree, M., Stewart, A., Silas, J., & Jones, A. (2018). The multidimensional assessment of interoceptive awareness, version 2 (MAIA-2). *PLoS ONE*, *13*(12), 1–12. 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.

Mendez, M. F. (2021). Frontotemporal dementia: A window to alexithymia. *Journal of Neuropsychiatry and Clinical Neurosciences*, *33*(2), 157–160. https://doi.org/10.1176/appi.neuropsych.20100252

Menon, V., & Uddin, L. Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Structure and Function*, *214*(5–6), 655–667.

Motley, M. T., & Camden, C. T. (1988). Facial expression of emotion: A comparison of posed expressions versus spontaneous expressions in an interpersonal communication setting. *Western Journal of Communication (Includes Communication Reports)*, *52*(1), 1–22.

Motomura, K., Terasawa, Y., Natsume, A., Iijima, K., Chalise, L., Sugiura, J., … Umeda, S. (2019). Anterior insular cortex stimulation and its effects on emotion recognition. *Brain Structure and Function*, *224*(6), 2167–2181. https://doi.org/10.1007/s00429-019-01895-9

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.

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.

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.

Neary, D., Snowden, J. S., Gustafson, L., Passant, U., Stuss, D., Black, S., … Albert, M. (1998). Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria. *Neurology*, *51*(6), 1546–1554.

Nord, C. L., & Garfinkel, S. N. (2022). Interoceptive pathways to understand and treat mental health conditions. *Trends in Cognitive Sciences*, *xx*(xx), 1–15. https://doi.org/10.1016/j.tics.2022.03.004

O'Callaghan, C., Bertoux, M., Irish, M., Shine, J. M., Wong, S., Spiliopoulos, L., … Hornberger, M. (2016). Fair play: Social norm compliance failures in behavioural variant frontotemporal dementia. *Brain*, *139*(1), 204–216. https://doi.org/10.1093/brain/awv315

Omar, R., Henley, S. M. D., Bartlett, J. W., Hailstone, J. C., Gordon, E., Sauter, D. A., … Warren, J. D. (2011). The structural neuroanatomy of music emotion recognition: evidence from frontotemporal lobar degeneration. *Neuroimage*, *56*(3), 1814–1821.

Ondobaka, S., Kilner, J., & Friston, K. (2017). The role of interoceptive inference in theory of mind. *Brain and Cognition*, *112*, 64–68. https://doi.org/10.1016/j.bandc.2015.08.002

Pasquier, F., & Petit, H. (1997). Frontotemporal dementia: its rediscovery. *European Neurology*, *38*(1), 1–6.

Paulus, C. (2009). Der Saarbrücker Persönlichkeitsfragebogen SPF (IRI) zur Messung von Empathie. *Psychometrische Evaluation Der Deutschen Version Des Interpersonal Reactivity Index*.

Petzschner, F. H., Weber, L. A., Wellstein, K. V., Paolini, G., Do, C. T., & Stephan, K. E. (2019). Focus of attention modulates the heartbeat evoked potential. *NeuroImage*, *186*(November 2018), 595–606. https://doi.org/10.1016/j.neuroimage.2018.11.037

Pick, A. (1892). Uber die Beziehungen der senilen Hirnatrophie zur Aphasie. *Prag Med Wochenschr*, *17*, 165–167.

Piguet, O., Hornberger, M., Mioshi, E., & Hodges, J. R. (2011). Behavioural-variant frontotemporal dementia: diagnosis, clinical staging, and management. *The Lancet Neurology*, *10*(2), 162–172.

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.

Pollatos, O., Kirsch, W., & Schandry, R. (2005). On the relationship between interoceptive awareness, emotional experience, and brain processes. *Cognitive Brain Research*, *25*(3), 948–962. https://doi.org/10.1016/j.cogbrainres.2005.09.019

Pollatos, O., & Schandry, R. (2004). Accuracy of heartbeat perception is reflected in the amplitude of the heartbeat-evoked brain potential. *Psychophysiology*, *41*(3), 476–482. https://doi.org/10.1111/1469-8986.2004.00170.x

Pollatos, O., Schandry, R., Auer, D. P., & Kaufmann, C. (2007). Brain structures mediating cardiovascular arousal and interoceptive awareness. *Brain Research*, *1141*, 178–187.

