There is More to a Heartbeat: The Role of Autonomic Regulation in the Psychophysiological Stability of the Self in Relation to the External Physical and Social World

Lilla Hodossy

Submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in the Department of Psychology

Royal Holloway, University of London

2019

Declaration of Authorship

I, Lilla Hodossy, hereby declare that this work, which is approximately 45623 words in length (excluding references), was carried out in accordance with the Regulations of the University of London. I declare that this submission is my own work, and to the best of my knowledge does not represent the work of others, published or unpublished, except where duly acknowledged in the text. No part of this thesis has been submitted for a higher degree at another university or institution.

Signed:

Date: 16/04/2020

Supervisor’s declaration:

I, Manos Tsakiris, hereby certify that the candidate has fulfilled the conditions of the Regulations appropriate for the degree of PhD at the University of London and that the candidate is qualified to submit this thesis in application for that degree.

Signature of supervisor:

Date: 16/04/20

For Anikó

Acknowledgments

First and foremost, I am grateful for the support of my supervisor, Prof Manos Tsakiris. You always had my back and inspired me to achieve things I did not think were possible before. Your support meant that I could freely explore a topic that stands so close to my heart. Thank you for being my mentor and good friend throughout these years and giving me the opportunity to do work that I can be proud of. This PhD would not have been possible without your exceptional supervision.

Second, huge thanks to the members of Lab of Action and Body (i.e. LAB – still a great pun) who treated me like family from the first day I arrived to work as a research assistant. I am especially thankful for Vivien Ainley to whom I could always turn for advice and kind words. I will never forget that much needed coffee break when we stopped to smell the Sweet Peas in the garden. Massive thanks to Lara Maister who showed me the beauty of statistics and R. We would talk about linear mixed effects models anywhere and anytime, let it be a ferry ride from a Greek island or a continuous flow of messages on Slack. Also, big thanks to Ruben Azevedo who tirelessly guided me through my first steps in programming. I will always remember that six-hours-long joint programming session which itself could have been a case study for this PhD.

Big thanks to my friends Hannah Connolly, Clare Lally, Zsuzsanna Döbröntei, Rachel Nesbit and Ben Tappin who made this PhD experience fun, and whose friendship was like therapy when times were hard. I will always remember those Wine-down Wednesdays, Car Club rides, board game nights and all our inside jokes.

Finally, I am grateful for my fun and geeky family for their unconditional support and love. Without my parents Ágnes Hodossy and Károly Hodossy and my brothers Bálint Hodossy, Ákos Hodossy I would not be the person I am today. I am thankful to my partner,

Zsolt Simon for lifting me up (sometimes literally) when I was down and sharing all the happiness and tears all these years brought us. This thesis is for all of you.

Table of Contents

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

Table of Figures

Chapter 1. General Introduction

Chapter 2. Methodology

Chapter 3. Study1 and Study 2

Chapter 4. Study 3

Chapter 5. Study 4

Chapter 6. General Discussion

Table of Tables

Chapter 1. General Introduction

Chapter 2. Methodology

Chapter 4. Study 3

Chapter 5: Study 4

Bibliography of Work Completed During PhD

Hodossy, L., & Tsakiris, M. (2020). Wearing your heart on your screen: Investigating congruency-effects in autonomic responses and their role in interoceptive processing during biofeedback. Cognition, 194, 104053. [Here Study 1 and Study2] https://doi.org/10.1016/j.cognition.2019.104053

Hodossy, L., & Tsakiris, M. (under review). Neurobehavioral Differences across Strategies Used in Cardiac Recognition. Journal of Experimental Psychology: General [Here Study 4]

Hodossy, L., & Tsakiris, M. (submitted). Standing your Ground through Change: Autonomic Regulation and Interpersonal Cardiac Synchrony during Cooperative and Competitive Social Interactions Introduction. Biological Psychology. [Here Study 3]

Maister, L., **Hodossy, L**., & Tsakiris, M. (2017). You fill my heart: Looking at one's partner increases interoceptive accuracy. Psychology of Consciousness: Theory, Research, and Practice, 4(2), 248. [Not in thesis]. https://doi.org/10.1037/cns0000110

Maister, L., **Hodossy, L**., Tsakiris, M., & Shinskey, J. (accepted). Self or (M)other? Infants' Sensitivity to Bodily Overlap with their Mother Reflects their Dyadic Coordination. Child Development. [Not in thesis]

Abstract

The experience of one's own body is dependent on the integration of signals originating both from the body's internal milieu and the outside world. During the processing and integration of these signals, the bodily self must maintain a fine balance between stability and malleability. Across 4 studies, this thesis explored the role of autonomic regulation in maintaining psycho-physiological stability. Specifically, the present studies addressed different aspects of the individuals' engagement in explicit inference and unconscious interoceptive inference about one's self in relation to the external world and others. The main theory of this thesis predicted that autonomic regulation could contribute to the psychophysiological stability of the self across multiple hierarchical levels. In all studies participants received either congruent or incongruent feedback of their cardiac activity, while the experimental conditions were optimised to the rationale of the corresponding study. According to the findings of Study 1 and 2, autonomic regulation does contribute to the integration of self-relevant information across different hierarchical levels resulting in physiological stability. Study 3 revealed that autonomic regulation and physiological synchrony could function as self-stability preserving processes during competition. Finally, going beyond non-conscious interoceptive inference, Study 4 used a novel cardiac recognition paradigm that employed a set of different measures that correspond to different facets of interoception (i.e. the ability to infer the state of one's body). Integrating behavioural, psycho-physiological and metacognitive measures, the findings of Study 4 suggest that people can use different strategies (potentially linked to different hierarchical levels in the brain) to reach valid inferences about their sensations. When taken together, the findings of the present thesis have important theoretical implications for predictive coding models of the self and social-affective neuroscience as they pave the way for a more direct application of abstract theories in experimental designs. The general discussion outlines a novel mechanistic framework of interpersonal congruency that could contribute to closing the arbitrary gap between individual and social approaches in experimental designs.

Chapter 1: General Introduction

1.1 The Complexity of Researching the Basis of the Self

One of the fundamental challenges of experimental psychology is to establish and prove a testable theoretical framework that explains how the experience of one's self comes about. One of the earliest suggestions of the core mechanisms providing the sense of self was by William James. James (1890) distinguished two aspects of the self, such as the self as *Me*, being the object of experience, and the self as *I*, reflecting a subjective experience originating from one's body. A new line of research highlighted the complicated computations that need to be resolved by the brain to create the bodily foundations of the self (e.g. Friston, 2009). Specifically, these theories investigate the intricate conscious and unconscious mechanisms that enable a stable and continuous experience of the self in an everchanging physical and social world (e.g. Sterling, 2012; 2014). What further complicates the computational challenges is that the continuous stream of information can arise from multiple external and bodily sources. The more studies started to investigate the integration of multisensory information (e.g. Aspell et al., 2013; Sel, Azevedo, & Tsakiris, 2017; Suzuki, Garfinkel, Critchley, & Seth, 2013) the clearer it became that autonomic responding and interoception (i.e. the ability to infer the state of one's body) provide cardinal underpinnings for the experience of one's self (e.g. Seth, Suzuki, Keisuke, & Critchley, 2012). Briefly, the overall aim of this thesis was two-fold. Using a predictive coding framework, it aimed to explore the links of autonomic responding (i) to psycho-physiological self-stability and then to (ii) self-other distinction on conscious and less conscious levels.

While each study of the present thesis emphasizes different aspects of how the bodily self comes about, this introductory chapter aims to put forward the overarching principles that are relevant throughout. First, I will present theories on predictive coding and interoception in the context of psycho-physiological stability of the self. The section after that will review studies revealing the cognitive and behavioural effects of social context with the aim to illustrate the importance of studying the self in relation to other people. The same section will consider studies that have relevance for understanding the contribution of interoception in social processes. To bring the individual and social processing even closer, the third section will discuss interpersonal autonomic responding in the form of physiological synchrony studies. After that, a biofeedback paradigm will be considered as a viable experimental approach for exploring the self in relation to the outside world and other people. Following the review of past findings on cardiac biofeedback, the final section will present the specific rationales behind all four studies of this thesis and conclude the introduction.

1.2 Bodily Foundation of the Self

There are two significant models depicting the ubiquitous role of interoception in self-awareness that received the extensive attention of the field (Damasio, 2010; Craig, 2010*)*. First, the subjective experience of the feeling that one's body exists was proposed to be underpinned by the continuous mapping of internal homeostatic states of the body (Damasio, 2010*)*. Damasio argues that the biological imperative to maintain the body within narrow homeostatic parameters that ensure its survival will eventually give rise to the experience of the continuity and stability of one's self through time (Damasio, 2003).

Second, it has been observed that interoceptive afferent information from the body is firstly conveyed via the thalamus will eventually reach the posterior insula cortex bilaterally, where interoceptive information is represented in somato-topographical and modality-specific way (Björnsdotter, Löken, Olausson, Vallbo, & Wessberg, 2009; Brooks, Zambreanu, Godinez, Craig, & Tracey, 2005). Craig (2010) in his model of *the sentient self* suggests that these anatomical features facilitate the re-mapping and integration of all body-related activity (i.e. activity from interoceptive, exteroceptive, vestibular, premotor and homeostatic systems). It is proposed that this afferent information is re-mapped into the anterior insula and is integrated with emotional and hedonic signals and eventually becomes available to consciousness (Craig, 2010).

1.3 Predictive Coding Concepts

1.3.1 Free Energy Principle and the hierarchical brain

To survive, a living organism needs to maintain its body within a narrow range of ‘desirable’ states by minimizing free energy in the system (Friston, 2009, 2010). It is proposed that this is achieved by minimising the sum of differences between the desired and the actual states of the body. Predictive coding builds on the premise that a living organism cannot objectively access the true state of its environment (Helmholtz, 1860). On the contrary, the organism needs to infer the hidden causes of the effects that the environment has on the organism’s internal neural milieu (Clark, 2013). According to current concepts in computational neuroscience, known as predictive coding (Friston, 2009), our brain is a “prediction engine” that generates probabilistic models of the world and even about the body itself. The core idea behind this theory is that the brain seeks to minimize surprise (i.e. free energy), which is the difference between the brain’s prior predictions and what the senses signal to the brain. These prior predictions relate to the origins of hidden events in the environment that could have caused a certain sensory input. If the sensory input and predictions are not compatible, prediction errors (PE-s) arise, and these are then passed onto a higher level in the cortical hierarchy for resolution. During the process of perceptual inference, a percept is formed when PE is minimized at all levels within the cortical hierarchy. PE-s can result in the organism updating their prior predictions or alternatively to engage in “*active inference*”. In the motor domain, active inference appears when the organism moves to gain more sensory input to confirm or update its predictions (Adams, Shipp, & Friston, 2012; Brown, Friston, & Bestmann, 2011). Movement gives rise to proprioceptive PE-s between the current and predicted or desired proprioceptive state of the organism. These proprioceptive PE-s provide motor control that are fulfilled at the lowest level by peripheral motor reflexes (Friston, 2013).

To summarise, according to the predictive coding theory the information flow within the brain's hierarchical network is two-way. First, "top- down" signals encode a prior probability distribution (prediction) while bottom up incoming sensory data are understood as a statistical likelihood. The PE between these two inputs is passed bottom-up to the next level of the hierarchy until it is fully explained away. When the prior gets updated, then this posterior descends top-down in the hierarchy and it will become the prior for that lower hierarchical level (Friston, 2013).

This general flow of information is linked to a neurobiologically microcircuit, in which deep (infra- granular) pyramidal cells receive PE-s from lower levels while superficial (agranular) cells encode predictions. This is assumed to be appear throughout the brain, for instance PE originating from the deep layer of the primary visual cortex (V1) will be integrated with a prediction originating from the subsequent layer of secondary visual cortex (V2). Given that higher level areas integrate increasingly multimodal and abstract inputs (e.g. multisensory parietal areas), the predictions and PE-s at each level will become more and more removed from of the original sensory stimuli and their fast temporal dynamics (Allen & Tsakiris, 2018). Furthermore, these higher levels were also proposed to be not only more abstract but also more intrinsically self – related. As free energy or unpredicted sensory input can also originate from the inside of one's body, interoception has been proposed to have a ubiquitous role in free energy minimisation by recent predictive coding models (Seth, Suzuki, & Critchley, 2012; Seth, 2013). Within such accounts, interoception has been referred to as "*interoceptive inference*" and is considered being directly analogous to the perceptual inference in exteroceptive systems (Seth et al., 2012; Seth, 2013). If interoceptive inference indicates that there are deviations from the desired inner state of the body, one way to resolve the consequent interoceptive PE-s would be to update interoceptive predictions about the internal state of the body (Friston, 2013; Seth et al., 2012). What makes interoceptive inference special and different from perceptual inference is that there is only a small range of deviations from the desirable interoceptive states that are compatible with life (Damasio, 2003). The sensory input that is relevant for the maintenance of an equilibrium of interoceptive states via homeostasis are represented and in a distinct pathway that leads to the insular cortex (Craig, 2010). This area is hidden inside the Sylvian fissure at the lateral sulcus which runs down the side of the brain bilaterally. While this is the first fold that occurs developmentally, evolutionarily it is considered as one of the last folds to appear. The inward folded morphology of the interoceptive cortex brings the parts of the brain closer that it needs to communicate with, saving valuable energy for the brain this way. The insular cortex evolved very quickly in the hominid line in the last six million years and its anterior part is not even present in primates. The anterior insula is the region is where all relevant and detailed interoceptive and exteroceptive information brought together and that generates a representation of one’s interoceptive self (“material me” or “subjective me”) (Craig, 2010).

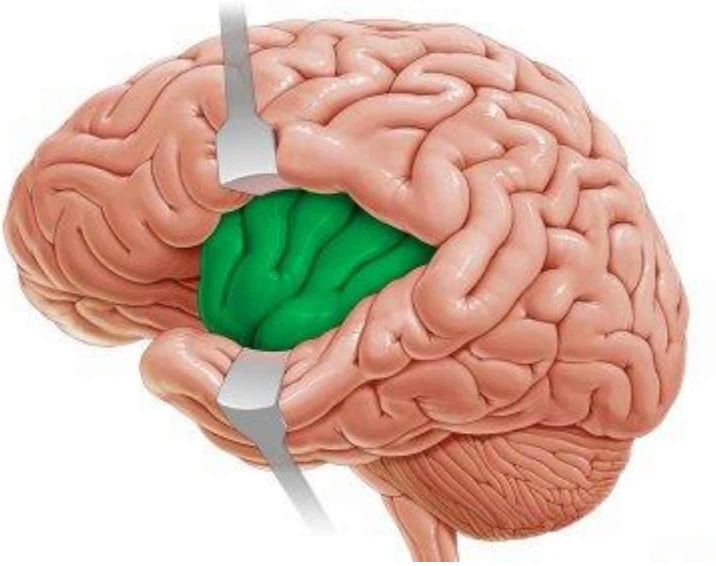


Figure 1. Illustration of the morphology of the insular cortex (here represented by the colour green). Original figure retrieved from kenhub.com.

Homeostasis is driven by adaptive reactions in response to perturbations in the environment (Pezzulo, Rigoli & Friston, 2015). Recently, these autonomic responses were proposed to have a function beyond sheer reactivity, such as the maintenance of stability through anticipation and change (Sterling, 2014), a process captured by the term *allostasis*. Allostasis contributes to stability via prospective control, in which temporary change in homeostatic set-points appear before the perturbation is present (Sterling, 2012). Neuroanatomically, homeostasis and allostasis can be linked to the anterior insula cortex and anterior cingulate cortex that respectively reflect the sensory, and the motor part of autonomic and emotional control.

Within the predictive coding framework, allostasis can be regarded as active inference (Barrett & Simmons, 2015; Seth, Suzuki, Critchley; 2012) through which agents can minimize PE-s by performing *"actions"* that bring sensory experiences closer to predictions (Friston, 2009; Feldman & Friston, 2010). To depict the action-like feature of allostasis, Seth and Tsakiris (2018) referred to autonomic changes as "interoactions". To illustrate active inference in the interoceptive domain Ainley and colleagues (2016) suggested the example of someone jumping in the pool where one's body temperature drops from the water. There PE-s can be resolved (making the water feel less cold) via interoceptive active inference by engaging autonomic reflexes (e.g. closing capillaries).

1.3.2 Precision

Precision plays a crucial role in determining the relative weight of the PE-s compared to the prior predictions, across every level of the hierarchy. Precision represents the reliability of the information that they convey (i.e. the noisier they are the less precision is assigned to them). Precision can be defined as the inverse variance (i.e. the uncertainty) of the probability distribution associated with the prediction or with the PE-s associated with the available sensory data (Friston, 2009). In other words, the brain makes both a “first-order prior prediction” about the percept that could explain the incoming sensory data but also a “second order” guess of the reliability (precision) of both the prior prediction and the PE-s (Brown et al., 2011; Hohwy, 2012). Precision therefore applies weighting to prior predictions and PE-s and determines their relative impact on the final percept. Precision is always an estimate that depends on context, previous learning and attention (Hohwy, 2012). To illustrate, if a PE is more precise than the prediction then the posterior prediction will be updated by the PE. In contrast, an imprecise PE would have only a little influence on the prediction. After this the updated or unchanged posterior prediction descends in the hierarchy where it becomes a prior prediction.

When linking precision to a cognitive process, attention and precision interact reciprocally. People divert their attention to signals that their brain’s estimate to be relatively precise (Ainley, Apps, Fotopoulou, & Tsakiris, 2016). But similarly, attending to a certain sensory channel will increase the precision of PE-s in that modality (Jiang, Summerfield, & Egner, 2013). For this reason, precision is crucial when selecting amongst information in a variety of modalities because the brain will prefer signals that are the most precise in the current context (Ainley, Apps, Fotopoulou, & Tsakiris, 2016).

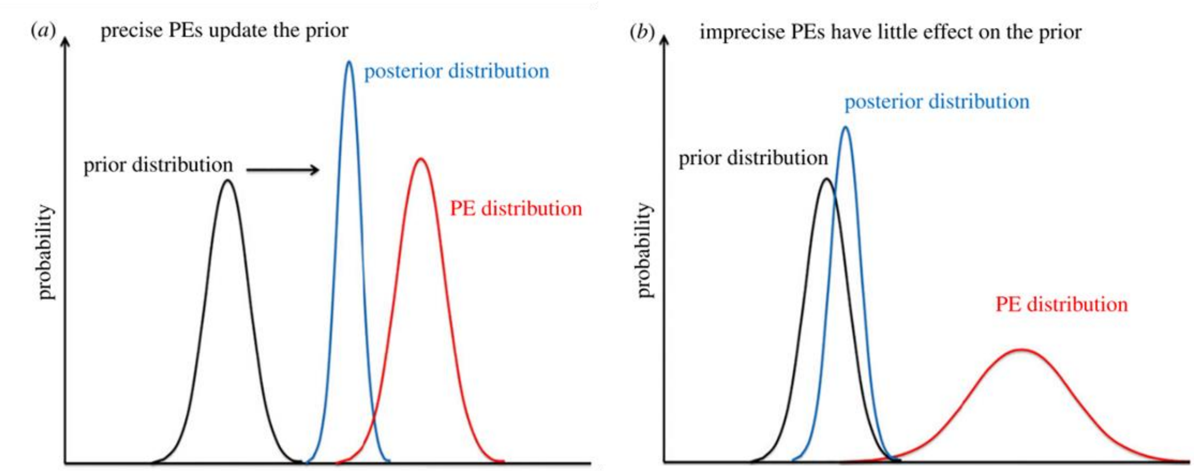


Figure 2. Figures (a) and (b) picturing the comparison between the top-down *prior* (here black line) and the bottom-up prediction error (PE, here red line). The widths of the Gaussian distributions correspond to their reliability. The narrower these distributions are the less variance is linked to them, meaning that their *precision* will be higher. The relative precision of the PE and the prior compared to one and other is crucial in determining the posterior (here blue line). (a) When the PE is more precise (narrower distribution) than the prior then the posterior will be shifted towards the PE (here to right side) meaning prior belief got updated by the incoming sensory. (b) By contrast, imprecise PE (wider distribution) will have little impact on the prior signals leaving them relatively unchanged. After this posterior descends in the hierarchy and it will become the prior for that lower hierarchical level. Original figure adapted from Ainley and colleagues (2016).

1.3.3 The Self within the Predictive Coding Model

Predictive coding framework can be applied to explain fundamental aspects of the self, such as body ownership and agency. Apps and Tsakiris (2014) were one of the firsts to propose a predictive coding self-model that they applied to the sense of body ownership. Their model is centred around the multisensory integration of bodily and exteroceptive (i.e. signals arising from the outside world) sensations. When considering the hierarchical organisation of the brain, their model can be placed at the lower levels, specifically between the integrative and sensorimotor layers. The functioning of their model can be illustrated by the Rubber Hand Illusion (RHI) where a participant would watch a rubber hand being stroked synchronously with their own hidden hand. This simultaneous stroking causes the rubber hand to be experienced as part of one's body (Tsakiris, 2010 for a review) and shows that predictions and PE-s are used to compute the probability that a body part belongs to the participant. Specifically, at the beginning of the illusion, PE is evoked by the fact that the person *feels* the brush stroking one location (i.e. one's own, not visible hand) and *sees* a brush stroking another location (i.e. on the rubber hand). This mismatch causes the PE because the brain predicts to both see *and* feel touch in the *same* location. The PE will be minimised by assuming the tactile experience is caused by what the person sees, which subsequently updates the prior model of the body, generating the percept that the rubber hand is indeed one's own hand. RHI suggests that one's body, when perceived exteroceptively, is processed in a probabilistic manner such as the body/body parts that are "the most likely to be me" (Apps & Tsakiris, 2014). To summarise, within an exteroceptive model of the self, body ownership seems highly malleable, as it is sensitive to the changes in perception of the body from the outside. However, the exteroceptive channel carries only one type of information available for self-awareness. As reviewed earlier, interoceptive signals were also proposed to contribute to the experience of self (Craig, 2002; Damasio, 2003). Going back to the RHI, the exteroceptive evidence suggesting that the rubber hand belongs to one's body will contradict the interoceptive signals associated with the actual hand - giving rise to interoceptive PE-s. This suggests that the information carried by the exteroceptive and interoceptive channels must be integrated to explain the remaining PE-s on higher multisensory level where vision has higher relative precision over touch. Addressing the dichotomy between the exteroceptive (Apps & Tsakiris, 2014; Tsakiris, 2010) and interoceptive models of the self (Seth, 2013) a new line of research appeared that simultaneously studies the exteroceptive and interoceptive aspects (e.g. Aspell et al., 2013; Salomon et al., 2017; Sel et al., 2017; Suzuki et al., 2013; Tsakiris, Jimenez, & Costantini, 2011).

There is a potential alternative explanation of the sense of self and its stability rooted more closely in the classical multisensory integration literature. Recent research assessing spatial judgments about multisensory stimuli suggests that humans integrate multisensory inputs in a statistically optimal manner (i.e. maximising precision), weighting each input by its normalised reciprocal variances (Ernst & Bülthoff, 2004). In other words, more weight will be assigned to a modality that is more appropriate for reaching a certain judgment (Welch & Warren, 1980). To date, multisensory studies that have tested the optimal integration model found that for spatial judgments vision dominates across modalities (Ley, Haggard & Yarrow, 2009). For this reason, the RHI illusion could be explained by bottom-up processes such as the visual system being better suited for deciding about the location of one’s hand. This would provide a more parsimonious explanation of the phenomenon than the predictive coding account, where RHI would be attributed to top- down effects from multisensory areas, updating the prediction of what and where one’s real hand is (Allen & Tsakiris, 2018). However, additional research in the area suggests that while bottom-up processes of visuo-tactile integration are necessary they are not enough to drive the illusion. Given that the RHI seems to be affected by the handedness identity of the viewed hand (e.g. using a left-handed rubber hand for a right-handed person) suggest that the mislocalization illusion is constrained by an abstract representation of a coherent body scheme. This suggests that the illusion is modulated by top-down influences originating from the representation of one’s own body (Tsakiris & Haggard, 2005b) revealing the importance of including such effects in a theoretical framework, favouring the application of predictive coding framework to classic multisensory integration theories.

Early studies that had a multisensory motivation explored the correlation between interoceptive accuracy (i.e. the ability to accurately monitor internal bodily signals) and the experience of RHI and revealed a negative correlation between the two constructs. It has been observed that people with lower interoceptive accuracy tend to experience a stronger RHI (Schauder, Mash, Bryant, & Cascio, 2015; Tsakiris et al., 2011). This finding suggests a dynamic and inverse relationship between interoceptive and exteroceptive channels, such as that in the absence of accurate interoception representation, exteroceptive information dominates the mental model of the self. The same shift in body part ownership were observed as in RHI when participants were looking at a virtual hand that gave pulses in synchrony with the heart’s beating (Suzuki et al, 2013). These findings were replicated on body parts but also with stimuli that has high relevance for one’s identity such as one’s whole body (Aspell et al., 2013) and face (Sel et al., 2016).

1.3.4 Stability within the Predictive Coding Framework

While one’s self is embedded in an everchanging physical and social world, the brain will continue to look for ways to maintain psychophycological stability, even if that requires change. Taken together, the findings from multisensory studies suggest that exteroceptive influences highlight the malleability of body awareness. In contrast, studies that also measured interoceptive accuracy suggest that interoceptive signals could serve the stability of the bodily self in relation to the exteroceptive information. While the presented studies did not measure autonomic regulation directly, their findings seem to support that homeostasis could have psychological consequences for the self. This would also mean that autonomic regulation has the potential to unify physiological and psychological aspects of selfhood (Allen & Tsakiris, 2018).

In this framework, on the physiological level, homeostasis ensures the stability of the organism while on a psychological level, autonomic signals would contribute to stability through being integrated with other signals. While the reviewed evidence on bodily illusions (e.g. RHI) suggests that exteroceptive information underpins the malleability of the mental model of the self, interoception and autonomic signals could provide a continuity and therefore psychophysiological stability. This means that regardless of the changes in the exteroceptive perception of the body, interoceptively one’s body will continue to feel like one’s own (Craig, 2010; Damasio, 2010). Specifically, given that autonomic signals can be considered as inherently self-related, when they get integrated with exteroceptive information they will provide a basic experience of selfhood and affect towards external stimuli - available for conscious experience (Allen & Tsakiris, 2018). While recent predictive coding theories seem to agree on the relevance of autonomic responses in the psychological stability of the self, they differ in terms of the specific function they assign to them. Specifically, autonomic responses were proposed as (i) the estimates of self- stability (Allen & Tsakiris, 2018), which could make that able to trace PEs themselves, but also as (ii) “interoactions”, that would signal interoceptive inference via the minimization of PE (Seth & Tsakiris, 2018). As described earlier, interoceptive PE-s could be minimized through autonomic reflexes (Petzschner, Weber, Gard, & Stephan, 2017; Pezzulo, 2014) like proprioceptive PE-s via motor reflexes. A way that autonomic reflexes could minimize PE is via adjusting the precision of interoceptive priors. Lowering the relative impact of interoceptive signals on perception enables the self to adapt to external stimuli whilst keeping its psychophysiological stability unperturbed.

1.3.5 Shortcomings of Predictive Coding Model

While the predictive coding framework provides an elegant way of modelling very complex physiological, neural and mental processes, the direct experimentation with its specific components is relatively unexplored. Specifically, hypotheses that link autonomic responses with the minimization of PE-s need to be tested directly. Furthermore, the majority of predictive coding theories discuss processes within an individual setting - with a few exceptions only. For instance, Friston and Frith (2015) extended the function of inference to cover others' mental states by using the same predictive processes as before, but this time for inferring other people's states of mind instead one's own experiences. The authors proposed that a by-product of generalized synchrony would inherently emerge when two observers are engaged in the modelling of each other's behaviour. The advantage of the extension of predictive coding theories to social settings is that it can elegantly bridge the potentially arbitrary gap between self- and social processing. Fittingly, Fotopoulou and Tsakiris (2016) put forward the theory that the bodily foundations of the core self are fundamentally shaped by early interactions with other people. In a commentary on the work by Fotopoulou and Tsakiris, Friston (2017) agreed that through mimicry and synchrony in joint settings the brain can acquire generative models about the other person by using the same predictive processes as they would for themselves. Friston (2017) suggested that for generalized synchrony to occur, the action–perception cycle needs to be closed in both directions. Friston endorses the conclusion by Tsakiris and Fotopoulou by arguing that interpersonal interactions are needed for shaping the mentalization of one's own internal bodily states and not the other way around. This would mean that predictive processes do not develop in an individual context and then get applied to others during mentalisation. Instead, predictions about the self would appear through internalising other people's mentalisation about one's own needs. Nonetheless, experiments directly testing interoceptive inference within a social context did not exist, which provided an interesting opportunity as a research objective for this PhD. The next section will review the main observed cognitive and behavioural effects associated with the social context and the link between interoception and social processes.

1.4 Social Context

While social context potentially has high relevance for studying interoceptive inference and self-stability this proposal has not been studied directly. To highlight the importance of experimentally manipulated social context, I will first review the abundance of evidence coming from joint action studies, suggesting that human behaviour and cognition cannot be fully understood when exclusively studied in an individual context (Sebanz, Bekkering, & Knoblich, 2006), a notion echoed by recent comprehensive theories on the development of the self (e.g. Atzil, Gao, Fradkin, & Barrett, 2018). Second, I will present the results from social cognition studies on the links between interoception and social processes suggesting that interoceptive inference could also contribute to the better understanding of others in relation to one’s self.

1.4.1 Joint Action Studies

Joint action can be regarded as any form of social interaction whereby two or more individuals coordinate their actions in space and time to bring about a change in the environment (Sebanz et al., 2006). The next section will briefly review findings of studies on the main mechanisms underlying joint action such as joint attention, action observation and shared task representation. Finally, I will present findings that highlight the impact of social context on the perception of one’s task and other people.

1.4.1.1 Joint Attention

Studies on joint attention suggest that directing one’s attention to the partner’s focus of attention provides a “perceptual common ground” in joint action (Sebanz et al., 2006). Supporting this, it has been observed that joint performance deteriorates when interaction partners cannot jointly attend to the same objects and events (Clark & Krych, 2004). In this study, pairs of participants were instructed to build Lego models together, with one person giving instructions and the other person building the model. There were three different groups where the builders' workspace was (i) visible to the director or (ii) not visible or where the (iii) instructions were delivered by an audiotape. The results revealed that the two partners took longer when directors could not see the builders' workspace, and they made more mistakes when the instructions were audiotaped.

1.4.1.2 Action Observation

Beyond joint attention, action observation can also support the understanding of others' action goals more directly, given that action perception and action production share common mechanisms (for review see van der Wel, Sebanz, & Knoblich, 2013). When people observe other's actions, their own motor system will become active, and the degree of this activation will depend on the similarity between the observer's own action repertoire and the observed actions (Calvo-Merino et al., 2004, Repp & Knoblich, 2004; Knoblich and Flach, 2003; Loula et al., 2005). This perception–action coupling was proposed to contribute to a more automatic side of action coordination called *emergent coordination* (Van der Wel, Sebanz, Knoblich, 2015). Findings of action observation are also relevant for understanding mechanisms underlying self-other distinction and self-recognition. It has been proposed that the sense of agency provides an inherent sense of self by determining whether a certain experience is the result of one's own actions or an external factor (Knoblich & Flach, 2003). Recognising one's self through agency was suggested to rely on the intact functioning of a comparative mechanism in the motor system enabling the distinction between self-generated and other generated events (Blakemore, Wolpert, & Frith, 2002). Specifically, it has been suggested that internal motor models use an efference copy to predict the results of one's own actions (e.g. Blakemore, Frith, & Wolpert, 1999). Subsequently, an internal "forward model" (Wolpert, 1997) compares the predicted/simulated sensory outcome of one's own actions with the received somatosensory feedback. Based on this, the agent would know that any sensory input that is not predicted from the motor command reflects an external event. In contrast, if there is no mismatch between the predicted and actual feedback, the action will be recognized as self-generated.

1.4.1.3 Shared Task Representation

Nonetheless, there are other, more intentional mechanisms than action observation that can facilitate the achievement of a planned joint goal, such as the shared task representation between the co-actors. Findings suggest that people spontaneously represent another actor's task in a social context, even when they are not instructed to do so (Elekes, Bródy, Halász, & Király, 2016; Welsh & McDougall, 2012) or even if it is detrimental to one's own individual performance (Sebanz et al., 2003; Welsh et al., 2005). Evidence for such representations originated from studies in which participants performed a classic Simon task (Simon, 1969) together with someone else (Sebanz, Knoblich, & Prinz, 2003, 2005). In the individual version of the task, a participant is presented with trials depicting an index finger wearing a coloured ring that could be either green or red. The participant's task is to indicate the ring's colour by pressing the corresponding either left or right button. What makes the task challenging is that the index finger could either point to the left or to the right. While the latter is a task-irrelevant feature, participants were observed to respond faster when the location of the required button corresponds with the pointing direction of the finger. In the joint version of the Simon task participants only do half of the task. Specifically, one participant is only responsible for pressing the right button if the ring had their assigned colour, while the other participant presses the left button if the ring has the other colour. In other words, the social context turns the Simon task into a standard go-no go task for the individuals. The findings of the joint version of the task indicated that participants still showed a Simon effect in contrast to the condition where participants performed go-no-go task by themselves. While boundary conditions exist for this effect to appear (e.g., Guagnano, Rusconi, & Umiltà, 2010) and the exact reasons for the joint Simon effect have been questioned (e.g., Dolk, Hommel, Prinz, & Liepelt, 2013), other experiments employing a similar rationale but a different design support the interpretation of a shared task representation in the joint Simon task (e.g., Atmaca, Sebanz, Prinz, & Knoblich, 2008; Schuch & Tipper, 2007).

1.4.1.4 Beyond Action-Related Effects of Social Context

In addition to action-related effects of social context, changes in the participants' visual perception were also reported if they believed that another person was looking at the same pictures as them (Richardson, Street, Tan, Hoover, & Ghane Cavanaugh, 2012). Specifically, participants were told they will be collaborating with another person in memorising and recognising positive and negative pictures. While participants could not see their partner, they were told on half the trials that they were looking at the same images as their partners, and on the other half that they were looking at different images. In the shared looking condition, the negative images were recognized faster than the positive ones, while in the individual condition the pattern was the opposite. In a follow-up study, participants looked at the negative pictures longer during the memorising stage only when they believed their partner had the same task as them (i.e. memorising or searching task on the same pictures). Based on these results the authors suggested that shared exposure by itself is not enough to produce the reported effects. Instead, participants must also believe that they are engaged in the same task when processing the shared stimuli. Furthermore, pro-social feelings were also reported to increase from two people attending to the same stimuli (Wolf, Launay, & Dunbar, 2016). Participants engaged in a simple reaction time task while sitting next to a partner performing the same task. In a joint attention condition, both participants attended to stimuli presented on the same half of a computer screen, while in a control condition, they attended to opposite sides of the computer screen. The authors observed higher ratings towards a partner on social bonding scale when attending the same stimuli.

Although interoceptive inference potentially contributes to the understanding of others in relation to one’s self, the reviewed joint action studies only consider cognitive and behavioural effects. While social context and interoception might seem like incompatible concepts on the surface, the next section will present studies that suggest the opposite, such as a bidirectional link between interoception and a variety of social processes (e.g. emotion recognition, socially induced affect, self-other differentiation).

1.4.2 Interoception in Social Processing

1.4.2.1 Psychological and Neural Correlates of Interoceptive Accuracy

The ability to access internal bodily states through interoception has been found to predict socially relevant aspects of human cognition and behaviour. Interoceptive accuracy[1] was also found to influence emotional sensitivity (Terasawa, Moriguchi, Tochizawa, & Umeda, 2014), measured by the emotion recognition of morphed faces ranging between a neutral and an emotional facial expression (i.e. anger, sadness, disgust and happy). Their results indicated that interoceptive accuracy was linked to the sensitivity to recognise emotions of others, especially for expressions of sadness and happiness.

---

[1] In all listed studies here interoceptive accuracy was measured by the Heartbeat Counting Task (Schandry, 1981) which requires the individual to mentally track their heart beating during short time windows with varying lengths then report the counted number of heartbeats.

Moreover, it was found that eye-contact-induced self-focus (Baltazar et al., 2014) enhanced the accuracy of participants' ability to rate the intensity of their physiological arousal (i.e. measured by skin conductance) in response to emotional pictures. Fittingly, Maister, Hodossy and Tsakiris (2017) reported that participants with low baseline interoceptive accuracy more accurately detected their heartbeats when looking at a picture of their partner's face. This suggest that exteroceptive information about one's *relational* self can improve state-like interoceptive accuracy, similarly as the reflection of one's self in the mirror (Ainley, Tajadura-Jiménez, Fotopoulou, & Tsakiris, 2012; Weisz, Balazs, & Adam, 1988) or one's own photograph (Maister & Tsakiris, 2014) did in the individual context.

When considering neural activities, there seem to be common brain areas involved both in interoception and social processes, further supporting a link between interoceptive capabilities on social inference. Adolfi and colleagues (2017) found that the anterior insula is involved in social cognition. The same area was found to be crucial in the conscious access of interoceptive processes (Craig, 2010). For this reason, interoception was suggested to be part of the core mechanisms underlying theory of mind as well as the ability to distinguish between the perspectives of the self and other people (Ondobaka, Kilner, & Friston, 2017).

These results together suggest a potential bidirectional link between social interactions and interoception, potentially facilitating the regulation of physiological responding and affect in social contexts (Löffler, Foell, & Bekrater-Bodmann, 2018).

1.4.2.2 Interoception and Self-Other Distinction

Interoceptive accuracy has also been investigated in the context of self-other distinction using the enfacement illusion. With the enfacement illusion, one can induce controlled changes in the representation of one's identity (Sforza, Bufalari, Haggard, & Aglioti, 2010; Tajadura-Jiménez, Longo, Coleman, & Tsakiris, 2012; Tsakiris, 2008) and quantify the contribution of exteroceptive self-models to self-awareness. In the improved version of the task the illusion is elicited by stroking the participant's face on the left side while seeing the face of an unfamiliar other person being stroked in synchrony, either in a specularly (like looking into a mirror) congruent or incongruent location. After this, participants are presented with a "face-morphing" movie depicting the face of an unfamiliar person's face being morphed into the participant's own face. Participants are required to stop the movie when the face starts to look more like one's self than the other person (Keenan, Ganis, Freund, & Pascual-Leone, 2000; Tajadura-Jiménez et al., 2012). The study by Tajadura-Jiménez and Tsakiris (2014) contrasted the exteroceptive and interoceptive models of the self by measuring the participant's interoceptive accuracy alongside the enfacement illusion. Their results revealed that lower levels of interoceptive accuracy were associated with larger changes in self-other boundaries caused by multisensory stimulation. In other words, participants with lower interoceptive accuracy incorporated more of the other person into their identity following congruent stroking. This study suggests that interoceptive inference and thus autonomic responses could also affect the perception of others in relation to one's self. However, to test that idea, one needs to trace autonomic responses while directly changing the correspondence between self and other by manipulating the social context – like Study 3 attempted in this thesis. In general, social neuroscience of self-other distinction would greatly benefit from including interoception in their designs. This approach could also help to further close the gap between individual and social processes fittingly with the suggestions by Friston (2017).

1.5 Interpersonal Physiology

1.5.1 Theoretical Importance

While theories differ in terms of whether social relationships enable the development of autonomic regulation (Fotopoulou & Tsakiris, 2016) or the other way around (Porges, 2007), they agree on the strong link between the two constructs. Interpersonal physiological processes (often measured by physiological synchrony) are considered as important factors in self-regulation in social settings, as they signal the way people adjust their own physiology to connect to others through influencing and being influenced (Palumbo et al., 2017). When considering synchrony with a domain general approach, Friston and Frith (2015) suggested that generalised synchronisation reflects the coupling between systems that try to predict each other's states while also being aware of their own states. A famous example of generalized synchrony was observed in pendulum clocks that became synchronized through the imperceptible motions travelling through the beam that the clocks shared (Huygens, 1673/1986). Turning from theories to empirical data, the next section will review findings on synchrony in the physiological domain as these findings have high relevance for interoceptive inference in social interactions.

1.5.2 Empirical Findings

Studies demonstrated that interpersonal regulatory processes play a substantial role in interpersonal interactions (Reeck, Ames, & Ochsner, 2016) although, these theories are largely limited to intentional social regulation (e.g., calming down a partner) leaving the regulatory processes of interpersonal autonomic *physiology* relatively unexplored. The main interest of most studies that *did* explore interpersonal autonomic physiology was to identify the different population groups where physiological synchrony (PS) is present between partners, rather than attempting to explicitly manipulate potential components of PS

(Palumbo et al., 2017). The few studies that investigated differences in experimental groups and conditions (for review see Table 1) mostly relied on participants from already existing close relationships, leaving only two studies where participants did not know each other before (Elkins et al., 2009; Mitkidis, McGraw, Roepstorff, & Wallot, 2015) – that will be described in more detail in the next paragraph. PS studies involving participants with an existing relationship (e.g. couples, friends, mother – child, relatives) looked at the effects of conflict or non-conflict conversation topics (Helm, Sbarra & Ferrer, 2014; Levenson & Gottman, 1983), completion of a fire walking ritual (Konvalinka et al., 2011) or different activities such as gazing or imitation (Ferrer & Helm, 2013; McAssey, Helm, Hsieh, Sbarra, & Ferrer, 2013; Helm et al., 2012) compared to baseline. The remaining studies were exploring the effects of a joint task (Suveg, Shaffer, & Davis, 2016; Montague, Xu & Chiou, 2014), competition (Chanel, Kivikangas & Ravaja, 2012; Järvelä, Kivikangas, Kätsyri, & Ravaja, 2013), performance level (Elkins et al., 2009), trust (Mitkidis et al., 2015) or the presence of a virtual team member (Järvelä et al., 2013). These findings have shown that the presence or absence of PS can be informative of the state of a relationship, whereas synchrony due to external variables could represent the shared levels of involvement. Studies overall revealed that the magnitude of PS does *not* function as a simple quantitative marker of a certain affective state as PS was present both during conflict (Levenson & Gottman, 1983), empathy (Marci, Ham, Moran, & Orr, 2007) or even in neutral settings (Ferrer & Helm, 2013). Unfortunately, none of these studies directly focused on the explanation of *why* these between group and conditional differences exist.

The results of one of the two studies that looked at experimentally induced group differences in people who previously did not know each other, indicated PS was higher in high performing groups (Elkins et al., 2009). Participants went through a military task in four-person teams where they had to move through a building while eliminating their opponents without shooting the non-combatants. Participants had a moderate level of experience with first person shooter style video games and no formal weapon training. The second study indicated higher levels of PS in groups that played an economic game they needed to trust their partner (Mitkidis et al., 2015). Participants engaged in a series of joint action tasks in pairs (i.e. building optimally performing and aesthetically pleasing cars from LEGO bricks). After each joint task participants belonging to the trust group also played the Public Goods Game (Ledyard, 1995) while the other group did not play this game. In the Public Goods Game, the total financial outcome is maximized if each participant contributes maximally while the individual outcome is maximized when participants do not contribute. The authors used this game to induce trust-relevant thoughts between participants. The Public Goods Game was played after each Lego building session during which participants indicated how much from the given money they wanted to contribute and how much they expect the other participants to invest. Heartrate measures were collected during the Lego building sessions. It is important to note that, participants only got feedback on how much money they earned from the Public Goods Game after the study was over. Results revealed that the economic game shifted participants' attention to the dynamics of the interaction, increased their overall heartrate and synchrony between the series of heart rates on a group level. Furthermore, the degree of heart rate synchrony was positively associated with participants' expectations regarding their partner’s investment in the economic game. Although further research is needed to adequately control for the contextual effects of social framing, these results (Elkins et al., 2009; Mitkidis et al., 2015) suggest that PS can potentially trace changes in the framing of the social context, regardless of people’s a priori relatedness.

1.5.3 Physiological Synchrony and Autonomic Regulation

It has been proposed that co-recording the level of autonomic regulation with PS would help interpreting the function of PS (Palumbo et al., 2017), potentially also resolving some of the contradictions in the existing findings. For instance, it is possible that PS in the sympathetic domain is more likely to appear during negative contexts, whereas positive contexts may result in the coupling of partners' parasympathetic changes (Palumbo et al., 2017). In support of this idea, results indicate that the magnitude of PS differs as arousal levels change (Creaven, Skowron, Hughes, Howard, & Loken, 2014; Ebisch et al., 2012; Ghafar-Tabrizi, 2008; Manini et al., 2013; Reed, Randall, Post, & Butler, 2013; Waters, West, & Mendes, 2014). Furthermore, Butler and Randall (2013) have suggested that autonomic regulation could contribute to one's capacity to optimally function as one unit with another person. However, their argument only considered baseline levels of autonomic regulation while ideally, one would need this measure to capture temporal, state-like changes too. Studies that specifically looked at synchronisation in autonomic regulation across condition used the measure of high-frequency heart rate variability (HF-HRV) - which was found to reflect phasic vagal impact upon heart and thus function as a reliable measure of state-like changes in autonomic regulation (Berntson, Cacioppo, & Grossman, 2007). In general, studies (Table 1) found higher PS in HF-HRV during conversation regardless of the valence of the topic (Helm et al., 2014) in higher performing groups (Elkins et al., 2009). PS in HF-HRV was also reported to be positively linked to empathy towards team members (Järvelä et al., 2014). Unfortunately, none of these studies reported the conditional changes on the levels of *individual* autonomic regulation. Studying individual and interpersonal changes in autonomic physiology within the same experiment could appeal to the suggestions by predictive coding theories and would have the potential to unify processes across social and individual settings (Friston & Frith, 2015).

Table 1.

Studies on cardiac PS exploring experimentally assigned groups or conditions

Study	Population	Conditions	Physiological measure, effect size	Results
Ferrer & Helm (2013)	couples	baseline v. gazing v. imitation	HR, (not enough information)	No change in PS across tasks on average.
Helm, Sbarra, & Ferrer (2014)	couples	baseline v. conversation (neutral v. positive v. negative)	HF-HRV, (not enough information)	↑ PS during conversation v. baseline, but not between conversion types.
Levenson & Gottman (1983)	couples	low-conflict v. conflict topic	IBI $\eta_p^2 = 0.60$ $\omega_p^2 = 0.44$ (medium)	↑ PS during high conflict v. low conflict topics.
McAssey, Helm, Hsieh, Sbarra & Ferrer (2013)	couples	baseline v. gazing v. imitation	HR, (not enough information)	↑ PS during imitation v. baseline. ↑ PS gazing v. baseline
Konvalinka et al. (2011)	firewalkers, relatives and strangers	baseline v. ritual	HR, $\eta_p^2 = 0.54$ $\omega_p^2 = 0.38$ (medium)	↑ PS ritual v baseline between relatives and firewalker.
Suveg, Shaffer, & Davis (2016)	mother and child	baseline v. joint task	IBI, $\eta_p^2 = 0.13$ $\omega_p^2 = 0.09$ (small)	PS was present between mothers and their children during a joint task, but not during baseline.
Helm, Sbarra, & Ferrer (2012)	friends	baseline v. gazing v. imitation	HR $\eta_p^2 = 0.01$ (not informative)	PS detected in all conditions even in baseline.

Study	Population	Conditions	Physiological measure, effect size	Results
Elkins et al. (2009)	team members (strangers)	high v. low performance teams	HF-HRV Cohen's d = 0.70 (medium)	↑ PS in high v. low performance group.
Mitkidis, McGraw, Roepstorff, & Wallot (2015)	team members (strangers)	trust game v. no trust game	HR, Cohen's d = 0.60 (medium)	↑ PS with trust game v. without trust game prior to cooperative task.
Montague, Xu, & Chiou (2014)	team members (N/2 = friends)	baseline v. virtual team task	IBI, HF-HRV, $\eta_p^2 = 0.13$ $\omega_p^2 = 0.09$ (small)	IBI PS ↑ at team trials v. baseline. HF-HRV PS ~ team performance.
Järvelä, Kivikangas, Kätsyri, & Ravaja (2014)	friends	competitive v. cooperative and presence v. absence of additional VTM	IBI, HF-HRV Cohen’s d > 0.80 (large)	IBI PS is highest during competition without a VTM. HF-HRV PS ~ empathy at TMs.
Chanel, Kivikangas, & Ravaja (2012)	friends	competitive v. cooperative game	IBI $\eta_p^2 = 0.12$ $\omega_p^2 = 0.10$ (small)	IBI PS ↑ during competition v. cooperation

Note. ↑ = increased level; ↓ = decreased level; PS = Physiological Synchrony, VTM = Virtual Team Member; TM = Team Member, HR = Heart Rate;IBI = Inter Beat Interval; HF-HRV = High Frequency-Heart Rate Variability. Guidelines (see more at Lakens, 2013) ω_p^2 > = 0.1 (small), 0.3 (medium), 0.5(large). Please note most studies used bespoke analysis therefore it is difficult to bring conclusion about the overall effect size regarding cardiac physiological synchrony.

1.6 Biofeedback

When exploring interoceptive inference in relation to the outside world and other people, biofeedback paradigms provide a viable experimental approach, as they can unambiguously represent and share internal signals that are normally only accessible to the individual. One of the most popular domains for biofeedback research has been the representation of heart rate (HR), which was mainly used in training people to the conscious control of their cardiac activity (for review see Williamson & Blanchard, 1979). Experiments using learned HR control have received the most attention by researchers during the 70-s. The basic procedural requirements in HR biofeedback training paradigms are to (i) provide participants with an exteroceptive signal that represents changes in HR and (ii) instruct them to control this exteroceptive stimulus (e.g. a bar changing its height depending on HR) in a certain way (e.g. move the bar up or down). While certain reviews of the field conclude that participants can reliably influence their HR-s using biofeedback (e.g. Williamson & Blanchard, 1979), in reality the results of classic biofeedback studies are inconsistent, and their methodology has several limitations (as reviewed below) leaving fundamental questions about this method unanswered.

1.6.1 Limitations of Past Biofeedback Research

Firs of all, it is unclear if the *combination* of biofeedback training and instructions is more effective than simply instructing subjects to control their heart, without any training (for review of these studies see White, Holmes, & Bennett, 1977). Furthermore, most previous studies did not account for the natural decrease in HR-s associated with the sheer adaptation to the situation. Some studies attempted to control for this adaptation by using the “shifting base-rate” approach (e.g. Bergman & Johnson, 1972; Blanchard, Scott, Young, & Edmundson, 1974; Blanchard & Young, 1972) where participants alternate between rest and the biofeedback task. Unfortunately, because the baseline to biofeedback comparison in these studies is confounded by *time*, this approach is still not an adequate control of the adaptation effect. This issue shows that the randomisation of task orders is crucial when looking for differences across biofeedback conditions. When there are no multiple conditions within the biofeedback, the focus being on the difference between baseline and HR regulation, then it is important to collect data about how participants' baseline changes naturally as time passes. This can be achieved simply by recording baseline multiple times without using regulation in between rest periods. Fittingly with these concerns, a well-powered study of N = 180 by White, Holmes and Bennett (1977) found that participants' regulation attempt in a biofeedback paradigm was not more efficient than the condition where participants did not receive any instructions. When instructed to *increase* HR, participants showed higher HR-s at the end of training than subjects instructed to decrease. However, this increase was still lower than their corresponding baseline. Furthermore, the authors found that instructions concerning the use of cognitions (i.e. think of either exciting or relaxing thoughts) did not facilitate changes in HR.

Nonetheless, biofeedback paradigms still seem to have the potential to help us explore empirically testable questions about the role of autonomic responses in interoceptive inference both in individual and interpersonal context. As studies in the past did not have the technical advances to acquire measures of autonomic regulation with such ease as nowadays there was little information whether HR biofeedback could affect state-like changes in autonomic regulation. When considering these limitations together, it was clear that the biofeedback paradigm itself needed extensive piloting and the main concept needed to be proved with a study before introducing more complex social manipulations to the design. For further methodological considerations about final biofeedback paradigm please refer to the *Methodology* section.

1.7 Rationale of the Present PhD Thesis

During the integration of signals arising from one' own body and the everchanging social and physical world outside, the brain and body needs to establish a fine balance between stability and adaptation. The present thesis intended to investigate physiological, neural and behavioural processes that contribute to the psychophysiological stability of the self by using an experimental approach that integrates the relevant aspect of three different fields (Table 2). The theoretical approach across all four studies presented in this thesis relies on the predictive coding theories using the measures from the field of interpersonal physiology while considering the necessity to explore the absence or presence of differences between individual and social contexts.

Table 2.

Integration of different fields contributing to the rationale of the present thesis

Field	Interpersonal physiology	Joint action	Predictive coding
Approach	Social-affective	Cognitive	Computational
Less explored, aspects	Mechanism, non-related participants	Physiology	Experimental application of theories
Prioritised context	Significant relationships	Social context	Individual
Main measure	Synchrony	Behavioural	Simulation outcomes

1.7.1 Overview of Present Studies

The overarching research question of the present thesis is whether and how the conscious and unconscious regulation of autonomic functions contribute to interoceptive inference of self and others. While recent theories emphasized the special role of autonomic regulation in interoceptive inference and psychophysiological stability no studies existed that directly tested this hypothesis. Furthermore, the exiting studies only had limited relevance for interoceptive inference within a social context. As this thesis assimilated the proposal that many psychological processes can be best understood in a social context (e.g. Sebanz et al., 2006; Atzil et al., 2018) it aimed to explore interoceptive inference across different social settings and test psychophysiological stability of the self always in relation to others. This PhD work used cardiac biofeedback as its main paradigm because (i) it can unambiguously represent cardiac signals and (ii) turn signals that are normally accessible only to the individual sharable with others. In all four studies participants received either congruent and incongruent feedback of their cardiac activity and instructions that were specific to the rationale of the corresponding study. Before moving onto more complex social designs, Studies 1 and 2 explored the function of conscious and unconscious autonomic regulation within an individual context. More precisely, Study 1 served as a proof of concept and tested if changes in autonomic regulation (i.e. measured by HF-HRV) could be indicative of self-stability related processes as suggested by the predictive coding framework. Participant were presented with a biofeedback that was either congruent or incongruent with their own cardiac activity. Furthermore, Study 1 also investigated if the ways in which participants were asked to engage with their biofeedback affected HF-HRV. In particular, participants were asked to either perform a stability neutral action (i.e. attending to certain visual features of the biofeedback) or stability facilitating behaviour (e.g. attempt to control their heartrate using the biofeedback). This manipulation aimed to test if the type of engagement had additional effects to biofeedback congruency. Study 2 wanted to investigate if the congruency-effect could be generalised to a higher hierarchical level - where multisensory information and beliefs get integrated. To reach that hierarchical level, Study 2 manipulated participants' beliefs about the biofeedback they were receiving being true or false. Study 3 investigated if the function of autonomic regulation could be generalised from an individual context to complex social settings, such as competition or cooperation. In other words, Study 3 wanted to explore the function of interoceptive inference in self-stability in situations when the self is in different relations to another person. With the competition condition, the other person was meant to be perceived as less congruent with self, while in the cooperation condition aimed to signal higher congruency between self and other. To summarise, this design tested if congruency effects could be generalised to an even higher hierarchical than is Study 2 – represented by changes in HF-HRV from the interaction between biofeedback congruency and self-other congruency. To capture interpersonal changes as well, Study 3 also measured the synchrony between participants' cardiac activity. Together, HF-HRV and HR synchrony aimed to bring the studying of individual and social processes closer to each other. Finally, going beyond unconscious physiological responding, Study 4 brought the focus on a more explicit and functional aspect of interoception as it investigated the consequences of three different strategies (i.e. Attend, Feel, Regulate) for inferences about the self. In that study participants had to decide if the cardiac biofeedback belonged to them or someone else. To make this decision participants were instructed to rely on their experience using a certain strategy for conscious inferences about the biofeedback. Learning from the behavioural results of Study 1, Study 4 optimised the biofeedback paradigm for measuring cardiac recognition by increasing the number of trials while also radically shortening the length of trials. The Attend strategy involved the participant engaging with an exteroceptive feature of the biofeedback while the Feel and Regulate strategies required the participant to use interoception. The Feel strategy aimed to capture the essence of classic measures of interoceptive accuracy where participants try to passively feel their heartbeats. In contrast, the Regulate strategy was a novel approach to cover a more functional active aspect of interoception, where interoceptive inference is based on whether the participant felt control over the biofeedback or not. While the focus of Study 4 remained on interoceptive inference, the Regulate conditions has analogies with research on action observation and agency that used the concept of efference copy to explain self-recognition. Similarly, to the efference copy that is proposed to predict the results of one's own actions (e.g. Blakemore et al., 2000), Study 4 used the concept of interoceptive predictions and PE-a to study self-recognition in the cardiac domain. During the Regulate strategy, interoceptive predictions could be compared to the visual outcomes represented by the biofeedback on the screen. When there is a mismatch between the interoceptive predictions and the perceived changes in the biofeedback, interoceptive PE-s could arise and the biofeedback would be regarded as belonging to one’s self. Study 4 employed measures from several domains of interoception to capture and integrate behavioural, cortical and metacognitive aspects.

To summarise, this thesis aimed to explore the links of autonomic responding (i) to self-awareness and then to (ii) self-other distinction on conscious and less conscious levels – interpreting the results within a predictive coding framework. The findings of all four studies have important theoretical implications for predictive coding models of the self and social cognitive neuroscience exploring interoceptive processes in a social context and self-other differentiation. Together these results pave the way for a more direct tracking of the subtle changes in psycho-physiological stability both in individual and social settings.

Chapter 2: Methodology

2.1 Autonomic Nervous System

Motor activity and patterned behaviours are controlled by the brain and are only possible when cells, tissues and organs of the body are maintained in an optimal level, enabling a continuous adjustment to varying internal and external demands placed on the organism (Jänig & McLachlan, 2013). The autonomic nervous system (ANS), is a division of the peripheral nervous system that largely unconsciously regulates the function of internal organs (Langley, 1921), such as the heart rate, digestion, respiratory rate, pupillary response, urination, and sexual arousal. Langley’s division of the ANS into sympathetic, parasympathetic and enteric nervous system (governing the function of the gastrointestinal tract) are now universally accepted (Jänig and McLachlan, 2013). In the following section I will describe the sympathetic and parasympathetic nervous system as they are relevant for the focus of this thesis. The parasympathetic nervous system is often described as the *rest and digest* system, but it also plays a major role in maintaining the physiological balance of the body in relation to its internal and external environment (Goldstein, 2013). On the other hand, the sympathetic nervous system promotes a *fight-or-flight* response (Cannon, 1929) which corresponds with arousal and energy generation, and the inhibition of digestion.

The sympathetic and parasympathetic systems not only differ in terms of their functions (i.e. engaging approximately the same organs with opposite effects), they also have anatomical differences (Jänig and McLachlan, 2013). For instance, the sympathetic ganglia are located near the spinal cord, while on the parasympathetic side, they’re close to the effector organs. Furthermore, the use of neurotransmitters in the two systems is similar, but not identical. For both, neurons release acetylcholine (Ach) in their preganglionic synapses (Hamilton, 1982). However, in the parasympathetic system, the postganglionic neurons also release ACh at their synapses with the effector organs, in contrast with the sympathetic system where the organs receive norepinephrine instead (Heilbronn & Bartfai, 1978). Finally, the main anatomical difference comes from the physical networks that the sympathetic and parasympathetic systems form. While the sympathetic nerves start from the thoracolumbar area of the spinal cord at the body's midsection, the nerves of the parasympathetic division are craniosacral, meaning they are neurons that begin at the cranial nerves or the sacral spinal cord (Jänig and McLachlan, 2013).

The cranial nerves run right from the brain almost straight to their effector organs without ever going through the spinal cord. There are 12 of these cranial nerves (Vilensky, Robertson, & Suarez-Quian, 2015) , and they differ depending on the kinds of neurons they contain such as motor fibers (carrying autonomic or voluntary motor information from the brain to the effector organs) or sensory fibers (carrying information from the sensory organs to the brain) and lastly the combination of both motor and sensory neurons. The main nerve of the parasympathetic nervous system is the tenth cranial nerve, called the vagus nerve (Goldstein, 2013). This long nerve stretches from near the brainstem to most of the visceral organs, innervating the heart, lungs, and the digestive system. The vagus nerves is bidirectional as it carries incoming sensory information from the peripheral system to the brain but also transmits motor instructions from the brain to the organs. The importance of the vagus nerve in the two-way heart-brain communication has been known for over 100 years (Lane et al., 2009). The heart is dually innervated by the autonomic nervous system meaning it is not only linked to the vagus nerve but also has connections to sympathetic nervous system (Berntson, Cacioppo, & Quigley, 1993). It has been revealed that autonomic control of dually innervated organs cannot be viewed as a continuum between parasympathetic and sympathetic dominance, as the two systems can vary reciprocally, independently or coactively (e.g. Berntson et al., 1991). Nonetheless there is a way to differentiate between sympathetic and parasympathetic effects on the heart. While the sympathetic effects happen on the time scale of seconds, the parasympathetic effects are much faster as they happen within milliseconds. For this reason, only the parasympathetic influences carried by the vagus nerve can lead to rapid changes in the beat-to-beat timing of the heart, also known as Heart Rate Variability. In other words, the rapid changes in heart rate are associated with both the mechanical and neural gating of vagal outflow (Lane et al., 2009).

2.2 Heart Rate Variability

2.2.1 Definition and Measurement

With technological advances, established standards and research guidelines (Malik, 1996; Berntson et al., 1997) heart rate variability (HRV) received high interest by the field of psychophysiology in the last twenty years. HRV represents the variability of successive inter-beat-intervals (IBI-s) which was consistently found a non-invasive index of the parasympathetic nervous system (Malik, 1996; Chapleau & Sabharwal, 2011) when registered and analysed under the right conditions. Measures of HRV are valuable for understanding the relationship between brain and body, given that the parasympathetic nervous system has been found to be relevant for self-regulation mechanisms with links to cognitive, affective, social processes, but also to general health (e.g. Porges, 2007; Thayer et al., 2009; McCraty & Shaffer, 2015). While nowadays it has become relatively easy to access and measure HRV, this simplicity often hides the complications that come with the interpretation of HRV. Even the term HRV introduces confusion as studies use a variation of terms to refer to HRV. As the vagus nerve is the is the main nerve of the parasympathetic nervous system (Brodal, 2010), HRV is often referred to as vagal tone, vagal control or vagal reactivity. However, as the response to certain conditions, rather than baseline, is the focus of most psychophysiological research (Laborde, Mosley, & Thayer, 2017), recent work has suggested not to use the phrase vagal *tone* in studies that look at conditional differences. Chambers and Allen (2007) also argued that vagal tone suggests the stability of individual differences and does not reflect the dynamic nature of the parasympathetic nervous system. Addressing the same conceptual concerns about vagal *tone*, Berntson, Cacioppo, Grossman (2007) suggested the term Respiratory Sinus Arrythmia (RSA) as a reflection of the *phasic* rather than *tonic* vagal impact upon the heart. RSA depicts the phenomenon where heart rate increases with inspiration and decreases with expiration (Eckberg & Eckberg, 1982). However, RSA depicts some of the > 70 variables that can be calculated from HRV analysis (Bravi et al., 2011) more closely than others which can lead to further confusion in the field. Table 1 includes some of the most popular measures of HRV with their physiological origins that are ought to have links to vagal control (being the focus of this PhD): RMSSD, pNN50, peak-valley, HF–HRV and LF/HF.

Table 1.

Main HRV parameters and their physiological origin. Original table adapted from Laborde, Mosley and Thayer (2017)

	Variable	Description	Physiological origin
Time-domain	RMSSD	Root mean square of successive differences	Vagal tone/control
	pNN50	Percentage of successive normal sinus IBI-s more than 50 ms	Vagal tone/control
	Peak-valley	Time-domain filter dynamically centred at the exact ongoing respiratory frequency	Vagal tone/control
Frequency-domain	HF	High frequencies	Vagal tone/control
	LF/HF	Low frequencies/high-frequencies ratio	Mix of sympathetic and vagal activity

Measures of HRV can be separated into three distinct categories: time-domain, frequency domain and non-linear measures (for examples on the letter see Laborde et al., 2017) based on their approach to quantify the variation of IBI-s during the time of the recording. Within the time-domain measure, the measure called root mean square of successive differences (RMSSD) is typically preferred (Otzenberger et al., 1998) to the percentage of successive normal sinus IBI-s that are more than 50 ms (pNN50). Beyond these traditional variables, there is peak-valley analysis - (Grossman et al., 1990; Lewis et al., 2012), which works as a time-domain filter that is dynamically centred at an exact ongoing respiratory frequency (Grossman et al., 1990). Peak-valley analysis stands the closest to the term RSA conceptually as it is based on the difference between the peaks appearing during inhalation and the valleys linked to exhalation. However, studies using peak-valley analysis control for respiration by instructing participants to breathe at a specific rate during the experiment (Grossman & Taylor, 2007). While the controlled-breathing procedure works fine for baseline measurements, it influences the outcome of a cognitive or affective task (Quintana & Heathers, 2014) as participants must consciously follow the pacing cue while doing the main task.

In the frequency-domain, the analysis requires the filtering of the main cardiac signal into different bands (for illustration see Figure 1). The high frequency band, specifically the frequencies between 0.15 and 0.40 Hz (Malik, 1996), is also frequently called the respiratory band because it corresponds to respiratory cycle-related heart rate variations (Eckberg & Eckberg, 1982). When breathing rates remain between nine cycles (0.15 Hz) and 24 cycles per minute (0.40 Hz) then HF-HRV reflects vagal control. These bands might be adjusted to special populations like children and infants who breathe faster at rest (Quintana et al., 2016).

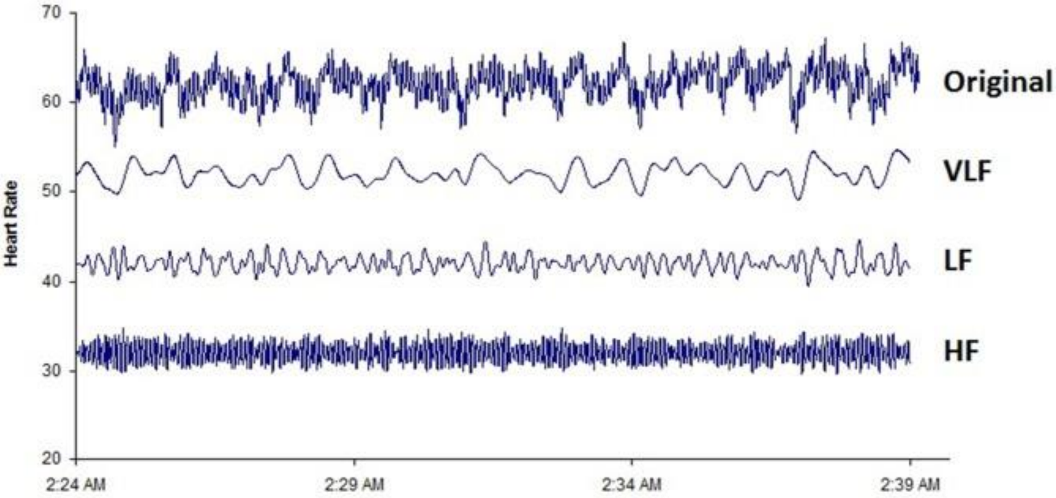


Figure 1. Original figure adapted from Shaffer et al. (2014) depicting a 15-minute-long HRV recording during baseline in a healthy individual. This figure illustrates the frequency-domain methods that uses filtering techniques to separate the original ECG waveform into three different frequency bands (VLF: very low frequency, LF: low frequency, and HF: high frequency).

While RMSSD and the HF-HRV are highly correlated (Kleiger et al., 2005), the present thesis decided to focus on HF-HRV as it has been reliably used during shorter periods (i.e. 2 - 5 mins) in psychophysiological studies (Camm et al., 1996). In contrast, RMSSD is the most accurate when calculated over 24 h (Shaffer & Ginsberg, 2017). More precisely, the normalised units (nu) of HF-HRV were used in the analyses as they provide a degree of interpretability between studies. They stay roughly the same regardless of the spectral methods, windowing methods or the type of algorithm used for calculation (Sandercock, 2007). The general formula used for calculating the normalized units of HF-HRV is the following:

$$100 \, * \, \frac{HF\ Power}{(Total\ Power - VLF\ Power)}$$

Before continuing with further methodological aspects of HRV it is important to consider another commonly used measure, that of low frequency to high frequency ratio

(LF/HF). In the past, it has been suggested that the lower frequencies (LF) of HRV could track influences associated with the sympathetic nervous system. Specifically, in around 65% of current HRV papers (Heathers, 2014) LF / HF is interpreted as a measure of sympatho-vagal balance, or in other words the measure of sympathetic to parasympathetic activation. Despite the enthusiasm of the field for this possibility, it has been repeatedly shown that approximately 90% of LF-HRV is parasympathetically mediated (e.g. it remains virtually intact after sympathetic blockades, for more on this see Eckberg, 1997; Goedhart, Willemsen, Houtveen, Boomsma, & De Geus, 2008; Heathers, 2012; Billman, 2013; Reyes del Paso, Langewitz, Mulder, Roon, & Duschek, 2013). For this reason, the interpretation as a measure of sympatho-vagal balance simply cannot hold. Researchers are strongly recommended to adopt HRV indices that clearly reflect the physiological system of vagal control such as RMSSD, peak-valley respiratory sinus arrythmia (RSA), and HF-HRV (Laborde, Mosley, Thayer, 2017).

2.2.2 Tonic and Phasic HRV

HRV can be considered tonic when HRV is measured only once at rest. Phasic HRV on the other hand represents the change in HRV between two time points or in response to a condition. Phasic HRV is also called as reactivity, stimulus-response, change or delta HRV and vagal withdrawal (when the change is a decrease) and tonic HRV is also known as resting or baseline HRV. Their distinction is important as there are reported differences between tonic and phasic HRV in terms of their links to psycho-physiological adaptivity. Specifically, a higher baseline HRV is considered beneficial for optimal functioning (Thayer et al., 2012) with a few exceptions, like people with eating disorders (Peschel et al., 2016). Hyperactivity of vagal *afferents* has been suggested as a pathophysiological component in bulimia nervosa that is relevant for the continuation of binge eating and vomiting (Faris et al., 2008). Similarly, as discussed earlier, hyperactivity of vagal *efferents* may also contribute to bulimic episodes via increased secretion of ghrelin, that could facilitate binge eating episodes. For phasic HRV the question of adaptivity is context dependent. For example, a decrease in HRV is adaptive when an individual is facing a physical or mental stressor that does not require the involvement of executive functions. A decrease here is ought to demonstrate the ability of the organism to gain the necessary energy to tackle the stressor (Neumann et al., 2004; Porges, 2007; Rottenberg et al., 2005; Lewis et al., 2007; Messerotti Benvenuti et al., 2015). In contrast, when the stressor requires the involvement of executive functions, then a decrease would be maladaptive, as showed experimentally (Marcovitch et al., 2010; Laborde et al., 2014; Park et al., 2014, Thayer et al., 2012). Related to this, it has been shown that tonic HRV is also linked to phasic HRV (Park et al., 2014), meaning that both tonic and phasic HRV values need to be measured for the completeness of the results. The recording of an accurate baseline is crucial and for phasic HRV measures the main suggestion is to use a non-task situation that provides the best control for the task during comparison (Quintana & Heathers, 2014). A baseline recording at the beginning of the experiment also has the potential to improve the quality of the data recorded through the rest of the experiment. This is simply because the baseline recording will help with acclimatizing participants to the recording environment, facilitating the potential anxiety and the increased attention to respiration and heart rate to fade out (Quintana et al., 2016).

2.2.3 Experimental Design

Given that there are high inter-individual differences in HRV and that HRV is a more reliable measure when tracing phasic changes (Berntson, 1997), within-subject designs are highly recommended when conducting experiments. Within-subject designs also weaken the effects of respiratory differences across participants, increase statistical power, and reduce the impact of external factors (Quintana & Heathers, 2014). However, as the same participant will do different task conditions it is crucial to eliminate and control for carry-over effects by fully randomising the trials when possible. Personally, I found that counterbalancing conditions might not be enough to control for unwanted orders effects (e.g. for more see Study 1) while full randomisation was efficient. Registering multiple baseline measures throughout the experiment might also be required when testing if a new task is prone to carry-over effects.

2.2.4 Respiration

The control of respiratory effects on HRV has been long debated within the field. What started this debate was that in certain circumstances HRV could be affected by respiratory depth, which is the volume of air taken into the lungs, (Hirsch & Bishop, 1981), and respiratory frequency, which is the number of breaths taken per minute (Brown et al., 1993; Houtveen et al., 2002). For these reasons, it has been proposed to “correct” HRV for respiration (Grossman, 1992) either during the experiment or after the experiment with relevant statistical analyses. However, the effect of respiratory depth has been shown to account for less than 5% of the variance in the several measures of HRV -with the exception of the peak-valley method which was 10% (Lewis et al., 2012). In contrast, there are certain assumptions (i.e. 9 – 24 cycles per minute) about the respiratory rate when measuring HF-HRV to trace vagal control effects (Malik, 1996; Berntson et al., 1997). In other words, respiratory rates below or above these limits would mean that HF-HR may no longer accurately represent vagal control. However, the crucial issue with automatically correcting for respiratory rate is that it can also remove variance associated with the *common neural origin* of HRV and respiration (for review see Thayer et al., 2011). In other words, correction would remove variability associated with the *neural control* over the heart – hampering the main interest of the study (for review see Larsen et al., 2010). Furthermore, it has been reliably shown, that the non-vagal effects of respiration on parasympathetic indices of HRV are minimal during resting state or non-demanding conditions (e.g., Larsen et al., 2010;

Bertsch et al., 2012). This means that under rest-like conditions HRV is recorded best with spontaneous breathing. Nonetheless, it is suggested to record respiratory rates and to ascertain that (i) participants were breathing within the required bounds and (ii) to test if respiratory rates are significantly different across conditions (Allen, Chambers &Towers, 2006).

2.3 Biofeedback

Biofeedback is an either explicit or implicit representation of a biological signal by using one or more modalities (e.g. Kuikkaniemi et al., 2010; Nacke et al., 2011). When explicit feedback is given, the participant receives a direct correlate of their biological signal at most times in the form of a visual, auditory or tactile feedback. On the other hand, during implicit biofeedback, the signal is not explicitly presented to the participant, but instead some other detail of the experimental conditions would change. An example for the latter could be a videogame where the access to bonus items depends on the participant’s momentary HR. The studies in this PhD thesis use explicit biofeedback to maximise the possibility of finding biofeedback related changes but also to simplify the design. While there are a handful of recent cardiac biofeedback studies with a relevant design published (Jones et al., 2015; Peira, Fredrikson & Pourtois, 2014; Peira, Pourtois & Fredrikson, 2013) most biofeedback studies were run during the 70-s (Locke, 1980). Past studies used different technologies, methodological standards and analytical frameworks making it hard to apply their procedure. As the biofeedback paradigm was designed to be at the core of all four studies part of this thesis it was important to keep the task simple and control for potential confounding effects that earlier studies may have missed. Firs of all, studies in the past did not have the technical advances to acquire measures HF-HRV with such ease as nowadays so it needed to be tested if this outcome measure can be affected with the paradigm (e.g. for review on biofeedback see Williamson & Blanchard, 1979). Second, it is unclear instructions and training on regulating one's heartrate are more effective than participants simply attending to the biofeedback (White et al., 1977). For this reason, Study 1 involved attention and regulation conditions and found no differences in HF-HRV between the two conditions. For this reason, the following studies (i.e. Study 2 and Study 3) only kept the more implicit attention condition through. Furthermore, most previous studies did not account for the natural decrease in HR-s associated with the sheer adaptation to the situation (White et al., 1977). To address this, Study 1 recorded baseline before every trial and counterbalanced whether participants started with regulation or simple attention to the biofeedback signal and within that it randomised whether participants received congruent or incongruent biofeedback. Study 1 revealed that counterbalancing was not enough to account for time related confounds (for more details on this please refer to the chapter with Study 1), therefore all subsequent studies fully randomised the presentation of its trials.

It is important to note that the type of biofeedback used in this thesis is both conceptually and visually different from recent studies using "heart rate variability biofeedback" (HRVB, e.g. Lehrer & Gevirtz, 2014). HRVB is a form of cardiorespiratory biofeedback training used in mainly clinical settings as an experimental treatment of conditions like asthma and depression (Gevirtz, 2013). The procedure is based on the phenomenon of RSA which is a heart pattern that occurs as the heart rate increases during inhalation and decreases during exhalation. HRVB consists of feeding back beat by beat heart rate data during slow breathing manoeuvres where the participant tries to maximize the peak and valley difference of RSA (for illustration see Figure 2). The breathing pattern that provides the highest peaks and lowest valleys is specific to every individual (but averages around a very slow breathing such as 6 breaths per minute with longer exhalations). Given HRVB requires a high involvement of the participant and is very demanding we decided to use a simpler cardiac biofeedback and measure HF-HRV rather than RSA to quantify autonomic regulation. This meant that the biofeedback can rely on a simpler presentation cardiac activity through signalling changes in heart rate and the appearance of heartbeats. Participants found the biofeedback easily understandable during the pilot phase of this paradigm.

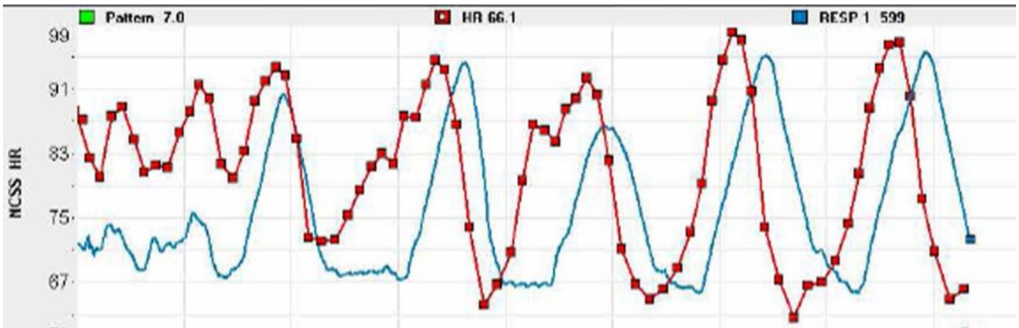


Figure 2. Original figure from Lehrer and Gevirtz (2014) depicting a heart rate variability biofeedback interface. The blue sine-wave-like pattern represents the participant’s breathing, while the red signal is associated with their heart rate. The aim of heart rate variability biofeedback is to match the two patterns by following a demanding breathing technique.

The following section will present the overarching characteristics of the cardiac biofeedback used across the four studies included in this PhD thesis, while also explaining the psycho-engineering principles certain features of the biofeedback. In all four studies an analogue output of inter beat intervals (IBI-s) were obtained live and recorded digitally on a PC into MATLAB (MathWorks, Sherborn, Mass., USA). Within MATLAB, a script using Cogent 2000 (http://www.vislab.ucl.ac.uk/Cogent2000) was created which provided the visual display of cardiac activity as a biofeedback. For a summary on the study specific alterations of the biofeedback stimulus please refer to Figure 3 below.

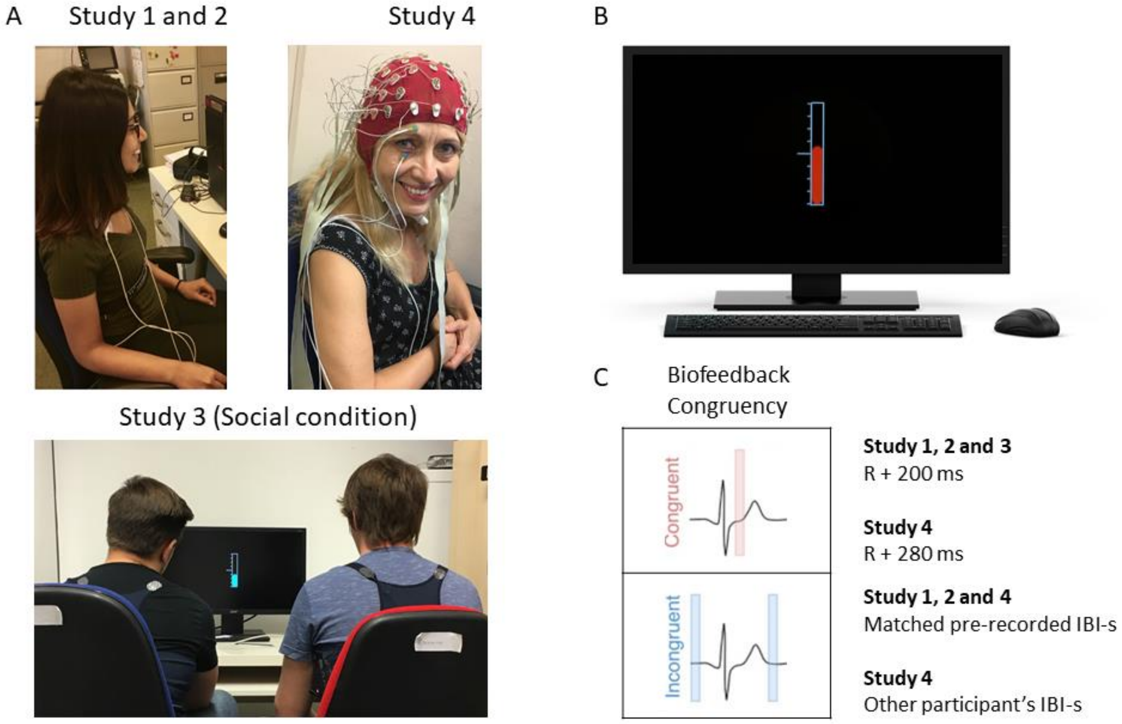


Figure 3. Illustration of participants and characteristics of biofeedback presentation across the four studies (for a video example please click here). (A) Participants completed all trials as individuals in Study 1 2 and 4, while Study 3 manipulated the social context where participants shared a screen in certain trials. Electrocardiogram (ECG) was recorded in all four studies, but Study 1, 2 and 3 also recorded respiration using either a respiratory belt (Study 1, 2) or a vest (Study 3) around participants' chest area. Study 4 also recorded participants' electroencephalogram (EEG) to capture the cortical processing of heartbeats. (B) Representation of the visual cardiac biofeedback stimulus which was specifically optimized for the rationale of every study. (C) Schematic depiction of congruent and incongruent biofeedback used with the alterations specified for each study. For the specific reason behind these changes please refer to the methods section of the corresponding study.

At beginning of every biofeedback trial, a middle size red bar appeared on the screen, representing the participant's baseline HR. As one's HR increased, the bar rose higher, and as the HR dropped the bar fell lower. Simultaneously to this, a more direct feedback of beat to beat changes was also provided, where a short pulse with a certain colour (specific to the certain study) also appeared on the bar either after 200 (Study 1, 2 and 3) or 280 ms (Study 4) after the R-wave peak. These small “delays” coincide with the time window (200 – 300 ms post R-wave) of peak systolic pressure, which is the period of maximum subjective perception of stimulus – heartbeat synchrony (Brener et al., 1993, Suzuki et al. 2013). The biofeedback was present on the screen for 5 minutes per trial at Study 1, 2 and 3, and for 10 seconds at Study 4. While most previous studies created asynchronous cardiac feedback by altering the participant’s own heart rate by making it either 30% slower or faster (e.g. 201 Suzuki et al., 2013). Unlike these studies, Studies 1, 2 and 4 used a database of incongruent IBI series collected from a completely different sample of participants who completed biofeedback task on a previous occasion. On the other hand, as participants did Study 3 in pairs, the incongruent biofeedback came from the live recording of the other dyad member’s cardiac activity. The reason for not using participants’ own cardiac activity was that during the piloting stage of the biofeedback paradigm, participants (N = 10) performed consistently below chance when trying to recognise their own biofeedback when they had to differentiate it from the altered heartbeats (Figure 4). In other words, there participants were more likely to respond incorrectly than correctly when identifying the source of the feedback - regardless if accuracy was quantified as proportion correct or response bias free d' (d prime). On the contrary, participants were more accurate when the incongruent feedback was based on cardiac data of another person (for test statistics see Table 2).

Table 2.

Test statistics on the pilot using two different types of incongruent feedback

Paired Samples T-Test

	t	df	p	Cohen's d	95% CI for Cohen's d	
					Lower	Upper
PC Other IBI - PC Altered own IBI	4.29	9	0.002	1.36	0.46	2.21
d' Other IBI - d' Altered own IBI	3.96	9	0.003	1.25	0.39	2.08

Note. Student's t-test. PC stands for proportion correct.

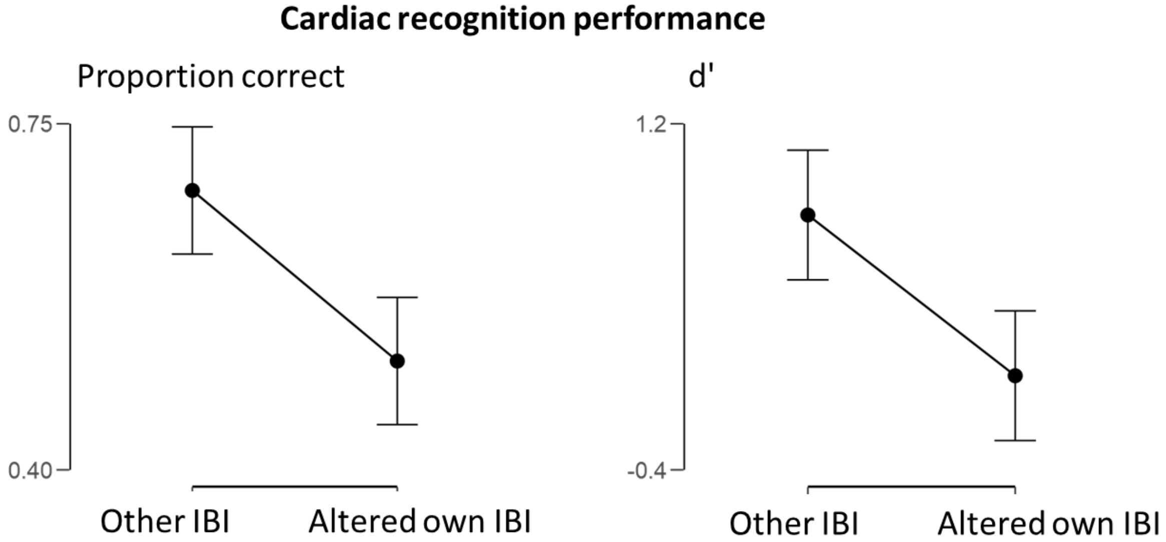


Figure 4. Differences between two types of incongruent biofeedback. Pilot data suggesting significant differences in cardiac recognition accuracy (both as proportion correct or response bias free d') when incongruent feedback used other people’s pre-recorded cardiac activity or the altered previous trial of the participant.

During the pilot phase, the biofeedback task was tested and fine-tuned to meet certain engineering features like *controllability*, *observability*, *linearity*, *temporality* and *precision* (Gaume, Vialatte, Mora-Sánchez, Ramdani, & Vialatte, 2016). The aspect of *controllability* was important for Studies 1 and 4 where participants were instructed to try and lower the bar by decreasing their heart rate. To achieve this, the movements of the biofeedback had to represent successful decreases in heart rate. As there is a natural variation of heart rate, the bar took three consecutive heartbeats to more smoothly update its height. Participants were also informed that successful regulation would still involve the up and down fluctuation of the bar to some extent. Therefore, a successfully regulated HR would be visually depicted as the bar fluctuating just under the middle line (meaning the HR is lower than at the beginning of the trial). To further strengthen the *controllability* aspect, the feedback was more sensitive to the changes in the lower ranges of heart rate. This was achieved by setting the minimum of the bar by subtracting a quarter of the baseline and creating the maximum value by adding half the baseline. The required change for every step was also relative to the baseline proceeding every trial. By doing so the biofeedback became more responsive to changes in the targeted direction and less sensitive to movement artefacts making the biofeedback more *precise*. Uniquely scaling the biofeedback to the participant also made the changes more *observable* to the participant. The re-calibration and rescaling of the biofeedback at the beginning of every trial also helped to address the *non-linearity* of the cardiac signal because the baseline HR could change during the experiment continuous it is important to rescale the biofeedback bar regularly. When participants received true, congruent feedback of their heart, the centre of the bar was calculated at the beginning of every trial either based on the average of 10 current consecutive IBI-s (Study 1, 2 and 3) or the previous trial (Study 4). This baseline was used to establish the centre of the feedback bar. This was achieved by making the centre and the scaling of the feedback relative to participant average heart rate during baseline and matching them with an incongruent feedback with an IBI series with a similar heart rate (Studies 1, 2 and 4). Furthermore, the biofeedback was continuous, which is more efficient than discrete, state-like feedback when addressing the issue of *temporality*. Finally, it was important to make the biofeedback task engaging or even fun, so participants stay motivated throughout the task. Based on the pilot and the personal conversation with the participants after completing the study, participants said they found the task somewhat challenging but more fun than usual psychological experiments.

2.4 Preregistration and Open Science

As the first two studies served as a proof concept, their hypotheses were more exploratory than the two studies following them. Study 3 and Study 4 had well-formed predictions about the outcome of corresponding experiments. Accordingly, both studies were preregistered on the Open Science Framework (OSF, https://osf.io/). This means, that the hypothesis of predicted outcome, sampling strategy, methodology, study design and analysis plan were specified a-priori to data collection. The main two benefits of preregistration are (i) preventing unreported intentional or unintentional flexibility in data analysis (Nosek et al., 2018), and (ii) hypothesizing after knowing the results (e.g., Kerr, 1998). These important issues need to be addressed as they violate the assumptions of null hypothesis significance testing such as the observed data should not condition the researchers on their analytic strategy or deriving their hypothesis.

Personally, preregistration helped me to be stricter about the sampling strategy and power calculations and placed my predictions under high scrutiny. For instance, when expecting an interaction, I specified the exact contrast and the direction of the effect. Also, accepting the potential that I might need to change certain aspects of the analysis after preregistration was humbling. Preregistration made me realise how important it is to be justifiably transparent at the moment of conceptualising the study. For instance, I remember being surprised that I did not predict a result (given it was not noted in the preregistration) that seemed obvious after looking at result. Without the preregistration, I would have been wrong to think that I had the prediction prior to the experiment. Furthermore, my experience with preregistration proved to me that, even after careful planning, analyses sometimes require amendments once the data collection started or as the data are in. While such deviations with the right reasons can be accepted, transparency in reporting will remain essential (ideally before going under the peer-review process). To give an example, one of the amendments of Study 4 was to restrain from calculating convoluted difference scores for the conditional difference in terms of heart beat evoked potential (HEP). Instead, following the suggestion of colleagues, Study 4 took on a more traditional approach. First, the analysis determined whether there are *any* differences at all between conditions on the amplitudes of HEP (using dependent samples F-statistic). Then with the given significant interaction, simple effects of the two independent variables were inspected in the traditional way. Specifically, the analysis included nine pair-wise comparisons (using specified dependent samples T-statistic) and corrected for multiple comparisons. For more on the applied amendments on Study 4 please see https://osf.io/s27vj/. I believe, transparency has the power to take away the risk of unbeneficial and arbitrary confidence in the writing of paper. To summarise, it is important to keep track of changes between preregistration and the final manuscript, but it is acceptable to conduct additional (or different) analyses if the preregistered analyses are suboptimal.