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

Price, C. J., & Hooven, C. (2018). Interoceptive awareness skills for emotion regulation: Theory and approach of mindful awareness in body-oriented therapy (MABT). *Frontiers in Psychology*, *9*, 798.

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*, 112–128. https://doi.org/10.1111/nyas.13915

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87.

Rascovsky, K., Hodges, J. R., Knopman, D., Mendez, M. F., Kramer, J. H., Neuhaus, J., … Hillis, A. E. (2011). Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia TT - Empfindlichkeit der revidierten Diagnosekriterien der Verhaltensvariante der frontotemporalen Demenz. *Brain*, *134*(9), 1–22. https://doi.org/10.1093/brain/awr179

Ratnavalli, E., Brayne, C., Dawson, K., & Hodges, J. R. (2002). The prevalence of frontotemporal dementia. *Neurology*, *58*(11), 1615–1621.

Ring, C., & Brener, J. (2018). Heartbeat counting is unrelated to heartbeat detection: A comparison of methods to quantify interoception. *Psychophysiology*, *55*(9), 1–10. https://doi.org/10.1111/psyp.13084

Ring, C., Brener, J., Knapp, K., & Mailloux, J. (2015). Effects of heartbeat feedback on beliefs about heart rate and heartbeat counting: a cautionary tale about interoceptive awareness. *Biological Psychology*, *104*, 193–198.

Rosen, H. J., Gorno–Tempini, M. L., Goldman, W. P., Perry, R. J., Schuff, N., Weiner, M., … Miller, B. L. (2002). Patterns of brain atrophy in frontotemporal dementia and semantic dementia. *Neurology*, *58*(2), 198–208.

Sato, W., Kochiyama, T., Yoshikawa, S., Naito, E., & Matsumura, M. (2004). Enhanced neural activity in response to dynamic facial expressions of emotion: an fMRI study. *Cognitive Brain Research*, *20*(1), 81–91.

Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review*, *69*(5), 379.

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

Schiffer, A.-M., & Schubotz, R. I. (2011). Caudate nucleus signals for breaches of expectation in a movement observation paradigm. *Frontiers in Human Neuroscience*, *5*, 38.

Schiffer, A. M., Ahlheim, C., Wurm, M. F., & Schubotz, R. I. (2012a). Surprised at all the

entropy: Hippocampal, caudate and midbrain contributions to learning from prediction errors. *PLoS ONE*, *7*(5). https://doi.org/10.1371/journal.pone.0036445

Schiffer, A. M., Ahlheim, C., Wurm, M. F., & Schubotz, R. I. (2012b). Surprised at all the entropy: Hippocampal, caudate and midbrain contributions to learning from prediction errors. *PLoS ONE*, *7*(5). https://doi.org/10.1371/journal.pone.0036445

Schoeller, F., Haar, A. J. H., Jain, A., & Maes, P. (2019). Enhancing human emotions with interoceptive technologies. *Physics of Life Reviews*, *31*, 310–319. https://doi.org/10.1016/j.plrev.2019.10.008

Schubotz, R. I. (2015). Prediction and expectation. In *A. W. Toga (Ed.), Brain mapping: An encyclopedic reference (Vol. 3, pp. 295–302). Oxford: Academic Press/Elsevier.*

Schulz, S. M. (2016). Neural correlates of heart-focused interoception: a functional magnetic resonance imaging meta-analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1708), 20160018.

Seeley, W. W. (2010). Anterior insula degeneration in frontotemporal dementia. *Brain Structure and Function*, *214*(5), 465–475.