Going beyond preregistration I aimed to further apply good research practices around reproducibility, transparency, and research data management. I decided to write all manuscripts as fully reproducible APA manuscripts in R (using the package *papaja* by Aust & Barth, 2016), where the analysis is embedded within the core text of the presented work. Beyond reproducibility and transparency, I found this approach exceptionally beneficial for three main reasons. First, as the manuscript was integrated with the analysis, the room for human error was reduced in the reporting of the results. Instead of manually copying test statistics and significance results into the reporting sentences, the manuscript directly called and printed the formatted values of the test outputs within the corresponding sentence. Second, rather than running the analysis once, I found myself double-checking the script for errors and tested for alternative hypotheses more easily – simply because the analysis was at hand. Furthermore, this approach made it much easier to come back to the work after receiving the reviews and address the reviewers' comments. By eliminating the obstacles between the analysis and the written work, in my opinion one can have higher confidence in their reported results.

Chapter 3: Study 1 and Study 2

Wearing your Heart on your Screen: Investigating Congruency-effects in Autonomic Responses and their Role in Interoceptive Processing during Biofeedback

3.1 Abstract

The experience of one's embodied sense of self is dependent on the integration of signals originating both from within and outwith one's body. During the processing and integration of these signals, the bodily self must maintain a fine balance between stability and malleability. Here we investigate the potential role of autonomic responses in interoceptive processing and their contribution to the stability of the bodily self. Using a biofeedback paradigm, we manipulated the congruency of cardiac signals across two hierarchical levels: (i) the low-level congruency between a visual feedback and participant's own cardiac signal and (ii) the high-level congruency between the participants' beliefs about the identity of the cardiac feedback and its true identity. We measured the effects of these manipulations on high-frequency heart rate variability (HF-HRV), a selective index of phasic vagal cardiac control. In Experiment 1, HF-HRV was sensitive to low-level congruency, independently of whether participants attempted to regulate or simply attend to the biofeedback. Experiment 2 revealed a higher-level congruency effect, as participants' prior veridical beliefs increased HF-HRV while when false they decreased HF-HRV. Our results demonstrate that autonomic changes in HF-HRV are sensitive to congruencies across multiple hierarchical levels. Our findings have important theoretical implications for predictive coding models of the self as they pave the way for a more direct way to track the subtle changes in the co-processing of the internal and external milieus.

Keywords: vagal control, interoception, biofeedback, predictive coding, multisensory, self

3.2 Introduction

Our body has an ever-present role in the perception of ourselves and the world around us. Although this permanence provides the experience of continuity, recent models of bodily self-awareness emphasize its constructed nature and explore the ways in which different signals from multiple sources are integrated across different hierarchical levels (Apps & Tsakiris, 2014; De Preester & Tsakiris, 2009; Friston, 2005; Seth, 2013). Tsakiris, Tajadura-Jiménez, and Costantini (2011) were among the first to show that external and internal bodily signals are integrated to provide a coherent, multi-sensory experience of one's own body. The Rubber Hand Illusion (Botvinick & Cohen, 1998; for review see Tsakiris, 2010) is a classic example of the *exteroceptive* channel's input on self-awareness by showing how the experience of body-ownership can be influenced by the perception of the body from the outside, using exteroception. Watching a rubber hand being stroked in synchrony with one's own hidden hand, the visible rubber hand will overrule the real hand and will be experienced as one's own body part. The Enfacement Illusion reveals similar effects on facial identity (Sforza, Bufalari, Haggard, & Aglioti, 2010; Tsakiris, 2008). In both cases the multi-sensory (visuo-tactile) integration aims at the resolution of inter-sensory conflicts to produce a coherent representation of the world and the body - even if that induces changes in the perception of self. The other channel of information available for self-awareness contains *interoceptive* information about the body, which originates from within one's body. Recent theories emphasize the central role of interoceptive processing in representing the core self, constructed by basic homeostatic processes and inputs from visceral organs (Craig, 2010; Damasio, 2010). In summary, even though both sources are essential in the construction of selfhood, the exteroceptive signals primarily foster the malleability, whilst interoceptive signals contribute towards the stability of self-awareness (Allen & Tsakiris, 2018). Moreover, recent evidence suggests that interoceptive and exteroceptive signals are not processed in isolation. Studies using biofeedback aimed to explore the *integration* of interoceptive and exteroceptive signals, by inducing *multi-sensory* stimulation that combines interoceptive and exteroceptive signals (Aspell et al., 2013; Azevedo, Ainley, & Tsakiris, 2016; Canales-Johnson et al., 2015; De Pascalis, Palumbo, & Ronchitelli, 1991; Pfeiffer & De Lucia, 2017; Salomon et al., 2016; Schandry & Weitkunat, 1990; Sel, Azevedo, & Tsakiris, 2017; Suzuki, Garfinkel, Critchley, & Seth, 2013). Specifically, all these studies used cardiac signals as interoceptive input in combination with a visual or auditory stimulus that could either be presented synchronously or asynchronously with cardiac systole. An effect of synchrony was revealed in many different domains such as the detection of heartbeats after biofeedback (Schandry & Weitkunat, 1990), cortical processing of cardiac signals measured by heartbeat evoked potentials (Pfeiffer & De Lucia, 2017; Schandry & Weitkunat, 1990; Sel et al., 2017), and insular activity (Salomon et al., 2016) - in most cases without any conscious awareness of these effects. Some of these studies suggest that the synchrony effect can be modulated by trait-like characteristics of interoception like baseline measures of heartbeat detection (Schandry & Weitkunat, 1990), interoceptive accuracy (Azevedo et al., 2016; Sel et al., 2017) and interoceptive learning abilities (Canales-Johnson et al., 2015), while others revealed null-results in this domain (De Pascalis et al., 1991). Synchrony effects were also prominent for the identification of self (Aspell et al., 2013; Suzuki et al., 2013) or with another person (Sel et al., 2017) - suggesting a transfer-effect to higher level cognitive and social domains. Visual signals that occur at cardiac frequency were also found to take longer to access visual awareness - probably signaling interoceptive sensory attenuation (Salomon et al., 2016). These results support the hypothesis that the processing of co-occurring exteroceptive and interoceptive signals is crucial for self-awareness. However, the question arises as to whether the integration of these signals is performed via temporal synchrony or more general perhaps *amodal congruencies* between the body and environment. In the study by Peira and colleagues (2014) the biofeedback represented changes in heart rate through color changes on the screen which they updated every half second. When heart rate accelerated the colour changed towards red, when it decelerated it changed towards green. Even though the true feedback did not provide exact temporal synchrony by capturing individual heartbeats of participants an effect resembling a synchrony effect was revealed. Participants were better at intentionally downregulating their heart rate during true biofeedback than during fake feedback, suggesting a more general *congruency* (although still temporally aligned) effect. In our study we expand on this idea and explore the potential role congruency on even higher hierarchical levels of the self-model.

Predictive Coding (PC) principles provide a suitable framework for considering the mechanisms underlying synchrony (and potentially congruency) effects and the processes enabling multi-sensory integration overall. According to the PC account, the Bayesian brain continuously generates probabilistic models about the sources of sensory inputs (Apps & Tsakiris, 2014; Friston, 2005; Seth, 2013) by comparing descending predictions or priors with ascending sensory inputs. Discrepancies between the estimated and the perceived world generate prediction errors (PE-s) that the brain attempts to minimize through either action altering sensory input (i.e. exteroceptive or interoceptive signals) or by updating predictions about the causes of sensory stimuli (i.e. body ownership). A supra modal self-model would arise from the integration of multiple predictions and PE-s on different hierarchical levels and across several sensory and motor domains. This requires a novel approach in experimental design by shifting the focus from exploring the circumstances inducing PEs as a result of (a)synchrony to the study of the mechanism itself. A more targeted investigation of PE requires the identification of an outcome variable that is associated with the integration processes *per se*. With our study, we attempted to fill this gap by exploring the integration of sensory inputs and predictions by studying the effect of congruencies across multiple hierarchical levels and by proposing a way to detect physiological responses involved in the *generation* and subsequent *minimization* of PE. It has been recently suggested that, at the psychological level, interoceptive autonomic signaling can be considered as a continuous estimate of self-stability, given its role in homeostasis and its inherent self-related nature (Allen & Tsakiris, 2018). In line with this control-oriented approach, Petzschner, Weber, Gard, and Stephan (2017) suggested that interoceptive PE-s could be minimized through autonomic reflexes. This would also mean that the discrepancies between stability estimations and stability relevant signals of the external would lead to changes in autonomic signaling. In other words, PE-s or their minimization could potentially be tracked via autonomic responses.

Among a wide range of physiological factors that determine the functioning of the heart, the autonomic nervous system (ANS) is the most prominent (Thayer, Ahs, Fredrikson, Sollers, & Wager, 2012). The ANS maintains internal homeostasis and promotes the adaptive flexibility of the nervous system, which is often quantified by measures of heart rate variability (HRV). In the past HRV was suggested to have the ability to index the brain’s capacity to integrate signals from the inside and outside of the body to support adaptive regulation (Thayer et al., 2012). Together with the previously discussed theories on the contribution of autonomic responses to predictive self-processes — such as the estimation of self-stability (Allen & Tsakiris, 2018) and the minimization of interoceptive PE-s (Petzschner et al., 2017) — we hypothesized that high frequency HRV (HF-HRV) could provide a more direct physiological outcome measure associated with PE-s (or their minimization). This idea could be tested with a design that addresses the responsiveness and sensitivity of HF-HRV to essential components of predictive self-processes. For HF-HRV to serve as a useful physiological marker or outcome measure of PE related processes, HF-HRV needs to be (i) sensitive to sensory inputs that are relevant for maintaining the stability of the self across (ii) multiple hierarchical levels, such as congruencies between interoceptive and exteroceptive signals or congruencies between more general beliefs and multi-sensory input. To test this hypothesis, we performed two experiments to directly measure changes in vagal control in response to congruencies or incongruencies arising from different levels of hierarchy. Experiments 1 and 2 both explore low-level congruency effects by using a cardio-visual biofeedback that is either congruent or incongruent with participants' cardiac activity. Experiment 1 also investigates the interaction between biofeedback congruency and the level of task involvement. Here participants either actively regulate (stability facilitating behavior) or perform an attention task related to the biofeedback (stability neutral behavior). Although recent studies on biofeedback (Peira et al., 2014; Peira, Pourtois, & Fredrikson, 2013) found a facilitating effect of regulation during congruent biofeedback, this effect did not replicate in a different context (Jones et al., 2015) Moreover, these studies did not include HF-HRV in their measures. In Experiment 2 we further explore higher level congruency by manipulating participants' prior beliefs about the ownership of the biofeedback signal, and therefore succeeded in inducing congruency or incongruency between multi-sensory biofeedback and participants' beliefs (i.e. they believed the feedback belongs to them or someone else), allowing us to test, for the first time to the best of our knowledge, a higher-level congruency effect on interoceptive PE-s.

Experiment 1

3.3 Methods

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study.

3.3.1 Participants

An a priori power analysis using G*Power (version 3.1.9.2; Faul, Erdfelder, Lang, & Buchner, 2007) suggested a sample size of 35 to achieve 80% power (with $\alpha = 0.05$) by estimating a medium effect size of ($f = 0.25$) given that most previous studies with a within subject design found middle to large effects of HF-HRV reactivity (Marci, Ham, Moran, & Orr, 2007; e.g. Rainville, Bechara, Naqvi, & Damasio, 2006). To be conservative we recruited N = 40 participants (5 males, $M_{AGE} = 20.98$, $SD_{AGE} = 3.70$) through the Psychology Subject Pool of Royal Holloway, University of London. Participants gave their informed consent, with approval by the Ethics Committee, Department of Psychology, Royal Holloway University of London. No participants had to be removed from the final sample.

3.3.2 Experimental Procedure

Experiment 1 had a repeated measures design (Figure 1A) with two conditions of interest: Biofeedback Task (referring to the way people engaged with the biofeedback signal i.e. Regulation or Attention) and Biofeedback Congruency (depicting the presence or lack of congruency between participants cardiac activity and visual feedback i.e. Congruent or Incongruent). On arrival participants were seated on a comfortable chair 60 cm from a monitor (56.5 x 33.5 cm). The experiment alternated between baseline and active task measures. During the baseline recording participants were instructed to keep their eyes open and breath normally and silently think about their day. After the baseline measure it was explained to the participants that they will be observing movements of a biofeedback bar representing either their own or someone else’s heart rate changes from a previous session. The instructions described the way they could interpret the movements of the bar: when the heart beats faster the bar will be taller, and when it beats slower the bar will be shorter although a general fluctuating motion is also to be expected. A yellow pulse would also appear with every heartbeat, but participants did not receive an explanation about the pulsing of the feedback bar. In the Regulation condition participants were instructed to attempt lowering the bar as much as possible via relaxation whilst keeping their eyes open and their breathing as normal as possible. Participants had the chance to freely experiment with the feedback bar for 1 minute before the first time they attempted to regulate. In the Attention condition participants were instructed to simply count the number of a randomly appearing green pulse and subsequently report it to the experimenter - which requires high attention to the feedback, but no intentional interoception or regulation. At the end of every task participants had to indicate whether they thought the feedback was representing their own or someone else's heart. At the beginning of each task participants received instructions specific to the Biofeedback Task condition they were completing. We organized the experiment into two blocks separated by a 5-minute-long break half-way through. The order of tasks was assigned to the participants prior to the experiment in a semi-randomized and counterbalanced way (Figure 1B). Participants could either start with the Attention or the Regulation condition - within which they were randomly presented with a congruent or incongruent feedback. After the break they continued the experiment with Biofeedback Task condition that is different from the one they started with. Participants completed 4 Biofeedback tasks and 4 baseline measures, each of them lasting 5 minutes (40 minutes in total).

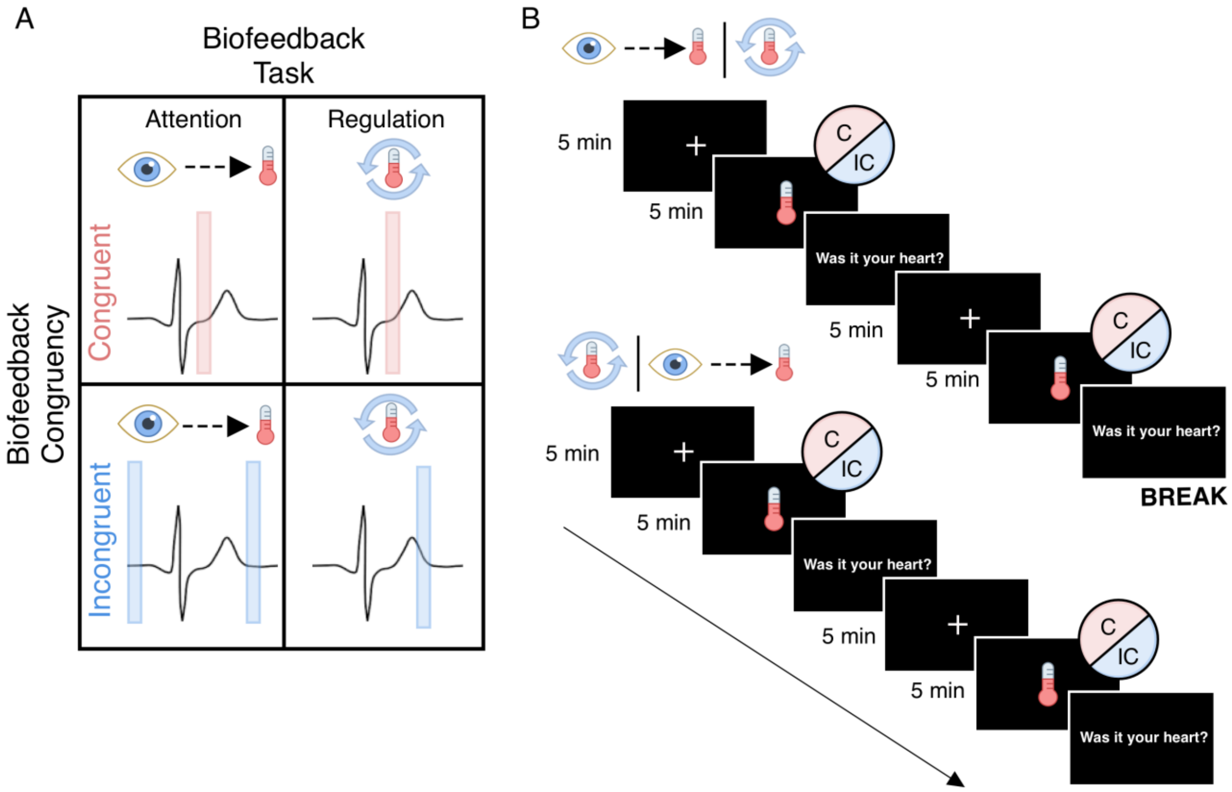


Figure 1. (A) Schematic representation of biofeedback paradigm in Experiment 1 consisting of the factors of Biofeedback Task (Attention or Regulation) and Biofeedback Congruency (Congruent or Incongruent). The eye symbol depicts the Attention condition whilst the heart symbol represents the Regulation condition of the Biofeedback Task factor. (B) Timeline of task execution, which includes two time series alternating between the biofeedback task and baseline (fixation cross). Note. C: Congruent, IC: Incongruent.

Participants received instantaneous and continuous feedback provided by a red bar that was changing across two dimensions simultaneously: in its height - whereby changes in height indicated changes in heart rate - and in its color - whereby pulses in yellow indicated individual heartbeats. Analogue output of inter beat intervals (IBI-s) was obtained online and recorded digitally on a PC into MATLAB (MathWorks, Sherborn, Mass., USA). Within MATLAB, a script was created for providing visual display to the subject during biofeedback exercises. To establish the center of the bar serving as a reference point 10 IBI-s were averaged prior to receiving any feedback. This value represented the middle point of potential values on the feedback bar. The parameters of the biofeedback bar were scaled to every individuals' baseline. To make the feedback more sensitive to the changes in the lower ranges of heart rate we set the minimum of the bar by subtracting the quarter of the participants initial heart rate while we created the maximum value by adding the half of the measured baseline. The required change for every step was also scaled accordingly the participants' baseline. The bar was set to the middle at the beginning of every task. Most previous studies created asynchronous feedback by changing the frequency of the participants own estimated heart-rate to be either 30% slower or faster (e.g. Suzuki et al., 2013). Unlike these studies, we used a database of incongruent IBI series ($N = 72$, $M_{IC\_IBI} = 779.89$, $SD_{IC\_IBI} = 142.03$) collected from a completely different sample of participants who completed the cardiac recognition task on a previous occasion. We decided to do so as in the piloting stage of this experiment we found that participants performed consistently below chance when trying to differentiate between congruent and incongruent feedback when presented with their own altered heartbeats. In other words, participants were more likely to respond incorrectly than correctly when identifying the source of the feedback. On the contrary, participants stood a higher chance to be accurate when the incongruent feedback was based on cardiac data of other individuals rather than their own. Given that cardiac recognition was of our interest in Study 1 we decided to use this database to create a task that is challenging yet accomplishable. The incongruent feedback was tailored for every participant by matching them with a similar, adjusted IBI series based on their average heart rate. We intended to keep the level of incongruency between conditions (and across participants) as constant as possible. We addressed this by adjusting, in every trial, for the percentage difference between the incongruent signal and the participant's own baseline. In the Attention condition the pulse appeared green following a randomized pattern (approximately 50% of times of all pulses).

3.3.3 Measures

Three disposable ECG electrodes were placed in a modified lead I chest configuration: two electrodes were positioned underneath the left and right collarbone and another on the participant's lower back on the left side. The ECG signal was recorded with a Powerlab 8/35 (Powerlab, ADInstruments, http://www.adinstruments.com/) using LabChart8 Pro software. The sampling rate was 1000 Hz and a hardware band-pass filter (Bio Amp 132) between 0.3 and 1000 Hz was applied. Heartbeats were detected online with a hardware-based function (fast output response), which identifies the ECG R-wave with a delay smaller than 1 ms (www.adinstruments.com/) by detecting when the amplitude exceeds an individually defined threshold. Internal lab reports using this method confirm that the software presenting the stimuli receives the transistor-transistor logic (TTL) pulse signaling a heartbeat and can process it within <2 ms. Every heart trace record was visually examined for artifacts and missing, or additional R-wave-induced beats were manually corrected if necessary. We analyzed the beat-to-beat interval variation of heartbeat traces using the HRV Add-On of LabChart8 Pro, which generates the Spectrum Plot (Frequency to Power) using the Lomb Periodgram Method (least-squares spectral analysis). Periodic components of heart rate variability aggregates in frequency bands. The respiratory frequency band is considered to range from 0.15 to 0.4 Hz in the high frequency band. We decided to use respiratory/high frequency heart rate variability as our main measure, because under appropriate recording and data processing conditions it reflects phasic vagal impact upon the heart (Berntson, Cacioppo, & Grossman, 2007) and it has been reliably used during shorter periods (i.e. 2 - 5 mins) at psychophysiological studies (Camm et al., 1996). We have specifically chosen the high frequency range instead of low-frequency (LF) or the LF / HF measure as LF reflects an indistinguishable mixture of sympathetic a parasympathetic influences rather than changes in vagal control only (Billman, 2013; e.g. Eckberg, 1997; Goedhart, Willemsen, Houtveen, Boomsma, & De Geus, 2008; Heathers, 2012; Reyes del Paso, Langewitz, Mulder, Roon, & Duschek, 2013). HF-HRV is a respiratory rate and depth dependent phenomenon and is uninterpretable in the absence of quantification of respiratory parameters. Respiratory rate (RR) is a stronger determinant of respiratory/high frequency HRV within typical breathing ranges than tidal volume (Berntson et al., 1997), therefore the administration of this parameter is fundamental. Confounds could arise if individual differences in respiration are present or there are differences across experimental conditions that push the respiratory power outside the analytical bandwidth (Berntson et al., 2007). For this reason, we registered and checked for the effects associated with the changes in respiratory rate in every condition across both studies. We used a respiratory belt transducer (ADInstruments, http://www.adinstruments.com/) to control for respiration. Given that the length of recording could affect the measures of HRV, we used the recommended 5 minutes epoch in every baseline and task, so we can relate our finding to most previous studies. We also recorded participants accuracy in recognizing the source of the feedback (Self or Other), although ideally more trials would be required for reliable measure of cardiac recognition.

3.4 Results

We used *R* (Version 3.5.1; R Core Team, 2018) for all our analyses. A test of normality was conducted for the dependent variable using the Shapiro-Wilks test and revealed that the assumption of normality was significantly violated ($p < .001$). The violation of normality is expected at measures of HRV and normally addressed by running the statistical analyses on the transformed value. We explored the distribution of different transformations with the *fitdistrplus* (Version 1.0.9; Delignette-Muller & Dutang, 2015) and *logspline* (Version 2.1.11; Kooperberg, 2018) packages. For further analyses on HF-HRV we chose the square root transformed values over the logarithmic one as the logarithmic transformation proved to be too strong a correction for the modest positive skew of the data. Descriptive statistics and confidence intervals are noted in text.

Before our main analysis we checked for potential carry-over effects. Data analysis revealed a carry-over effect depending on the order of the Biofeedback Task conditions β = 5.57, [CI] = 1.16 – 9.99, p = .013, R^2_{MARGINAL} = 0.06, $R^2_{\mathrm{CONDITIONAL}}$ = 0.25. This meant that we could only keep the first block of Experiment 1 as the effects associated with our manipulation and the carry-over effects are inseparable in the second block. As a result, Biofeedback Task became a between-subjects variable, which probably introduced limitations of power for this factor (Biofeedback Congruency remained a within-subject factor and powered-enough). The rest of results presented from Experiment 1 are only using data from the first half of the study.

Experiment 1 had one dependent variable: HF-HRV (nu) and two categorical predictors: Biofeedback Task (1 = Attention; 2 = Regulation); Biofeedback Congruency (1 = Congruent; 2 = Incongruent). Respiratory rate and baseline HF-HRV and recognition accuracy were coded as covariates and included in the model only when significantly improving the model fit (also testing for potential interaction between a certain covariate and our main predictors). We selected the optimal model by using *buildmer* (Version 1.0; Voeten, 2019) which can perform backward stepwise elimination based on the change in the set criterion (AIC in our case). We defined the maximal model as:

$\text{HF-HRV}_{\text{SQUARE\_ROOT}}$ ~ Biofeedback Congruency*Biofeedback Task + $\text{HF-HRV}_{\text{BASELINE}}$ + Respiratory Rate + (1|PPT)

The expression outside the parentheses indicates fixed effects while the expression inside the random effects defined in the model (i.e. the intercept over participants) – for more on random effects please refer to the Results section under Experiment 2. The model that was providing the best fit with our data was the following:

$\text{HF-HRV}_{\text{SQUARE\_ROOT}}$ ~ Biofeedback Congruency + Biofeedback Task + $\text{HF-HRV}_{\text{BASELINE}}$

fixedWe ran a linear regression for our main statistical analysis – using *stats* (Version 3.5.1; R Core Team, 2018) and relevant test-statistic were gathered by using *sjPlot* (Version 2.5.0; Lüdecke, 2018b) and *sjmisc* (Version 2.7.4; Lüdecke, 2018a) R packages. Our results revealed that HF-HRV (nu) was significantly higher in the Congruent condition (M_C = 37.73, SD_C = 17.46) than in the Incongruent conditions (M_{INC} = 35.07, SD_{INC} = 19.29) β = -0.58, [CI] = -1.10 – -0.06, p = .030, R^2 = 0.488, R^2_{adjusted} = 0.47 when baseline HF-HRV was included in the model β = 1.00, [CI] = 0.76 – 1.24, p = < .001(Figure 2). Even though the optimal model includes the Biofeedback Task as a factor its effect was non-significant p = .151 Results are depicted by raincloud plots (Allen, Poggiali, Whitaker, Marshall, & Kievit, 2018).

When analyzing the accuracy of cardiac recognition, we see that 71.25% of probability of correctly identifying Biofeedback Congruency across all conditions. Fitting a logistic regression on the binary values of accuracy did not reveal a significant interaction nor main effects of Biofeedback Congruency and Biofeedback Task OR = 0.29, [CI] = 0.04 – 2.08, p = .224, $R^2_{\text{Cox\&Snell}}$ = 0.03, $R^2_{\text{Nagelkerke}}$ = 0.04.

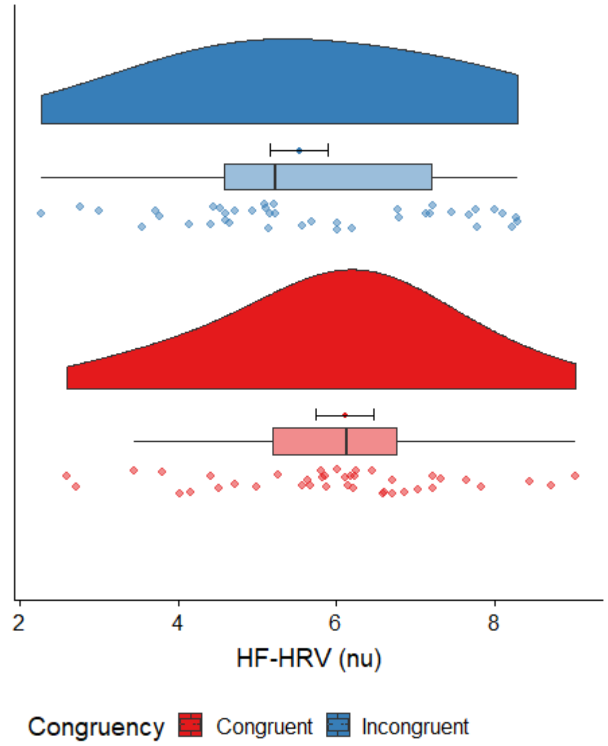


Figure 2. Effect of lower level Biofeedback Congruency (Congruent vs Incongruent) on the square root transformed HF-HRV (nu) values. The raincloud plot provides data distribution, the central tendency by boxplots and the jittered presentation of our raw data. Error bars indicate 95% confidence intervals around the estimates of the linear mixed effects model.

3.5 Discussion of Experiment 1.

We observed changes in HF-HRV associated with the integration of exteroceptive and interoceptive signals on a lower sensory level, but the level of task involvement (i.e. Attention or Regulation) did not have an additional effect. Receiving incongruent visual feedback with one's own cardiac activity was associated with a lower level of HF-HRV when compared to congruent feedback. These results indicate that differences in HF-HRV can potentially serve as an index of PE-s as it is sensitive to multisensory congruencies. It is important to note that the detected carry-over effects associated with the task order potentially makes the null finding of task involvement inconclusive. Although keeping only the first half of the experimental session was methodologically the right choice, it probably introduced power issues in terms of detecting the effects of task involvement. Nonetheless, similar to our null result, the study by Jones and colleagues (2015) found no difference in regulation performance whilst receiving true or fake feedback which might question the potential facilitating effect of an increased level of task involvement. However, further research is needed to understand whether this effect requires certain circumstances to be present, or it is indeed non-existing. For instance, the carry-over effect from Experiment 1 implies that the task involvement effect could be more prominent *after* rather than *during* biofeedback. Another option is that behavioral regulation only has an effect if arousing stimuli are co-presented with the feedback - similarly to the design of Peira and colleagues' (2013) in which participants were presented with arousing pictures during biofeedback. Finally, it is possible that the levels of task involvement were not distinct enough in our design. Asking participants to increase their heart rate might provide a better contrast to down-regulation than the attention condition.

Having established a low-level congruency effect (known as synchrony effect in previous studies), we next investigated whether this congruency effect is generalizable to higher-levels, which would suggest a more amodal role in hierarchical predictive processing. Specifically, we were interested whether the manipulation of prior belief would influence the effects of feedback congruency.

Experiment 2

3.6 Methods

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study.

3.6.1 Participants

An a priori power analysis using G*Power (version 3.1.9.2; Faul et al., 2007) suggested a sample size of 35 to achieve 80% power (with α = 0.05) by estimating a medium effect size of (f = 0.25). We recruited N = 40 (9 males, M_{AGE} = 21.60, SD_{AGE} = 5.29) participants through the Lab of Action and Body Database. To further increase the reliability of our measures, participants completed every task twice (in a completely randomized order). Participants gave their informed consent, with approval by the Ethics Committee, Department of Psychology, Royal Holloway University of London. No participants had to be removed from the final sample.

3.6.2 Experimental procedure

Experiment 2 had a repeated measures design (Figure 3A) with two conditions of interest: Belief (referring to participants' beliefs on the ownership of the feedback i.e. Self or Other) and Biofeedback Congruency (depicting the presence or lack of congruency between participants cardiac activity and visual feedback i.e. Congruent or Incongruent).

To increase reliability of our measures and to reduce proneness to carry-over effects we improved the design from Experiment 1. Most importantly, participants engaged with the biofeedback signal only through attention in every task (and not through regulation). Also, participants completed every condition twice and the order presentation was fully randomized prior to the experiment. On arrival participants were seated on a comfortable chair 60 cm away from the screen (56.5 x 33.5 cm). Again, participants alternated between baseline and active task measures. During the baseline recording participants were instructed to keep their eyes open and breath normally and silently think about their day. After the baseline measure participants received instructions about the way the biofeedback bar works in identical way as in Experiment 1. Participants' beliefs were manipulated by the instructions at the beginning of each task. Participants were told that the biofeedback belonged to them or someone else. Unbeknown to the participants these beliefs could either match the Biofeedback Congruency condition or not. The repetition of tasks increased the time to complete study (from 45 to 90 minutes), therefore the experiment was conducted over two separate days. Baseline HF-HRV was recorded before every active task (Figure 3B). The stimuli in this experiment were identical to those in Experiment 1.

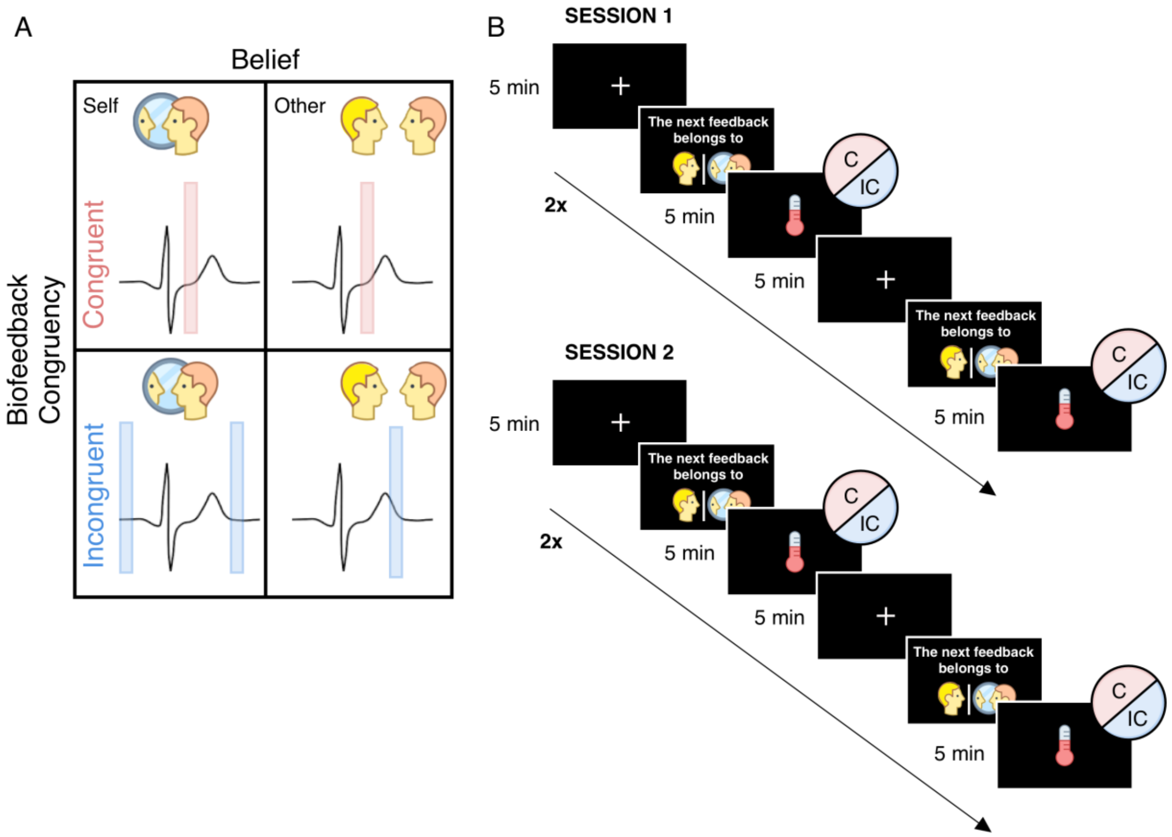


Figure 3. (A) Schematic representation of the biofeedback paradigm in Experiment 2 outlining the factors of Belief (Self or Other) and Biofeedback Congruency (Congruent or Incongruent). (B) Timeline of task execution during Biofeedback using a completely randomized pattern. Each biofeedback task was preceded by a baseline recording of HRV and the instructions on the identity of the forthcoming biofeedback. Note. BL: Baseline, BF: Biofeedback, C: Congruent, IC: Incongruent, S: Self, O: Other.

3.6.3 Measures

Apart from some exception we used the same measures in this experiment as in Experiment 1. We used the same respiratory belt transducer (ADInstruments, http://www.adinstruments.com/) to control for respiration, but due to equipment failure we had to replace the respiratory band also resulting in losing 1.77% of our respiratory rate data. The missing values for these measures were interpolated with the most recent non-missing value also known as last observation carried forward (LOCF) method. It has been recently suggested that water consumption could provoke changes in cardiovagal outflow and should be controlled during experimentation (Heathers et al., 2018). To address this issue, we contacted our participants prior to the experiment and instructed them to avoid extensive water consumption (more than a glass of water) within 1.5 hours prior to their appointment but also recorded their self-reports of actual water intake within the specified time. To make sure participants engaged with the biofeedback on an appropriate level throughout the whole task Experiment 1 conceptualized participants attention level by looking at their performance in counting green pulses. Attention scores were calculated for each trial with the following formula, where scores closer to 1 represent better performance:

$$1 - \frac{|recorded\ green\ pulses\ -\ counted\ green\ pulses|}{recorded\ green\ pulses}$$

3.6.4 Debriefing

To understand if participants detected or had any suspicion about the belief manipulation, at the end of the whole experimental session, we asked them whether there was something that stood out for them at any point in the study. If the participant's response indicated suspicion regarding the instructions on the ownership of the feedback, then the participant was given a timeline of the experiment on which they had to mark the beginning of this impression. Only one participant expressed suspicion about the study, but it was unrelated to our belief manipulation.

3.7 Results

Given that the focus of interest was the potential interaction between Biofeedback Congruency and Belief, but also to emphasize the state-like nature of our measure we considered the change HF-HRV (nu) from baseline as our primary dependent variable. Experiment 2 had two predictors: Belief (ownership of signal: 1 = Self; 2 = Other); Biofeedback Congruency (1 = Congruent; 2 = Incongruent). We chose to model our data with a Gaussian distribution and linear mixed effects as the change scores seemed to follow normality (p = .058). We tested for the effects of water consumption, the level of attention, respiratory rate, task order and repetition and baseline HF-HRV - included in the model only when significantly improving the model fit. Note that it is good practice to check for baseline covariation effects even when the analysis focuses on the change from baseline, as it provides a more precise measure of an effect than an analysis without baseline adjustment (CHMP, 2015). We applied linear mixed models for our statistical analysis - using *lme4* (Version 1.1.17; Bates, Mächler, Bolker, & Walker, 2014). Mixed effects modelling is particularly useful in within-subject designs where each subject has several measurements resulting in correlated errors for those measurements (Baayen, Davidson, & Bates, 2008). The solution to this problem is to let each subject have their own personal intercept (and/or slope) randomly deviating from the mean intercept as the errors around the personal regression lines this way will be uncorrelated. Reported p-values were computed via Wald-statistics approximation (treating t as Wald z). We selected the optimal model by using *buildmer* (Version 1.0; Voeten, 2019) which can perform backward stepwise elimination based on the change on a set criterion (AIC in our case). We defined the maximal model as:

$\text{HF-HRV}_{\text{CHANGE}}$ ~ Biofeedback Congruency*Belief + $\text{HF-HRV}_{\text{BASELINE}}$ + Water Consumption + Attention + Respiratory Rate + (1|PPT)

The model that was providing the best fit with our data based on the AIC values was the following:

$\text{HF-HRV}_{\text{CHANGE}}$ ~ Biofeedback Congruency*Belief + $\text{HF-HRV}_{\text{BASELINE}}$ + Respiratory Rate + (1|PPT)

The expression outside the parentheses indicates fixed effects while inner expression depicts the random effects in the model (i.e. the intercept over participants). Results revealed a significant interaction between Biofeedback Congruency and Belief β = 7.33, [CI] = 2.55 – 12.12, p = .003, R^2_{MARGINAL} = 0.34, $R^2_{\text{CONDITIONAL}}$ = 0.62 when including baseline HF-HRV (nu) β = -9.90, [CI] = -11.58 – -8.23, p = < .001 and respiratory rate p = .057 in the model (Figure 4A). To further investigate this interaction simple effects analysis was run with *phia* (Version 0.2.1; De Rosario-Martinez, 2015) across the levels of the factors (Biofeedback Congruency and Belief) in our fitted model. In the analysis Bonferroni corrections were applied for multiple comparisons when exploring simple effects of interaction. Results revealed a significant difference in the changes of HF-HRV (nu) between Incongruent ($M_{\text{INC\_SELF}}$ = -2.74, $SD_{\text{INC\_SELF}}$ = 15.05) and Congruent conditions ($M_{\text{C\_SELF}}$ = 2.33, $SD_{\text{C\_SELF}}$ = 15.00) when participants were told that they are looking at their own cardiac feedback, χ^2 = 7.70, p = .011. This can be considered as a replication of the lower level congruency effect identified by Experiment 1. In contrast with this when participants believed that the feedback was representing someone else's prerecorded cardiac activity there was no effect of Feedback Congruency χ^2 = 2.15, p = .285. There was a significant simple effect of Belief resulting in a difference between the Self ($M_{\text{INC\_SELF}}$ = -2.74, $SD_{\text{INC\_SELF}}$ = 15.05) and Other conditions ($M_{\text{INC\_OTHER}}$ = 1.92, $SD_{\text{INC\_OTHER}}$ = 12.56) when receiving incongruent feedback χ^2 = 5.64, p

= .035, but not during congruent feedback χ^2 = 3.52, *p* = .121 (M_{C_OTHER} = -1.39, SD_{C_OTHER} = 15.7).

This interaction effect can also be framed as higher-level or meta Congruency occurring between Belief and lower-level Congruency (i.e. Congruent with belief = when believing that Congruent feedback belongs to Self or when believing that Incongruent feedback belongs to Other; whilst Incongruent with belief = when believing Congruent feedback belongs to Other or Incongruent feedback belongs to Other). With this approach we see a significant main effect of meta Congruency between the Incongruent (M_{IC_HIGHER} = -2.06, SD_{IC_HIGHER} = 15.35) and Congruent conditions (M_{C_HIGHER} = 2.12, SD_{C_HIGHER} = 13.80) β = 3.67, *CI* = 1.28 – 6.05, *p* = .003, $R^2_{MARGINAL}$ = 0.34, $R^2_{CONDITIONAL}$ = 0.62 (Figure 4B). These results indicate that HF-HRV can be conceptualized as a generalizable index of PE-s in a hierarchical predictive model of the self as it is sensitive to the integration of different sources of information and their (in)congruency across multiple hierarchical levels.

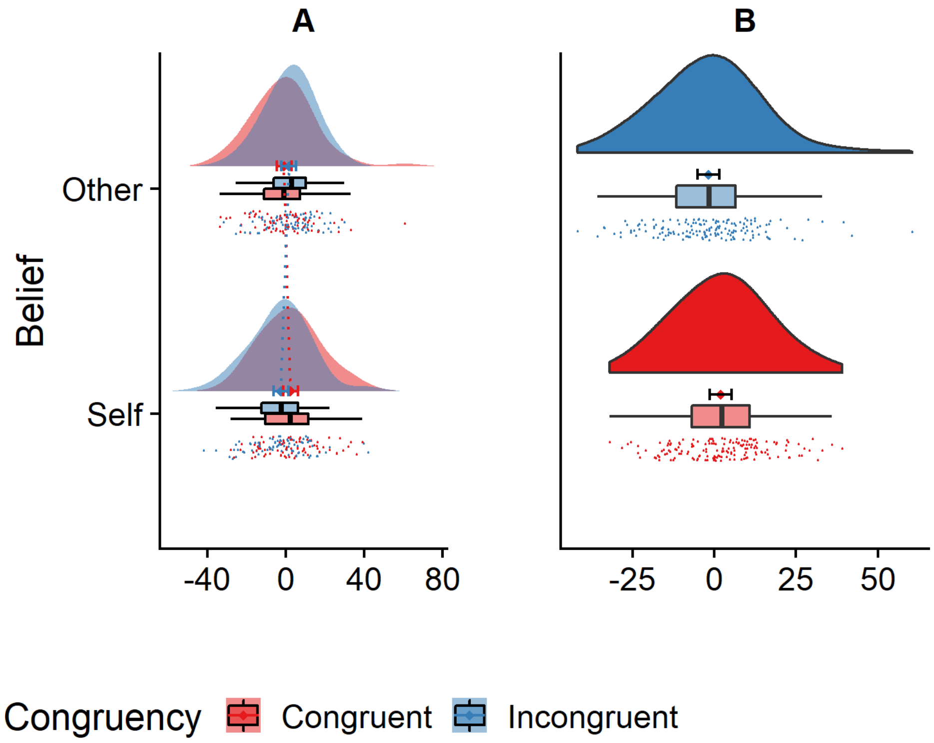


Figure 4. Higher-level congruency effect on the changes in HF-HRV (nu) from baseline depicted as (A) an interaction between Biofeedback Congruency (Congruent vs Incongruent) and Belief (Other vs Self) and as a (B) main effect where higher-level Congruency is coded as a single predictor (Congruent with Belief vs Incongruent with Belief). The raincloud plot provides data distribution, the central tendency by boxplots and the jittered presentation of our raw data. Error bars indicate 95% confidence intervals around the estimates of the linear mixed effects model. Random intercept models include baseline HF-HRV (nu) and respiratory rate as a covariate.

3.8 General Discussion

The integration of signals arising from within and outwith one's body has a primary role in self-awareness. Studies on multisensory integration (Aspell et al., 2013; e.g. Botvinick & Cohen, 1998; Salomon et al., 2016; Sel et al., 2017; Sforza et al., 2010; Suzuki et al., 2013) found evidence for both the stability and malleability of the self - mostly in the form of synchrony effects. Across two experiments, we tested whether the previously reported synchrony effects could be generalized to more abstract levels as higher-level congruency effects. Given that autonomic responses were recently suggested to reflect estimates of self stability (Allen & Tsakiris, 2018), we used an index of vagal control (i.e. the changes in HF-HRV) as a measure of congruency effects. We observed that the changes in HF-HRV were predicted by differences in congruency on both lower and higher hierarchical levels of self-processing. Specifically, Experiment 1 revealed lower HF-HRV during incongruent feedback when compared to congruent feedback. Given that low-level congruency was induced by temporal alignment across cardiac and visual domains our result from Experiment 1 corresponds to the synchrony effects reported by previous studies (e.g. Aspell et al., 2013; Salomon et al., 2016; Suzuki et al., 2013). However, to emphasize the similarities in the mechanism across different hierarchical levels we use the term low-level congruency to describe this effect. Providing support to our hypothesis of a higher-level congruency effect, Experiment 2 identified an interaction between participants' beliefs and low-level congruency of the biofeedback signal. More precisely, when participants received biofeedback that was incongruent with their belief the change in vagal control was significantly lower than in the condition when their beliefs were veridical. Experiment 2 also measured the level of attention directed at the biofeedback through quantifying performance of counting green pulses. We found that all participants performed well - suggesting that they engaged with the task and sustained their attention at a good level throughout. Our results have important implications for the predictive models of the self. Earlier models (Apps & Tsakiris, 2014; Seth, 2013; Tsakiris, 2010) were focusing on the apparent differences between different sources (i.e. exteroceptive and interoceptive) of self-relevant information, whilst novel approaches emphasize the integration of these signals - which is proposed to be reflected by the balance between stability and adaptation (Allen & Tsakiris, 2018; Seth & Tsakiris, 2018).