Seeley, W. W., Carlin, D. A., Allman, J. M., Macedo, M. N., Bush, C., Miller, B. L., & DeArmond, S. J. (2006). Early frontotemporal dementia targets neurons unique to apes and humans. *Annals of Neurology*, *60*(6), 660–667. https://doi.org/10.1002/ana.21055

Seeley, W. W., Crawford, R., Rascovsky, K., Kramer, J. H., Weiner, M., Miller, B. L., & Gorno-Tempini, M. L. (2008). Frontal paralimbic network atrophy in very mild behavioral variant frontotemporal dementia. *Archives of Neurology*, *65*(2), 249–255. https://doi.org/10.1001/archneurol.2007.38

Segaert, K., Weber, K., de Lange, F. P., Petersson, K. M., & Hagoort, P. (2013). The suppression of repetition enhancement: A review of fMRI studies. *Neuropsychologia*, *51*(1), 59–66. https://doi.org/10.1016/j.neuropsychologia.2012.11.006

Seth, A. (2014). The cybernetic bayesian brain - from interoceptive inference to sensorimotor contingencies. *Open MIND*, *35*, 1–24. https://doi.org/10.15502/9783958570108

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. J. (2016). Active interoceptive inference and the emotional brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1708). 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*, *3*(JAN), 1–16. https://doi.org/10.3389/fpsyg.2011.00395

Sherrington, C. (1906). The integrative action of the nervous system. In *New Haven: Yale University Press.*

Silveri, M. C. (2007). Frontotemporal dementia to Alzheimer's disease. *Dialogues in Clinical Neuroscience*, *9*(2), 153–160.

Skiba, R. M., & Vuilleumier, P. (2020). Brain Networks Processing Temporal Information in Dynamic Facial Expressions. *Cerebral Cortex*, *30*(11), 6021–6038. https://doi.org/10.1093/cercor/bhaa176

Smith, S. D., Nadeau, C., Sorokopud-Jones, M., & Kornelsen, J. (2021). The Relationship Between Functional Connectivity and Interoceptive Sensibility. *Brain Connectivity*.

Snowden, J. S., Austin, N. A., Sembi, S., Thompson, J. C., Craufurd, D., & Neary, D. (2008). Emotion recognition in Huntington's disease and frontotemporal dementia. *Neuropsychologia*, *46*(11), 2638–2649.

Sokoloff, P., & Schwartz, J.-C. (2002). The Dopamine D 3 Receptor and Its Implication in Neuropsychiatry Disorders and Their Treatments. In *Dopamine in the CNS I* (pp. 185–221). Springer.

Sowden, S., Brewer, R., Catmur, C., & Bird, G. (2016). The specificity of the link between alexithymia, interoception, and imitation. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(11), 1687–1692. https://doi.org/10.1037/xhp0000310

Spence, M. L., Dux, P. E., & Arnold, D. H. (2016). Computations underlying confidence in visual perception. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(5), 671.

Stephan, K. E., Manjaly, Z. M., Mathys, C. D., Weber, L. A. E., Paliwal, S., Gard, T., … Seth, A. K. (2016). Allostatic self-efficacy: a metacognitive theory of dyshomeostasis-induced fatigue and depression. *Frontiers in Human Neuroscience*, *10*, 550.

Strange, B. A., Duggins, A., Penny, W., Dolan, R. J., & Friston, K. J. (2005). Information theory, novelty and hippocampal responses: Unpredicted or unpredictable? *Neural Networks*, *18*(3), 225–230. https://doi.org/10.1016/j.neunet.2004.12.004

Sturm, V. E., Sible, I. J., Datta, S., Hua, A. Y., Perry, D. C., Kramer, J. H., … Rosen, H. J. (2018). Resting parasympathetic dysfunction predicts prosocial helping deficits in behavioral variant frontotemporal dementia. *Cortex*, *109*, 141–155.

Sueyoshi, T., Sugimoto, F., Katayama, J., & Fukushima, H. (2014). Neural correlates of error processing reflect individual differences in interoceptive sensitivity. *International Journal of Psychophysiology*, *94*(3), 278–286.

Summerfield, C., Trittschuh, E. H., Monti, J. M., Mesulam, M.-M., & Egner, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. *Nature Neuroscience*, *11*(9), 1004–1006.

Terasawa, Y., Fukushima, H., & Umeda, S. (2013). How does interoceptive awareness interact with the subjective experience of emotion? An fMRI Study. *Human Brain Mapping*, *34*(3), 598–612. https://doi.org/10.1002/hbm.21458

Terasawa, Y., Kurosaki, Y., Ibata, Y., Moriguchi, Y., & Umeda, S. (2015). Attenuated sensitivity to the emotions of others by insular lesion. *Frontiers in Psychology*, *6*, 1314.

Terasawa, Y., Moriguchi, Y., Tochizawa, S., & Umeda, S. (2014). Interoceptive sensitivity predicts sensitivity to the emotions of others. *Cognition and Emotion*, *28*(8), 1435–1448. https://doi.org/10.1080/02699931.2014.888988

Terasawa, Y., Motomura, K., Natsume, A., Iijima, K., Chalise, L., Sugiura, J., … Umeda, S. (2021). Effects of insular resection on interactions between cardiac interoception and emotion recognition. *Cortex*, *137*, 271–281. https://doi.org/10.1016/j.cortex.2021.01.011

Terasawa, Y., Shibata, M., Moriguchi, Y., & Umeda, S. (2013). Anterior insular cortex mediates bodily sensibility and social anxiety. *Social Cognitive and Affective Neuroscience*, *8*(3), 259–266.

Todorovic, A., van Ede, F., Maris, E., & de Lange, F. P. (2011). Prior expectation mediates neural adaptation to repeated sounds in the auditory cortex: An MEG study. *Journal of Neuroscience*, *31*(25), 9118–9123. https://doi.org/10.1523/JNEUROSCI.1425-11.2011

Trempler, I., Schiffer, A.-M., El-Sourani, N., Ahlheim, C., Fink, G. R., & Schubotz, R. I. (2017). Frontostriatal contribution to the interplay of flexibility and stability in serial prediction. *Journal of Cognitive Neuroscience*, *29*(2), 298–309.

Tsakiris, M., & Critchley, H. (2016). Interoception beyond homeostasis: Affect, cognition and mental health. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1708). https://doi.org/10.1098/rstb.2016.0002

Van Den Houte, M., Vlemincx, E., Franssen, M., Van Diest, I., Van Oudenhove, L., & Luminet, O. (2021). The respiratory occlusion discrimination task: A new paradigm to measure respiratory interoceptive accuracy. *Psychophysiology*, *58*(4), e13760.

Van den Stock, J., De Winter, F.-L., de Gelder, B., Rangarajan, J. R., Cypers, G., Maes, F., … Vandenbulcke, M. (2015). Impaired recognition of body expressions in the behavioral variant of frontotemporal dementia. *Neuropsychologia*, *75*, 496–504.

Van den Stock, J., & Kumfor, F. (2019). Behavioural variant frontotemporal dementia: At the interface of interoception, emotion and social cognition? *Cortex*, *115*, 335–340. https://doi.org/10.1016/j.cortex.2017.08.013

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.

Wang, Xiaoqin, Tan, Y., Van den Bergh, O., von Leupoldt, A., & Qiu, J. (2020). Intrinsic functional brain connectivity patterns underlying enhanced interoceptive sensibility. *Journal of Affective Disorders*, *276*, 804–814.

Wang, Xingchao, Wu, Q., Egan, L., Gu, X., Liu, P., Gu, H., … Gao, Z. (2019). Anterior insular cortex plays a critical role in interoceptive attention. *Elife*, *8*, e42265.

Weickert, T. W., Leslie, F., Rushby, J. A., Hodges, J. R., & Hornberger, M. (2013). Probabilistic association learning in frontotemporal dementia and schizophrenia. *Cortex*, *49*(1), 101–106.