When interpreting our results within the PC framework we suggest that the participants' cardiac activity and their beliefs were used to generate predictive models of the timing of pulses and the movements of the biofeedback bar. The brain continuously estimates the sources of sensory input by comparing top-down predictions (priors) about sensory events and bottom-up sensory input. Incongruencies give rise to prediction errors (PE-s) that are passed upward to higher hierarchical cortical levels - that encode more abstract, supramodal representations - until they are resolved. In our study, the low-level congruency effect and the associated PE would arise from the multisensory level when unimodal exteroceptive (i.e. stability relevant visual signals) and interoceptive priors (i.e. stability estimations) get integrated. In comparison, a higher-level congruency effect could be generated by the mismatch between participants' beliefs and the multisensory biofeedback input. When PE is minimized a percept is formed that can lead to the attribution of the origin of biofeedback;

specifically, when PE cannot be minimized sufficiently then the biofeedback would be attributed to someone else. Proprioceptive PE-s can be minimized through action (i.e. active inference, Friston, Daunizeau, & Kiebel, 2009) engaging reflex arcs. Aligned with our findings, it has been suggested that *interoceptive* PE-s could be minimized through *autonomic* reflexes (Petzschner et al., 2017; Pezzulo, 2014). A way that autonomic reflexes could minimize PE is via adjusting the precision of interoceptive priors. Lowering the relative impact of interoceptive signals on perception enables the self to adapt to external stimuli whilst keeping its stability unperturbed. Given that autonomic responses could not only signal the *minimization* of PE but potentially be the error signals themselves, their exact interpretation will depend on the experimental design at hand. In our case, autonomic responses are more likely to represent *interoactions* (i.e. minimization of PE, Seth & Tsakiris, 2018) given that they arise in response to incongruencies - in contrast with a design that would focus on the effects of manipulating the autonomic responses themselves.

In our two studies the primary focus was on the physiological responses rather than the explicit self-recognition measures. However, future studies could adapt our methods, paradigms and task instructions (without the framing we used in the higher-level congruency manipulation) and ask participants whether the biofeedback originates from the self or from others, as in the design by Azevedo, Aglioti, and Lenggenhager (2016). Past research using behavioral measures revealed contradictory evidence regarding the link between interoceptive abilities (such as heartbeat detection or interoceptive accuracy) and biofeedback. While heartbeat perception seems to improve post-biofeedback (Schandry & Weitkunat, 1990), others found no difference between good or bad heartbeat perception groups in heart rate control performance (e.g. De Pascalis et al., 1991). The findings on heartbeat evoked potential are more consistent, suggesting that the neural response associated with the attention directed to one’s heartbeat is affected by the synchrony of the feedback

(Pfeiffer & De Lucia, 2017; Schandry & Weitkunat, 1990; Sel et al., 2017). However, more research is needed to understand the way low-level congruency influences interoception and whether it could be detected behaviorally. In line with our study it will be interesting to test whether state-like changes of interoception (measured by trial-by-trial cardiac recognition) are modulated by the autonomic response to the biofeedback signal. Given that the measures of HRV require a longer time window (i.e. at least 2 to 5 min) alternative indices of autonomic responses could be considered when optimizing the design of the task on cardiac recognition (e.g. measuring the pre-ejection period).

To conclude, we adopted a novel approach in our experimental design investigating the psychophysiological stability and adaptability of the self by shifting the focus from the contributing factors to the *integration process* itself. Across two experiments, we show that autonomic responses are sensitive to congruencies and incongruencies arising from the integration of sensory signals and predictions across multiple hierarchical levels. This finding provides supportive evidence for the role of autonomic responses in interoceptive processing as stability estimations that are engaged in the minimization of PE.

Chapter 4: Study 3

Standing your Ground through Change: Autonomic Regulation and Interpersonal Cardiac Synchrony during Cooperative and Competitive Social Interactions

Introduction

4.1 Abstract

One of the fundamental challenges of brains and organisms in general is to strike the right balance between stability and adaption in an everchanging environment. Stability of one's core self has been proposed to be constructed by internal homeostatic states. Recent predictive coding theories of self-processing proposed that autonomic regulation is closely linked to the integration of predictions and sensory input across multiple hierarchical levels. However, it is unknown whether their function can be extended to social interactions, which we aimed to test with a novel joint biofeedback paradigm. We explored the effects of low-level Biofeedback Congruency (i.e. Congruent or Incongruent with one's own heart) across different Social Context settings (i.e. Individual, Cooperation, Competition). To capture the effects at the intra- and inter-personal levels, we analyzed high-frequency heart rate variability (HF-HRV) and heart rate coupling between dyad members. Extending previous findings, congruent biofeedback resulted in an increase in HF-HRV. Moreover, autonomic regulation was influenced by the Social Context. In particular, HF-HRV was higher and heart rate coupling between dyad members was lower during Competition than in the Individual condition. We argue that these physiological responses reflect stability preserving processes, potentially via increased self-other differentiation on a psychophysiological level. Our results have relevant implications for predictive coding models of the self and for testing the importance of social context on physiological regulation and by extension the regulation of affect.

Keywords: Heart rate variability; Self; Self-other congruency; Biofeedback; Predictive coding; Social context; Interoception

4.2 Introduction

While one’s self is embedded in an everchanging physical and social world, the brain will continue to look for ways to maintain stability even if that requires change. According to predictive coding approaches (Friston, 2009), the brain achieves stability by minimizing differences between perceived and predicted sensory input - also known as prediction errors (PE-s). It has been proposed that stability originates in the continuous mapping of internal homeostatic states of the body (Damasio, 2010). Homeostasis is maintained by the autonomic nervous system (Thayer et al., 2012) via adaptive reactions in response to perturbations (Pezzulo, Rigoli, & Friston, 2015). Going beyond the described reactivity, recent theories of autonomic responding focus on the concept of proactive allostasis, such as maintaining stability through anticipation and change (Sterling, 2014). Allostasis contributes to stability via prospective control, in which temporary change in homeostatic set-points appear before the perturbation is present (Sterling, 2012). Given their role in homeostasis and allostasis, autonomic signaling was proposed to function as a continuous estimate of self-stability (Allen & Tsakiris, 2018) or even as “interoactions” - minimizing PE arising from different hierarchical levels (Seth & Tsakiris, 2018). Supporting this proposal, a recent study by Hodossy and Tsakiris (2020) revealed that autonomic regulation is indeed sensitive to perturbations arising from lower and more abstract hierarchical levels in the form of incongruencies. In the hierarchically organized brain, the higher a level is the further the information is from its original unisensory input and has higher relevancy for the global model of one’s self (Allen & Tsakiris, 2018). Within the predictive coding framework, when a PE error cannot be fully explained it will ascend to a higher level. The study by Hodossy and Tsakiris (2020) investigated a higher hierarchical level by influencing the validity of participants’ beliefs. There, participants could either be right or wrong to think that a cardio-visual feedback belonged to them or someone else. The authors found a significant difference when participants belief was congruent with the validity of the feedback signal and when it was incongruent. The same congruency effect appeared on a lower level when the visual signal could either be congruent or incongruent with participants’ cardiac activity. These results suggest that autonomic signals have the potential to trace PE related processes more closely, at least up to a level where multisensory information and beliefs get integrated. However, the *extent* of the generalizability of autonomic signals to social settings remains a question. It would be important to see whether the function of autonomic regulation for psychophysiological stability can be extended into social interactions, a premise that has appeared earlier for instance in the polyvagal theory (Porges, 2007). In this theory Porges (2001) reviews evidence in support of the adaptive significance of physiological responses for social behavior. According to the polyvagal theory, there are three phylogenetic shifts in the neural regulation of the autonomic nervous system, each associated with a different behavioral strategy. The last phylogenetic shift, which is unique to mammals, can be characterized with the output of the vagus nerve that can rapidly regulate cardiac activity to foster engagement and disengagement with the environment. As the vagus is linked to the cranial nerves it can regulate social engagement via facial expression and vocalization. Taking the rationale of Porges forward, Feldman (2007) suggests that physiological synchrony enables infants to engage in temporally matched social interactions. This theory considers time as a central parameter of the emotion and communication systems making the early experience of physiological synchrony critical for achieving interpersonal intimacy.

It is not the first time when autonomic regulation was proposed to track integration processes within the brain (Thayer et al., 2012). The amygdala and the medial prefrontal cortex (mPFC) are both part of the brain’s “core integration” system, because of their use of internal and external information, and their link to regulating peripheral physiology and behavior. Building on this, Thayer and colleagues (2012) have suggested that heart rate variability (HRV) could serve as an output of neural integration and can track one's capacity to effectively function in a complex environment. To study such level of complexity, autonomic signals (measured by HRV) need to be considered within a richer, social context. It has been suggested that the evolution of the autonomic nervous system was central to the development of self-regulation and social engagement (Porges, 2007). The conclusions of a recent meta-analysis (Holzman & Bridgett, 2017) provided support for the proposed association between autonomic functioning and top-down self-regulation. Based on these results, autonomic regulation could indeed be sensitive to changes in self-representations in relation to the social context where one finds herself. Findings on the psychological correlates of state-like changes in HRV are informative for predictions about the potential effects of different types of social context on autonomic regulation. To date, studies reported increased levels of HRV during successful performance on emotion regulation tasks (Butler, Wilhelm, & Gross, 2006; Ingjaldsson, Laberg, & Thayer, 2003; Smith et al., 2011) and after receiving oxytocin intranasally (Kemp et al., 2012). On the other hand, it has been observed that HRV decreases during stress (for meta-analysis see Kim, Cheon, Bai, Lee, & Koo, 2018), episodes of hopelessness (Schwarz, Schächinger, Adler, & Goetz, 2003), sustained attention (Suess, Porges, & Plude, 1994), when performing a task that involves executive functions (Hansen, Johnsen, & Thayer, 2003) and as memory-load becomes higher (Backs & Ryan, 1992). Within a dyadic context, negative social interactions (e.g. using Trier Social Stress Task, exclusion or public speaking) seem to reduce HRV, in contrast, interactions of a positive or neutral valence leave HRV unchanged from baseline (for review see Shahrestani, Stewart, Quintana, Hickie, & Guastella, 2015).

Joint action studies provide an abundance of evidence for the importance of social context when studying human behavior and cognition (Sebanz, Bekkering, & Knoblich, 2006). Joint action can be regarded as any form of social interaction whereby two or more individuals coordinate their actions in space and time to bring about a change in the environment (Sebanz et al., 2006). Several conscious and preconscious mechanisms were proposed through which social context could influence individuals including shared task representation across co-actors. Evidence for such mechanism originated from studies where participants performed a classic Simon task (Simon, 1969) together with someone else (Sebanz, Knoblich, & Prinz, 2003, 2005). In the original version, the participant’s task is to indicate a ring’s colour by pressing a button on the left or on the right side. The task becomes challenging from a task-irrelevant feature which is the pointing direction of the index finger (i.e. pointing to the left or to the right). Participants were observed to respond faster when the location of the required button press corresponded with the pointing direction of the finger (Simon, 1969). In the joint version, each participant is only responsible for pressing one of the buttons - turning the original task into a go-no-go task. When doing the go-no-go version together, participants still showed a Simon effect in contrast to the condition where participants performed the go-no-go version by themselves. While boundary conditions exist for this effect to appear (e.g. Guagnano, Rusconi, & Umiltà, 2010) and the exact reasons for the joint Simon effect have been questioned (e.g., Dolk, Hommel, Prinz, & Liepelt, 2013), other experiments employing a similar logic supports the interpretation of shared task representation (e.g. Atmaca, Sebanz, Prinz, & Knoblich, 2008; Schuch & Tipper, 2007). While social context is expected to activate shared action representations in general (Ruys & Aarts, 2010), in a competitive relationship task co-representation was found to be lower (De Cremer & Stouten, 2003). The authors suggested this effect could potentially signal that co-actors see themselves as less similar to each other in the competition condition than in the cooperation condition. Fittingly with the similarity explanation, the social Simon effect (SSE) was reported to be larger when participants think of their co-actor as more human like (Müller et al., 2011; Stenzel et al., 2014).

Building on these behavioural results, it is possible that self-stability is also affected by the social context. Within the predictive coding framework, the other person could be considered more congruent with the predictions of one’s self in a cooperative than in a competitive context. The perceived congruency between Self and other could in turn influence how other related perturbances are processed. Such proposal extends the recent suggestion that individual predictive processes could be applied to cover others’ mental states (Friston & Frith, 2015) by specifying that such process would require congruency between Self and other. Specifically, this rationale would imply that individual predictive processing could stay intact in a social context where Self is congruent with another person (e.g. cooperative situations). This is an intriguing possibility as it could capture a physiological mechanism bridging physical and psychological distance between people and tap into more global predictions about one’s Self. To summarize, in this study we wanted to explore the effect of different social settings, such as competition and cooperation. While both types of social context are expected to activate shared action representations (Ruys & Aarts, 2010), based on DeCremer and Stouten (2003) we predict that the cooperation context would lead to higher levels of task co-representation and thus self-other congruency than the competition condition. By establishing self-other congruency via synchrony, one’s interoceptive self could provide a cardinal source to understand the hidden states of others - a crucial aspect of mentalising. Putting forward the idea of embodied mentalising Fotopoulou and Tsakiris (2017) argue that the most minimal aspects of selfhood are fundamentally shaped by embodied interactions with other people starting in early infancy. Interpersonal congruency (or generalised synchrony - Friston & Frith, 2015) is in line with studies showing higher levels of shared task representation when the co-actor is more human like, which emphasizes the importance of the perceived similarity between people (Müller et al., 2011; Stenzel et al., 2014). Furthermore, it has been reported that the attribution of emotional and mental states to others seems to be influenced by egocentric tendencies (e.g. Royzman, Cassidy, & Baron, 2003). As egocentric tendencies could hinder the understanding of other people’s emotions, interpersonal congruency could potentially facilitate empathy for others. According to Hoemann and Feldman Barrett (2019) cognition, emotion and perception are considered as the same domain-general process but would differ in terms of what underlying process are prioritised by the brain. Interpersonal congruency could manifest as mentalising (i.e. cognition) when the brain prioritises mental contents and processes. In contrast when interoceptive states and changes are prioritised (i.e. emotion) then interpersonal congruency would contribute towards empathising.

Friston and Frith (2015) further proposed that generalized synchrony will inherently emerge as a byproduct when two observers are engaged in the modelling of each other’s behavior. Findings on interpersonal dynamics to date have shown that physiological synchrony (PS) can indeed be informative about the state of a relationship and shared levels of task-involvement (Palumbo et al., 2017). However, most studies mainly aimed to identify PS in certain populations and activities, rather than attempting to explicitly manipulate potential components of the mechanism. The presence of PS was observed in couples, friends, mother - children dyads when compared to baseline (Ferrer & Helm, 2013; Helm, Sbarra, & Ferrer, 2012; McAssey, Helm, Hsieh, Sbarra, & Ferrer, 2013). State-like changes in PS were also reported when comparing conflict to non-conflict conversation topics (Helm, Sbarra, & Ferrer, 2014; Levenson, Gottman, Robert W. Levenson, & John M. Gottman, 1983; McAssey et al., 2013) and also during activities like gazing or imitation when compared to baseline (Ferrer & Helm, 2013; Helm et al., 2012; McAssey et al., 2013). The remaining studies were exploring the effects of a joint task (Montague, Xu, & Chiou, 2014; Suveg, Shaffer, & Davis, 2016), competition (Chanel, Kivikangas, & Ravaja, 2012; Järvelä, Kivikangas, Kätsyri, & Ravaja, 2014), performance level (Elkins et al., 2009), trust (Mitkidis, McGraw, Roepstorff, & Wallot, 2015) or the presence of a virtual team member (Järvelä et al., 2014). We will describe studies looking at the effects of joint task and competition briefly with more detail as they have some relevance to the present experiment. It is important to note that these studies had participants who were somehow related to each other, being either mother and child (Suveg et al., 2016) or friends (or whole sample Chanel et al., 2012; Järvelä et al., 2014; i.e. half of the sample at Montague et al., 2014). Nonetheless, Suveg and colleagues (2016) observed that PS was present during a joint drawing task, but not during baseline, although the authors did not compare baseline and joint task directly. In contrast, Montague and colleagues (2014) reported higher levels of during virtual team trials than during baseline. Chanel and colleagues (2012) and Järvelä and colleagues (2014) both found higher levels of PS when playing a competitive video game than in the cooperative condition, which was the highest when it was only the two competitors present without virtual allies (Järvelä et al., 2014). Beyond those four studies, overall it was revealed, that the magnitude of PS does not function as a simple quantitative marker of affiliation of the quality of the interaction nor it is associated with a certain affective state, as PS can appear both during conflict (Levenson et al., 1983), empathy (Marci, Ham, Moran, & Orr, 2007) but even in neutral settings (Ferrer & Helm, 2013). It has been proposed that meaningful conclusions about PS would require the co-registration of autonomic regulation (Palumbo et al., 2017). We decided to include both PS and autonomic regulation in our measures because together they have the potential to reveal valuable information about the changes in physiological stability and relatedness both on the individual and dyadic levels.

To summarize, the present study aimed to explore the effects of low-level Biofeedback Congruency across different social context settings on participants' own autonomic regulation and their physiological synchronization with others. We preregistered our hypotheses which we deposited on the Open Science Framework (to access project use the following link: https://tinyurl.com/y6mfs734). Below we present the summary of our hypotheses about the expected physiological differences across conditions.

We hypothesized an interaction between Biofeedback Congruency and Social Context on the measure of high frequency heart rate variability (HF-HRV), which is a selective index of phasic vagal cardiac control. Specifically, we predicted the simple effect of Social Context on HF-HRV to arise in opposite direction during *Congruent* biofeedback (i.e. *Cooperation* > *Competition*) and *Incongruent* biofeedback (i.e. *Cooperation* < *Competition*) biofeedback. This model could be described as higher-level congruency between the Self-Other congruency and low-level Biofeedback Congruency.

We expected a main effect of Social Context on PS, more precisely on heart rate coupling within dyads. Specifically, we predicted both the *Cooperation* and *Competition* conditions to have higher levels of cardiac synchrony when compared to the *Individual* condition (i.e. *Cooperation* > *Individual*, *Competition* > *Individual*), but we did not expect the *Cooperation* and *Competition* conditions to be significantly different from each other (i.e. *Cooperation* = *Competition*). Please note that with our design the factor of Biofeedback Congruency is not meaningful for the analysis on PS.

4.3 Methods

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study.

4.3.1 Participants

Sample size was calculated a-priori to the study using the ANOVApower package by Lakens and Caldwell (2019), which is a simulation-based power-analysis method for repeated measures designs. For the power calculation the means and standard deviation of a similar design by (Hodossy & Tsakiris, 2020) were used. This analysis revealed that 80 participants (40 dyads) would be enough to reach 82.02% power with a Cohen's f = 0.33 for the interaction analysis between Social Context and Biofeedback Congruency. To be conservative, we recruited a total of N = 82 healthy participants (53 females; M_{AGE} = 23.74, SD_{AGE} = 5.64) through the Psychology Subject Pool of Royal Holloway, University of London. Participants did not know each other prior to the experiment We excluded 1 dyad from the final analysis due to cardiovascular drift hampering the validity of live biofeedback of heart rate changes, leaving N = 80 participants in the final sample. Participants gave their informed consent, with approval by the Ethics Committee, Department of Psychology, Royal Holloway University of London.

4.3.2 Procedure

Our experiment had a single-blinded, repeated measures design with different number of categorical factors for the specific dependent variable (i.e. autonomic regulation or PS). Autonomic regulation was analyzed at the individual level and had two categorical factors such as Biofeedback Congruency (the congruency between participants own cardiac activity and the visual signal i.e. *Congruent* and *Incongruent* Biofeedback) and Social Context (i.e. *Individual, Cooperation, Competition*). In our original plan we optimized our approach to detect a potential interaction between Social Context and Biofeedback Congruency. For this reason, we initially planned to use the *Individual* condition as a covariate rather than a separate level (leaving us with a 2x2 design) when analyzing autonomic regulation. However, as the interaction term was not improving the model fit significantly, we decided to turn to an approach that is better suited to the actual design of the experiment (i.e. 3x2 design). Without the inclusion of the *Individual* level in the Social Context factor, we found the results could be misleading. For more on model selection please refer to the Results section of the manuscript.

When measuring the effects on PS we could keep our planned approach exploring the effect of one categorical independent variable Social Context with three levels (i.e. *Individual*, *Cooperation*, *Competition*). On the level of the dyad, Biofeedback Congruency cannot be used as a factor because only one dyad member receives *Congruent* feedback during the trial. For more on this please see below.

On arrival, participants were seated on comfortable chairs, 60 cm away from their monitor (Acer V226HQL, 21.5“) in a dimly lit, sound-attenuated room. The two monitors were placed on the opposite sides of the room, which had an opaque screen in the middle - blocking the view from one side of the room to the other. After the completion of the consent form, three disposable electrocardiogram (ECG) electrodes were placed in a modified lead I chest configuration. Specifically, two electrodes were placed underneath the left and right collarbones, while the third was attached to the participants' left lower back. First, we recorded participants' baseline cardiac activity for 5 minutes, while they were sitting in silence with their eyes open, looking at a black screen. After this, the experimenter explained to the participants that in every trial both members of the dyad will be presented with the same cardiac feedback (for more on the biofeedback please see the Stimuli section) belonging to one of the dyad-members. In other words, within the same trial, one participant received the *Congruent* biofeedback (representing their own cardiac activity) and the other participant received the *Incongruent* biofeedback (representing the other person's cardiac activity). Previously was shown that participants' beliefs about the biofeedback mattered (Hodossy & Tsakiris, 2020) therefore we decided to explicitly tell participants whose cardiac feedback they will be seeing next. Participants were told that their task will be to watch this feedback, and continuously count the number of times that their preassigned colour (either green or blue) appeared on the screen. They were informed that they can earn more money beyond the default payment for the participation if they perform well. Depending on their own and -in some conditions- the other participant's counting accuracy, participants received additional points for each trial. In the *Individual* condition participants had their own separate screens at different parts of the room and were scored based on their own performance only. For the Social Context condition, participants were told how their performance will influence each other's scores in the next trial and were moved to the appropriate location of the room. Thus, in both the *Cooperation* and *Competition* conditions participants shared the screen and influenced each other's scores. In the *Cooperation* condition, participants were scored as a team, both receiving a score based on their average performance. In the *Competition* condition, the person who counted more accurately (even if it was only a 1% difference) received the points of both participants, leaving the other person with no points. To ensure that participants moved around the room equally, participants shared the screen of the person who was getting the *Congruent* feedback in that trial. At the end of each trial, participants were asked to write the counted number of their preassigned colour on separate pieces of paper which they handed to the experimenter, who recorded their response. The experiment consisted of 6 fully randomized 5-min-long trials. Participants were only given feedback on their performance by the end of the whole study.

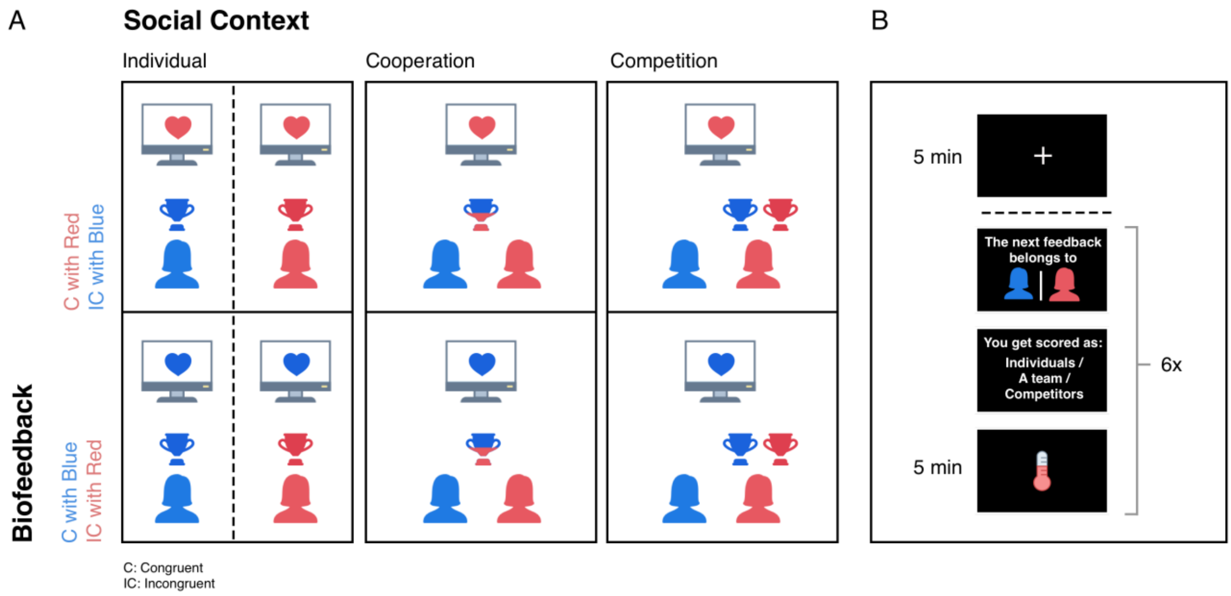


Figure 1. (A) Schematic representation of the joint biofeedback paradigm. The 2 x 3 repeated measures design consisted of the factors of Social context (i.e. Individual, Cooperation, Competition) and Biofeedback Congruency (i.e. Congruent, Incongruent). During the Congruent trials the biofeedback was linked to one of the dyad members’ cardiac systole. (B) Timeline representing the progression of trials that started with a single, 5-minute-long baseline measure - registered at the beginning of the whole experiment.

4.3.3 Stimuli

Analogue output of inter beat intervals (IBI-s) were obtained online and recorded digitally on a PC into MATLAB (MathWorks, Sherborn, Mass., USA). Within MATLAB, a script using the Cogent 2000 (http://www.vislab.ucl.ac.uk/Cogent2000) was created to provide the visual display of cardiac activity as a biofeedback. At beginning of the biofeedback, a middle size red bar appeared on the screen, representing one’s baseline heart rate. As one’s heart rate increased, the bar became taller, and as the heart rate dropped the bar followed it by becoming shorter. Simultaneously to this, a more direct feedback of beat to beat changes was also provided, where a short pulse with a random colour also appeared on the bar 200ms after the R-wave peak. This small delay coincides with the time window (200 - 300 ms post R-wave) of peak systolic pressure, which is the period of maximum subjective perception of stimulus - heartbeat synchrony (Brener, Liu, & Ring, 1993; Suzuki, Garfinkel, Critchley, & Seth, 2013). Each pulse could take one of the following 6 colors (represented with the corresponding code on the RGB scale inside the parentheses): orange (1, 0.5, 0), yellow (1, 1, 0), green (0, 0.8, 0), cyan (0, 1, 1), blue (0, 0.5, 1), magenta (1, 0, 1) -with the default colour of the bar being red (1, 0, 0). Participants were given an example of all possible colors as a reference before every trial and were pointed out the colour of the pulse they will be counting in the experiment (either blue or green). Although the presentation of these colors was fully randomized, the target colors such as green and blue were assigned to a double probability to appear. We made this decision to keep participants engaged with the counting, prompting attention to the biofeedback throughout the task. At the beginning of every trial, the average inter-beat-interval (IBI) of 10 consecutive heartbeats were measured of the participant who was receiving the *Congruent* biofeedback in that trial. This baseline was used to establish the center of the feedback bar. To make the feedback more sensitive to the changes in the lower ranges of heart rate (and less sensitive to movement artifacts in the higher range) we set the minimum of the bar by subtracting a quarter of the baseline while we created the maximum value by adding half the baseline. The required change for every step was also relative to the baseline proceeding every trial.

4.3.4 Data analysis

4.3.4.1 High frequency - heart rate variability

The ECG signal was recorded with a Powerlab 8/35 (Powerlab, ADInstruments, http://www.adinstruments.com/) using LabChart8 Pro software. The sampling rate was 1000 Hz and a hardware band-pass filter (Bio Amp 132) between 0.3 and 1000 Hz was applied.

Heartbeats were detected online with a hardware-based function (fast output response), which identifies the ECG R-wave with a delay smaller than 1 ms (www.adinstruments.com/) by detecting when the amplitude exceeds an individually defined threshold. HF-HRV was calculated using the HRV Add-On of LabChart8 Pro, which generates the Spectrum Plot (Frequency to Power) using the Lomb Periodgram Method (least-squares spectral analysis). Periodic components of HRV aggregates in frequency bands - the high frequency band ranging from 0.15 to 0.4 Hz. We decided to use HF-HRV as our outcome measure, because under the right conditions it reliably reflects phasic vagal impact (parasympathetic activation) upon the heart (Berntson, Cacioppo, & Grossman, 2007) even during shorter, 2 to 5 -minute-long periods (Camm et al., 1996).

4.3.4.2 Cross recurrence quantification analysis

While repeated evidence suggests that PS can be detected using more traditional correlational analyses and group-level trends of synchrony, the associated results may be too superficial to identify the complex, time-dependent components. These techniques use averaging to determine similarity between signals along an entire epoch and ignore dynamic changes that take place throughout. It has been proposed that keeping data on the level of single dyads and using an analysis that is time-varying has the greatest potential to track dynamic processes involved in PS (Palumbo et al., 2017). For this reason, we chose the cross-recurrence quantification analysis (CRQA) method (Shockley, Santana, & Fowler, 2003) which does not assume the stationarity of the data (i.e. the statistical properties of the process staying the same over time) and retains single-dyad-level information. Moreover, CRQA can capture many properties of the heart rate dynamics that would otherwise be lost due to averaging in more traditional correlation analysis.

4.3.4.3 Cross recurrence plots

A cross-recurrence plot for dyads consists of a symmetrical square matrix with a time-series of dyad members along both the x and y axes. Here the time series consist of heart rates sampled in every 300 ms throughout each 5-min-long trial (based on the suggestions by Wallot, Fusaroli, Tylén, & Jegindø, 2013). The plot creates a black dot every time when a phase space trajectory goes through approximately the same region at both dyad members (for more on the method please refer to Marwan, Wessel, Meyerfeldt, Schirdewan, & Kurths, 2002). These cross-recurrent plots represent the core of the analysis as they provide ways to identify and quantify both global and small-scale structures on the plots. The phase space was reconstructed by using the time-delay method (Takens, 1981), which requires the estimates of embedding dimension and time delay (Marwan, Carmen Romano, Thiel, & Kurths, 2007). First, we standardized the heart rates of everyone, and estimated the embedding parameters of delay and dimension for the whole sample. Following the tutorials of Wallot (2017) and Wallot & Leonardi (2018), we used the false nearest neighbors algorithm to estimate the embedding dimension (resulting in 7 on the average for the whole sample). Then we estimated the time delay parameter using the first local minimum of the mutual average information function of the time- series (resulting in 9 on the average for the whole sample). After this, we used these average time delay and embedding parameters when running CRQA for every dyad, which ensures a common standard throughout our dataset (Wallot & Leonardi, 2018). The final required parameter, radius was chosen separately for every dyad by incrementally increasing its value until recurrence rate the (percentage of dark dots in the recurrence plot) yielded 2% (Marwan et al., 2007). Generally speaking, a smaller radius means that a smaller interval is needed for the time series to be recurrent. By keeping the recurrence rate fixed we can compare our dyads based on other CRQA metrics.

4.3.4.4 Quantification

To quantify the similarity of participants' heart rates, we performed the quantification part of CRQA [Shockley2003]. While there are several CRQA components (e.g. recurrence, determinism, laminarity, entropy) capturing different aspects of synchrony, they are all somewhat related to one another. To be considerate about the number of metrics in our analysis we decided to focus on the measure of %Det (i.e. determinism) as it implies the strength of the coupling between dyad members. Specifically, if the recurrence metric is kept the same in the sample then higher levels of %Det imply stronger coupling. %Det is calculated as the percentage of recurrent points that form lines that are parallel to the central diagonal line in the recurrence plot. Nonetheless, we recommend researchers the exploration of the other CRQA metrics as it can offer invaluable information which can be of interest for future studies. To clarify, CRQA applies the same methods as the recurrence quantification analysis (RQA). The only difference is that while CRQA explores the shared dynamics of two systems (i.e. x axis is one person and y axis is another person on the plot), RQA represents the dynamics of a single system (i.e. both axes are the same exact time series that belongs to one person).

4.4 Results

4.4.1 Autonomic regulation

Our main behavioral outcome variable was phasic vagal input on the heart measured by HF-HRV. This analysis had two categorical predictors: Social Context (i.e. *Individual* vs. *Competition* vs. *Cooperation*); Biofeedback Congruency (i.e. *Congruent* vs. *Incongruent*). We used *R* (Version 3.5.1; R Core Team, 2018) for our analyses. Specifically, we selected the optimal model by using the *buildmer* (Version 1.0; Voeten, 2019) which can perform backward stepwise elimination, based on the change in the set criterion (AIC in our case). For linear mixed effects modeling we used the package *lme4* (Version 1.1.17; Bates, Mächler, Bolker, & Walker, 2015). Relevant test-statistic were gathered by using *sjPlot* (Version 2.5.0; Lüdecke, 2018b) and *sjmisc* (Version 2.7.4; Lüdecke, 2018a).

Mixed effects modelling is particularly useful in repeated measures designs where each subject has several observations that involves correlated errors for those measurements (Baayen, Davidson, & Bates, 2008). To resolve this one can let each subject to have their own personal intercept (and/or slope), randomly deviating from the mean intercept (and/or slope). As a result of using this approach the errors around the personal regression lines will be uncorrelated. Another reason for using mixed effects modeling instead of repeated measures ANOVA is the presence of an inherent hierarchy in our observations (i.e. participants are grouped within dyads) making it important to test for the presence of nested random effects. Reported p-values were computed via Wald-statistics approximation (treating t as Wald z). We tested for the covariate effects of Gender, Water consumption, Order of conditions, BMI, Respiratory rate, Baseline HF-HRV, %Det and Attention. We interpolated 3,75% of missing Respiratory rate data, which occurred because the vests measuring respiration did not fit 3 of our participants. Attention was quantified as the following:

$$1 - \frac{|presented\ pulses\ -\ counted\ pulses|}{presented\ pulses}$$

Covariates (centered when continuous) were included in the optimal model only if they improved the model fit significantly based on the AIC values. We defined the maximal model as:

$$HF - HRV \sim Biofeedback\ Congruency\ *\ Social\ Context\ +\ Baseline\ HF - HRV \\ +\ Respiratory\ Rate\ +\ Attention\ +\ BMI\ +\ \%Det\ +\ Water \\ +\ Order\ +\ Gender\ +\ (1|ID/Dyad)$$

The optimal model providing the best fit with our data was the following:

$$HF - HRV \sim Biofeedback\ Congruency\ +\ Social\ Context\ +\ Baseline\ HF - HRV \\ +\ BMI\ +\ Water\ +\ Order\ +\ (1|ID)$$

The expression outside the parentheses indicates fixed effects, while the expression inside reflects the random effects defined in the model (which here is the random intercept for each participant). To check if the assumptions of linear regression were held, we visually inspected the plot of residuals and fitted values where no fitted pattern was revealed (linearity confirmed) and the residuals had a similar amount of deviation from our predicted values (homoskedasticity confirmed). We also investigated the Q-Q plot and histogram of the residuals, which both indicated normality. When running the optimal model, we found a replication of the previously reported (Hodossy & Tsakiris, 2020) a significant decrease in HF-HRV during receiving *Incongruent* cardiac feedback (M_{IC} = 40.88, SD_{IC} = 18.42) when compared to the *Congruent* condition (M_{C} = 43.15, SD_{C} = 19.64) (β = -2.27, [CI] = -3.95 – -0.59, *p* = .008) In the same model, results revealed a significant increase in HF-HRV in the *Competition* condition ($M_{COMPETITION}$ = 44.13, $SD_{COMPETITION}$ = 19.51) when compared to the *Individual* condition ($M_{INDIVIDUAL}$ = 40.34, $SD_{INDIVIDUAL}$ = 18.31) β = 3.69, [CI] = 1.62 – 5.75, *p* = .001, $R^2_{MARGINAL}$ = 0.52, $R^2_{CONDITIONAL}$ = 0.76. There was no significant difference between the *Individual* and *Cooperation* ($M_{COOPERATION}$ = 41.57, $SD_{COOPERATION}$ = 19.24) conditions (*p* = .564), and the difference between *Competition* and *Cooperation* did not survive the corrections *p* = .087) (Figure 2). The presented p values of Social context were

Bonferroni corrected for three comparisons. For results on the covariates and uncorrected test statistics please refer to Table 1.

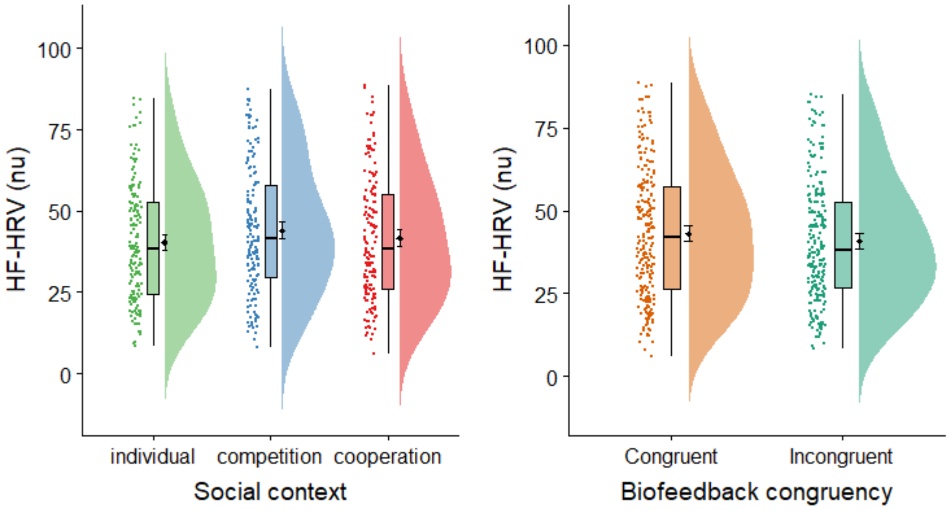
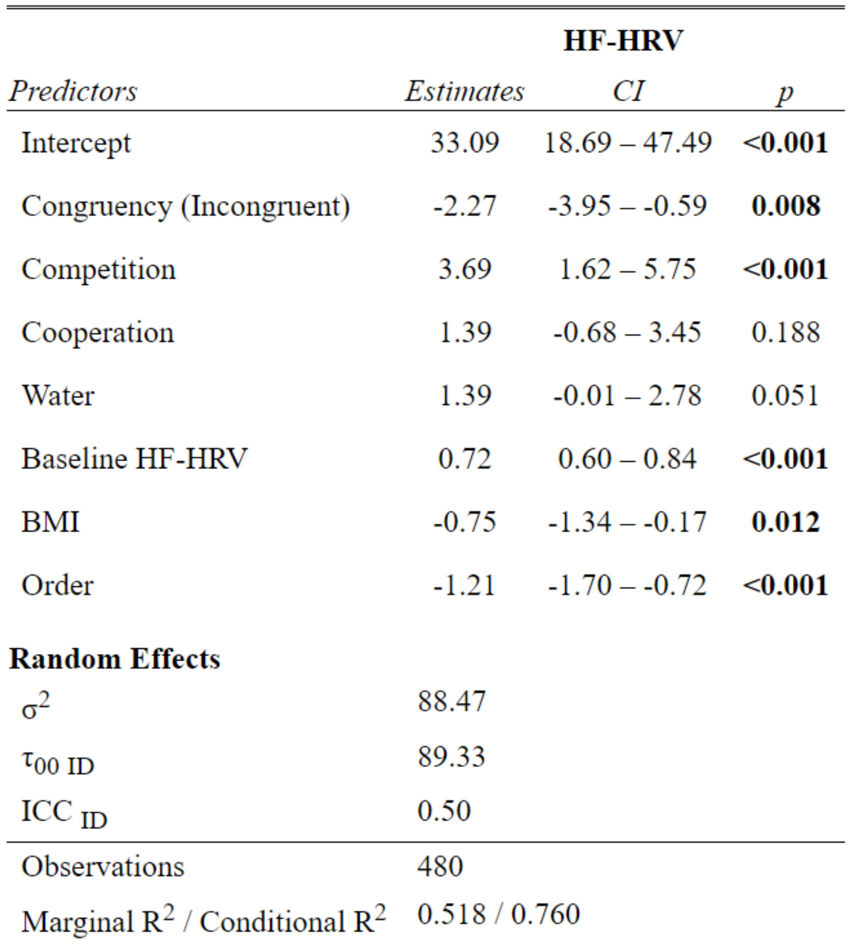


Figure 2. Effect of Social Context and Biofeedback Congruency on HF-HRV. The raincloud plots (Allen, Poggiali, Whitaker, Marshall and Rogier, 2018) provide data distribution, the central tendency by boxplots and the jittered presentation of our raw data. Error bars indicate 95% confidence intervals around the estimates of the linear mixed effects model. The random intercept model includes baseline HF-HRV (nu), water consumption, order of tasks and BMI as covariates.

Table 1.

Linear mixed effects model on HF-HRV (with the unadjusted significance values)



		HF-HRV	
Predictors	*Estimates*	*CI*	*p*
Intercept	33.09	18.69 – 47.49	**<0.001**
Congruency (Incongruent)	-2.27	-3.95 – -0.59	**0.008**
Competition	3.69	1.62 – 5.75	**<0.001**
Cooperation	1.39	-0.68 – 3.45	0.188
Water	1.39	-0.01 – 2.78	0.051
Baseline HF-HRV	0.72	0.60 – 0.84	**<0.001**
BMI	-0.75	-1.34 – -0.17	**0.012**
Order	-1.21	-1.70 – -0.72	**<0.001**
Random Effects			
σ^2	88.47		
τ_{00} ID	89.33		
ICC $_{\text{ID}}$	0.50		
Observations	480		
Marginal R^2 / Conditional R^2	0.518 / 0.760		

To test for overall differences in Arousal and Attention, we repeated the same analysis on Heart rate, and Attention but found that neither Biofeedback Congruency nor Social Context improve the model-fit significantly during model selection. While participants were only provided with feedback about their performance at the end of the study, we wanted check if there were any differences in HF-HRV between the winning and losing trials in the *Competition* condition. We created a two-level categorical factor Result (i.e. *Win*, *Lose*), but again, when estimating the components of the optimal model Result factor did not got selected by *buildmer* as it did not improve the model-fit significantly.

4.4.1 Physiological synchrony

Second, we explored the effects of Social Context on heart rate synchrony for every dyad in each condition. To quantify the interpersonal dynamics of heart rates, we produced cross-recurrence plots for each dyad and calculated their CRQA metrics (following Wallot & Leonardi, 2018) using the *crqa* R package (Version 1.0.9; Coco & Dale, 2014). Figure 3 represents three cross-recurrence plots of one dyad across the three social context conditions: *Individual*, *Cooperation* and *Competition*. These cross-recurrence plots consist of the recurrences between the dyad members' heart rate data series, each axis belonging to one of the dyad members. The global differences between the cross-recurrence plots of the *Competition* and the two other Social Context conditions stand out. In the *Competition* condition, there is a more graded diffusion of points throughout more regions of the plot. These large-scale patterns, or in other words the typology of the (cross)recurrence plot (Marwan et al., 2007) can be either homogenous, periodic, drift, or disrupted. The visualization of such typologies allows for their classification. For the *Competition* and *Cooperation* conditions crossrecurrence plots resemble homogenous typology, while the fading to the upper left and lower right corners in the *Individual* condition suggest a drift (Figure 3). Furthermore, it is visible that the *Competition* condition has fewer diagonal lines (i.e. diagonal to the line of identity, which would appear in the middle) than the *Cooperation* and *Individual* conditions.