Weissman, D. H., Roberts, K. C., Visscher, K. M., & Woldorff, M. G. (2006). The neural bases of momentary lapses in attention. *Nature Neuroscience*, *9*(7), 971–978. https://doi.org/10.1038/nn1727

Werner, N. S., Kerschreiter, R., Kindermann, N. K., & Duschek, S. (2013). Interoceptive awareness as a moderator of affective responses to social exclusion. *Journal of Psychophysiology*, *27*(1), 39–50. https://doi.org/10.1027/0269-8803/a000086

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

Wood, A., Rychlowska, M., Korb, S., & Niedenthal, P. (2016). Fashioning the Face: Sensorimotor Simulation Contributes to Facial Expression Recognition. *Trends in Cognitive Sciences*, *20*(3), 227–240. https://doi.org/10.1016/j.tics.2015.12.010

Zaki, J. (2014). Empathy: a motivated account. *Psychological Bulletin*, *140*(6), 1608.

Zaman, J., Van de Pavert, I., Van Oudenhove, L., & Van Diest, I. (2020). The use of stimulus

perception to account for variability in skin conductance responses to interoceptive stimuli. *Psychophysiology*, *57*(3), 1–8. https://doi.org/10.1111/psyp.13494

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*(May), 12–17. https://doi.org/10.1016/j.biopsycho.2018.06.006

Zhou, J., Greicius, M. D., Gennatas, E. D., Growdon, M. E., Jang, J. Y., Rabinovici, G. D., … Seeley, W. W. (2010). Divergent network connectivity changes in behavioural variant frontotemporal dementia and Alzheimer's disease. *Brain*, *133*(5), 1352–1367. https://doi.org/10.1093/brain/awq075

Abbreviations

ACC	anterior cingulate cortex
AD	Alzheimer's disease
AIC	anterior insula cortex
BOLD	blood-oxygen-level-dependent
BPQ	Body Perception Questionnaire
bvFTD	behavioural variant frontotemporal dementia
dlPFC	dorsolateral prefrontal cortex
EEG	electroencephalogram
fMRI	functional magnetic resonance imaging
FTD	frontotemporal dementia
HDT	heartbeat detection task
IAcc	interoceptive accuracy
IAw	interoceptive awareness
IFG	inferior frontal gyrus
IS	interoceptive sensibility
MAIA-2	Multidimensional Assessment of Interoceptive Awareness, Version 2
nfvPPA	non-fluent/agrammatic aphasia
OFC	orbitofrontal cortex
PPA	primary progressive aphasia
SCR	skin conductance response
svPPA	semantic primary progressive aphasia
VENs	von Economo neurons

List of Figures

Figure 1. A (simplified) summary of the proposed neuroanatomy for integrating interoceptive information……20

Curriculum Vitae

Personal Details

Name	Amelie Marleen Hübner
Date of birth	08/12/1992
Place of birth	Hameln

Education

PhD Psychology Westfälische Wilhelms-Universität, Münster	10/2018 - present
Qualification as a Psychological Psychotherapist with focus on behavioural therapy Institute for Psychological Psychotherapy at Westfälische Wilhelms-Universität, Münster (IPP)	10/2017 - 10/2021
Master of Science, Psychology Julius-Maximillians-University Würzburg	10/2015 - 09/2017
Bachelor of Science, Psychology Julius-Maximillians-University Würzburg	10/2012 – 09/2015
Law studies Georg-August-University Göttingen	10/2011 – 09/2012
Abitur Albert-Einstein-Gymnasium, Hameln	2011

Work Experience

Research Assistant Department of Psychology University of Münster	10/2018 - present
Psychological Psychotherapist Institute for Psychological Psychotherapy at Westfälische Wilhelms-Universität, Münster (IPP)	04/2019 - present
Clinical Psychologist Department of depression & gerontopsychiatry LWL-Clinic Lengerich	10/2017 – 10/2018