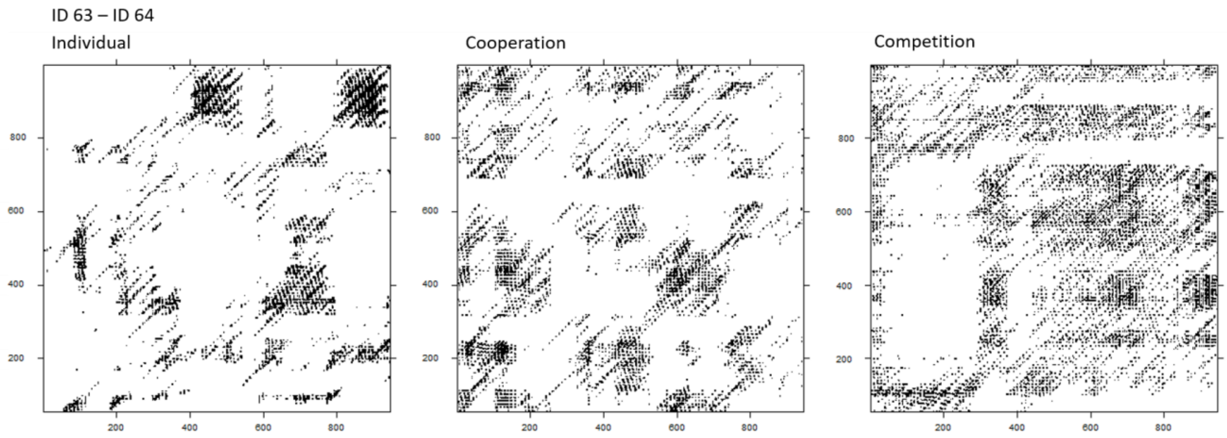


Figure 3. Cross-recurrence plots of the Individual, Cooperation and Competition conditions within the same dyad keeping the recurrence rate at 2%. The x (here ID 63) and y (here ID 64) axes represent time series of normalized heart rates, sampled at every 300 ms.

After quantifying the small-scale differences on the recurrence plots for each dyad using CRQA we analyzed the differences in %Det scores across conditions by running a random intercept model that also included the normalized baseline %Det as a covariate. We tested for the assumptions of linear regression (which also apply to the linear mixed effects models) using the same technique as before and found that none of the assumptions were violated. Given that only one participant of the dyad received *Congruent* feedback in each trial, this analysis only had one categorical predictor: Social context (i.e. *Individual*, *Competition* and *Cooperation*). After applying Bonferroni corrections for three comparisons, the results revealed a significant decrease in %Det in the *Competition* ($M_{\text{COMPETITION}}$ = 56.76, $SD_{\text{COMPETITION}}$ = 8.99) condition when compared to the *Individual* condition ($M_{\text{INDIVIDUAL}}$ = 60.32, $SD_{\text{INDIVIDUAL}}$ = 7.72) (β = -3.57, [CI] = -5.37 – -1.76, *p* = < .001, R^2_{MARGINAL} = 0.29 $R^2_{\text{CONDITIONAL}}$ = 0.53), when including baseline %Det in the model (β = 0.47, [CI] = 0.31 – 0.63, *p* = < .001). There were no significant differences between the *Competition* and *Cooperation* ($M_{\text{COOPERATION}}$ = 58.26, $SD_{\text{COOPERATION}}$ = 8.20) conditions (*p* = .313) and the difference between *Cooperation* and the *Individual* condition did not survive the corrections for multiple comparisons (*p* = .076).

While the differences in synchrony across conditions are meaningful regardless of the overall presence of synchrony, we wanted to check if the dyads in the original sample differed from randomly assigned pairs of dyads. To create the control group, we created new dyads by allocating participants randomly to one-another by making sure that participants did not get assigned to their own heart rate time series from another condition. We ran a linear regression with a between subject categorical predictor of *Group* denoting if the dyad members were randomly assigned or not (i.e. *Original* vs. *Random*) on the %Det metric. The analysis revealed no significant difference between the original and random dyads (*p* = .738), suggesting the lack of synchrony of the original dyads on a group level.

4.5 Discussion

To maintain a stable experience of one's self, the brain and body need to predict and adapt to incongruencies and perturbations emerging from the internal as well as the external social and physical environment (Pezzulo et al., 2015, Sterling (2014)). Recent predictive coding theories about the Self have suggested that autonomic signals play a ubiquitous role in estimating (Allen & Tsakiris, 2018) or eventually creating the required psycho-physiological stability via the minimization of PE-s (Petzschner, Weber, Gard, & Stephan, 2017). As a consequence, autonomic signals need to be sensitive and generalizable to higher-order self-representations, therefore be sensitive to PE-s arising from incongruencies between abstract predictions about the self and complex multisensory input. We here tested that hypothesis by considering the modulation of autonomic signals within a social context, namely in a joint biofeedback paradigm designed to affect different degrees of Self-Other integration (De Cremer & Stouten, 2003) using competitive and cooperative settings. Specifically, we aimed to explore the changes in the previously described low-level congruency/synchrony effect (e.g. Hodossy & Tsakiris, 2020; Salomon et al., 2016, Suzuki et al. (2013)) across different Social Context conditions (i.e. *Individual*, *Cooperation* and *Competition*) on participants' autonomic regulation and PS.

While we did not observe the hypothesized interaction between Biofeedback Congruency and Social Context, HF-HRV was significantly higher during *Competition* than in the *Individual* condition (i.e. *Competition* > *Individual*). We did not find the same pattern for the *Cooperation* condition (i.e. *Individual* = *Cooperation*) nor was there a significant difference between the *Competition* and *Cooperation* conditions (i.e. *Competition* = *Cooperation*). Replicating the results by Hodossy and Tsakiris (Hodossy & Tsakiris, 2020), we found a significantly lower level of HF-HRV when receiving *Incongruent* cardiac feedback in comparison to the *Congruent* condition. When analyzing the strength of the dynamical coupling between the dyad members' heart rate series, we found a significantly lower level of PS during the *Competition* than in the *Individual* condition (i.e. *Competition* < *Individual*). There were no significant differences between the *Individual* and *Cooperation* conditions (i.e. *Individual* = *Cooperation*) nor between the *Cooperation* and *Competition* (i.e. *Cooperation* = *Competition*) conditions in terms of PS. It is important to note that, Social Context and Biofeedback Congruency did not have any effects on participants heart rate, suggesting no significant differences in arousal across conditions (Pollatos, Matthias, & Schandry, 2007).

Physiological outcome measures are often assumed to reflect certain mental states (Porges, 2007). It is beneficial to consider our results within the context of previously reported psychological correlates of state-like changes in HRV. For example, during successful performance on emotion regulation tasks (Butler et al., 2006; Ingjaldsson et al., 2003; Smith et al., 2011) HRV increases, while the experience of stress is associated with a decrease (for meta-analysis see Kim et al., 2018). *Competition*, in the presence of another person, was reported both to decrease (i.e. playing a car-racing gaming Veldhuijzen Van Zanten et al., 2002) but, in a way comparable to the present findings, to also increase participants' parasympathetic activity (i.e. tie knotting game while keeping the conversation, Danyluck & Page-Gould, 2019). Such conflicting findings could be attributed to participants' awareness of their relative performance throughout the task. During car-racing, participants know about their performance in relation to the other person throughout the duration of the task. In contrast, in our biofeedback paradigm and the knot-tying task, participants only learn about their overall performance at the end of the whole experiment. Explicit knowledge of outcome in a *Competition* context could give rise to positive/negative affect associated with the experience of winning and losing respectively which, without successful emotion regulation, could lead to decreased HRV (Butler et al., 2006; Ingjaldsson et al., 2003; Smith et al., 2011). In contrast, under conditions where outcome performance during the task is not available, such as in our task, one can more closely monitor social context related, instead of affect-related, physiological outcomes. Danyluck and Page-Gould (2019) suggested that increases in HRV could arise from participants mutually dividing their attention over their social engagement and the task. However, this interpretation does not seem to be plausible in the context of the present experiment, as based on previous research, HRV is expected to decrease as attention shifts from focused to divided (Backs & Ryan, 1992) or requires higher working memory performance (Hansen et al., 2003). Given that we observed HF-HRV to change in the opposite direction (i.e. an increase not a decrease), it is unlikely that participants' attention got more divided. Furthermore, our results revealed that task-related attention (measured by the performance of counting randomly presented pulses) did not change across conditions. The absence of change in attention also invalidates the explanation that competitors would disengage from one another and be more focused on their own task.

A suitable cognitive mechanism associated with the increase in HF-HRV and the decrease in PS during *Competition* could be the changes in the shared representation of the co-actor's task. Previously, participants were found to share each other's action representations when jointly performing a task (Sebanz et al., 2006), regardless if they had a competitive or cooperative relationship established (Ruissen & De Bruijn, 2016). However, shared task representation was observed to be lower in competitive than in cooperative relationships potentially via the co-actors perceiving themselves as less similar to one another (Ruissen & De Bruijn, 2016). When applying the terminology of the present work the proposed similarity interpretation can be considered as congruency between self and other. Therefore, the decreases in interpersonal congruency could be interpreted, especially in a competitive context, as a mechanism for preserving self-stability and the self-other boundaries. Being perceived as incongruent with one's self in the context of *Competition*, the brain could predict the other person as a potential perturbance in one's self-stability. Within a predictive coding framework, the increase in HF-HRV could be considered as an indication of successful allostasis (Sterling, 2012), which prepares the body for a future perturbation (e.g. here winning/losing against the other person). Unfortunately, we did not record participant's predictions about their loss and win for each *Competition* trial. Nonetheless, future studies should test this prediction by including perceived wins and losses in their design.

Fittingly with the self-stability interpretation, it has been reported that *Competition* is linked to an increased activation of the mPFC (Decety, Jackson, Sommerville, Chaminade, & Meltzoff, 2004). This brain area has crucial role within the brain's core integration system of which HRV was proposed to serve as an output tracking one's capacity to effectively function in a complex environment (Thayer et al., 2012). On the physiological level, an effective stability preserving mechanism could correspond to the presence of self-regulatory processes, here reflected by the increases in HF-HRV (Butler et al., 2006; Ingjaldsson et al., 2003; Smith et al., 2011) and decreased PS across partners.

It is important to address that HF-HRV and PS did not change in the *Cooperation* condition when compared to the *Individual* condition, which can have two potential explanations. Fittingly with our null-results, it has been reported that shared task representation (measured by SSE) at co-actors in a cooperative setting did not become larger compared to the *Individual* condition (Ruissen & De Bruijn, 2016) which has been attributed to humans' default tendency to cooperate (Bowles, Gintis, Bowles, & Gintis, 2013). Participant's default tendency to cooperate tendency could explain the lack of difference in HF-HRV between our *Individual* and *Cooperation* conditions, although we are mindful of interpreting null effects. An alternative explanation could be that higher levels of shared task-representation and thus self-other congruency might only occur when there is a stable association between people's identities (Constable, Elekes, Sebanz, & Knoblich, 2019) like in preexisting close, significant relationships (for review see Palumbo et al., 2017) or during face-to-face communication (Bolis & Schilbach, 2018; Danyluck & Page-Gould, 2019). To summarize, the absence of a stable association between Self and Other could also explain why we could not detect an interaction between Biofeedback Congruency and Social Context. Using the suggested methods above, future studies could experiment with different ways to manipulate Self-Other integration between dyad members.

To conclude, we investigated the role of autonomic regulation to maintain physiological stability under different Social Context conditions. Our novel joint biofeedback paradigm aimed to test whether the function of autonomic regulation can be extended to the Self in social interactions. While we could not affect the integration process to the hypothesized extent (i.e. induce higher levels of self-other congruency in the *Cooperation* condition), our findings are still relevant for understanding how autonomic regulation participates in social interactions. In addition to the replication of the low-level congruency effect on HF-HRV, we also showed that autonomic regulation is sensitive to changes in social context. We argue that the increase in HF-HRV and the decrease in PS during *Competition* reflect stability preserving processes potentially via increased self-other differentiation during *Competition* on a cognitive-level.

Chapter 5: Study 4

Neurobehavioral Differences across Strategies Used in Cardiac Recognition

5.1 Abstract

The maintenance of psycho-physiological stability requires the ability to infer the state of our body (interoception) and to predict its future evolution. Yet standard measures of interoception do not have this functional approach as they are typically limited to the conscious perception of single heartbeats. We here present a new biofeedback paradigm to explore the effect of three strategies (i.e. exteroceptive, active or passive interoceptive) on interoceptive inference – defined here as the ability to recognize one's own heart. We observed an increase of cardiac recognition and a more pronounced cortical processing of heartbeats across both interoceptive strategies as compared to the exteroceptive one. We also observed the highest level of metacognition at the active, control-based interoceptive strategy. Strategy-specific cardiac recognition was linked to the modulation of cortical processing of heartbeats, exclusively in the passive interoceptive condition. We suggest that while both active and passive strategies increase the precision of the interoceptive channel, they exert distinct influences on different levels of the interoceptive hierarchy.

Keywords: interoception, strategies, metacognition, biofeedback, predictive coding

5.2 Introduction

The same external world surrounds everyone, yet its experience will always be *unique* to the individual, because it is anchored to what one feels, thinks and does in the moment. It has been proposed that the subjectivity of this experience is underpinned by the continuous mapping of internal homeostatic states of the body (Damasio, 2010). In an ever-changing physical and social environment, the brain and body need to find ways to maintain psycho-physiological stability to preserve one's viability and self-continuity at a psychophysiological level. The maintenance of a stable internal environment is realized on various hierarchical levels (Seth & Tsakiris, 2018), via homeostatic (reactive) and allostatic (prospective) control (Khalsa et al., 2017). On a lower hierarchical level, the autonomic nervous system supplies homeostasis and adaptive response to stress or more generally perturbations. Its functioning is mainly non-conscious as it operates through reflex-like *autonomic responses* (Pezzulo, Rigoli & Friston, 2015). In contrast, allostasis is a proactive control process that is aimed at minimizing energy costs by adaptively anticipating future needs of the body – oftentimes manifesting in conscious regulatory *actions* (Paulus, Feinstein & Khalsa, 2019).

Both control mechanisms necessitate the ability to distinguish between exteroceptive (originating from the external world) and interoceptive (arising from our body) sources of sensations. Evidence from multisensory studies suggests that interoceptive and exteroceptive signals are not processed in isolation (e.g. Pfeiffer & De Lucia, 2017; Salomon et al., 2016; Suzuki, Garfinkel, Critchley & Seth, 2013), which aligns with current concepts from computational neuroscience on the "Bayesian brain" (e.g. Allen and Friston, 2018). Within this framework the brain is assumed to actively construct generative models by integrating different sources of inputs. The appeal of predictive coding (PC) theories is that they offer an elegant description for the intertwined processes of inference, learning and prediction. The core principle in PC proposes that an organism is driven to minimize 'free energy', which is the sum of differences between the actual and predicted sensory input (Friston, 2009). The brain's *prior* predictions about the causes of its sensations are updated by *prediction errors* (PE-s), which are that part of the sensory data that the prior cannot explain. In perceptual inference, the flow of information for probabilistic predictions is top-down through hierarchical brain pathways. In contrast, PE-s are passed upward to hierarchically higher cortical levels – encoding abstract, supramodal representations - until they are resolved.

Priors, predictions and PE-s are characterized by their respective precision (Friston, 2009) that represents the reliability of the information. Therefore, precise PE-s will update less precise priors. In contrast, precise priors will be less likely to lead to an update. In other words, in any moment, context or modality, precision will determine if PE-s *or* priors will have bigger role in the resulting percept (Ainley, Apps, Fotopoulou & Tsakiris, 2016). Several studies within the PC framework support the view that the optimization of precision-weighting happens through attention-gating (Vossel et al., 2014). This means that attention directed at a certain exteroceptive or interoceptive channel will increase its relative precision and thus the impact of associated PE-s arising from that channel (Feldman & Friston, 2010).

While most interoceptive signals support homeostasis without the need for awareness, we are also capable of consciously attending to our interoceptive sensations. Research on interoceptive accuracy (IAcc) – the ability to precisely and correctly monitor changes in the internal body state – has been mainly studies in the context of cardioception. The reason behind this is practical as heartbeats represent discreet physiological events whose conscious perception can be easily quantified. It has recently been proposed that individual differences in IAcc depend on the *precision* in interoceptive systems (Ainley et al., 2016). According to this argument, people with high IAcc are able, by attention, to prioritize interoception over other sensory channels and thus adjust the relative precision of their interoceptive priors and PE-s.

Of relevance, the heartbeat evoked potential (HEP) – an electrophysiological brain response reflecting the cortical processing of individual heartbeats – has been found to be more pronounced in people with high IAcc (Pollatos & Schandry, 2004). Focusing on the potential underlying mechanism, HEP is proposed to reflect the precision-weighted PE of every heartbeat (Ainley et al., 2016). A recent study by Petzschner and colleagues (2019) provides support for this interpretation. Their results revealed that the amplitude of HEP is indeed sensitive to attentional differences between exteroceptive (attention to white noise) and interoceptive conditions (focus on heartbeats).

There are two standard IAcc tasks currently in use. In the *Heartbeat Discrimination* task (Whitehead, Drescher, Heiman, & Blackwell, 1977) individuals report the perceived synchrony of their heartbeats with a series of external stimuli (usually auditory cues). In contrast, *Heartbeat Counting* (Schandry, 1981) requires the individual to mentally track their heart beating during short time windows and report the counted number of heartbeats. Although these tasks are not flawless (for a summary see Paulus et al., 2019) the research on conscious cardiac perception remains highly relevant, as it provides the means to explore how the brain-body communication is achieved through inference. Understanding how interoceptive inference takes place is crucial because for the attainment of optimal control, the brain requires a model of the external world and the ability to make inferences and then choose suitable (not necessarily motor) actions in advance of any potential instability (Sterling, 2014).

Unfortunately, the aim to sustain optimal control has not been an integral part of the IAcc tasks described above, due to their assumptions about the way people process their cardiac signal while interacting with the world. Specifically, such tasks assume that people can (i) consciously perceive individual heartbeats and (ii) use this single heartbeat-related sensory input for making inferences. These assumptions do not reflect the functional and ecological role that interoception plays, namely the ability to monitor changes of internal body state (Khalsa et al., 2017) and to predict their future (Pezzulo et al., 2015). The present study aimed at bridging the gap between measure and definition and sought to go beyond these limiting assumptions to explore a more functional aspect of interoception. Using a biofeedback paradigm, we tested if the accuracy of interoceptive inference (here the ability to correctly recognize biofeedback representing one’s own cardiac activity) differs across different strategies used when engaging with one’s ongoing cardiac activity. We choose three strategies that all involve a combination of interoceptive and exteroceptive elements but differ in the levels of engagement they require and the feature of the biofeedback they emphasize. Specifically, we were interested in how these different strategies could inform the process of inference - measured by explicit cardiac recognition. The first, neutral strategy *Attend* aims to bring participant’s focus to exteroceptive characteristics of the cardiac biofeedback signal by asking participants to attend for randomly appearing pulses while looking at their cardiac biofeedback. The second strategy *Feel* relies on *passive* interoception, where participants attempt to feel random pulses while looking at their cardiac biofeedback. And lastly, the third strategy *Regulate* has an *active*, control-oriented approach to interoception as participants attempt to regulate their own interoceptive signal while looking at their cardiac biofeedback. The reason for including the *Regulate* condition was to emphasize the *function* of interoception, which is to maintain homeostatic/allostatic control (Stephan et al., 2016a). Specifically, we propose that a control-based inference (*Regulate*) has the potential to track interoception in a manner that is more relevant to anticipatory control (i.e. allostasis), as it requires the attention *to* but also the control of bodily states and their development, rather than the simple perception of individual heartbeats (*Feel*).

Interoception has several facets related to sensing, interpreting, and integrating information about the state of inner body systems (Khalsa et al., 2017). As recent studies have highlighted the importance of visceral precision on metacognition (Allen et al., 2016; Hauser et al., 2017), we also decided to explore strategy-related changes of interoceptive insight (i.e. correspondence between subjective experience and behaviour) with our cardiac recognition paradigm.

To summarize, here we investigate the consequences of three different strategies (i.e. *Attend, Feel, Regulate*) for inferences about the self, in relation to the external world, with an integrative approach. Our study uses a novel cardiac recognition paradigm and employs measures from several domains of interoception to capture behavioural, cortical and metacognitive aspects. Specifically, we contrast exteroception and the traditional passive interoceptive approach with an active regulation strategy, with which we aim to harness the functionality of interoception related to the monitoring and prediction of *changes* of broader internal states. We preregistered our hypotheses under the Preregistration Challenge by the Open Science Framework which can be viewed at https://tinyurl.com/y33bgvcd. Below we present the summary of our hypotheses about the expected neurobehavioral differences across strategies.

1. Participants' performance on the cardiac recognition task will not have a ceiling or floor effect, therefore will be able to detect individual differences.
2. Strategies will differ in their ability to inform accurate cardiac recognition on a behavioral level, leading to better performance in the cardiac recognition task (represented by higher d' values) following the pattern of increasing levels of engagement across strategies (i.e. *Attend* < *Feel* < *Regulate*).
3. There will be an interaction between the three distinct strategies and lower-level biofeedback congruency (i.e. synchrony) which will affect HEP amplitudes, again in the direction if increasing levels of engagement (i.e. *Attend* < *Feel* < *Regulate*).

In addition, we also employed exploratory analyses on the differences in metacognitive abilities across the three distinct strategies (i.e. *Attend*, *Feel*, *Regulate*) and the link between behavioral measures of strategy-specific cardiac recognition performance and corresponding modulation of HEP amplitudes.

5.3 Methods

We report how we determined our sample size, all data exclusions, all manipulations and all measures in the study.

5.3.1 Participants

As our design involved a combination of behavioural and neural measures, we carefully considered our sample size and the number of trials from several angles before running our experiment. When conducting mixed effects analysis Brysbaert and Stevens (2018) suggests 1600 trials across all participants to reach good levels of power. To avoid fatigue, we maximized the total number of trials for each participant at 156 trials, taking approximately 1 hour (i.e. 26 trials/condition; for more information see Figure 2B), which also fits the requirements of a signal detection task (Macmillan & Creelman, 2004). In terms of the EEG data, the unit of the analysis are the epochs around individual heartbeats. With an average of 60 BPM and 26 ten-second-long trials we can expect around 260 epochs in every condition, which fits the recommendations for ERP studies by Boudewyn, Luck, Farrens, and Kappenman (2018). To be conservative, we recruited a total of N = 34 healthy participants (14 females; $M_{AGE} = 28.71$, $SD_{AGE} = 8.71$) through the Psychology Subject Pool of Royal Holloway, University of London. Participants gave their informed consent, with approval by the Ethics Committee, Department of Psychology, Royal Holloway University of London.

5.3.2 Procedure

Our experiment had a 2 x 3 repeated measures design (Figure 1) with 2 conditions of interest: Strategy (referring to the instructions participants received about the strategy they should use for cardiac recognition i.e. *Attend*, *Feel* or *Regulate*) and Biofeedback Congruency (the congruency between participants own cardiac activity and the visual signal i.e. *Congruent* and *Incongruent* biofeedback).

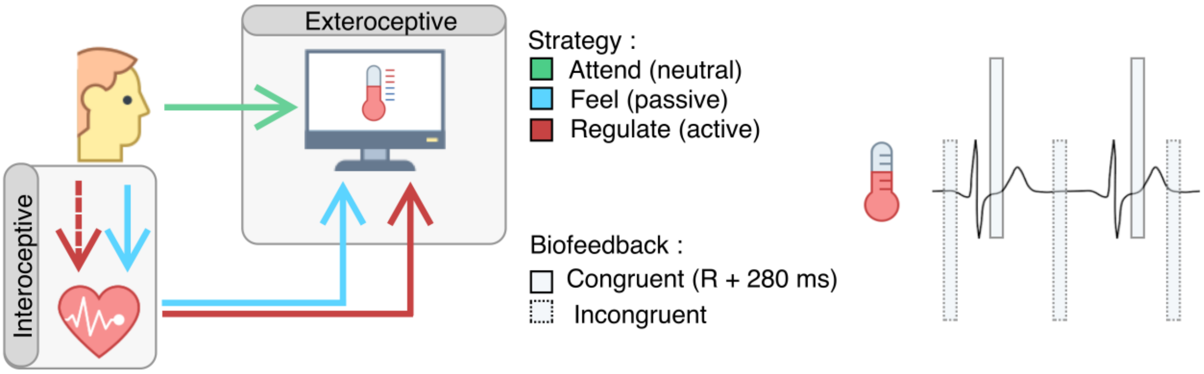
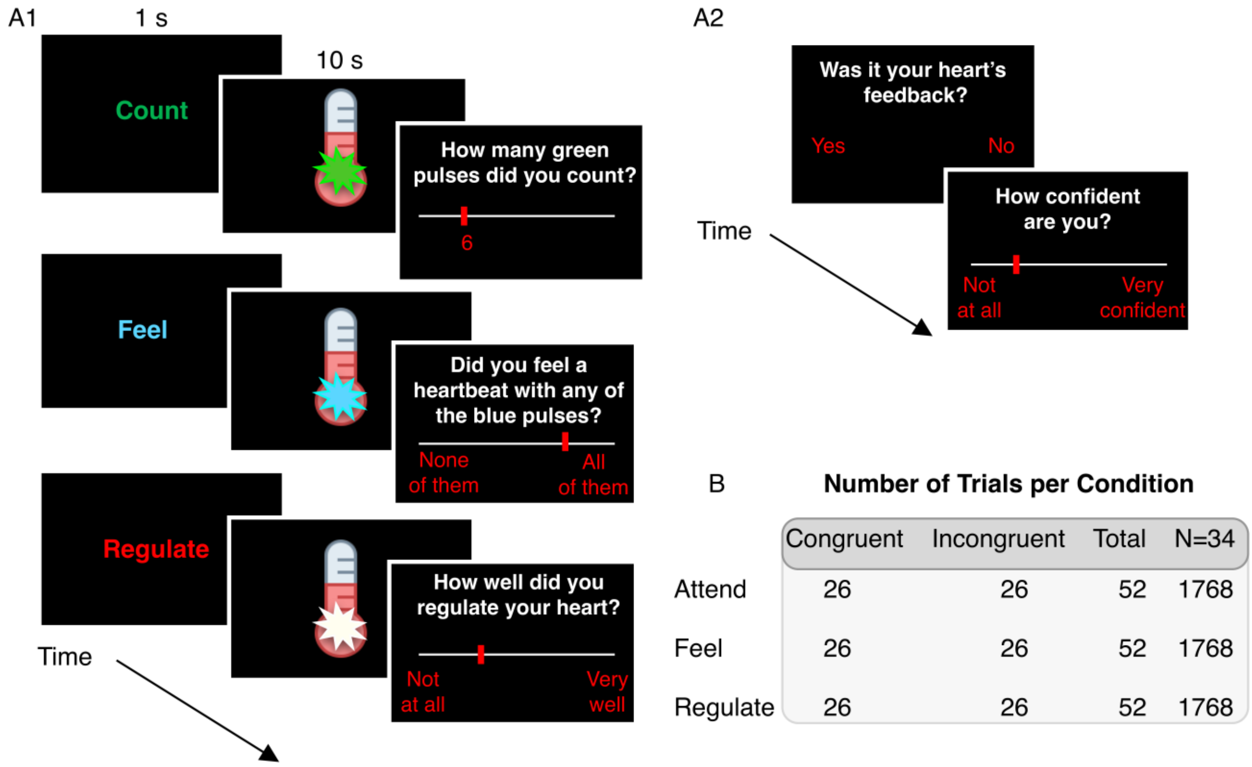


Figure 1. Schematic representation of the cardiac recognition paradigm. The 2 x 3 repeated measures design consisted of the factors of Strategy to use for cardiac recognition (exteroceptive: *Attend*; interoceptive: *Feel*, *Regulate*) and Biofeedback Congruency (*Congruent*, *Incongruent*). During the Congruent trials feedback presentation was linked to participants’ cardiac systole. The style of the arrow represents the passive/active (full/dashed line) stages of the participant’s engagement with the cardiac/feedback signal when using a particular Strategy.

On arrival participants were seated on a comfortable chair 55 cm from a CRT monitor (19.6 x 19.7 inches, Sony CPD-E530) in a dimly lit, sound-attenuated room. First a standard 3-lead electrocardiogram (ECG) was attached to participants’ chest, then we measured their baseline high frequency heart rate variability (HF-HRV) for 5 minutes, while sitting in silence with their eyes open, looking at a black screen. This was followed by the practice version of the experiment, where the experimenter answered the participants’ questions, made sure that all instructions were clear, and the participant had the chance to practice each of the three strategies *once* – all these steps taking approximately 15min in total. After the practice session, participants were equipped with the EEG electrode cap as well as the external electrodes. The experimenter explained to the participants the reason why artifacts are an issue for EEG analysis (i.e. blinks, saccades, and clench of jaw) and the ways they could contribute clean data, by asking participants to demonstrate such artifacts and showing them the resulting EEG signal. Participants were encouraged to time their blinking for the response phase only and to avoid any movements during the 10-second-long biofeedback presentation.

The experiment contained 156 randomly presented trials with a length of approx. 15-20 s (i.e. 10 seconds of biofeedback presentation and an unlimited response phase). At the beginning of every trial, a colour-coded word appeared on the screen for one second, representing the strategy participants should use to interact with their cardiac signals and/or the subsequent biofeedback (for details of the biofeedback see the Stimuli section). The words were colour-coded (i.e. "Count" – green, "Feel" – blue and "Regulate" – red) to guide participants attention to the feature of the biofeedback signal that was crucial for the upcoming strategy. In the *exteroceptive* trials (i.e. the *Attend* condition which was signalled by the *green* word "Count") participants were asked to count randomly presented *green* pulse(s). During the *interoceptive* trials participants were required to engage with the biofeedback through the context of their own cardiac signals. Specifically, when presented with the *blue* word "Feel", participants had to track whether they felt a heartbeat in time with the randomly presented *blue* pulse(s). When presented with the word "Regulate" written in *red*, participants were instructed to focus only on the up-and-down movements of the *red* bar – although to ensure compatibility with other conditions the pulsing was still present. In this trial their task was to attempt to lower the biofeedback bar, by reducing their own heartrate while breathing normally. After the strategy-prompting word participants were presented with the biofeedback for 10 seconds (for more on the stimulus itself please refer to the *Stimuli* section). Across all three strategies, participants were instructed to avoid explicitly thinking if they are seeing their own or someone else's biofeedback and simply focus on applying the strategy that was assigned to the current trial. Once the biofeedback disappeared from the screen, participants answered a strategy-specific control question to ensure that the right strategy was used in that trial. In the *Attend* condition, participants input the exact counted number of counted green pulses using a sliding scale once feedback had concluded. For the *Feel* trials, individuals indicated if they felt *any* of the blue pulses, by using a continuous sliding scale with the endpoints "None of them" and "All of them". In the *Regulate* condition, participants reported how well they thought they regulated their *own* heart (not the biofeedback bar), by using a sliding scale with endpoints: "Not at all" and "Very well" (Figure 2.A1).

The remaining two questions were the same for all three conditions. Participants first had to report if they thought the feedback represented their own heart (i.e. "Yes" or "No"). Participants could take as long as they wished when responding and received no feedback on their accuracy. As learned in practice phase, participants would use their strategy-specific experience to make their decision on cardiac-recognition. For the *Feel* trials they would report the feedback as their own if they felt at least one heartbeat in time with any of the blue pulses. In contrast, during the Regulate condition participants would indicate the feedback as their own if they felt that the movements of the feedback bar were responding to them. Lastly, for the *Attend* trials participants were instructed to simply *guess* whose feedback they have seen. Finally, participants reported their confidence in their cardiac recognition decision, by using a slider on a visual analogue scale with the endpoints "Not at all confident" and "Very confident" (Figure 2.A2). Unbeknown to the participants, the positions of this scale corresponded to values ranging between 1 and 50. The task took approximately 1 hour to complete, including a 10 min break half-way through.



Number of Trials per Condition

	Congruent	Incongruent	Total	N=34
Attend	26	26	52	1768
Feel	26	26	52	1768
Regulate	26	26	52	1768

Figure 2. Experimental procedure of the cardiac recognition task. (A1) Timeline of Strategy dependent stimuli presentation and strategy-specific questions. (A2) Questions on cardiac recognition and participants' subjective confidence were presented after every trial for all conditions. (B) Summary table of the number of behavioural trials within each condition and across the whole sample.

5.3.3 Stimuli

Analogue output of inter beat intervals (IBI-s) was obtained online and recorded digitally on a PC into MATLAB (MathWorks, Sherborn, Mass., USA). Within MATLAB, a script was created to provide the visual display to the subject during biofeedback exercises. During the 10-second-long biofeedback phase of each trial, participants received an instantaneous and continuous feedback of cardiac activity across two dimensions simultaneously. In general, as one's heart rate increased the bar became taller, and as the heart rate dropped the bar followed it by becoming shorter. To provide a more direct feedback of beat to beat changes, a short yellow pulse also appeared on the bar 280 ms after the R-wave peak, which coincides with the time window (200–300 ms post R-wave) of peak systolic pressure, which is the period of maximum subjective perception of heartbeats (Brener et al., 1993, Suzuki et al. 2013). This latency was chosen to ensure a sufficiently long, analysable epoch of the HEP that did not coincide with the visual evoked potential induced by the pulse. At approximately 50% of all pulses within each trial, the pulses changed from the default yellow to a different colour following a randomized pattern in their order. The alternative colour corresponded to the Strategy condition in the following way: *Attend* – green, *Feel* – blue and *Regulate* – white.

The feedback bar was set to the middle at the beginning of every trial. To establish the centre of the feedback bar for the first congruent trial, 10 inter-beat-intervals (IBI-s) were averaged prior to the beginning of the task. Afterwards, for the congruent trials, the value representing the middle reference point of the bar was extracted and updated, based on the participants *real* IBI-s from the previous trial (in any condition). In contrast, for the reference point of the Incongruent trials we used IBI-s from the previous Incongruent trial. In this way, we could ensure that the parameters of the biofeedback bar were continuously scaled. To make the feedback more sensitive to the changes in the lower ranges of heart rate (and less sensitive to movement artifacts) we set the minimum of the bar by subtracting a quarter of the baseline while we created the maximum value by adding half the baseline. The required change for every step was also relative to the participants' baseline (or the baseline of the incongruent IBI series).

Most previous studies created asynchronous feedback by changing the frequency of the participant's own estimated heartrate to be either 30% slower or faster (e.g. Suzuki et al., 2013). Unlike these studies, we used a database of incongruent IBI series (N = 72, M_{IC_IBI} = 779.89, SD_{IC_IBI} = 142.03) from a database of people who completed the same task on a different occasion. We decided to use this database instead of participants' own altered cardiac activity from previous trials, based on our pilot data because in the piloting stage of this experiment we found that participants performed consistently below chance when trying to differentiate between congruent and incongruent feedback (without using any strategies) when presented with their own altered heartbeats. In other words, participants were more likely to respond incorrectly than correctly when identifying the source of the feedback. By contrast, participants had a higher chance of being accurate when the incongruent feedback was based on the cardiac data of other individuals rather than their own. A potential explanation for this is that the difference between signal (congruent) and noise (incongruent trials) increased when using other people's cardiac signal, resulting in a paradigm that was accomplishable yet still challenging. The incongruent feedback was tailored for every participant by matching it with the most similar IBI series from our database, based on their average baseline IBI. We intended to keep the level of incongruency between conditions (and across participants) as constant as possible. We addressed this by adjusting, in every trial, for the percentage difference between the incongruent signal and the participant's own baseline. To introduce some noise, half of the incongruent trials were presented 15% slower, while the other half were 15% faster than the original – the order of which was randomly generated.

5.3.4 EEG and ECG Recording

EEG was recorded with Ag-AgCl electrodes from 64 active scalp electrodes mounted on an elastic electrode cap, according to the International 10/20 system, using ActiveTwo system (AD-box) and Actiview software (BioSemi; 512Hz sampling rate; band- pass filter 0.16-100Hz (down 3 dB); 24 bit resolution). Electrodes were referenced to the Common Mode Sense (CMS) and Driven Right Leg (DRL) electrodes and re-referenced to the average reference offline. ECG signal was recorded with a standard 3-lead ECG attached to participants' chest (Powerlab, ADInstrumens, www.adinstruments.com) which we used for sending out triggers for MATLAB in the cardiac recognition task and through MATLAB to the Actiview. The onset of the R-waves coming from the 6 conditions were marked in the EEG recording using 6 different triggers. In addition to four external electrodes recording eye movement artifacts, another external electrode was attached to the participants left sternum to provide a clear ECG trace for cardiac artifact detection. The average number HEP epochs in Congruent conditions were $M_{\text{ATTEND}} = 317.26$, $SD_{\text{ATTEND}} = 50.27$, $M_{\text{FEEL}} = 313.94$, $SD_{\text{FEEL}} = 49.81$, $M_{\text{REGULATE}} = 314.29$, $SD_{\text{REGULATE}} = 50.20$, while in the Incongruent conditions were $M_{\text{ATTEND}} = 310.65$, $SD_{\text{ATTEND}} = 53.94$, $M_{\text{FEEL}} = 309.65$, $SD_{\text{FEEL}} = 53.37$, $M_{\text{REGULATE}} = 306.50$, $SD_{\text{REGULATE}} = 53.10$. Importantly, there were no significant differences in the number of heartbeats between conditions $F(2,198) = 0.02$, $p = .980$.

5.3.5 EEG Data Analysis

Offline EEG pre-processing was performed using the BrainVision Analyzer software (Brain Products, Munich, Germany). The continuous EEG data was filtered with a bandpass filter of 0.1–30 Hz (24 dB/oct) and a 50 Hz notch filter. Then we applied Independent Component Analysis (ICA) on resampled data (250Hz) to remove ocular and cardiac-field artifacts (CFA-s) (Terhaar, Viola, Bär, & Debener, 2012) based on their timing, topographical and physical characteristics. The ICA method is of high efficiency in the removal of CFA-s from the EEG signal (e.g. Park, Correia, Ducorps, and Tallon-Baudry, 2014; Terhaar et al., 2012; Luft & Bhattacharya, 2015). After the ICA, the EEG signal was segmented into 600 ms epochs, starting 150 ms before the R-wave onset using the R-wave onset related triggers sent by MATLAB (i.e. -200 ms to 400 ms epochs around the R-wave). Segments were then baseline corrected using an interval from −150 to −50 ms before the R-wave onset to avoid the inclusion of artifacts related to R-wave rising edge (Canales-Johnson et al., 2015) and late components of visual evoked responses to the pulsing stimulus of the immediately preceding trial. Semiautomatic artifact rejection was combined with visual inspection for all participants. Epochs exceeding a voltage step of 200 µV/200 ms, a maximal allowed difference of 250 µV/200 ms, amplitudes exceeding ±250 µV, and low activity less than 0.5 µV /50 ms were rejected from analyses. There were no significant differences in included epochs between conditions ($p = .980$). These segments then were referenced to the arithmetic average and then grand averaged within every condition.

With a distribution of frontal-to-parietal, heartbeat evoked potential (HEP) has higher amplitudes over the right hemisphere (Dirlich, Vogl, Plaschke, & Strian, 1997; Kern, Aertsen, Schulze-Bonhage, & Ball, 2013; Pollatos & Schandry, 2004; Schulz et al., 2015). The polarity of the HEP varies with the task, region and latency analyzed (Canales-Johnson et al., 2015; Couto et al., 2013; Gray et al., 2007). Following Sel, Azevedo, and Tsakiris (2017), our analysis had 6 regions of interests (for ROI-s see Figure 6). To estimate the group level effects of Strategy and feedback Congruency on mean HEP voltages, Monte-Carlo random cluster-permutation method was implemented in FieldTrip (Maris & Oostenveld, 2007). When testing the relationship between interoceptive Strategy and specific HEP amplitudes and behavioral measures, we used the difference score of Congruency amplitudes (i.e. *Feel* (C-IC) – *Attend* (C-IC) and *Regulate* (C-IC) – *Attend* (C-IC)). To produce these difference values, subtraction was performed on the time-locked grand averages, keeping individuals distinct. When comparing different conditions at a neural level we used the absolute measure of HEP amplitudes.

The Monte-Carlo cluster-based permutation test corrects for multiple comparisons in space and time, which is cardinal issue for a multidimensional data such as an EEG trace. Using this method, first all samples that show a significant ($p < .05$) relationship with the independent variable were identified and clustered following spatiotemporal adjacencies.

Following this, cluster-level statistics were produced based on the sum of all the test statistic values within each cluster. Then, through a high number of random shuffling and resampling (10000 repetitions in our case), Monte-Carlo permutation calculated the probability of achieving the cluster-level statistic by chance only. Spatiotemporal clusters that resulted in a Monte - Carlo corrected p-value of less than the critical alpha level of .025 (necessary when running two tailed tests expecting either positive/negative clusters) were interpreted as 'significant'.

5.3.4 Data analysis

5.3.4.1 Cardiac recognition

One of the most commonly studied features of interoception interoceptive sensitivity reflects the ability to precisely and accurately monitor changes of internal body state (Khalsa et al., 2017). We quantified sensitivity in our cardiac recognition task by a signal detection metric: d' ("d-prime", see Macmillan & Creelman, 2004), which is commonly used in heartbeat detection tasks (e.g. Khalsa, Rudrauf and Tranel, 2009). Signal detection theory provides a general framework to describe and study decisions that are made in uncertain or ambiguous situations. This framework can describe the link between a physical stimulus and its subjective, psychological effect – here the visual biofeedback signal and the recognition of one's own cardiac activity represented by such signal. The benefit of signal detection measures is that they can differentiate response bias from one's true ability to discriminate between signal (i.e. congruent feedback) and noise (i.e. incongruent feedback) – which simpler measures of accuracy (e.g. proportion correct) cannot. Our study calculated d' by using the difference between participants' normalized hit rate (proportion of correct trials when the biofeedback represented one's own heart) and normalized false alarm rate (proportion of correct responses when the biofeedback represented someone else's heart):

$d' = z(hit\ rate) - z(false\ alarm\ rate)$

To summarize, d' represents the distance between the signal (hit rate) and noise (false alarm rate), where larger distance (i.e. larger d') represents better sensitivity. A potential limitation to the application of signal detection to cardiac signals is that the difference between signal (i.e. congruent cardiac feedback) and noise (i.e. incongruent cardiac feedback) might be small. For this reason, the probability of correctly differentiating the signal from noise could be at chance and lower than chance. The present study aimed to tackle this issue by using incongruent IBI series from a database of previously recorded from other people instead of altering the frequency of the participant's own heartrate (e.g. Suzuki et al., 2013).

5.3.4.2 Metacognition

The correspondence between objective interoceptive abilities and the self-reported *evaluation* of this performance (e.g. confidence) – also known as interoceptive insight – has attracted interest over recent years, as it provides the potential to tap into metacognitive aspects of interoception (Khalsa et al., 2017). Previous studies have found an association between IAcc and reports of self-evaluation (Khalsa et al., 2008) and confidence measured by type 2 AUROC for those individuals who had high IAcc (for issues with this method see Fleming & Lau, 2014; Garfinkel, Seth, Barrett, Suzuki, & Critchley, 2015). Our study employed a different, focused approach to explore the link between confidence and accuracy. In our analysis, we used Confidence Accuracy (CA) calibration, which measures the relationship between categorical levels of confidence and the binary measure of accuracy (Mickes, 2015). By simply regressing accuracy on confidence and plotting their relationship one can already gain interesting insights on metacognition. Moreover, it is possible to quantify such confidence – accuracy relationship by statistics commonly used in eyewitness research (for more see Brewer & Wells, 2006). Here we use the normalized resolution index (NRI) which provides a quantitative index of the ability to use confidence categories to effectively distinguish when an event occurs (i.e. feedback of own heart) and when it does not (i.e. feedback of someone else's heart) (Petrusic & Baranski, 1997). The NRI is calculated as:

$$\left[\frac{1}{n}\sum_{j=1}^{J} n_j(a_j - a)^2\right] \Big/ a(1 - a)$$

where a_j denotes the proportion of correct responses at confidence level j and a denotes overall mean accuracy. NRI ranges from 0 (no discrimination) to 1 (perfect discrimination). Given that NRI can be interpreted as eta-square (Petrusic & Baranski, 1997) – which is directly related to Cohen's f, cutoffs for NRI values can also be created (small: .010, medium: .059, large: .138) (Brewer & Wells, 2006). Note that CAC requires a high number of trials in general but that the separation of confidence judgments into more/fewer levels will also affect the reliability of the analysis (i.e. higher number of confidence bins will need more trials in order to be reliable).

5.4 Results

5.4.1 Behavioral analysis: d'

Our main behavioral outcome variable was the strategy-related cardiac recognition measured by d'. As d' already inherently captures the congruency to incongruency relation, our experiment only had one predictor on this level of the analysis: Strategy (i.e. engagement with the biofeedback signal: 1 = *Attend*; 2 = *Feel*, 3 = *Regulate*). We chose to model our d' data with a linear model as the d' values followed a Gaussian distribution (Shapiro-Wilks test *p* = .190). Descriptive statistics and confidence intervals are noted in text. We excluded from analysis those Congruent trials (1.3 % of our data) where technical difficulties lead to undetected heartbeats and disruption of congruent feedback.

We used R (Version 3.5.1; R Core Team, 2018) for our analyses. Specifically, we selected the optimal model by using the *buildmer* package (Version 1.0; Voeten, 2019) which can perform backward stepwise elimination, based on the change in the set criterion (AIC in our case). For linear mixed effects modeling we used the package *lme4* (Version 1.1.17; Bates, Mächler, Bolker, & Walker, 2015). Relevant test-statistic were gathered by using *sjPlot* (Version 2.5.0; Lüdecke, 2018b) and *sjmisc* (Version 2.7.4; Lüdecke, 2018a). Mixed effects modelling is particularly useful in within-subject designs where each subject has several measurements resulting in correlated errors for those measurements (Baayen, Davidson, & Bates, 2008). The solution to this problem is to let each subject have their own personal intercept (and/or slope), randomly deviating from the mean intercept, as the errors around the personal regression lines will be uncorrelated using this approach. Reported p-values were computed via Wald-statistics approximation (treating t as Wald z). We tested for the effects of task performance (unrelated to cardiac recognition), average heart rate, the average change in heart rate from baseline to task and baseline HF-HRV. Task performance measures were quantified as the following:

$$Task\ Performance\ _{(ATTEND)} = 1 - |\ (target\ pulses - reported\ pulses)\ |\ /\ target\ pulses$$

$$Task\ Performance\ _{(FEEL\ CORRECT)} = value\ associated\ with\ the\ position\ on\ scale * 2\ /\ 100$$

$$Task\ Performance\ _{(FEEL\ INCORRECT)} = 1 - (value\ associated\ with\ the\ position\ on\ scale * 2\ /\ 100)$$

$$Task\ Performance\ _{(REGULATE)} = mean\ IBI\ _{(TRIAL)} - mean\ IBI\ _{(PREV.TRIAL)}\ /\ mean\ IBI\ _{(PREV.TRIAL)}$$

Centered covariates were included in the final model only if they significantly improved the model fit. We defined the maximal model as:

$$d \sim Strategy\ +\ Task\ performance\ +\ Baseline\ HF - HRV_{BASELINE}\ +\ HR_{CHANGE}\ +\ HR\ +\ (1|ID)$$

Based on the AIC values, the final model which provided the best fit with our data was the following:

$d \sim Strategy\ +\ (1|ID)$

The expression outside the parentheses indicates fixed effects, while the expression inside reflects the random effects defined in the model (i.e. the intercept over participants).

Results revealed a significant increase of d' in comparison to the *Attend* condition (M_{ATTEND} = 0.27, SD_{ATTEND} = 0.45) in both the *Feel* (M_{FEEL} = 0.49, SD_{FEEL} = 0.58) β = 0.22, [CI] = 0.04 – 0.40, *p* = .017 and the *Regulate* conditions ($M_{REGULATE}$ = 0.58, $SD_{REGULATE}$ = 0.68) β = 0.31, [CI] = 0.13 – 0.49, *p* = .001, $R^2_{MARGINAL}$ = 0.05, $R^2_{CONDITIONAL}$ = 0.59 without being significantly different from each other (*p* = .350) (Figure 3). Results are depicted by raincloud plots (Allen, Poggiali, Whitaker, Marshall, & Kievit, 2018). These results remain significant after Bonferroni correction for three comparisons. A negative score of d' indicates a performance that is worse than chance (i.e. participants cannot discriminate congruent feedback from incongruent), which might hamper the interpretation of results at those participants. For this reason, we ran the same analysis excluding participants with negative d' in any of the three conditions (N=20) but found the same - even stronger - results. In this subsample, both the *Feel* β = 0.29, [CI] = 0.06 – 0.53, *p* = .014 and *Regulate* β = 0.39, [CI] = 0.16 – 0.62, *p* = .001, $R^2_{MARGINAL}$ = 0.12, $R^2_{CONDITIONAL}$ = 0.38 conditions were associated with higher d' than the *Attend* condition - without differing from each other (*p* = .416). It is important to note that heart rate remained the same across conditions (even in the Regulate condition), meaning that the observed effects were driven by differences in the type of engagement with the biofeedback signal rather than changes in physiological states.

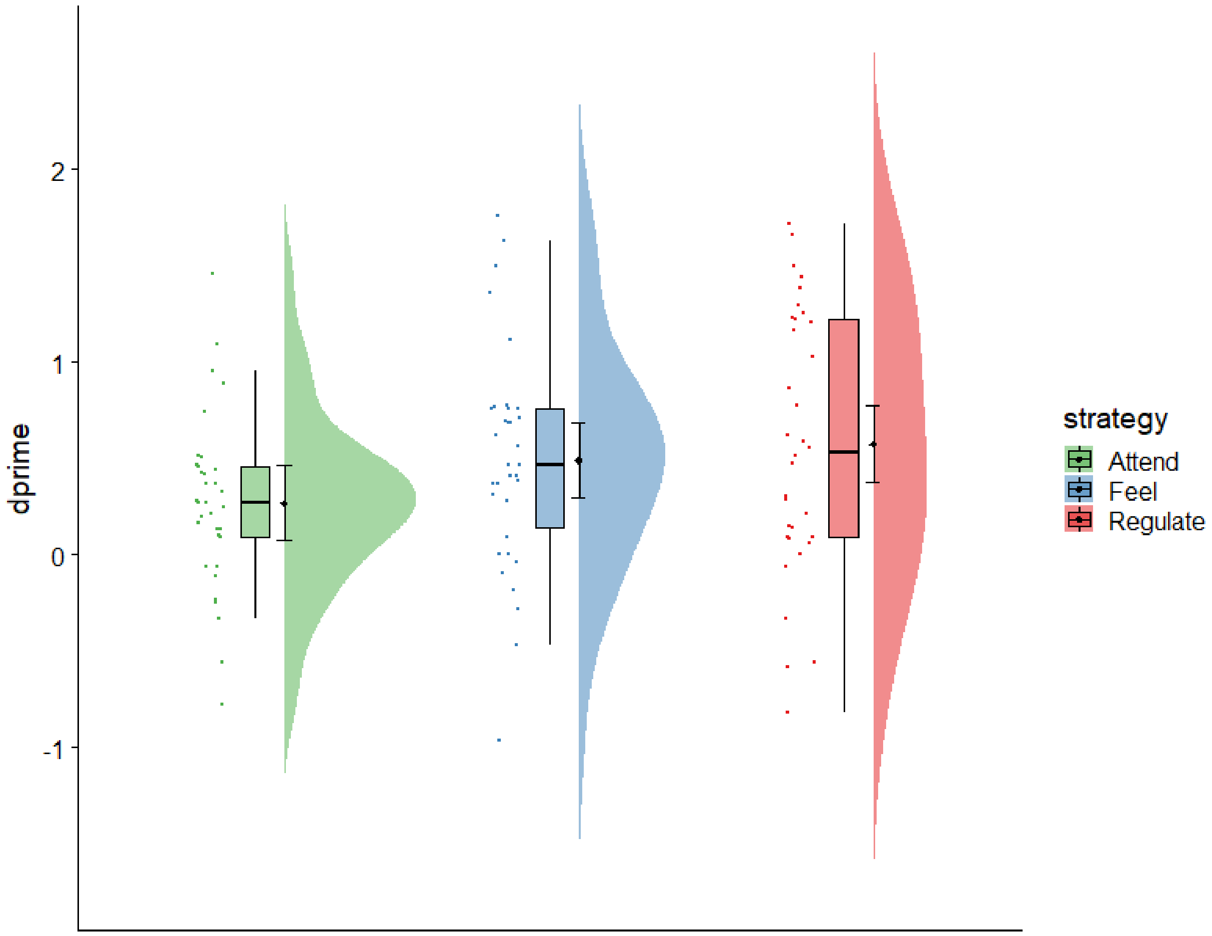


Figure 3. Strategy related effects on participants' cardiac recognition measured by d'. The raincloud plots of d' provide information showing data distribution, the central tendency (by boxplots) and the jittered presentation of our raw data. Error bars indicate 95% confidence intervals around the estimates of the linear mixed effects model with a random intercept.

Table 1.

Summary table of relevant signal detection measures int the function of strategy

Strategy	Hit	False alarm	Miss	Correct rejection	Response bias
Attend	14.79 (4.63)	12.35 (5.04)	10.56 (4.41)	13.65 (5.04)	-0.03 (0.12)
Feel	16.41 (4.37)	11.85 (3.68)	9.15 (4.5)	14.15 (3.68)	-0.08 (0.17)
Regulate	16.41 (5.44)	11.53 (4.45)	8.68 (5.23)	14.44 (4.52)	-0.15 (0.27)

Note. Mean (SD)

5.4.2 Response bias

Non-parametric estimate of *response bias* (B”D) was also measured, where the 0 indicates no bias, positive numbers represent conservative bias (i.e. a tendency to answer 'no'), negative numbers represent liberal bias (i.e. a tendency to answer 'yes'). Participants were liberal across all *Strategy* conditions (see Table 1), but after fitting a linear regression with a random intercept (the same model that we used at d') on B”D with the function of *Strategy*, the model revealed that participants in the *Regulate* were significantly more liberal than in the *Attend* condition β = -0.13, [CI] = -0.21 – -0.04, p = .005 (survives Bonferroni correction for three comparisons), but there was no difference between Attend and Feel p = .241, and Feel and Regulate conditions p = .107 R^2_{MARGINAL} = 0.06, $R^2_{\text{CONDITIONAL}}$ = 0.17. These results also confirm the importance to use bias-free measures of accuracy.

5.4.3 Confidence accuracy calibration

To gather further understanding of the link between self-reported confidence and cardiac recognition accuracy, we ran an exploratory Confidence Accuracy (CA) calibration analysis for which we used the beta R package *legalPsych* (Version 3; Van Boeijen & Saraiva, 2018). The main part of this analysis is simply plotting the proportion correct of cardiac recognition for each level of confidence – classically ranging between 0% - 100% and separated into bins of 10% increases or collapsed within wider ranges (Figure 4A). First, we ran a generalized linear mixed model, which is an extension of linear mixed models allowing response variables to have different distributions. Given the binary nature of our trial level outcome variable (i.e. accuracy with levels 0 = Incorrect, 1 = Correct), we fitted a random intercept model with a binomial distribution and a logit link. We found a positive – but in terms of effect size rather small – link between Accuracy reported levels of Confidence, Odds Ratio (OR) = 1.08, [CI] = 1.05 – 1.10, p = < .001, R^2_{MARGINAL} = 0.01, $R^2_{\text{CONDITIONAL}}$ = 0.05. To reduce noise, we collapsed Confidence into two categories (i.e. Low: 0% – 50% and High: 60% – 100%) and plotted proportion correct against confidence for every Strategy (Figure 4B). To quantify the differences in confidence-accuracy relationship across strategies, NRI was calculated for every individual. When contrasting different levels of Strategy in a linear mixed effects analysis, we found a significant increase in NRI for the *Regulate* ($M_{REGULATE}$ = 0.13, $SD_{REGULATE}$ = 0.18) condition when compared to the *Attend* (M_{ATTEND} = 0.05, SD_{ATTEND} = 0.07) β = 0.08, [CI] = 0.02 – 0.14, *p* = .006 and *Feel* (M_{FEEL} = 0.07, SD_{FEEL} = 0.12) conditions β = -0.06, [CI] = -0.11 – 0.001, *p* = .046 (n.s. after Bonferroni correction), $R^2_{MARGINAL}$ = 0.06, $R^2_{CONDITIONAL}$ = 0.21, while there was no difference between the *Attend* and *Feel* conditions (*p* = .459). For descriptive statistics see Table 1.

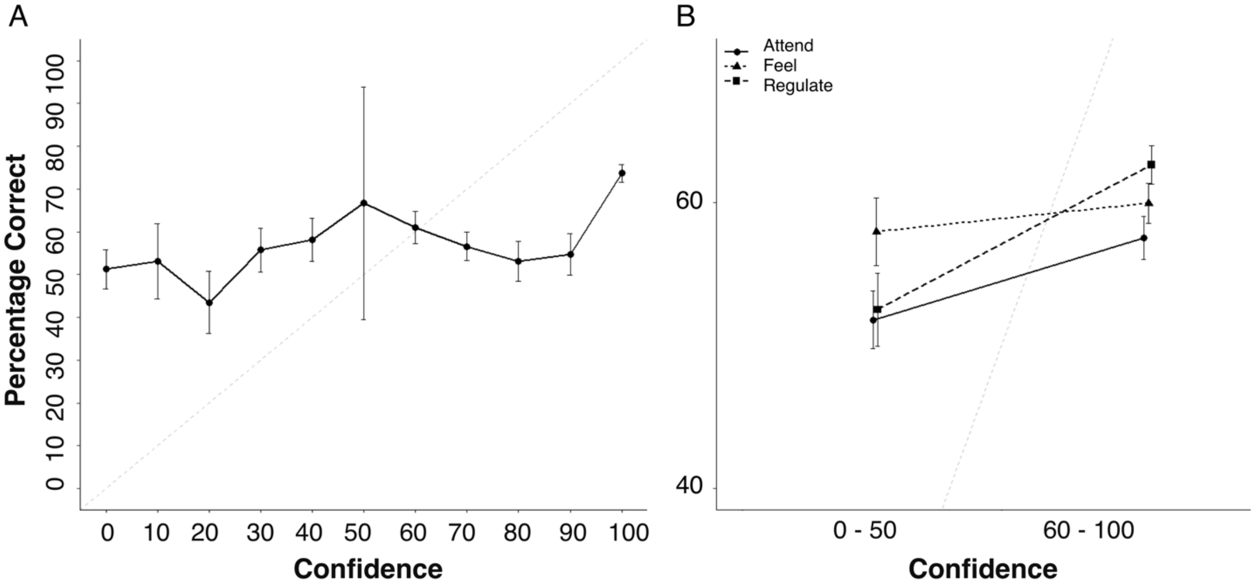


Figure 4. Results of confidence accuracy calibration analysis depicting a (A) positive linear relationship between the binned values of confidence and accuracy in general and (B) within certain Strategies. The diagonal line represents perfect calibration between accuracy. Error bars indicate 95% confidence intervals.

To summarize the behavioural results, participants' performance on the cardiac recognition task did not have a ceiling or floor effect. Specifically, cardiac recognition measured by d' revealed a significant difference between interoceptive strategies (both *Feel* and *Regulate*) and exteroception (*Attend*) but not between the different types of interoceptive strategies (i.e. *Feel* vs *Regulate*). Accuracy measured by proportion correct positively mapped onto self-reported confidence across all conditions. In comparison to the other two strategies (both *Attend* and *Feel*), participants' metacognition (i.e. accuracy and confidence association) was higher during *Regulate* trials where the instruction was to use control-based inference for cardiac recognition.

Table 2

Descriptive statistics of correct and incorrect response for low and high levels of confidence

Strategy	Levels %	Mean Confidence	Incorrect	Correct	Total	Proportion Correct
Regulate	0-50	26.14	179	198	377	0.53
Regulate	60-100	79.02	475	797	1272	0.63
Attend	0-50	25.26	292	314	606	0.52
Attend	60-100	77.51	447	606	1053	0.58
Feel	0-50	25.88	182	251	433	0.58
Feel	60-100	78.44	490	734	1224	0.60

Note: The total number of trials might differ slightly across conditions as some trials were excluded as the recording equipment occasionally missed heartbeats (approx. 1% of trials).

5.4.4 Cluster-based permutation analysis on HEP amplitudes

Prior to the main analysis on HEP-s we inspected the distribution of EEG amplitudes within the time window of interest (i.e. 200 - 300 ms after the R-wave onset) and looked for outliers. We used the multivariate model approach for outlier identification as declaring an observation as an outlier based on a just one feature could lead to misleading inferences. Four influential participants were identified based on the amount of impact their data points had on the predicted outcome - represented by the Cook's distance (Cook, 1977). We decided to remove these participants as they had more than one datapoints where Cook's distance was four times greater than the mean, leaving us with a sample of N = 30.

We used MATLAB (Version R2019a; MathWorks) with the toolbox *FieldTrip* (Version fieldtrip-lite-20190403; Maris and Oostenveld, 2007) for our analyses applying cluster-based permutation and the external functions *cbrewer* and *boundedline* for plotting results. Given that our main interest on this level of the analysis was the potential interaction effect between Strategy (1 = *Attend*, 2 = *Feel*, 3 = *Regulate*) and Biofeedback Congruency (1 = *Congruent*, 2 = *Incongruent*), first we needed to determine whether there are *any* differences between conditions on the amplitudes of HEP. For this reason, we specified a dependent samples F-statistic to be calculated for each sample, in each random reshuffling of the data. This analysis revealed a significant modulation of the HEP amplitude, as indicated by a significant positive cluster (F_{SUM} = 400.48, *p* = .024) between 232-280 ms within the right-frontal ROI (specifically electrodes AF4, F4). To investigate the simple effects of Strategy and Biofeedback Congruency in this interaction, we ran nine pair-wise comparisons (now specified with dependent samples T-statistic) at the right-frontal ROI. In the latency range from 200 to 300 ms post R-peak, the cluster-based permutation test revealed a significant positive difference between the *Attend* and *Feel* conditions during Incongruent biofeedback (T_{SUM} = 141.13, *p* = .003). In this latency range, the difference was globally pronounced over all sensors of this ROI within the whole pre-set latency range. Similarly, the *Regulate* condition was significantly different from the *Attend* strategy within the Incongruent feedback condition in a positive direction (T_{SUM} = 69.24, *p* = .002), which effect was most pronounced at the latency between 204-268 ms at the AF4, F4 electrodes. All statistic reported survived Bonferroni correction for 9 comparisons (Figure 5 A and C).

In addition, to ensure that the observed HEP differences between conditions cannot be explained by differences in the ECG signal, we analysed the ECG trace following the same protocol as in the HEP analysis reported above. The results of the cluster-based permutation test on the ECG did not reveal any significant cluster of significant interactions at $p < .050$ (Figure 5 B).

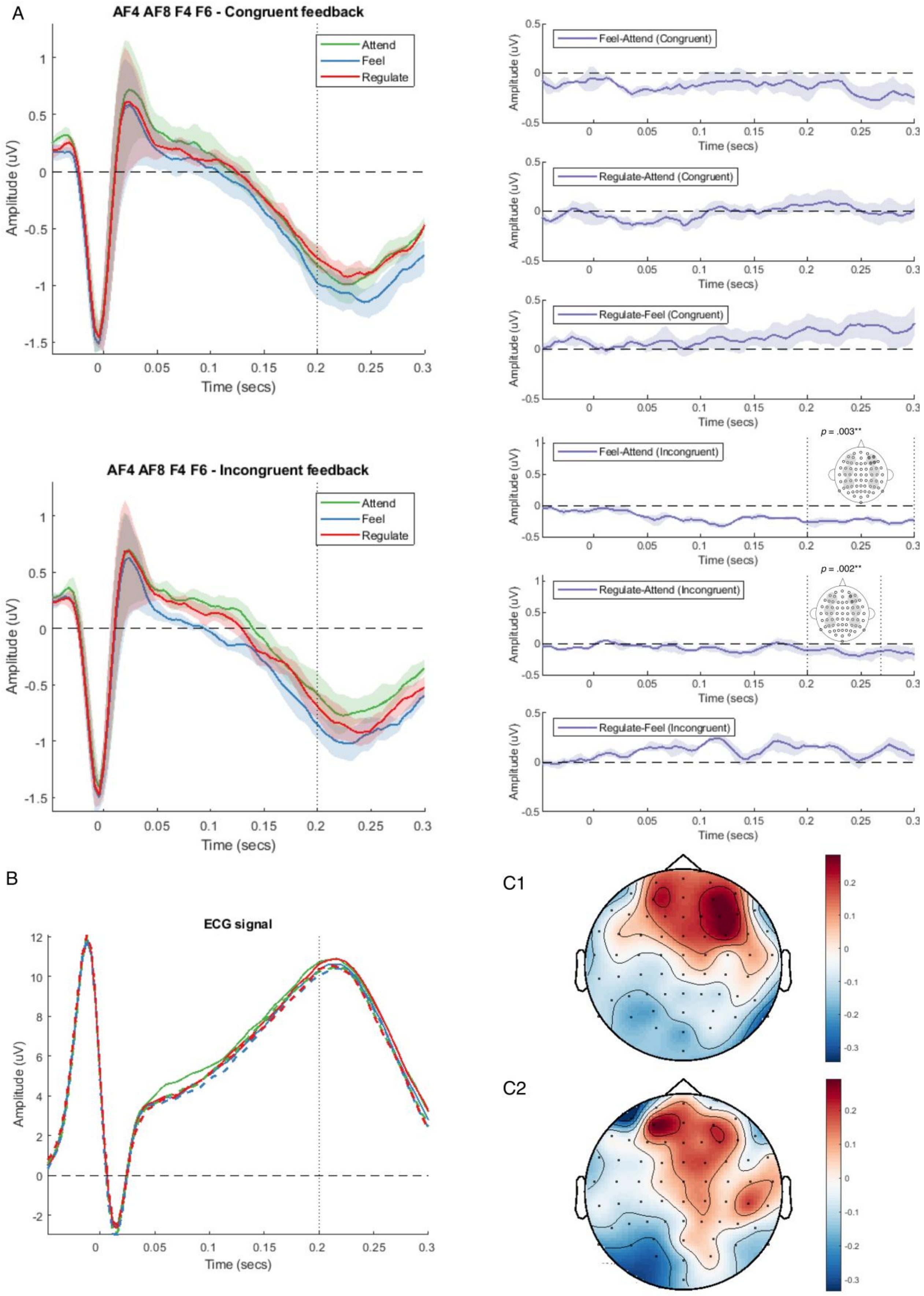

A
AF4 AF8 F4 F6 - Congruent feedback
Attend
Feel
Regulate
Amplitude (uV)
Time (secs)
AF4 AF8 F4 F6 - Incongruent feedback
Attend
Feel
Regulate
Amplitude (uV)
Time (secs)
Feel-Attend (Congruent)
Regulate-Attend (Congruent)
Regulate-Feel (Congruent)
Feel-Attend (Incongruent)
p = .003**
Regulate-Attend (Incongruent)
p = .002**
Regulate-Feel (Incongruent)
Amplitude (uV)
Time (secs)
B
ECG signal
Amplitude (uV)
Time (secs)
C1
C2
0.2
0.1
0
-0.1
-0.2
-0.3

Figure 5. (A) Presentation of Heartbeat Evoked Potentials (HEP-s) and their difference waves depicting the effect of condition over the right frontal ROI within the a-priori set latency of 200-300 ms during the presentation of biofeedback stimuli (N = 30, Monte-Carlo cluster analysis, $F_{SUM} = 400.48$, $p = .024$). For the two significant pairwise comparisons we also denoted the electrodes and latencies where the effect was the most pronounced. (B) Average ECG signal across all conditions (dashed lines represent the Incongruent while full line indicate the Congruent biofeedback condition). The shaded areas around mean amplitudes indicate 95% confidence intervals. (C) Topographical representation of positive right frontal clusters during Incongruent feedback when comparing the *Attend* condition to (C1) *Feel* and (C2) *Regulate* conditions. For the topographical plots, amplitudes were averaged within the time window where the effect on the cluster was most the pronounced. Colour bars show Monte-Carlo cluster statistic (t).

To see the psychological relevance of neural responses we also tested a potential link between Strategy-specific cardiac recognition and strategy-specific modulation of HEP amplitudes (Figure 6). To match HEP with d' - which inherently captures the congruency to incongruency relation - first we calculated *Congruency difference* amplitudes for every Strategy condition, by subtracting the Incongruent condition from the Congruent condition. Then to fully separate Strategy related effects from attentional processes, we subtracted the Congruency difference amplitudes in the *Attend* condition from the *Feel* and *Regulate* conditions (Figure 6B). To mirror this on a behavioural level, we subtracted d' scores in the *Attend* condition respectively from the other two interoceptive Strategy conditions (i.e. *Feel* and *Regulate*). After this, we performed a regression analysis between the two types of interoceptive d'-s and the HEP differences, using the same cluster-based permutation technique as before. Based on the results of our previous interaction analysis we selected the a-priori latency where HEP differences were the strongest (i.e. 232-280 ms), with the right-frontal area as our ROI. The analysis revealed a significant positive relationship between strategy specific HEP difference and d' difference in the *Feel* condition ($T_{SUM} = 24.64$, $p = .019$), but not in the Regulate condition. This effect was the most pronounced over electrodes AF8, F4, F6 within the time window of 272-284 ms after the R peak and survives Bonferroni correction for 2 comparisons.

To summarize the neural findings, the use of interoceptive strategies compared to exteroception for cardiac recognition was associated with differences in HEP amplitude over the right-frontal area within the latency of 200-300 ms. The cortical processing of cardiac signals during incongruent feedback was more pronounced during interoception – regardless of the strategy they were instructed to use to interact with their cardiac signals. We also found that strategy-specific difference HEP reflected individuals' strategy-specific cardiac recognition performance when participants focused on individual heartbeats (*Feel*) but not during regulation (*Regulate*). Specifically, larger *Feel* strategy related HEP difference amplitudes were positively linked to the ability to differentiate feedback representing one's own heart from someone else's.

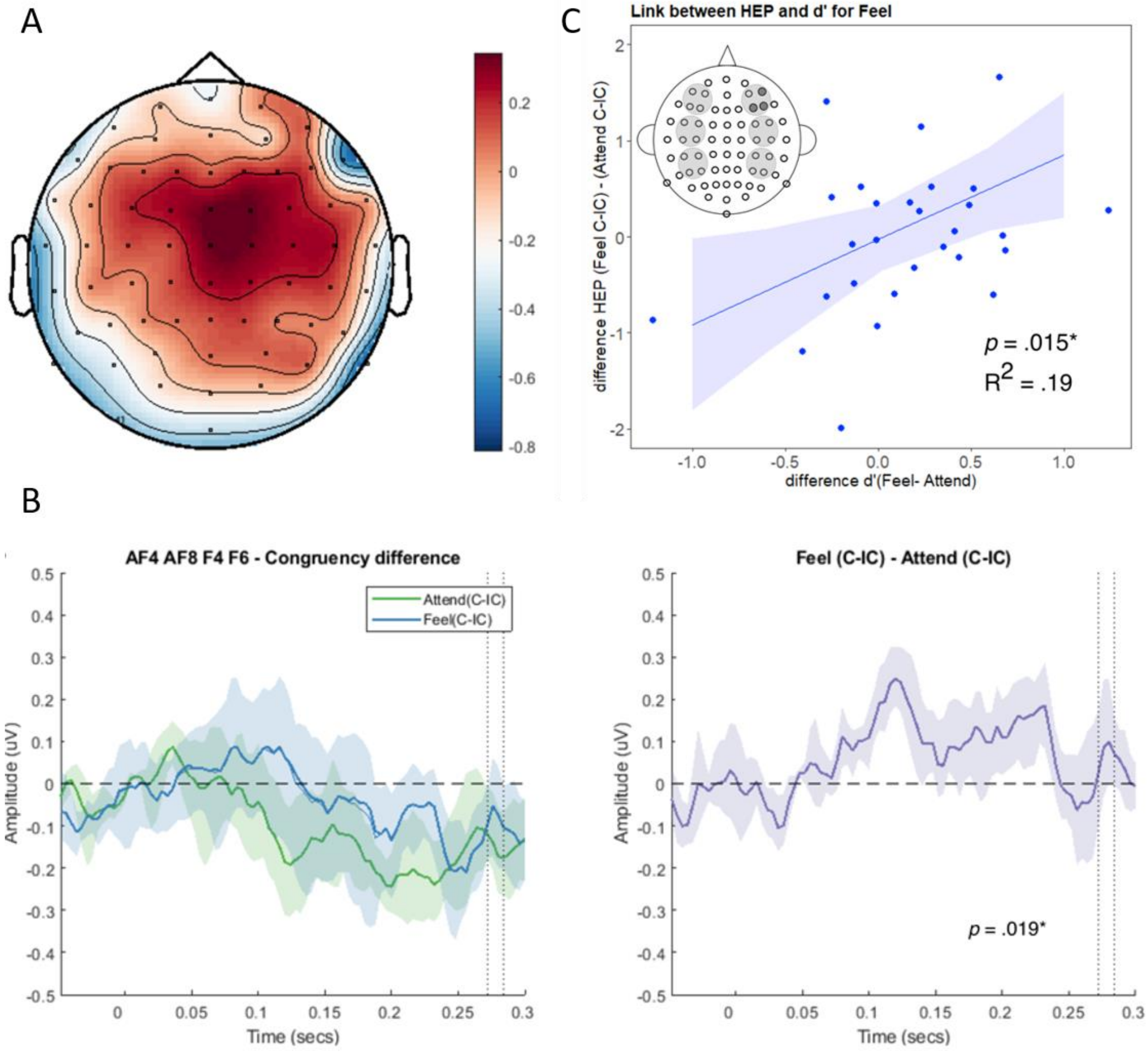


Figure 6. The amplitude of strategy-specific Heartbeat Evoked Potential (HEP) difference is related to interoceptive strategy-specific sensitivity. (A) Topographical plots depict the amplitude differences that were used in the regression (*not* the spatial effects associated with the regression itself). (B) Congruency difference amplitudes and their difference wave across certain strategies. Shaded area represents 95% confidence interval of fitted regression line. (C) For illustration purposes parametric linear regression lines were plotted using subject-wise average signal over the electrodes (dark shaded circles on layout map) and within the latency (dashed lines on amplitude plots) where the relationship was the strongest (identified by the Monte-Carlo cluster-based permutation).

5.5 Discussion

The function of interoception is to maintain psycho-physiological stability (Khalsa et al., 2017) which requires the ability to infer the state of our body and to predict its future (Pezzulo et al., 2015). However, the standard measures of interoception are distant from this functional definition as they simply test the perception of single heartbeats – leaving the interpretation of participants' performance in these tasks limited to a lower, sensory level (Schandry, 1981; e.g. Whitehead et al., 1977). With the intention to get closer to the functional role of interoception and to contrasts different ways of engaging with changes in one's physiological state, we measured participants' cardiac recognition, metacognitive abilities and associated neural responses across three different strategies. These strategies all involved a combination of interoceptive and exteroceptive elements, but they emphasized different features of the biofeedback and required different levels of engagement in an increasing order of (i) *neutral-exteroceptive (Attend),* (ii) *passive-interoceptive (Feel)* and (iii) *active-interoceptive (Regulate).*

We observed an increase in cardiac recognition for both the *Feel* and *Regulate* strategies when compared *Attend*, but there were no differences between the *Feel* and *Regulate* strategies. A similar pattern of results was revealed at the neural level, as we found that the cortical processing of cardiac signals (measured by the modulation of HEP amplitudes) during incongruent feedback was more pronounced at both interoceptive strategies (*Feel* or *Regulate*) when compared to exteroceptive *Attend.* According to these findings, our hypothesis about the gradual increase of cardiac recognition across strategies (i.e. *Attend* < *Feel* < *Regulate*) was *partly* confirmed. Insofar both *Regulate* and *Feel* resulted in significantly more accurate cardiac recognition compared to *Attend,* whereas no significant differences were observed between *Regulation* and *Feel* (i.e. *Attend* < *Feel, Attend* < *Regulate, Feel* = *Regulate*). This pattern suggests that there are alternative ways of engaging with interoceptive signals and that these may be equally beneficial for inferences about the self in relation to the external world.

When applying the PC framework (e.g. Allen & Friston, 2018) to our results, we suggest that the passive (*Feel*) and the active (*Regulate*) interoceptive strategies both increased sensitivity via the optimization of precision-weighting. Specifically, attention directed to the interoceptive channel (here to one’s own heart) would increase the relative precision (i.e. the impact of associated PE-s) arising from that channel (Feldman & Friston, 2010), leading to more accurate inferences about the sources of the biofeedback signal. The brain evaluates the relative precision of PE-s on multiple hierarchical levels (Brown, Adams, Parees, Edwards, & Friston, 2013; Hohwy, 2012). While both the *Feel* and *Regulate* interoceptive strategies appeared to increase the precision of the interoceptive channel, it is possible that they exert their influence at different levels of the hierarchy. The experience of single heartbeats (*Feel*) is likely to affect precision on a lower level, as it is a more strictly sensory strategy that operates on a narrower temporal scale. In contrast, the control-oriented, active strategy (*Regulate*) may increase precision higher up the hierarchy, as it requires a more abstract representation of internal states over a wider temporal window than a single heartbeat. The finding that healthy people do not perceive their heartbeats *consciously* most of the time (Ádám, 1998) supports this argument. In comparison, people are more likely to become aware of perturbations in their physiological states as the optimal behavior differs between balanced and imbalanced states. For instance, a physiological state characterized by a vagal withdrawal (i.e. imbalanced state) supports mobilization responses (i.e. fight and flight), while increased vagal control (i.e. balanced state) is associated with the appearance of spontaneous social engagement behaviours (Porges, 2007). Considering the above arguments and noting that PC unifies action, interoception and homeostatic/allostatic control, we suggest that while both *Feel* and *Regulate* strategies may increase the precision of the interoceptive channel, the *Regulate* strategy relates to a more *functional* aspect of interoception. Specifically, it can provide a more direct access to the estimates of bodily states –an essential information for maintaining homeostatic/allostatic control (Stephan et al., 2016b).

This argument is further supported by our findings on metacognition (i.e. the association between accuracy and confidence). Interestingly, across conditions we found the highest interoceptive metacognition/insight (Khalsa et al., 2017) during the *Regulate* strategy. Stephan and colleagues (2016b) postulate that the performance of the interoceptive circuit is monitored by a higher metacognitive layer, potentially in the anterior prefrontal cortex. This metacognitive layer encodes and updates beliefs about the brain’s capacity to regulate bodily states, with the resulting representation of one’s own self-efficacy. Considering that *Regulate* proved to be distinctly different from *Attend* and *Feel* strategies in terms of metacognition, the results seem to support its higher relevance for allostatic control. In addition, strategy-specific cardiac recognition sensitivity (d') was linked to the modulation of strategy-specific HEP differences exclusively in the *Feel* condition and not when using the *Regulate* strategy. The observed dissociation between *Feel* and *Regulate* strategies is understandable, as only in the *Feel* condition were participants instructed to use single heartbeat-based experience for cardiac recognition. Remaining careful when interpreting a null-result, we suggest that while both the *Feel* and *Regulate* strategies can facilitate sensitivity on a behavioural level, control-based inference (Regulate) may rely on a different process than the cortical processing of single heartbeats. To test this suggestion, ideally, one would need to identify a cortical response that maps onto the performance in cardiac recognition under the *Regulate* condition. P300 is a promising candidate to track such link to cardiac recognition, given it is thought to reflect higher-order perceptual processing of *motivationally* relevant input (e.g. Cuthbert, Schupp, Bradley, Birbaumer, & Lang, 2000; Schupp et al., 2004). Given that the highest level of metacognition was observed when using the *Regulate* strategy for cardiac recognition, there could be a link between the motivationally relevant processing of a stimulus (here heartbeat feedback) and cardiac recognition performance in the *Regulate* condition. The presence of such correspondence would further support the argument about the *Regulate* strategy capturing a more *functional* aspect of interoception than the perception of single heartbeats (*Feel*).

While the core idea behind our strategy manipulation was to influence the engagement with the cardiac signal rather than actual physiological states, it is important to address the fact that participants failed to decrease their heartrate in the *Regulate* condition. It is possible that longer periods are needed for self-induced heartrate regulation to take an effect, but it also might be that voluntary regulation of heartrate simply cannot be achieved in this form. Supporting the latter, a well-powered study of N = 180 by White, Holmes and Bennett (1977) found that participants' regulation attempts were not more efficient in the reduction of heartrate than a condition where participants simply attended to the biofeedback. As we had a novel paradigm, we looked for an unambiguous way to ensure that control-oriented engagement was achieved in the *Regulate* condition. Asking participants to regulate communicates this aim more clearly than the instruction to just attend to the changes in the biofeedback bar. However, it is possible that simply asking participants to focus on the changes and try to match their physiological state to the changes of the biofeedback would lead to similar effects as our instruction to *Regulate*.

Our findings have important implications for future research. First, we need to critically evaluate the underlying assumptions that certain tasks and measures make about interoception. To achieve this, we must gain a better insight about the way people engage with their internal states in real life. In other words, it is important that we study interoception during the modelling of realistic contexts such as social interactions and perturbations, where interoception has true meaning and significance for the individual. This includes but is not limited to modelling real-life stressful scenarios (e.g. job interview), health-related behaviours (e.g. attending one's own body with the aim to decide if they are feeling ill) and social interactions that require the understanding and communication of one's subjective experience with others. In general, future research would greatly benefit from the application of a more *functional* approach to interoception. This necessitates studying the ability to monitor and control internal bodily states and their changes in individuals who are *embedded* in the social and physical world surrounding them.

To conclude, we adopted a novel approach on cardiac recognition by exploring a more functional aspect of interoceptive engagement through inferences about the self in relation to the external world. Across behavioural, neural and metacognitive domains we found evidence that an active control-oriented strategy increases cardiac recognition sensitivity to the same extent as the focus on individual heartbeats. Based on our results showing that metacognition was highest when using a control-oriented approach for cardiac recognition, we suggest that both the passive and active strategies increase precision within the interoceptive channel, but they exert their influence on different levels of the hierarchy.

Chapter 6: General Discussion

6.1 Overview of Main Rationales and Theoretical Background

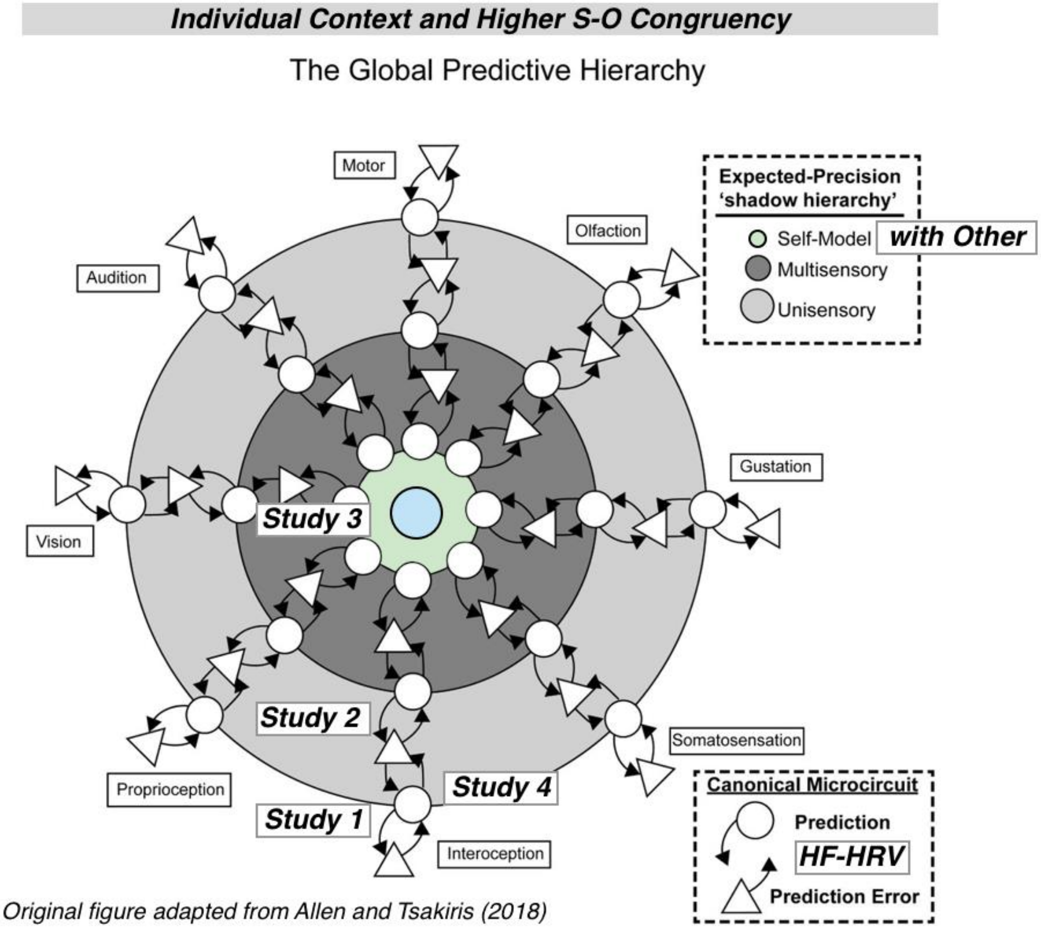
This thesis presented a series of empirical investigations on the role of autonomic regulation in balancing the psychophysiological stability and malleability of one's own self in relation to the external world and others. Fittingly with the framework of current predictive coding model of the self (Allen & Tsakiris, 2018; Seth & Tsakiris, 2018) on the physiological level this thesis considers self-stability in terms of autonomic regulation that ensures homeostasis for the organism. On the other hand, this the defines the psychological level of self-stability by the integration of autonomic and external signals. As autonomic signals can be considered as inherently self- related, when they get integrated with exteroceptive information they will provide a basic continuous experience of selfhood and affect - available for conscious experience (Allen & Tsakiris, 2018). Recent predictive coding theories differ in terms of the specific function they assign to autonomic regulation. Specifically, autonomic responses were proposed to function both as the estimates of self-stability (Allen & Tsakiris, 2018), but also as "interoactions" (Seth & Tsakiris, 2018). Autonomic regulation in the present studies is more likely to represent interoactions that minimise surprise across different hierarchies (Seth & Tsakiris, 2018), given that they arise in *response* to incongruencies and changes in the social context. While the two interpretations do not exclude one another, to investigate autonomic responses as the estimate of surprise would require the manipulation of autonomic responses themselves – which was not the scope of this thesis. Nonetheless, Studies 1, 2 and 3 found that autonomic regulation indeed responds to (in)congruencies arising from multiple hierarchical levels within the individual, but also from (in)congruencies between self and other. The directionality of the change (increase v. decrease in HF-HRV) across conditions needs further interpretation that were specifically discussed within the empirical chapter on the corresponding study. However, this discussion aims to provide a more comprehensive framework to create expectations regarding the changes in autonomic regulation under different situations (for more on this please refer to the *Interpersonal Congruency Framework* section).

Going beyond unconscious interoceptive inference, this thesis also explored how engaging in different ways with one's own body could facilitate conscious inferences about the body. Beyond their own specific rationales, an overarching theme across all experiments was to create paradigms that can more directly trace different components of interoceptive inference both in an individual and social context. Such approach was much needed in the field as most work related to predictive coding theories remain solely theoretical, simulation based or only testing basic components of the mechanisms. This PhD work contributes to filling this gap between a classic empirical approach and abstract theoretical models.

Briefly, predictive coding theories provide a computational framework that intends to explain how the brain guides action and creates meaning from sensation (e.g. Friston, 2010). One of the main principles of predictive coding is that the brain optimises the use of energy while keeping the body's physiology in balance. To maximise efficiency, the brain needs to accurately infer the causes of the sensations it receives from the outside (exteroceptive) and visceral (interoceptive) sensory channels. The body and past experience serve as a model of the world that the brain can use to predict sensations and actions that can most optimally fit those sensations. The brain's predictions about the causes of its sensations are updated by prediction errors (PE-s) which represent the difference between predictions and actual sensory information. PE-s are passed upward to hierarchically higher cortical levels until they can be explained away. Predictions and PE-s also have reliability attached to them represented as precision. Their relative precision determines if PE-s or predictions contribute more to the current sensation or action (Ainley et al., 2016).

CHAPTER 6: GENERAL DISCUSSION

To test the hypothesis that interoceptive inference plays a key role in both the physiological and psychological stability of the self, the present thesis used a variation of the same novel biofeedback (i.e. here the visual representation of cardiac activity) paradigm across all four studies. By using this task, it was possible to specifically target processes that require the integration of interoceptive and exteroceptive stimuli. With a proof of concept approach, Study 1 tested if HF-HRV could serve as a psychophysiological stability related outcome measure. Specifically, Study 1 considered synchrony between biofeedback and actual cardiac activity as low level congruency. This low-level congruency effect is analogous to the synchrony effects that were previous reported in multisensory integration studies (e.g. Salomon et al., 2016; Suzuki et al., 2013). Furthermore, Study 1 explored if stability neutral or stability facilitating behaviours has any additional effects to biofeedback congruency. Study 2 addressed the generalisability of the congruency effect to a higher hierarchical level of the brain, tracing congruencies between predictions and multisensory stimuli. By shifting the context from individual to a more social interacting setting, Study 3 aimed to move even higher up in the hierarchy than in Study 2 and explore stability relevant processes across different relation between the self and other. Specifically, Study 3 wanted to see if the effect of low-level biofeedback congruency changed across different levels of self-other congruency (competition vs. cooperation). Finally, Study 4 explored the role of individual’s engagement in cardiac recognition. Shifting the focus from PE minimization and unconscious interoceptive inference, Study 4 became more concentrated on the relative precision of the interoceptive and exteroceptive channels in the formation of conscious precepts. For a visual summary on where the rationales of the four different studies could be placed within the predictive coding approach using the hierarchical Self-model by Allen and Tsakiris (2018) please see Figure 1. In the following sections, I will first present a (i) brief summary of the main results of all four studies, following with a (ii) critical discussion of limitations and suggested directions for future research including a proposal for an integrative framework across Studies 1, 2 and 3, then finally I close this thesis with (iii) the overall conclusions.



Original figure adapted from Allen and Tsakiris (2018)

Figure 1. Allocation of study rationales presented in this thesis that were addressing different hierarchical levels on the Dynamic Self-Model (Allen & Tsakiris, 2018). The original model has three circles representing three increasingly abstract levels of hierarchies of the brain. The centre of the hierarchy, where multisensory PE-s and global self-predictions get integrated, was adapted from the original figure to highlight the original proposal of Study 3. This modification depicts a social situation where another person can be integrated with one's self (represented by the green and blue circles respectively) allowing the usual flow of predictive processes to stay intact. In other words, when interpersonal congruency is high the stability of

the global self-model will stay intact. This figure also notes the function of HF-HRV in stability estimation/preservation that is closely linked to PE-s within the canonical microcircuit. For model on lower self-other congruency see Figure 2 and the *Critical Analysis of Findings and Future Directions* section. *Notes*: S = Self, O = Other.

6.2 Summary of Findings

Study 1 explored the difference in HF-HRV and cardiac recognition between congruent and incongruent biofeedback but also the effects of stability facilitating (i.e. regulation) and neutral behaviour (i.e. attention). While there was no difference between the attempt to regulate or simply attend the biofeedback, HF-HRV was significantly lower when receiving incongruent than congruent feedback. The analysis also revealed that participants' performance was at ceiling in cardiac recognition, suggesting the need for adjustments in the paradigm to study conscious inferences. These results indicated that changes in autonomic regulation measured by HF-HRV can potentially serve as an index of PE related processes as they are sensitive to multisensory (in)congruencies.