Supervisor for university degrees

Bachelor thesis Psychology (Rieke Elfering, Pauline Fasthoff, Umay Kuyumcu)

Master thesis Psychology (Corinna Gietmann, Bella Bezverni)

Teaching Experience

Seminar „Biological psychology: What moves us - pain, stress, sexuality, emotion, motivation.” Westfälische-Wilhelms-Universität Münster	Summer term 2019 Summer term 2020 Winter term 2020/2021
Seminar „Experimental Research Internship - Affective Prediction.” Westfälische-Wilhelms-Universität Münster	Summer term 2021 Winter term 2021/2022

Declarations

Declarations by the candidate documenting **open-science activities** and **on the consideration of ethical aspects** as part of the doctoral process and disclosure of **personal contribution** for manuscripts completed by two or more authors (cumulative dissertations)

Doctoral candidate: Amelie Marleen Hübner

Title of dissertation: A Role of Interoception for Emotion Inference: Behaviour and Neurophysiology in Healthy Individuals and Frontotemporal Dementia

1. Documentation of **open-science activities**

Manuscript 1

	yes	no	If yes, please specify the source
Pre-registration		X	
Publication of data		X	
Publication of analysis scripts		X	
Publication of materials		X	
Open access publication		X	

Manuscript 2

	yes	no	If yes, please specify the source
Pre-registration		X	
Publication of data		X	
Publication of analysis scripts		X	
Publication of materials		X	
Open access publication		X	

Manuscript 3

	yes	no	If yes, please specify the source
Pre-registration		X	
Publication of data		X	
Publication of analysis scripts		X	
Publication of materials		X	
Open access publication		X	

2. Declaration on **the consideration of ethical aspects**

Study number	**Source (manuscript / chapter of dissertation):** e.g. study 1 in paper 2, study 1 described in chapter 4	**Was the study reviewed by an ethics commission?**	
		yes	**no**
1	Manuscript 1	X	
2	Manuscript 2	X	
3	Manuscript 3	X	

3. Declaration of one's **personal contribution** to the submitted academic manuscripts by two or more authors

Manuscript 1

Title	Interoceptive sensibility predicts the ability to infer others' emotional states		
Author(s)	Amelie M. Hübner, Ima Trempler, Corinna Gietmann, & Ricarda I. Schubotz		
Publication status:	not yet submitted		(please mark with X)
	submitted		
	in review		
	in revision		
	accepted		
	published	X	
Journal	Plos one		
Year of publication	2021		

Description of your own contribution in the case of joint authorship:

- partly responsible for the study's conception and design
- collecting, processing, analysing and interpreting data
- mainly responsible for drafting and revising the manuscript
- marked as corresponding author

Manuscript 2

Title	Selective Impairment of Emotion Inference in Behavioural Variant Frontotemporal Dementia? A Pilot Study		
Author(s)	Amelie M. Hübner, Ima Trempler, Nadine Kloth, Thomas Duning, Andreas Johnen & Ricarda I. Schubotz		
Publication status:	not yet submitted		(please mark with X)
	submitted		
	in review	X	
	in revision		
	accepted		
	published		
Journal	Plos one		
Year of publication			

Description of your own contribution in the case of joint authorship:

- partly responsible for the study's conception and design
- collecting, processing, analysing and interpreting data
- mainly responsible for drafting and revising the manuscript
- marked as corresponding author

Manuscript 3

Title	Interindividual Differences in Interoception Modulate Behavior and Brain Responses in Emotional Inference.		
Author(s)	Amelie M. Hübner, Ima Trempler, & Ricarda I. Schubotz		
Publication status:	not yet submitted		(please mark with X)
	submitted		
	in review	X	
	in revision		
	accepted		
	published		
Journal	NeuroImage		
Year of publication			
Description of your own contribution in the case of joint authorship: - partly responsible for the study's conception and design - collecting, processing, analysing and interpreting data - mainly responsible for drafting and revising the manuscript - marked as corresponding author			

Place, Date

Signature of the doctoral candidate