Capitalising on these findings Study 2 aimed to check the generalisability of HF-HRV to a higher level where multisensory information and more abstract beliefs get integrated. In addition to receiving congruent or incongruent biofeedback participants were guided to think correctly or erroneously, that the feedback belonged to them or someone else. These two manipulations together could create meta-congruency such as the congruency of participant's belief with the true congruency of the biofeedback. According to the findings, meta-congruency had the same effects on HF-HRV as on a lower level. In other words, changes in autonomic regulation could indeed reflect psychophysiological stability related processes, functioning as *interoactions* that minimise surprise across different hierarchies (Seth & Tsakiris, 2018).

Study 3 tested whether the effects of (in)congruent feedback of HF-HRV could be affected by differences in social context, specifically across individual, socially competitive and socially cooperative settings. This study was inspired by the findings of the joint action literature (e.g. Sebanz et al., 2006) suggesting the presence of the co-representation of the co-actor's task within a social context (Ruissen and De Bruijn, 2016). The social context manipulation aimed to create different levels of shared task representation and thus different levels of self-other congruency. In this design the competition condition would represent a lower, cooperation a higher and individual a neutral level of congruency between the self and another person. The core idea behind this study was to see if there is an interaction between different levels of self-other congruency and low-level biofeedback congruency. To capture interpersonal physiological effects, cardiac coupling was also measured alongside HF-HRV. The results revealed no interaction between social context and biofeedback congruency potentially due to the absence of a stable self-other association in the cooperation condition. However, the findings are still relevant for understanding the contribution of autonomic regulation to social interactions. In addition to replicating the low-level biofeedback congruency effect (also detected by Study 1 and Study 2), HF-HRV and cardiac coupling was sensitive to changes in the social context. Specifically, an increase in HF-HRV and a decrease in cardiac coupling was detected during competition when compared to the individual context. On a cognitive level, these findings can be interpreted as stability preserving physiological processes potentially via increased level of self-other differentiation during competition.

Following up from Study1, Study 4 optimised the biofeedback task and the experimental design for cardiac recognition. The trials were shortened to 10 seconds and the number of trials were increased to a level which gave better reliability and higher variation across individuals in terms of their cardiac recognition performance. The main rationale was to test the effects of three different engagement strategies (i.e. attending external features, feeling heartbeats or regulating heart rate) on participants’ sensitivity to identify their own cardiac feedback. Beyond cardiac recognition performance, Study 4 also measured heart beat evoked potentials (HEP, thought to reflect attention allocated to heartbeats) and metacognition (here the experience of certainty in one’s cardiac recognition performance). The results revealed a better cardiac recognition and a more pronounced cortical processing of heartbeats during the regulation or heartbeat feeling strategies when compared to the exteroceptive attention strategy. However, the two interoceptive strategies also differed to some extent. Specifically, metacognition was the highest when using the regulation strategy for cardiac recognition. In contrast, strategy specific cardiac recognition was only linked to the modulation of HEP when attempting to feel heartbeats.

The main take-away across studies presented in this thesis is that autonomic regulation does have the potential to trace integration processes within the brain (Thayer et al., 2012) and to function as a domain general process facilitating the communication across different hierarchies in the brain. Furthermore, an outcome measure of autonomic regulation (here HF-HRV) seem to be able to address some of the methodological gaps that the testing of predictive coding principles would require. Specifically, across studies 1-3, HF-HRV presented itself as an empirical means to track changes in PE related processes.

While the present thesis did not involve studies focusing on the localization of said effects in the brain, neuroanatomical relevance of the reported results could still be discussed with caution. Our brains are designed to optimize utilization of energy, not only in the care and maintenance of our bodies via the homeostatic autonomic systems but also in the behaviours that we perform (Craig, 2002). This is the main principle that brings Studies 1-2-3 closer with Study 4 which aimed to focus on participants regulation related actions and their neural signature. Emotional and regulation related behaviours provide shortcuts to an efficient response (sometimes in anticipation) to different situations. Resonating with this idea, Thayer and colleagues (2012) have suggested that HRV could serve as an output of neural integration and can track one’s capacity to effectively function in a complex social environment. The insular cortex, amygdala and the medial prefrontal cortex (mPFC) are all part of the brain’s “core integration” system, because of their use of internal and external information, and their link to regulating peripheral physiology and behavior. Specifically, the anterior insula was proposed to be the main hub of integration that generates a representation of the subjective, interoceptive self (Craig, 2002). Function-wise the *posterior* insula can be considered the interoceptive cortex as it gets activated by a whole range of interceptive stimuli (e.g. cooling, heating, pricking). However, only when the interoceptive information reaches the *anterior* insula is when the representation of bodily and mental states brought together resulting in a conscious interoceptive experience. Study 1, 2 and 3 could potentially be linked to the activity of anterior insular cortex—the sensory part of emotional control – as all three involved simple attention to interoceptive stimuli. Beyond the anterior insula, Study 4 but specifically its Regulate condition could also activate the anterior cingulate cortex (ACC) – functioning as the motor part of emotional control – because that condition required the participant’s *active* engagement with the biofeedback signal. More precisely, the ventral part of ACC is involved in emotional tasks, and it is directly connected to the amygdala, nucleus accumbens, insular cortex and hypothalamus. In contrast, the dorsal part, is mainly involved in cognitive tasks and is directly connected to the prefrontal and parietal cortices, that handle motor and sensorial conflict resolution and coordination (Bush, Luu & Posner, 2000). Stephan and colleagues (2016b) postulate that the performance of the interoceptive circuit is monitored by a higher metacognitive layer, in the anterior prefrontal cortex. This metacognitive layer encodes and updates beliefs about the brain’s capacity to regulate bodily states, with the resulting representation of one’s own self-efficacy. Considering that Regulate proved to be distinctly different from Attend and Feel strategies in terms of metacognition, the results seem to support that it might involve activation of the AC and prefrontal cortices beyond the insular cortex.

In terms of laterization it has been found that the right side of the anterior insula is larger than the left to a small degree and that it controls energy utilization of the sympathetic division of the autonomic nervous system (Craig, 2002). On the other hand, ethological evidence now seems to indicate that the left half of the brain across species, is more involved in calm, routine and affiliative behaviour which physiologically is associated with the functioning of the parasympathetic system (Rogers, 2013). This would mean that we could potentially see a difference in activity in the left insular cortices when comparing congruent to incongruent conditions across all 4 studies – accompanying the physiological differences in vagal control (i.e. parasympathetic dominance) reported in this thesis.

6.3 Critical Analysis of Findings and Future Directions

The following section aims to avoid repetition with the respective *Discussion* sections of each chapter by focusing on broader themes interlinking the four studies and simultaneously outlining directions for future research. First, Studies 1-3 together suggest that autonomic regulation measured by HF-HRV is sensitive to information conveyed by contextual changes that is relevant for self-stability. Furthermore, the three studies found evidence for the hypothesized role of autonomic signals in multimodal and multilevel integration across different hierarchical levels that are also assumed to map onto different cortical hierarchies (Allen & Tsakiris; 2018, Seth & Tsakiris, 2018; Thayer et al., 2012). When considering their precise role within the integration process, autonomic signals were both proposed to be estimates of self-stability or in other words PE-s themselves (Allen & Tsakiris, 2018) or alternatively “interoactions” that minimize PE-s (Seth & Tsakiris, 2018). While these options do not necessarily exclude one another, our experimental designs so far support the latter, given that HF-HRV was measured in *response* to incongruencies and to changes in social context. Future studies could elegantly test the estimation hypothesis with a design that manipulated the autonomic responses themselves. A possible implementation for the proposed manipulation is to use transcutaneous vagal nerve stimulation (tVNS) (Clancy et al., 2014) or apply beta blockers (Cook et al., 1991; Niemelä, Airaksinen & Huikuri, 1994). Both methods have been shown to increase HRV in healthy subjects. Study 4 also has important implications for predictive coding theories. Specifically, its results indicated that the precision (i.e. reliability) of the interoceptive signals can be affected across different hierarchical levels mirrored by differences in metacognition. The findings confirmed that precision can be assigned to different hierarchical levels when making inferences about the sources of one’s experience.

The studies covered by this thesis will hopefully motivate more research in this area with a greater emphasis on a more mechanistic and applied approach within the predictive coding framework. By assigning outcome measures and conditions to certain theoretical components, terms like PE or precision will stop being abstract concepts only. This approach fits into the traditions of experimental psychology and even has the potential to inform theory development with further discoveries. However, it is important to acknowledge the challenges of which only some can be compensated by adjusting the methodology. For instance, after running Study 1 it became clear that cardiac recognition and autonomic regulation cannot be optimally co-measured within the same paradigm. While HF-HRV ideally requires at least 2 minutes of recording (Berntson et al., 2007; Camm et al., 1996) the length of trials in classic decision-making paradigms operate on the scale of seconds or even milliseconds. This difference in timescales is purely related to the computation requirements of HF-HRV and it is not reflective of the phenomenon itself. In other words, while the physiological changes reflecting parasympathetic influence happen fast (Porges, 2007), but it is their accurate and meaningful quantification that requires a longer period. Without the co-registration of autonomic signals and cardiac recognition, Study 4 could only focus on conscious inferences without tracing interoceptive inferences via the changes in autonomic regulation. A solution for this issue could be to adapt measures of autonomic regulation to non-linear, dynamic measures of HRV like the recurrence quantification analysis (RQA, e.g. Marwan, 2002). To clarify, RQA applies the same methods as the cross-recurrence quantification analysis (CRQA) used in Study 3 to measure the coupling of time series of participants' heart rates. The only difference is that while CRQA explores the shared dynamics of two systems (i.e. x axis is one person and y axis is another person on the plot), RQA represents the dynamics of a single system (i.e. both axes are the same exact time series that belongs to one person). RQA can capture many properties of the heart rate dynamics that would otherwise be lost due to averaging. In other words, RQA relies on variability that is often viewed as unwanted noise from the perspective of linear statistics without making any assumptions about the underlying structure of the system. It has been shown that non-linear HRV parameters rose significantly corresponding to stress (Mohr et al., 2002), exercise (Wallot, Fusaroli, Tylen, & Jegindo, 2013), parasympathetic blockade by atropine (Dabire et al., 1998, Gonzalez, Cordero, Feria, & Pereda, 2000), whilst HRV decreased after propranolol inducing unspecific b-sympathetic blockade (Gonzalez et al., 2000). It was also suggested that different RQA values could trace different aspects of the stress response, such as quantitative changes or subjective experienced-based changes in the stress load (Mohr et al., 2001). RQA can also trace shifts between one *physiological state* to another which could be a valuable method for the adaptation of predictive coding theories to experimental designs.

By noting the epochs of an event or the appearance of a perturbation one can more closely explore state-like shifts in cardiac dynamics.

If one wanted to adopt a more classical approach in experimentation with multiple shorter trials, another possibility would be to use *ultrashort* measures of HRV (~ 10-second-long). This proposal is based on the observation that even a *single* 10 s (standard ECG) recording is sufficient to yield a valid RMSSD (time-domain measure of HRV) measurement, and their reliably would improve with multiple 10s recordings (Munoz et al., 2015). However, as promising this is, the field needs more studies using this methodology to prove the reliability of these findings. Nonetheless, future research on HRV might challenge more and more the 5 minutes gold standard of HRV recording. In any case, researchers would need to carefully justify their choice of periods duration and acknowledge that lengths of measurement less than 5 minutes could hamper the comparability of results across studies and laboratories. Regardless, more experiments with the specific aim to compare different lengths of recordings need to be acquired first, before treating ultrashort recordings as a guideline.

While both Studies 3 and 4 aimed to take on a more functional and ecological approach, they are still only early attempts to transition from a strict lab-based environment to the richness and diversity of social contexts that people normally find themselves in. The relevance of such approach is even higher when studying allostasis or homeostasis. These mechanisms no longer consider the individual as passive nor isolated from their social and physical environment. Instead, they emphasize that individuals adapt to the changes in the environment and actively shape their surroundings to maintain psychophysiological stability (e.g. Porges, 2007). Recent theories emphasize the fundamental role of early relationships in the development of physiological and behavioural regulation. Specifically, Fotopoulou and Tsakiris (2017) proposed that even the experience of one’s bodily self is shaped by embodied interactions with others starting in early infancy. The authors argue that through such ineractions a developing organism can learn to mentalize their homeostatic regulation and eventually build a good enough model of their own body. Eventually these mental models will represent the need to maintain physiological states within a given dynamic range even when internal or external perturbations arise. Fittingly with this, Atzil and colleagues (2018) go against the classic assumption that humans are born with a brain system that innately enables social affiliation. Alternatively, they argue that social affiliation is rooted in allostasis as social animals cannot survive alone. In summary, these theories further emphasize the importance of improving the ecological validity of future experiments that wish to study how autonomic regulation is embedded in social settings.

6.4 Interpersonal Congruency Framework

Study 3 found that autonomic regulation was sensitive to changes in the framing of the social context and proposed self-other congruency as an underlying cognitive mechanism with the function to maintain psychophysiological stability of the self. Elaborating on this idea, the next section proposes a novel mechanistic framework of *interpersonal congruency*, linking self-other congruency to the concept of psychophysiological self-stability and its measurement as traceable changes in HF-HRV (Figure 2). Within the interpersonal congruency framework, self-other congruency can be defined as the perceived similarity across different domains between individuals in a social context. Interpersonal congruency is in line with studies showing higher levels of shared task representation when the co-actor is more human like, which emphasizes the importance of the perceived similarity between people (Müller et al., 2011; Stenzel et al., 2014). Within the *interpersonal congruency* approach, communication and physiological entrainment/mapping could be considered as a means of increasing or decreasing congruency between people on multiple modalities. Communication within the interpersonal congruency framework is considered verbal, through which individuals can share their perceptions, concepts and emotional states with each other. The degree of interpersonal congruency conveyed by communication could be quantified by Conceptual similarity (CS) and the congruency conveyed via interoceptive processes (physiological mapping) could be quantified by Physiological synchrony (PS). Beyond verbal communication and physiological entrainment other domains like action coordination could also be included in the framework. However, as the present PhD mainly operated on the conceptual and physiological level the described model will not include the motor action side at this stage, although - again - their implementation is entirely plausible. Illustrating how the motor domain and actions could be incorporated into the interpersonal congruency framework we can turn to the model by Friston and Frith (2015) which considers communication in the context of active inference. Active inference employs the Bayesian brain with an embodied setting and generates action to minimise uncertainty about their causes. By widening up communication with the inclusion active inference, communication becomes inherently embodied and enactivist in nature. However, for the sake of simplicity, onwards from here the present work will mean verbal communication when using the term communication.

While these theoretical components (i.e. CS, PS) could look distinct on the surface, they should inherently be considered as the manifestation of the same process. Predictive coding theories dissolve the artificial boundaries across emotion, cognition and perception, unifying them within a domain-general process by which the brain creates meaning of the world and guides actions. Using the internal cognitive model of the world (Buckner, 2012; Hassabis & Maguire, 2009), with the body’s internal milieu (Barrett & Simmons, 2015; Garfinkel, Seth, Barrett, Suzuki, & Critchley, 2015), the brain can use past statistical regularities to predict which sensations are most probable in the future. According to Hoemann and Feldman Barrett (2019) an experience will manifest as cognition when the brain prioritises mental contents and processes. On the other hand, when the brain prioritises bodily changes the experience will manifest as an emotion. In other words, cognition, emotion and perception would be still considered as the same domain-general process but would differ what underlying process are prioritised by the brain.

It has been observed that emotion categories are initially used more broadly but over time their use gets narrower, suggesting the concepts becoming more refined (e.g. Widen & Russell, 2003, 2008). Language also structures both individual and shared experience as emotions can be shared through language, allowing predictions to be collectively constructed (e.g. Rimé, 2007, 2009). These premises are consistent with the *interpersonal congruency* framework where communication could be used to distance or draw closer the meaning of the world created and shared by the individuals. Similarly, on the physiological level, similarities or differences in individuals' physiology could facilitate the association or dissociation of their subjective experiences about the world. This idea relies on the premise that even if the same external world surrounds everyone, the corresponding experience will be anchored to what one feels, thinks and does in the moment. It has been proposed that the subjectivity of one's experience is underpinned by the continuous mapping of internal homeostatic states of the body (Damasio, 2010). This means that via physiological mapping/entrainment the interoceptive subjectivity of the two people could become more similar, therefore their experiences would become more similar as well. In contrast, when the self wants to dissociate from the other person then the inverse of the mechanism could be used to decouple the congruency between individual's interoceptive subjectivity. Whether individuals would aim for higher or lower interpersonal congruency could depend on the social context or their past and predicted experiences about the world in relation to the other person. The interpersonal congruency framework also fits with the proposals of Friston and Frith (2015) where they extend the function of predictive processes to cover others' mental states by inferring the states of mind in which the observers themselves are. Specifically, they suggested that a generalized synchrony would inherently emerge when an observer is modelling the behaviour of another person who is modelling the observer. The advantage of the extension of predictive coding theories to social settings is that it can elegantly bridge the potentially arbitrary gap between individual and social context.

In the following section I will review the way interpersonal congruency could relate to predictions about self-stability and autonomic regulation. I will hypothesize two alternative courses of progression for autonomic regulation, depending on the perceived degree of interpersonal congruency. Please note that this framework does not aim to cover situations where individuals could have mismatched perceptions or predictions about their interpersonal congruency. Higher self-other congruency could appear in relationships where there is a stable association between people, which could correspond to higher levels of CS and/or PS on the measurement level. The presence of interpersonal congruency could signal the unperturbed stability of the self in relation to the other person, leaving HF-HRV, at this stage, unaffected. Beyond this point predictive processing could continue in the same way as in the individual context (Figure 1) fittingly with arguments by Friston and Frith (2015). On the other hand, when self-other congruency is low, like during competition (e.g. in Study 3) the self would be perceived as incongruent with the other, reflected in decreases in CS and/or PS. Interpersonal incongruency could be associated with the anticipation of other-related perturbance of self-stability. From this point, individuals’ autonomic regulation would depend on whether the predicted sensation has a cognitive or emotional concept already attached to them or not. When past experiences of the category of sensation is found to fit the predicted sensory array, the brain can use this category as its best guess of what caused a given sensory input and what should be done about it (Hoemann & Feldman Barrett, 2019). In the present framework, this is represented as *model-free* (category/concept unrepresented) and *model-based* (category/concept exists) regulation. The differentiation between model-free and model-based control resonates with the proposed bases for intrapersonal emotion regulation by Etkin, Büchel and Gross (2015). According to the authors, model-free control is characterized by its responsiveness to environmental events within a limited set of potential stimuli and responses. In such cases, behaviour is guided solely by experienced PE-s (without the need for a priori knowledge/concepts) and is therefore computationally fast but not very flexible. In other words, the organism - without knowing better - could rely on autonomic reactivity associated with stress responding represented by the decreases in HF-HRV. In this state the system could be considered dysregulated and “locked in” to a pattern (Thayer et al., 2012). On the other hand, model-based control is characterized by application of rule-based decision-making and dynamic computation of optimal actions and be less efficient computationally (Etkin et al., 2015). In the predictive coding framework, model-based control would be when the brain has an internal model or concept representing the individual’s previous experience with the context of the sensory input. Model-based control is particularly useful when the a priori knowledge has a shortcut to a decision (Etkin et al., 2015). When linking it to an outcome variable, model-based regulation could be associated with increases in HF-HRV as the organism knows how to regulate optimally within such context. This argument is supported by reports on increased HRV during successful performance on emotion regulation tasks (Butler et al., 2006; Ingjaldsson et al., 2003; Smith et al., 2011). Having both model-free and model-based regulation when interpersonal-congruency is also gives the freedom for autonomic regulation and PS to vary independently form one another. This means that low PS would not always be associated with increased HF-HRV as we have seen in Study 3. Nonetheless, when applying this framework on Study 3, decreased PS would signal lower interpersonal congruency. Progressing within the same framework, increased HF-HRV in Study 3 could represent the existence of mental models for competition situation and the successful use of model-based regulation.

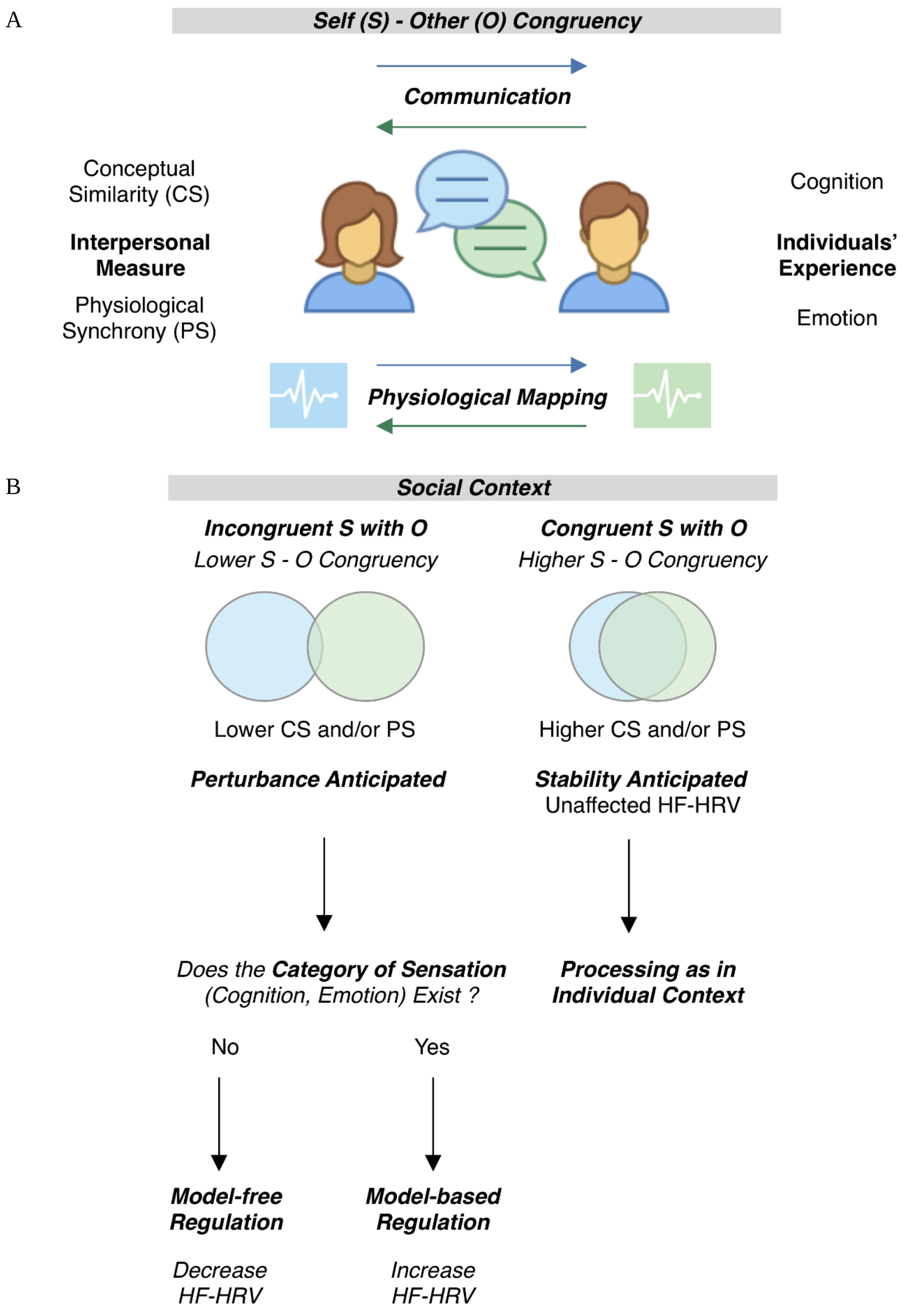


Figure 2. Interpersonal congruency framework. (A) Schematic representation of processes (i.e. Communication, Physiological mapping/entrainment) through which interpersonal

congruency can be compared and maintained or established. (B) Illustration of predictive stability relevant mechanisms in social contexts with lower and higher levels of interpersonal congruency.

The first step to directly test the proposed framework would be to see if interpersonal congruency affected egocentric tendencies. It has been reported that the attribution of emotional and mental states to others seems to be influenced by egocentric tendencies (e.g. Royzman, Cassidy, & Baron, 2003), which in turn could hinder the understanding of other people's emotions. The egocentricity bias is most pronounced when other people's emotions are incongruent with one's own and can be measured by Emotional Egocentric Bias (EEB) task (Silani, Lamm, Ruff, & Singer, 2013). In each trial of the audio-visual version of the task (AV-EEB, Von Mohr, Finotti, Villani, & Tsakiris, 2019) participants are listening to pleasant (e.g. a baby laughing) or unpleasant sounds (e.g. a baby crying) while being simultaneously presented with a picture on the screen depicting what the other person is listening to. After each trial, participants rate the pleasantness of their or other participant's experience while listening to the sound. The EEB score is based on the difference between congruent and incongruent trials after separating trials to ratings on own and other's experience. The proposed experiment could manipulate the interpersonal congruency prior to the AV-EEB task and look at differences in EB between dyads with low and high interpersonal congruency. To achieve high or interpersonal congruency, the study could have a pre-screening questionnaire ask for ratings on topics that people have strong opinions about (e.g. politics, environmental issues, social media). Based on the pre-screening, then participants could be grouped into dyads with potentially high and low interpersonal congruency. Supporting this idea, Study 3 suggest that the experience with face-to-face communication could affect the perception of the other person. For this reason, on the day of the experiment, dyad members could also discuss the topic they had low or high agreement on. After the discussion AV-EEB could be registered. The interpersonal congruency framework would suggest a lower egocentricity bias at the dyads with high interpersonal congruency when compared to the dyads with low interpersonal congruency. To achieve higher statistical power, I would suggest each individual participant to be paired into both low and high congruency dyads.

The second step in testing could be to see if the categories of sensation would really affect autonomic regulation within a social context. Emotional granularity (EG) refers to individual differences in the specificity of one's emotional experiences and concepts (e.g. Barrett, 2004) which could be used to check how refined are the categories people have for their emotions. When measuring EG, participants get prompts or are invited to the lab to report their emotional experiences several times and across a variety of circumstances. The quantification of EG requires the summary of correlations among the used emotion terms across measurement times (Barrett, 2004). The person who always reports the same levels of anger and sadness will produce a $r = 1$ between these states. This example would suggest low granularity as there is no distinction between these emotions for the individual. The proposed study would record EG and individual's opinions about certain conversational topics prior to the experiment. After this, dyads with low interpersonal congruency could be created and matched in terms of their average levels of EG. During the experiment, participants would be asked to discuss some of the topics they disagree on while measuring their cardiac activity. The hypothesis based on the interpersonal congruency framework would predict that couples with lower levels of EG would show lower levels of HF-HRV during conversation.

A related question could also consider whether interpersonal congruency improves through interoceptive self-awareness. Specifically, if two people are experiencing the same physiological patterns, would one's own interoceptive abilities also improve interpersonal congruency? In a recent theoretical work (Palmer & Tsakiris, 2018) interoception was proposed to play a critical role in social cognition. Specifically, Palmer and Tsakiris (2018) suggested interoceptive accuracy appears to stabilise the mental representation of one’s self as distinct from others. In their theory, for tasks with a low level self-other distinction, participants with lower interoceptive accuracy would display greater influences coming from the other person. In comparison, when a social task requires a high level of self-other distinction, people with higher interoceptive accuracy would be able to understand the emotional state of others more. The reason behind these predictions is that better interoceptive accuracy is likely to be associated with a stable representation of the bodily self, preventing the blurring of self and other resulting in improved performance. However, this question requires a reliable measure of interoceptive accuracy (correct and precise monitoring of interoceptive signals) but the most popular measure of this feature (i.e. heartbeat counting task, Schandry, 1981) has recently received wide criticism (Ring & Brener, 2018, Zamariola, Maurage, Luminet & Corneille, 2018). Ring and Brener (2018) found that heartbeat counting task scores were unrelated to the heartbeat detection task scores although these two performances supposed to trace the same process. Furthermore, Zamariola and colleagues (2018) put forward the following four key criticisms regarding the heartbeat tracking task. Firstly, they note that the heartbeat tracking performance largely depends on participants under-reporting the number of heartbeats. Secondly, they propose when accumulated across trials, the number of actual and reported heartbeats should be linked to one another. Yet, in their sample, they found only a weak overall positive correlation in this regard. Furthermore, they argue that the correlation between the total number of actual and reported heartbeats should increase with interoceptive accuracy, but they found no such rising linear trend. Thirdly, they suggest that a valid measure of IAcc should not be linked to one’s heart rate, yet they found a negative correlation between the two. Finally, the authors report that participants showed a tendency to a poorer performance on longer trials of the heartbeat counting task.

Nonetheless, once this criticism is addressed either with a new method or a defence of the old methods (Ainley, Tsakiris, Pollatos, Schulz & Herbert, 2019) testing the involvement of interoceptive accuracy in interpersonal congruency would be valuable.

Conclusions

During the integration of signals arising from one' own body and the everchanging outside world, the brain and body keep a fine balance between stability and change. Within a predictive coding framework, this PhD work investigated neural, physiological and behavioural processes that could contribute to the psychophysiological stability of the bodily self. Across 4 studies, this thesis explored the role of autonomic regulation and interoceptive strategies both in unconscious interoceptive inference and explicit inference. Using different versions of a cardiac biofeedback paradigm, participants received congruent and incongruent feedback of their cardiac activity and study-specific instruction. The findings of Study 1, 2 and partly 3, provide evidence that autonomic regulation (measured by HF-HRV) contributes to the psychophysiological stability of the self across multiple hierarchical levels. HF-HRV was not only sensitive to perturbance arising from the level of multisensory integration and beliefs, it was also influenced by changes in the social context. Based on the finding of Study 3, changes in autonomic regulation but also in physiological synchrony could signal a stability preserving process during competition. Finally, going beyond unconscious physiological responding, Study 4 used a novel cardiac recognition paradigm and employed measures to capture several features of interoception. Integrating behavioural, cortical and metacognitive aspects, the findings of Study 4 suggest that people can use different strategies to reach valid inferences about their sensations – again on different hierarchical levels. Together, these findings have important theoretical implications for predictive coding models of the self and social neuroscience as they pave the way for a more direct application of theories in experimental designs. In line with future directions, the general discussion proceeded to outline a novel mechanistic framework that aimed to bridge the arbitrary gap between experiments within an individual and social context.

References

Adams, R. A., Shipp, S., & Friston, K. J. (2013). Predictions not commands: Active inference in the motor system. https://doi.org/10.1007/s00429-012-0475-5

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

Ainley, V., Apps, M. A., 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., Tajadura-Jiménez, A., Fotopoulou, A., & Tsakiris, M. (2012). Looking into myself: Changes in interoceptive sensitivity during mirror self-observation. *Psychophysiology*, *49*(11), 1672–1676. https://doi.org/10.1111/j.1469-8986.2012.01468.x

Ainley, V., Tsakiris, M., Pollatos, O., Schulz, A., Herbert, B. M., Ainley, V., ... & Tsakiris, M. Comment on “Zamariola et al.,(2018), Interoceptive Accuracy Scores are Problematic: Evidence from Simple Bivariate Correlations”-The Empirical Data Base, the Conceptual Reasoning and the Analysis behind this Statement are Misconceived and do not Support the Authors’ Conclusions.

Allen, J. J. B., Chambers, A. S., & Towers, D. N. (2007). The many metrics of cardiac chronotropy: A pragmatic primer and a brief comparison of metrics. *Biological Psychology*, *74*(2), 243–262. https://doi.org/10.1016/j.biopsycho.2006.08.005

REFERENCES

Allen, M., & Tsakiris, M. (2018). The body as first prior: Interoceptive predictive processing and the primacy of self-models. The Interoceptive Mind from Homeostasis to Awareness, 27–45.

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. https://doi.org/10.7554/eLife.18103.001

Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R., & Kievit, R. (2019). Raincloud plots: a multi-platform tool for robust data visualization. Wellcome open research, 4. https://doi.org/10.7287/peerj.preprints.27137v1

Apps, M. A. J., & Tsakiris, M. (2014). The free-energy self: A predictive coding account of self-recognition. *Neuroscience and Biobehavioral Reviews, 41*, 85–97. https://doi.org/10.1016/j.neubiorev.2013.01.029

Aspell, J. E., Heydrich, L., Marillier, G., Lavanchy, T., Herbelin, B., & Blanke, O. (2013). Turning Body and Self Inside Out: Visualized Heartbeats Alter Bodily Self-Consciousness and Tactile Perception. *Psychological Science, 24*(12), 2445–2453. https://doi.org/10.1177/0956797613498395

Atmaca, S., Sebanz, N., Prinz, W., & Knoblich, G. (2008). Listening to humans walking together activates the social brain circuitry. *Social Neuroscience, 3*(3-4), 410–420. https://doi.org/10.1080/17470910801900908

Atzil, S., Gao, W., Fradkin, I., & Barrett, L. F. (2018). Growing a social brain. *Nature Human Behaviour, 2*(9), 624–636. https://doi.org/10.1038/s41562-018-0384-6

Aust, F., & Barth, M. (2018). *{papaja}: {Create} {APA} manuscripts with {R Markdown}*. Retrieved from https://github.com/crsh/papaja

REFERENCES

Azevedo, R. T., Aglioti, S. M., & Lenggenhager, B. (2016). Participants' above-chance recognition of own-heart sound combined with poor metacognitive awareness suggests implicit knowledge of own heart cardiodynamics. *Scientific Reports*, *6*, 26545. https://doi.org/10.1038/srep26545

Azevedo, R. T., Ainley, V., & Tsakiris, M. (2016). Cardio-visual integration modulates the subjective perception of affectively neutral stimuli. *International Journal of Psychophysiology*, *99*, 10–17. https://doi.org/10.1016/j.ijpsycho.2015.11.011

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Backs, R. W., & Ryan, A. M. (1992). Multimodal Measures of Mental Workload during Dual-Task Performance: Energetic Demands of Cognitive Processes. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *36*(18), 1413–1417. https://doi.org/10.1177/154193129203601811

Baltazar, M., Hazem, N., Vilarem, E., Beaucousin, V., Picq, J. L., & Conty, L. (2014). Eye contact elicits bodily self-awareness in human adults. *Cognition*, *133*(1), 120–127. https://doi.org/10.1016/j.cognition.2014.06.009

Barrett, L. F. (2004). Feelings or words? Understanding the content in self-report ratings of experienced emotion. *Journal of Personality and Social Psychology*, *87*(2), 266–281. https://doi.org/10.1037/0022-3514.87.2.266

Barrett, L. F., & Simmons, W. K. (2015). Interoceptive predictions in the brain. https://doi.org/10.1038/nrn3950

REFERENCES

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bergman, J. S., & Johnson, H. J. (1972). Sources of Information Which Affect Training and Raising of Heart Rate. *Psychophysiology*, *9*(1), 30–39. https://doi.org/10.1111/j.1469-8986.1972.tb00740.x

Berntson, G. G., Bigger, J. T., Eckberg, D. L., Grossman, P., Kaufmann, P. G., Malik, M., et al. (1997). Heart rate variability: origins, methods, and interpretive caveats. Psychophysiology 34, 623–648.

Berntson, G. G., Cacioppo, J. T., & Grossman, P. (2007). Whither vagal tone. *Biological Psychology*, *74*(2), 295–300. https://doi.org/10.1016/j.biopsycho.2006.08.006

Berntson, G. G., Cacioppo, J. T., & Quigley, K. S. (1991). Autonomic determinism: The modes of autonomic control, the doctrine of autonomic space, and the laws of autonomic constraint. *Psychological Review*, *98*(4), 459–487. https://doi.org/10.1037/0033-295X.98.4.459

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

Berntson, G. G., Thomas Bigger, J., Eckberg, D. L., Grossman, P., Kaufmann, P. G., Malik, M., … Van Der Molen, M. W. (1997). Heart rate variability: Origins methods, and interpretive caveats. https://doi.org/10.1111/j.1469-8986.1997.tb02140.x

Bertsch, K., Hagemann, D., Naumann, E., Schachinger, H., & Schulz, A. (2012). Stability of heart rate variability indices reflecting parasympathetic activity. Psychophysiology 49, 672–682. https://doi.org/10.1111/j.1469-8986.2011.01341.x

Billman, G. E. (2013). The LF/HF ratio does not accurately measure cardiac sympatho-vagal balance. *Frontiers in Physiology*, *4,* 26. https://doi.org/10.3389/fphys.2013.00026

Björnsdotter, M., Löken, L., Olausson, H., Vallbo, Å., & Wessberg, J. (2009). Somatotopic organization of gentle touch processing in the posterior insular cortex. *Journal of Neuroscience*, *29*(29), 9314–9320. https://doi.org/10.1523/JNEUROSCI.0400-09.2009

Blakemore, S. J., Frith, C. D., & Wolpert, D. M. (1999). Spatio-temporal prediction modulates the perception of self-produced stimuli. *Journal of Cognitive Neuroscience, 11*(5), 551–559. https://doi.org/10.1162/089892999563607

Blakemore, S. J., Wolpert, D. M., & Frith, C. D. (2002). Abnormalities in the awareness of action. https://doi.org/10.1016/S1364-6613(02)01907-1

Blanchard, E. B., & Young, L. D. (1972). The relative efficacy of visual and auditory feedback for self-control of heart rate. The Journal of general psychology, 87(2), 195-202.

Blanchard, E. B., Scott, R. W., Young, L. D., & Edmundson, A. E. D. (1974). Effect of Knowledge of Response on the Self-Control of Heart Rate. *Psychophysiology*, *11*(3), 251–264. https://doi.org/10.1111/j.1469-8986.1974.tb00542.x

Bolis, D., & Schilbach, L. (2018). ‘I Interact Therefore I Am’: The Self as a Historical Product of Dialectical Attunement. *Topoi*, 1–14. https://doi.org/10.1007/s11245-018-9574-0

Botvinick, M., & Cohen, J. (1998, February). Rubber hands ’feel’ touch that eyes see [8]. Nature Publishing Group. https://doi.org/10.1038/35784

Boudewyn, M. A., Luck, S. J., Farrens, J. L., & Kappenman, E. S. (2018). How many trials does it take to get a significant ERP effect? It depends. *Psychophysiology*, *55*(6). https://doi.org/10.1111/psyp.13049

Bowles, S., Gintis, H., Bowles, S., & Gintis, H. (2013). *A Cooperative Species: Human Reciprocity and Its Evolution*. Princeton University Press.

Bravi, A., Longtin, A., & Seely, A. J. (2011). Review and classification of variability analysis techniques with clinical applications. https://doi.org/10.1186/1475-925X-10-90

Brener, J., Liu, X., & Ring, C. (1993). A method of constant stimuli for examining heartbeat detection: Comparison with the Brener-Kluvitse and Whitehead methods. *Psychophysiology*, *30*(6), 657–665. https://doi.org/10.1111/j.1469-8986.1993.tb02091.x

Brewer, N., & Wells, G. L. (2006). The confidence-accuracy relationship in eyewitness identification: Effects of lineup instructions, foil similarity, and target-absent base rates. *Journal of Experimental Psychology: Applied*, *12*(1), 11–30. https://doi.org/10.1037/1076-898X.12.1.11

Brodal, P. (2010). The Central Nervous System – Structure and Function. New York, NY: Oxford University Press.

Brooks, J. C., Zambreanu, L., Godinez, A., Craig, A. D., & Tracey, I. (2005). Somatotopic organisation of the human insula to painful heat studied with high resolution functional imaging. *NeuroImage*, *27*(1), 201–209. https://doi.org/10.1016/j.neuroimage.2005.03.041

Brown, H., Adams, R. A., Parees, I., Edwards, M., & Friston, K. (2013). Active inference, sensory attenuation and illusions. *Cognitive Processing*, *14*(4), 411–427. https://doi.org/10.1007/s10339-013-0571-3

Brown, H., Friston, K., & Bestmann, S. (2011). Active inference, attention, and motor preparation. *Frontiers in Psychology*, *2*(SEP). https://doi.org/10.3389/fpsyg.2011.00218

Brown, T. E., Beightol, L. A., Koh, J., & Eckberg, D. L. (1993). Important influence of respiration on human R-R interval power spectra is largely ignored. J. Appl. Physiol. (1985) 75, 2310–2317.

Brysbaert, M., & Stevens, M. (2018). Power Analysis and Effect Size in Mixed Effects Models: A Tutorial. *Journal of Cognition*, *1*(1), 1–20. https://doi.org/10.5334/joc.10

Buckner, R.L. (2012). The serendipitous discovery of the brain’s default network. NeuroImage, 62(2), 1137–45.

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. Trends in cognitive sciences, 4(6), 215-222.

Butler, E. A., & Randall, A. K. (2013). Emotional coregulation in close relationships. *Emotion Review*, *5*(2), 202–210. https://doi.org/10.1177/1754073912451630

Butler, E. A., Wilhelm, F. H., & Gross, J. J. (2006). Respiratory sinus arrhythmia, emotion, and emotion regulation during social interaction. *Psychophysiology*, *43*(6), 612–622. https://doi.org/10.1111/j.1469-8986.2006.00467.x

Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: An fMRI study with expert dancers. *Cerebral Cortex*, *15*(8), 1243–1249. https://doi.org/10.1093/cercor/bhi007

Camm, A. J., Malik, M., Bigger, J. T., Breithardt, G., Cerutti, S., Cohen, R. J., & Lombardi, F. (1996). Heart rate variability: standards of measurement, physiological interpretation and clinical use.Task Force of the European Society of Cardiology and the

North American Society of Pacing and Electrophysiology. *Circulation*, *93*(5), 1043–1065. https://doi.org/10.1161/01.CIR.93.5.1043

Canales-Johnson, A., Silva, C., Huepe, D., Rivera-Rei, Á., Noreika, V., Del Carmen Garcia, M., … Bekinschtein, T. A. (2015). Auditory feedback differentially modulates behavioral and neural markers of objective and subjective performance when tapping to your heartbeat. *Cerebral Cortex*, *25*(11), 4490–4503. https://doi.org/10.1093/cercor/bhv076

Cannon, W. B. (1929). Organization for Physiological Homeostasis. *Physiological Reviews*, *9*(3), 399–431. https://doi.org/10.1152/physrev.1929.9.3.399

Chambers, A. S., & Allen, J. J. (2007). Cardiac vagal control, emotion, psychopathology, and health. https://doi.org/10.1016/j.biopsycho.2006.09.004

Chanel, G., Kivikangas, J. M., & Ravaja, N. (2012). Physiological compliance for social gaming analysis: Cooperative versus competitive play. *Interacting with Computers*, *24*(4), 306–316. https://doi.org/10.1016/j.intcom.2012.04.012

Chapleau, M. W., & Sabharwal, R. (2011). Methods of assessing vagus nerve activity and reflexes. *Heart Failure Reviews*, *16*(2), 109–127. https://doi.org/10.1007/s10741-010-9174-6

CHMP. (2015). *Committee for Medicinal Products for Human Use (CHMP) Guideline on adjustment for baseline covariates in clinical trials*. Retrieved from www.ema.europa.eu/contact

Clancy, J. A., Mary, D. A., Witte, K. K., Greenwood, J. P., Deuchars, S. A., & Deuchars, J. (2014). Non-invasive Vagus nerve stimulation in healthy humans reduces sympathetic nerve activity. *Brain Stimulation*, *7*(6), 871–877. https://doi.org/10.1016/j.brs.2014.07.031

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204. https://doi.org/10.1017/S0140525X12000477

Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, *50*(1), 62–81. https://doi.org/10.1016/j.jml.2003.08.004

Coco, M. I., & Dale, R. (2014). Cross-recurrence quantification analysis of categorical and continuous time series: An R package. *Frontiers in Psychology*, *5*(JUN). https://doi.org/10.3389/fpsyg.2014.00510

Constable, M. D., Elekes, F., Sebanz, N., & Knoblich, G. (2019). Relevant for us? We-prioritization in cognitive processing. *Journal of Experimental Psychology: Human Perception and Performance*, *Advance on*. https://doi.org/10.1037/xhp0000691

Cook, J. R., Bigger Jr, J. T., Kleiger, R. E., Fleiss, J. L., Steinman, R. C., & Rolnitzky, L. M. (1991). Effect of atenolol and diltiazem on heart period variability in normal persons. Journal of the American College of Cardiology, 17(2), 480-484.

Cook, R. D. (1977). Detection of Influential Observation in Linear Regression. *Technometrics*, *19*(1), 15. https://doi.org/10.2307/1268249

Couto, B., Salles, A., Sedeño, L., Peradejordi, M., Barttfeld, P., Canales-Johnson, A., … Ibanez, A. (2013). The man who feels two hearts: The different pathways of interoception. *Social Cognitive and Affective Neuroscience*, *9*(9), 1253–1260. https://doi.org/10.1093/scan/nst108

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

Craig, A. D. (2010, June). The sentient self. Springer-Verlag. https://doi.org/10.1007/s00429-010-0248-y

Creaven, A. M., Skowron, E. A., Hughes, B. M., Howard, S., & Loken, E. (2014). Dyadic concordance in mother and preschooler resting cardiovascular function varies by risk status. *Developmental Psychobiology*, *56*(1), 142–152. https://doi.org/10.1002/dev.21098

Cuthbert, B. N., Schupp, H. T., Bradley, M. M., Birbaumer, N., & Lang, P. J. (2000). Brain potentials in affective picture processing: Covariation with autonomic arousal and affective report. *Biological Psychology*, *52*(2), 95–111. https://doi.org/10.1016/S0301-0511(99)00044-7

Dabiré, H., Mestivier, D., Jarnet, J., Safar, M. E., & Chau, N. P. (1998). Quantification of sympathetic and parasympathetic tones by nonlinear indexes in normotensive rats. *American Journal of Physiology - Heart and Circulatory Physiology*, *275*(4 44-4). https://doi.org/10.1152/ajpheart.1998.275.4.h1290

Damasio, A. (2003). Feelings of emotion and the self. In *Annals of the new york academy of sciences* (Vol. 1001, pp. 253–261). https://doi.org/10.1196/annals.1279.014

Damasio, A. (2010). *Self comes to mind: Constructing the conscious brain*. London: William Heinemann.

Danyluck, C., & Page-Gould, E. (2019). Social and Physiological Context can Affect the Meaning of Physiological Synchrony. *Scientific Reports, 9*(1). https://doi.org/10.1038/s41598-019-44667-5

De Cremer, D., & Stouten, J. (2003). When do people find cooperation most justified? The effect of trust and self-other merging in social dilemmas. *Social Justice Research, 16*(1), 41–52. https://doi.org/10.1023/A:1022974027023

De Pascalis, V., Palumbo, G., & Ronchitelli, V. (1991). Heartbeat perception, instructions, and biofeedback in the control of heart rate. *International Journal of Psychophysiology*, *11*(2), 179–193. https://doi.org/10.1016/0167-8760(91)90010-U

De Preester, H., & Tsakiris, M. (2009). Body-extension versus body-incorporation: Is there a need for a body-model? *Phenomenology and the Cognitive Sciences*, *8*(3), 307–319. https://doi.org/10.1007/s11097-009-9121-y

De Rosario-Martinez, H. (2015). *phia: Post-Hoc Interaction Analysis*. Retrieved from https://cran.r-project.org/package=phia

Decety, J., Jackson, P. L., Sommerville, J. A., Chaminade, T., & Meltzoff, A. N. (2004). The neural bases of cooperation and competition: An fMRI investigation. *NeuroImage*, *23*(2), 744–751. https://doi.org/10.1016/j.neuroimage.2004.05.025

Delignette-Muller, M. L., & Dutang, C. (2015). {fitdistrplus}: An {R} Package for Fitting Distributions. *Journal of Statistical Software*, *64*(4), 1–34. Retrieved from http://www.jstatsoft.org/v64/i04/

Dirlich, G., Vogl, L., Plaschke, M., & Strian, F. (1997). Cardiac field effects on the EEG. *Electroencephalography and Clinical Neurophysiology*, *102*(4), 307–315. https://doi.org/10.1016/S0013-4694(96)96506-2

Dolk, T., Hommel, B., Prinz, W., & Liepelt, R. (2013). The (Not So) social Simon effect: A referential coding account. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(5), 1248–1260. https://doi.org/10.1037/a0031031

Ebisch, S. J., Aureli, T., Bafunno, D., Cardone, D., Romani, G. L., & Merla, A. (2012). Mother and child in synchrony: Thermal facial imprints of autonomic contagion. *Biological Psychology*, *89*(1), 123–129. https://doi.org/10.1016/j.biopsycho.2011.09.018

Eckberg, D. L. (1997). Sympathovagal balance: A critical appraisal. *Circulation*, *96*(9), 3224–3232. https://doi.org/10.1161/01.CIR.96.9.3224

Eckberg, D. L., & Eckberg, M. J. (1982). Human sinus node responses to repetitive, ramped carotid baroreceptor stimuli. *American Journal of Physiology - Heart and Circulatory Physiology*, *11*(4). https://doi.org/10.1152/ajpheart.1982.242.4.h638

Elekes, F., Bródy, G., Halász, E., & Király, I. (2016). Enhanced encoding of the co-actor's target stimuli during a shared non-motor task. *Quarterly Journal of Experimental Psychology*, *69*(12), 2376–2389. https://doi.org/10.1080/17470218.2015.1120332

Elkins, A. N., Muth, E. R., Hoover, A. W., Walker, A. D., Carpenter, T. L., & Switzer, F. S. (2009). Physiological compliance and team performance. *Applied Ergonomics*, *40*(6), 997–1003. https://doi.org/10.1016/j.apergo.2009.02.002

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. Trends in cognitive sciences, 8(4), 162-169.

Etkin, A., Buechel, C., & Gross, J. J. (2015). The neural bases of emotion regulation. *Nature Reviews Neuroscience*, *16*(11), 693–+. https://doi.org/10.1038/nrn4044

Faris, P. L., Hofbauer, R. D., Daughters, R., VandenLangenberg, E., Iversen, L., Goodale, R. L., … Hartman, B. K. (2008). De-stabilization of the positive vago-vagal reflex in bulimia nervosa. *Physiology and Behavior*, *94*(1), 136–153. https://doi.org/10.1016/j.physbeh.2007.11.036

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Feldman, H., & Friston, K. J. (2010). Attention, Uncertainty, and Free-Energy. *Frontiers in Human Neuroscience*, *4*, 215. https://doi.org/10.3389/fnhum.2010.00215

Feldman, R. (2007). Parent–infant synchrony and the construction of shared timing; physiological precursors, developmental outcomes, and risk conditions. Journal of Child psychology and Psychiatry, 48(3-4), 329-354. https://doi.org/10.1111/j.1469-7610.2006.01701.x

Ferrer, E., & Helm, J. L. (2013). Dynamical systems modeling of physiological coregulation in dyadic interactions. *International Journal of Psychophysiology*, *88*(3), 296–308. https://doi.org/10.1016/j.ijpsycho.2012.10.013

Ferri, F., Ardizzi, M., Ambrosecchia, M., & Gallese, V. (2013). Closing the Gap between the Inside and the Outside: Interoceptive Sensitivity and Social Distances. *PLoS ONE*, *8*(10), e75758. https://doi.org/10.1371/journal.pone.0075758

FitzGerald, T. H., Schwartenbeck, P., Moutoussis, M., Dolan, R. J., & Friston, K. (2015). Active inference, evidence accumulation, and the urn task. *Neural Computation*, *27*(2), 306–328. https://doi.org/10.1162/NECO_a_00699

Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience*, *8*, 443. https://doi.org/10.3389/fnhum.2014.00443

Fotopoulou, A., & Tsakiris, M. (2017). Mentalizing homeostasis: The social origins of interoceptive inference. *Neuropsychoanalysis*, *19*(1), 3–28. https://doi.org/10.1080/15294145.2017.1294031

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. (2009). The free-energy principle: a rough guide to the brain? *Trends in Cognitive Sciences*, *13*(7), 293–301. https://doi.org/10.1016/J.TICS.2009.04.005

Friston, K. (2013). Consciousness and hierarchical inference. *Neuropsychoanalysis*, *15*(1), 38–42. https://doi.org/10.1080/15294145.2013.10773716

Friston, K. J. (2017). Self-evidencing babies: Commentary on “mentalizing homeostasis: The social origins of interoceptive inference” by Fotopoulou & Tsakiris. Taylor; Francis Ltd. https://doi.org/10.1080/15294145.2017.1295216

Friston, K. J., Daunizeau, J., & Kiebel, S. J. (2009). Reinforcement learning or active inference? *PLoS ONE*, *4*(7), e6421. https://doi.org/10.1371/journal.pone.0006421

Friston, K., & Frith, C. (2015). A Duet for one. *Consciousness and Cognition*, *36*, 390–405. https://doi.org/10.1016/j.concog.2014.12.003

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

Gaume, A., Vialatte, A., Mora-Sánchez, A., Ramdani, C., & Vialatte, F. B. (2016). A psychoengineering paradigm for the neurocognitive mechanisms of biofeedback and neurofeedback. *Neuroscience and Biobehavioral Reviews*, *68*, 891–910. https://doi.org/10.1016/j.neubiorev.2016.06.012

Ghafar-Tabrizi, R. (2008). Reversal theory and physiological linkage in the low-conflict and high-conflict mother-daughter dyadic interactions. *Contemporary Psychology*, *3*(1), 62–75.

Goedhart, A. D., Willemsen, G., Houtveen, J. H., Boomsma, D. I., & De Geus, E. J. (2008). Comparing low frequency heart rate variability and preejection period: Two sides of a different coin. *Psychophysiology*, *45*(6), 1086–1090. https://doi.org/10.1111/j.1469-8986.2008.00710.x

Goldstein, D. S. (2013). Differential responses of components of the autonomic nervous system. In *Handbook of clinical neurology* (Vol. 117, pp. 13–22). Elsevier B.V. https://doi.org/10.1016/B978-0-444-53491-0.00002-X

González, J. J., Cordero, J. J., Feria, M., & Pereda, E. (2000). Detection and sources of nonlinearity in the variability of cardiac R-R intervals and blood pressure in rats. *American Journal of Physiology - Heart and Circulatory Physiology*, *279*(6 48-6), 3040–3046. https://doi.org/10.1152/ajpheart.2000.279.6.h3040

Gray, M. A., Taggart, P., Sutton, P. M., Groves, D., Holdright, D. R., Bradbury, D., … Critchley, H. D. (2007). A cortical potential reflecting cardiac function. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(16), 6818–6823. https://doi.org/10.1073/pnas.0609509104

Grossman, P. (1992). Respiratory and cardiac rhythms as windows to central and autonomic biobehavioral regulation: selection of window frames, keeping the panes clean and viewing the neural topography. Biol. Psychol. 34, 131–161.

Grossman, P., & Taylor, E. W. (2007). Toward understanding respiratory sinus arrhythmia: Relations to cardiac vagal tone, evolution and biobehavioral functions. *Biological Psychology*, *74*(2), 263–285. https://doi.org/10.1016/j.biopsycho.2005.11.014

Grossman, P., Beek, J. van, & Wientjes, C. (1990). A Comparison of Three Quantification Methods for Estimation of Respiratory Sinus Arrhythmia. *Psychophysiology*, *27*(6), 702–714. https://doi.org/10.1111/j.1469-8986.1990.tb03198.x

Guagnano, D., Rusconi, E., & Umiltà, C. A. (2010). Sharing a task or sharing space? On the effect of the confederate in action coding in a detection task. *Cognition*, *114*(3), 348–355. https://doi.org/10.1016/j.cognition.2009.10.008

Hamilton, S. L. (1982). The structure of the nicotinic acetylcholine receptor. *Progress in Clinical and Biological Research*, *79*, 73–85. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7045887

Hansen, A. L., Johnsen, B. H., & Thayer, J. F. (2003). Vagal influence on working memory and attention. *International Journal of Psychophysiology*, *48*(3), 263–274. https://doi.org/10.1016/S0167-8760(03)00073-4

Hassabis, D., Maguire, E.A., 2009. The construction system of the brain. Philos. Trans. R. Soc. Lond. B: Biol. Sci. 364, 1263–1271, http://dx.doi.org/10.1098/rstb.2008. 0296.

Hauser, T. U., Allen, M., Purg, N., Moutoussis, M., Rees, G., & Dolan, R. J. (2017). Noradrenaline blockade specifically enhances metacognitive performance. *eLife*, *6*, e24901. https://doi.org/10.7554/elife.24901

Heathers, J. A. (2014). Everything Hertz: methodological issues in short- term frequency-domain HRV. Front. Physiol. 5:177. https://doi.org/10.3389/fphys.2014. 00177

Heathers, J. A. (2012). Sympathovagal balance from heart rate variability: an obituary. *Experimental Physiology*, *97*(4), 556–556. https://doi.org/10.1113/expphysiol.2011.063867

Heathers, J., Quintana, D., Angus, D., Krygier, J., Kemp, A., & de Rosnay, M. (2018, February 27). Water consumption as a source of error in the measurement of heart rate variability. https://doi.org/10.31219/osf.io/83exy

Heilbronn, E., & Bartfai, T. (1978). Muscarinic acetylcholine receptor. https://doi.org/10.1016/0301-0082(78)90011-4

Helm, J. L., Sbarra, D. A., & Ferrer, E. (2014). Coregulation of respiratory sinus arrhythmia in adult romantic partners. *Emotion (Washington, D.C.)*, *14*(3), 522–531. https://doi.org/10.1037/a0035960

Helm, J. L., Sbarra, D., & Ferrer, E. (2012). Assessing cross-partner associations in physiological responses via coupled oscillator models. *Emotion*, *12*(4), 748–762. https://doi.org/10.1037/a0025036

Helmholtz, H. V. (1860). 1962. Treatise on physiological optics, 3, 1.

Hirsch, J. A., & Bishop, B. (1981). Respiratory sinus arrhythmia in humans: how breathing pattern modulates heart rate. Am. J. Physiol. 241, H620–H629.

Hodossy, L., & Tsakiris, M. (2020). Wearing your heart on your screen: Investigating congruency-effects in autonomic responses and their role in interoceptive processing during biofeedback. *Cognition*, *194*, 104053. https://doi.org/10.1016/j.cognition.2019.104053

Hoemann, K., & Feldman Barrett, L. (2019). Concepts dissolve artificial boundaries in the study of emotion and cognition, uniting body, brain, and mind. *Cognition and Emotion*, *33*(1), 67–76. https://doi.org/10.1080/02699931.2018.1535428

Hohwy, J. (2012). Attention and conscious perception in the hypothesis testing brain. *Frontiers in Psychology*, *3*(APR), 96. https://doi.org/10.3389/fpsyg.2012.00096

Holzman, J. B., & Bridgett, D. J. (2017). Heart rate variability indices as bio-markers of top-down self-regulatory mechanisms: A meta-analytic review. *Neuroscience and Biobehavioral Reviews*, *74*, 233–255. https://doi.org/10.1016/j.neubiorev.2016.12.032

Houtveen, J. H., Rietveld, S., & de Geus, E. J. (2002). Contribution of tonic vagal modulation of heart rate, central respiratory drive, respiratory depth, and respiratory frequency to respiratory sinus arrhythmia during mental stress and physical exercise. Psychophysiology 39, 427–436.

Huygens, C. (1673/1986). The pendulum clock or geometrical demonstrations concerning the motion of pendula as applied to clocks. Ames: Iowa State University Press (R.J. Blackwell, Trans.).

Ingjaldsson, J. T., Laberg, J. C., & Thayer, J. F. (2003). Reduced heart rate variability in chronic alcohol abuse: Relationship with negative mood, chronic thought suppression, and compulsive drinking. *Biological Psychiatry*, *54*(12), 1427–1436. https://doi.org/10.1016/S0006-3223(02)01926-1

James, W. (1890). The Principles of Psychology (New York: Henry Holt & Co.).

Jänig, W., & McLachlan, E. M. (2013). Neurobiology of the autonomic nervous system. Autonomic failure, 21-34.

Järvelä, S., Kivikangas, J. M., Kätsyri, J., & Ravaja, N. (2014). Physiological Linkage of Dyadic Gaming Experience. *Simulation and Gaming*, *45*(1), 24–40. https://doi.org/10.1177/1046878113513080

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

John O. Ledyard, Public goods: a survey of experimental research, in: John H. Kagel, Alvin E. Roth (Eds.), Handbook of Experimental Economics, Princeton University Press, Princeton, NJ 1995, pp. 111–194.

Jones, C. L., Minati, L., Nagai, Y., Medford, N., Harrison, N. A., Gray, M., … Critchley, H. D. (2015). Neuroanatomical substrates for the volitional regulation of heart rate. *Frontiers in Psychology*, *6*(MAR), 1–14. https://doi.org/10.3389/fpsyg.2015.00300

Keenan, J. P., Ganis, G., Freund, S., & Pascual-Leone, A. (2000). Self-face identification is increased with left hand responses. *Laterality: Asymmetries of Body, Brain and Cognition*, *5*(3), 259–268. https://doi.org/10.1080/713754382

Kemp, A. H., Quintana, D. S., Kuhnert, R. L., Griffiths, K., Hickie, I. B., & Guastella, A. J. (2012). Oxytocin Increases Heart Rate Variability in Humans at Rest: Implications for Social Approach-Related Motivation and Capacity for Social Engagement. *PLoS ONE*, *7*(8), e44014. https://doi.org/10.1371/journal.pone.0044014

Kern, M., Aertsen, A., Schulze-Bonhage, A., & Ball, T. (2013). Heart cycle-related effects on event-related potentials, spectral power changes, and connectivity patterns in the human ECoG. *NeuroImage*, *81*, 178–190. https://doi.org/10.1016/j.neuroimage.2013.05.042

Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. Personality and Social Psychology Review, 2(3), 196-217.

Khalsa, S. S., Adolphs, R., Cameron, O. G., Critchley, H. D., Davenport, P. W., Feinstein, J. S., … Paulus, M. P. (2017). Interoception and Mental Health: a Roadmap. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *0*(0). https://doi.org/10.1016/j.bpsc.2017.12.004

Khalsa, S. S., Rudrauf, D., Damasio, A. R., Davidson, R. J., Lutz, A., & Tranel, D. (2008). Interoceptive awareness in experienced meditators. *Psychophysiology*, *45*(4), 671–677. https://doi.org/10.1111/j.1469-8986.2008.00666.x

Kim, H. G., Cheon, E. J., Bai, D. S., Lee, Y. H., & Koo, B. H. (2018, March). Stress and heart rate variability: A meta-analysis and review of the literature. Korean Neuropsychiatric Association. https://doi.org/10.30773/pi.2017.08.17

Kleiger, R. E., Stein, D. S., & Bigger, M. D. (2005). Heart rate variability: measurement and clinical utility. Ann. Noninvasive Electrocardiol. 10, 88–101. https://doi.org/10.1111/j.1542-474X.2005.10101.x

Knoblich, G., & Flach, R. (2003). Action identity: Evidence from self-recognition, prediction, and coordination. In *Consciousness and cognition* (Vol. 12, pp. 620–632). Academic Press Inc. https://doi.org/10.1016/S1053-8100(03)00070-9

Konvalinka, I., Xygalatas, D., Bulbulia, J., Schjødt, U., Jegindø, E.-M., Wallot, S., … Roepstorff, A. (2011). Synchronized arousal between performers and related spectators in a fire-walking ritual. *Proceedings of the National Academy of Sciences of the United States of America*, *10*(20), 8514–8519. https://doi.org/10.1073/pnas.1016955108

Kooperberg, C. (2018). *logspline: Routines for Logspline Density Estimation*. Retrieved from https://cran.r-project.org/package=logspline

Kuikkaniemi, K., Laitinen, T., Turpeinen, M., Saari, T., Kosunen, I., Ravaja, N., 2010. The influence of implicit and explicit biofeedback in first-person shooter games. CHI ’10 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Brains and Brawn. ACM, Atlanta, GA, USA, pp. 859–868.

Laborde, S., Mosley, E., & Thayer, J. F. (2017). Heart rate variability and cardiac vagal tone in psychophysiological research - Recommendations for experiment planning, data analysis, and data reporting. Frontiers. https://doi.org/10.3389/fpsyg.2017.00213

Laborde, S., Raab, M., & Kinrade, N. P. (2014). Is the ability to keep your mind sharp under pressure reflected in your heart? Evidence for the neurophysiological bases of decision reinvestment. Biol. Psychol. 100, 34–42. https://doi.org/10.1016/j.biopsycho.2014.05.003

Lane, R. D., McRae, K., Reiman, E. M., Chen, K., Ahern, G. L., & Thayer, J. F. (2009). Neural correlates of heart rate variability during emotion. *NeuroImage*, *44*(1), 213–222. https://doi.org/10.1016/j.neuroimage.2008.07.056

Langley JN (1921). The Autonomic Nervous System. W. Heffer and Sons, Cambridge, England.

Larsen, P. D., Tzeng, Y. C., Sin, P. Y., & Galletly, D. C. (2010). Respiratory sinus arrhythmia in conscious humans during spontaneous respiration. Respir. Physiol. Neurobiol. 174, 111–118. https://doi.org/10.1016/j.resp.2010. 04.021

Lehrer, P. (2013). How Does Heart Rate Variability Biofeedback Work? Resonance, the Baroreflex, and Other Mechanisms. *Biofeedback*, *41*(1), 26–31. https://doi.org/10.5298/1081-5937-41.1.02

Lehrer, P. M., & Gevirtz, R. (2014). Heart rate variability biofeedback: How and why does it work? *Frontiers in Psychology*, *5*(JUL), 1–9. https://doi.org/10.3389/fpsyg.2014.00756

Levenson, R. W., Gottman, J. M., (1983). Marital Interaction: Physiological Linkage and Affective Exchange. *Journal of Personality and Social Psychology*, *45*(3), 587–597. https://doi.org/10.1037/0022-3514.45.3.587

Lewis, G. F., Furman, S. A., McCool, M. F., & Porges, S. W. (2012). Statistical strategies to quantify respiratory sinus arrhythmia: Are commonly used metrics equivalent? Biol. Psychol. 89, 349–364. https://doi.org/10.1016/j.biopsycho.2011. 11.009

Lewis, G. F., Furman, S. A., McCool, M. F., & Porges, S. W. (2012). Statistical strategies to quantify respiratory sinus arrhythmia: Are commonly used metrics equivalent? *Biological Psychology*, *89*(2), 349–364. https://doi.org/10.1016/j.biopsycho.2011.11.009

Lewis, M. J., Kingsley, M., Short, A. L., & Simpson, K. (2007). Rate of reduction of heart rate variability during exercise as an index of physical work capacity. Scand. J. Med. Sci. Sports 17, 696–702. https://doi.org/10.1111/j.1600-0838.2006. 00616.x

Ley, I., Haggard, P., & Yarrow, K. (2009). Optimal integration of auditory and vibrotactile information for judgments of temporal order. Journal of Experimental Psychology: Human Perception and Performance, 35(4), 1005.

Locke, S. E. (1980). Biofeedback and the Modification of Behavior. *American Journal of Psychiatry*, *137*(10), 1290–1291. https://doi.org/10.1176/ajp.137.10.1290

Löffler, A., Foell, J., & Bekrater-Bodmann, R. (2018). Interoception and Its Interaction with Self, Other, and Emotion Processing: Implications for the Understanding of Psychosocial Deficits in Borderline Personality Disorder. Current Medicine Group LLC 1. https://doi.org/10.1007/s11920-018-0890-2

Loula, F., Prasad, S., Harber, K., & Shiffrar, M. (2005). Recognizing people from their movement. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(1), 210–220. https://doi.org/10.1037/0096-1523.31.1.210

Lüdecke, D. (2018a). sjmisc: Data and Variable Transformation Functions. *Journal of Open Source Software*, *3*(26), 754. https://doi.org/10.21105/joss.00754

Lüdecke, D. (2018b). *sjPlot: Data Visualization for Statistics in Social Science*. https://doi.org/10.5281/zenodo.1308157

Luft, C. D. B., & Bhattacharya, J. (2015). Aroused with heart: Modulation of heartbeat evoked potential by arousal induction and its oscillatory correlates. *Scientific Reports*, *5*(1), 15717. https://doi.org/10.1038/srep15717

Macmillan, N. A., & Creelman, C. D. (2004). *Detection Theory: A User's Guide: 2nd edition* (pp. 1–445). Psychology Press. https://doi.org/10.4324/9781410611147

Maister, L., & Tsakiris, M. (2014). My face, my heart: Cultural differences in integrated bodily self-awareness. *Cognitive Neuroscience*, *5*(1), 10–16. https://doi.org/10.1080/17588928.2013.808613

Maister, L., Hodossy, L., & Tsakiris, M. (2017). You Fill My Heart: Looking at One's Partner Increases Interoceptive Accuracy. *Psychology of Consciousness (Washington, D.C.)*, *4*(2), 248–257. https://doi.org/10.1037/cns0000110

Malik, M. (1996). Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *European Heart Journal*, *17*(3), 354–381. https://doi.org/10.1111/j.1542-474X.1996.tb00275.x

Manini, B., Cardone, D., Ebisch, S. J., Bafunno, D., Aureli, T., & Merla, A. (2013). Mom feels what her child feels: Thermal signatures of vicarious autonomic response while watching children in a stressful situation. *Frontiers in Human Neuroscience*, *7*(JUN). https://doi.org/10.3389/fnhum.2013.00299

Marci, C. D., Ham, J., Moran, E., & Orr, S. P. (2007). Physiologic correlates of perceived therapist empathy and social-emotional process during psychotherapy. *Journal of Nervous and Mental Disease*, *195*(2), 103–111. https://doi.org/10.1097/01.nmd.0000253731.71025.fc

Marcovitch, S., Leigh, J., Calkins, S. D., Leerks, E. M., O’Brien, M., & Blankson, A. N. (2010). Moderate vagal withdrawal in 3.5-year-old children is associated with optimal performance on executive function tasks. Dev. Psychobiol. 52, 603–608. https://doi.org/10.1002/dev.20462

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Marwan, N., Carmen Romano, M., Thiel, M., & Kurths, J. (2007, January). Recurrence plots for the analysis of complex systems. https://doi.org/10.1016/j.physrep.2006.11.001

Marwan, N., Wessel, N., Meyerfeldt, U., Schirdewan, A., & Kurths, J. (2002). Recurrence-plot-based measures of complexity and their application to heart-rate-variability data. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, *66*(2). https://doi.org/10.1103/PhysRevE.66.026702

McAssey, M. P., Helm, J., Hsieh, F., Sbarra, D. A., & Ferrer, E. (2013). Methodological advances for detecting physiological synchrony during dyadic interactions. *Methodology*, *9*(2), 41–53. https://doi.org/10.1027/1614-2241/a000053

McCraty, R., & Shaffer, F. (2015). Heart rate variability: New perspectives on physiological mechanisms, assessment of self-regulatory capacity, and health risk. https://doi.org/10.7453/gahmj.2014.073

Messerotti Benvenuti, S., Mennella, R., Buodo, G., & Palomba, D. (2015). Dysphoria is associated with reduced cardiac vagal withdrawal during the imagery of pleasant scripts: evidence for the positive attenuation hypothesis. Biol. Psychol. 106, 28–38. https://doi.org/10.1016/j.biopsycho.2014.11.017

Mickes, L. (2015). Receiver operating characteristic analysis and confidence-accuracy characteristic analysis in investigations of system variables and estimator variables that affect eyewitness memory. *Journal of Applied Research in Memory and Cognition*, *4*(2), 93–102. https://doi.org/10.1016/j.jarmac.2015.01.003

Mitkidis, P., McGraw, J. J., Roepstorff, A., & Wallot, S. (2015). Building trust: Heart rate synchrony and arousal during joint action increased by public goods game. *Physiology and Behavior*, *149*, 101–106. https://doi.org/10.1016/j.physbeh.2015.05.033

Mohr, E., Langbein, J., & Nürnberg, G. (2002). Heart rate variability: a noninvasive approach to measure stress in calves and cows. Physiology & Behavior, 75(1-2), 251-259.

Montague, E., Xu, J., & Chiou, E. (2014). Shared experiences of technology and trust: An experimental study of physiological compliance between active and passive users in technology-mediated collaborative encounters. *IEEE Transactions on Human-Machine Systems*, *44*(5), 614–624. https://doi.org/10.1109/THMS.2014.2325859

Müller, B. C., Kühn, S., Van Baaren, R. B., Dotsch, R., Brass, M., & Dijksterhuis, A. (2011). Perspective taking eliminates differences in co-representation of out-group members' actions. *Experimental Brain Research*, *211*(3-4), 423–428. https://doi.org/10.1007/s00221-011-2654-7

Munoz, M. L., van Roon, A., Riese, H., Thio, C., Oostenbroek, E., Westrik, I., et al. (2015). Validity of (ultra-)short recordings for heart rate variability measurements. PLoS ONE 10:e0138921. https://doi.org/10.1371/journal.pone. 0138921

Nacke, L.E., Kalyn, M., Lough, C., Mandryk, R.L., 2011. Biofeedback Game Design: Using Direct and Indirect Physiological Control to Enhance Game Interaction. Proceedings of the CHI 2011 Conference on Human Factors in Computing Systems. ACM, Vancouver, BC, Canada, pp. 103–112.

Neumann, S. A., Sollers, J. J., Thayer, J. F., & Waldstein, S. R. (2004). Alexithymia predicts attenuated autonomic reactivity, but prolonged recovery to anger recall in young women. Int. J. Psychophysiol. 53, 183–195.

Niemelä, M. J., Airaksinen, K. E., & Huikuri, H. V. (1994). Effect of Beta-blockade on heart rate variability in patients with coronary artery disease. *Journal of the American College of Cardiology*, *23*(6), 1370–1377. https://doi.org/10.1016/0735-1097(94)90379-4

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(11), 2600–2606. https://doi.org/10.1073/pnas.1708274114

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

Otzenberger, H., Gronfier, C., Simon, C., Charloux, A., Ehrhart, J., Piquard, F., & Brandenberger, G. (1998). Dynamic heart rate variability: A tool for exploring sympathovagal balance continuously during sleep in men. *American Journal of Physiology - Heart and Circulatory Physiology*, *275*(3 44-3). https://doi.org/10.1152/ajpheart.1998.275.3.h946

Palmer, C. E., & Tsakiris, M. (2018). Going at the heart of social cognition: is there a role for interoception in self-other distinction? *Current Opinion in Psychology*, *24*, 21–26. https://doi.org/10.1016/j.copsyc.2018.04.008

Palumbo, R. V., Marraccini, M. E., Weyandt, L. L., Wilder-Smith, O., McGee, H. A., Liu, S., & Goodwin, M. S. (2017). Interpersonal Autonomic Physiology: A Systematic Review of the Literature. *Personality and Social Psychology Review*, *21*(2), 99–141. https://doi.org/10.1177/1088868316628405

Park, G., Vasey, M. W., Van Bavel, J. J., & Thayer, J. F. (2014). When tonic cardiac vagal tone predicts changes in phasic vagal tone: the role of fear and perceptual load. Psychophysiology 51, 419–426. https://doi.org/10.1111/psyp.12186

Park, H. D., Correia, S., Ducorps, A., & Tallon-Baudry, C. (2014). Spontaneous fluctuations in neural responses to heartbeats predict visual detection. *Nature Neuroscience*, *17*(4), 612–618. https://doi.org/10.1038/nn.3671

Paulus, M. P., Feinstein, J. S., & Khalsa, S. S. (2019). An Active Inference Approach to Interoceptive Psychopathology. *Annual Review of Clinical Psychology*, *15*, 97–122. https://doi.org/https://doi/10.1146/annurev-clinpsy-050718-095617

Peira, N., Fredrikson, M., & Pourtois, G. (2014). Controlling the emotional heart: Heart rate biofeedback improves cardiac control during emotional reactions. *International Journal of Psychophysiology*, *91*(3), 225–231. https://doi.org/10.1016/j.ijpsycho.2013.12.008

Peira, N., Pourtois, G., & Fredrikson, M. (2013). Learned Cardiac Control with Heart Rate Biofeedback Transfers to Emotional Reactions. *PLoS ONE*, *8*(7). https://doi.org/10.1371/journal.pone.0070004

Peschel, S. K., Feeling, N. R., Vogele, C., Kaess, M., Thayer, J. F., & Koenig, J. (2016). A meta-analysis on resting state high-frequency heart rate variability in bulimia nervosa. Eur. Eat. Disord. Rev. 24, 355–365. https://doi.org/10.1002/erv. 2454

Petrusic, W. M., & Baranski, J. V. (1997). Context, feedback, and the calibration and resolution of confidence in perceptual judgments. *American Journal of Psychology*, *110*(4), 543–572. https://doi.org/10.2307/1423410

Petzschner, F. H., Weber, L. A., Gard, T., & Stephan, K. E. (2017, September). Computational Psychosomatics and Computational Psychiatry: Toward a Joint Framework for Differential Diagnosis. Elsevier. https://doi.org/10.1016/j.biopsych.2017.05.012

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

Pezzulo, G. (2014). Why do you fear the bogeyman? An embodied predictive coding model of perceptual inference. *Cognitive, Affective and Behavioral Neuroscience*, *14*(3), 902–911. https://doi.org/10.3758/s13415-013-0227-x

Pezzulo, G., Rigoli, F., & Friston, K. (2015). Active Inference, homeostatic regulation and adaptive behavioural control. *Progress in Neurobiology*, *134*, 17–35. https://doi.org/10.1016/j.pneurobio.2015.09.001

Pfeiffer, C., & De Lucia, M. (2017). Cardio-audio synchronization drives neural surprise response. *Scientific Reports*, *7*(1), 14842. https://doi.org/10.1038/s41598-017-13861-8

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., Matthias, E., & Schandry, R. (2007). Heartbeat perception and P300 amplitude in a visual oddball paradigm. *Clinical Neurophysiology*, *118*(10), 2248–2253. https://doi.org/10.1016/j.clinph.2007.06.057

Porges, S. W. (2001). The polyvagal theory: phylogenetic substrates of a social nervous system. International journal of psychophysiology, 42(2), 123-146. https://doi.org/10.1111/j.1469-7610.2006.01701.x

Porges, S. W. (2007a). A phylogenetic journey through the vague and ambiguous Xth cranial nerve: A commentary on contemporary heart rate variability research. *Biological Psychology*, *74*(2), 301–307. https://doi.org/10.1016/j.biopsycho.2006.08.007

Porges, S. W. (2007b). The polyvagal perspective. *Biological Psychology*, *74*(2), 116–143. https://doi.org/10.1016/j.biopsycho.2006.06.009

Rogers, L. J., Vallortigara, G., & Andrew, R. J. (2013). Divided brains: the biology and behaviour of brain asymmetries. Cambridge University Press.

Quintana, D. S., & Heathers, J. A. (2014). Considerations in the assessment of heart rate variability in biobehavioral research. Frontiers Research Foundation. https://doi.org/10.3389/fpsyg.2014.00805

Quintana, D. S., Alvares, G. A., & Heathers, J. A. (2016). Guidelines for reporting articles on psychiatry and heart rate variability (GRAPH): recommendations to advance research communication. Transl. Psychiatry 6:e803. https://doi.org/10.1038/tp. 2016.73

R Core Team. (2018). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/

Rainville, P., Bechara, A., Naqvi, N., & Damasio, A. R. (2006). Basic emotions are associated with distinct patterns of cardiorespiratory activity. *International Journal of Psychophysiology*, *61*(1), 5–18. https://doi.org/10.1016/j.ijpsycho.2005.10.024

Reeck, C., Ames, D. R., & Ochsner, K. N. (2016). Special Focus on Emotion The Social Regulation of Emotion: An Integrative, Cross-Disciplinary Model. https://doi.org/10.1016/j.tics.2015.09.003

Reed, R. G., Randall, A. K., Post, J. H., & Butler, E. A. (2013). Partner influence and in-phase versus anti-phase physiological linkage in romantic couples. *International Journal of Psychophysiology*, *88*(3), 309–316. https://doi.org/10.1016/j.ijpsycho.2012.08.009

Repp, B. H., & Knoblich, G. (2004). Perceiving action identity: How pianists recognize their own performances. *Psychological Science*, *15*(9), 604–609. https://doi.org/10.1111/j.0956-7976.2004.00727.x

Reyes del Paso, G. A., Langewitz, W., Mulder, L. J., Roon, A. van, & Duschek, S. (2013). The utility of low frequency heart rate variability as an index of sympathetic cardiac tone: A review with emphasis on a reanalysis of previous studies. *Psychophysiology*, *50*(5), 477–487. https://doi.org/10.1111/psyp.12027

Richardson, D. C., Street, C. N., Tan, J. Y., Kirkham, N. Z., Hoover, M. A., & Cavanaugh, A. G. (2012). Joint perception: Gaze and social context. *Frontiers in Human Neuroscience*, (JULY). https://doi.org/10.3389/fnhum.2012.00194

Rimé, B. (2007). Interpersonal emotion regulation. In J. J. Gross (Ed.), Handbook ofemotion regulation (pp. 466–485). New York: Guilford.

Rimé, B. (2009). Emotion elicits the social sharing of emotion: Theory and empirical review. Emotion review, 1(1), 60-85.

Rottenberg, J., Salomon, K., Gross, J. J., & Gotlib, I. H. (2005). Vagal withdrawal to a sad film predicts subsequent recovery from depression. Psychophysiology 42, 277–281. https://doi.org/10.1111/j.1469-8986.2005.00289.x

Royzman, E. B., Cassidy, K. W., & Baron, J. (2003). “I Know, You Know”: Epistemic Egocentrism in Children and Adults. https://doi.org/10.1037/1089-2680.7.1.38

Ruissen, M. I., & De Bruijn, E. R. (2016). Competitive game play attenuates self-other integration during joint task performance. *Frontiers in Psychology*, *7*(MAR). https://doi.org/10.3389/fpsyg.2016.00274

Ruys, K. I., & Aarts, H. (2010). When competition merges people’s behavior: Interdependency activates shared action representations. *Journal of Experimental Social Psychology*, *46*(6), 1130–1133. https://doi.org/10.1016/j.jesp.2010.05.016

Salomon, R., Ronchi, R., Dönz, J., Bello-Ruiz, J., Herbelin, B., Martet, R., … Blanke, O. (2016). The Insula Mediates Access to Awareness of Visual Stimuli Presented Synchronously to the Heartbeat. *The Journal of Neuroscience*, *36*(18), 5115–5127. https://doi.org/10.1523/JNEUROSCI.4262-15.2016

Sandercock, G. (2007). Normative values, reliability and sample size estimates in heart rate variability. Clin. Sci. 113, 129–130. https://doi.org/10.1042/CS20070137

Schandry, R. (1981). Heart Beat Perception and Emotional Experience. *Psychophysiology*, *18*(4), 483–488. https://doi.org/10.1111/j.1469-8986.1981.tb02486.x

Schandry, R., & Weitkunat, R. (1990). Enhancement of heartbeat-related brain potentials through cardiac awareness training. *International Journal of Neuroscience*, *53*(2-4), 243–253. https://doi.org/10.3109/00207459008986611

Schauder, K. B., Mash, L. E., Bryant, L. K., & Cascio, C. J. (2015). Interoceptive ability and body awareness in autism spectrum disorder. *Journal of Experimental Child Psychology*, *131*, 193–200. https://doi.org/10.1016/j.jecp.2014.11.002

Schuch, S., & Tipper, S. P. (2007). On observing another person ' s actions : *Perception & Psychophysics*, *69*(5), 828–837.

Schulz, A., Ferreira de Sá, D. S., Dierolf, A. M., Lutz, A., Dyck, Z. van, Vögele, C., & Schächinger, H. (2015). Short-term food deprivation increases amplitudes of heartbeat-evoked potentials. *Psychophysiology*, *52*(5), 695–703. https://doi.org/10.1111/psyp.12388

Schupp, H. T., Cuthbert, B. N., Bradley, M. M., Hillman, C. H., Hamm, A. O., & Lang, P. J. (2004). Brain processes in emotional perception: Motivated attention. *Cognition and Emotion*, *18*(5), 593–611. https://doi.org/10.1080/02699930341000239

Schwarz, A. M., Schächinger, H., Adler, R. H., & Goetz, S. M. (2003). Hopelessness is associated with decreased heart rate variability during championship chess games. *Psychosomatic Medicine*, *65*(4), 658–661. https://doi.org/10.1097/01.PSY.0000075975.90979.2A

Sebanz, N., Bekkering, H., & Knoblich, G. (2006, February). Joint action: Bodies and minds moving together. https://doi.org/10.1016/j.tics.2005.12.009

Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: Just like one's own? *Cognition*, *88*(3). https://doi.org/10.1016/S0010-0277(03)00043-X

Sebanz, N., Knoblich, G., & Prinz, W. (2005). How two share a task: Corepresenting stimulus-response mappings. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(6), 1234–1246. https://doi.org/10.1037/0096-1523.31.6.1234

Sel, A., Azevedo, R. T., & Tsakiris, M. (2017). Heartfelt Self: Cardio-Visual Integration Affects Self-Face Recognition and Interoceptive Cortical Processing. *Cerebral Cortex*, *27*(11), 5144–5155. https://doi.org/10.1093/cercor/bhw296

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., & Tsakiris, M. (2018). Being a Beast Machine: The Somatic Basis of Selfhood. *Trends in Cognitive Sciences*, *22*(11), 1–30. https://doi.org/10.1016/j.tics.2018.08.008

Seth, A. K., Suzuki, K., & Critchley, H. D. (2012). An interoceptive predictive coding model of conscious presence. *Frontiers in Psychology*, *3*(JAN), 395. https://doi.org/10.3389/fpsyg.2011.00395

Sforza, A., Bufalari, I., Haggard, P., & Aglioti, S. M. (2010). My face in yours: Visuo-tactile facial stimulation influences sense of identity. *Social Neuroscience*, *5*(2), 148–162. https://doi.org/10.1080/17470910903205503

Shaffer, F., & Combatalade, D. C. (2013). Don’t add or miss a beat: a guide to cleaner heart rate variability recordings. Biofeedback 41, 121–130. https://doi.org/10.5298/1081-5937-41.3.04

Shahrestani, S., Stewart, E. M., Quintana, D. S., Hickie, I. B., & Guastella, A. J. (2015). Heart rate variability during adolescent and adult social interactions: A meta-analysis. *Biological Psychology*, *105*, 43–50. https://doi.org/10.1016/j.biopsycho.2014.12.012

Shockley, K., Santana, M. V., & Fowler, C. A. (2003). Mutual Interpersonal Postural Constraints are Involved in Cooperative Conversation. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(2), 326–332. https://doi.org/10.1037/0096-1523.29.2.326

Simon, J. R. (1969). Reactions toward the source of stimulation. *Journal of Experimental Psychology*, *81*(1), 174–176. https://doi.org/10.1037/h0027448

Smith, T. W., Cribbet, M. R., Nealey-Moore, J. B., Uchino, B. N., Williams, P. G., MacKenzie, J., & Thayer, J. F. (2011). Matters of the Variable Heart: Respiratory Sinus Arrhythmia Response to Marital Interaction and Associations With Marital Quality. *Journal of Personality and Social Psychology*, *100*(1), 103–119. https://doi.org/10.1037/a0021136

Stenzel, A., Dolk, T., Colzato, L. S., Sellaro, R., Hommel, B., & Liepelt, R. (2014). The joint Simon effect depends on perceived agency, but not intentionality, of the alternative action. *Frontiers in Human Neuroscience*, *8*(AUG). https://doi.org/10.3389/fnhum.2014.00595

Stephan, K. E., Binder, E. B., Breakspear, M., Dayan, P., Johnstone, E. C., Meyer-Lindenberg, A., … Friston, K. J. (2016a). Charting the landscape of priority problems in psychiatry, part 2: Pathogenesis and aetiology. *The Lancet Psychiatry*, *3*(1), 84–90. https://doi.org/10.1016/S2215-0366(15)00360-0

Stephan, K. E., Manjaly, Z. M., Mathys, C. D., Weber, L. A. E., Paliwal, S., Gard, T., … Petzschner, F. H. (2016b). Allostatic Self-efficacy: A Metacognitive Theory of Dyshomeostasis-Induced Fatigue and Depression. *Frontiers in Human Neuroscience*, *10*, 550. https://doi.org/10.3389/fnhum.2016.00550

Sterling, P. (2012). Allostasis: A model of predictive regulation. *Physiology and Behavior*, *106*(1), 5–15. https://doi.org/10.1016/j.physbeh.2011.06.004

Sterling, P. (2014). Homeostasis vs allostasis implications for brain function and mental disorders. *JAMA Psychiatry*, *71*(10), 1192–1193. https://doi.org/10.1001/jamapsychiatry.2014.1043

Suess, P. E., Porges, S. W., & Plude, D. J. (1994). Cardiac vagal tone and sustained attention in school-age children. *Psychophysiology*, *31*(1), 17–22. https://doi.org/10.1111/j.1469-8986.1994.tb01020.x

Suveg, C., Shaffer, A., & Davis, M. (2016). Family stress moderates relations between physiological and behavioral synchrony and child self-regulation in mother-preschooler dyads. *Developmental Psychobiology*, *58*(1), 83–97. https://doi.org/10.1002/dev.21358

Suzuki, K., Garfinkel, S. N., Critchley, H. D., & Seth, A. K. (2013). Multisensory integration across exteroceptive and interoceptive domains modulates self-experience in the rubber-hand illusion. *Neuropsychologia*, *51*(13), 2909–2917. https://doi.org/10.1016/j.neuropsychologia.2013.08.014

Tajadura-Jiménez, A., & Tsakiris, M. (2014). Balancing the “inner” and the “outer” self: Interoceptive sensitivity modulates self-other boundaries. *Journal of Experimental Psychology: General*, *143*(2), 736–744. https://doi.org/10.1037/a0033171

Tajadura-Jiménez, A., Longo, M. R., Coleman, R., & Tsakiris, M. (2012). The person in the mirror: Using the enfacement illusion to investigate the experiential structure of self-identification. *Consciousness and Cognition*, *21*(4), 1725–1738. https://doi.org/10.1016/j.concog.2012.10.004

Takens, F. (1981). Detecting strange attractors in turbulence. In (pp. 366–381). https://doi.org/10.1007/bfb0091924

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

Terhaar, J., Viola, F. C., Bär, K. J., & Debener, S. (2012). Heartbeat evoked potentials mirror altered body perception in depressed patients. *Clinical Neurophysiology*, *123*(10), 1950–1957. https://doi.org/10.1016/j.clinph.2012.02.086

Thayer, J. F., Ahs, F., Fredrikson, M., Sollers, J. J., & Wager, T. D. (2012). A meta-analysis of heart rate variability and neuroimaging studies: Implications for heart rate variability as a marker of stress and health. *Neuroscience and Biobehavioral Reviews*, *36*(2), 747–756. https://doi.org/10.1016/j.neubiorev.2011.11.009

Thayer, J. F., Hansen, A. L., Saus-Rose, E., & Johnsen, B. H. (2009). Heart rate variability, prefrontal neural function, and cognitive performance: The neurovisceral integration perspective on self-regulation, adaptation, and health. https://doi.org/10.1007/s12160-009-9101-z

Thayer, J. F., Loerbroks, A., & Sternberg, E. M. (2011). Inflammation and cardiorespiratory control: the role ofthe vagus nerve. Respir. Physiol. Neurobiol. 178, 387–394. https://doi.org/10.1016/j.resp.2011.05.016

Tsakiris, M. (2008). Looking for myself: Current multisensory input alters self-face recognition. *PLoS ONE*, *3*(12), 4040. https://doi.org/10.1371/journal.pone.0004040

Tsakiris, M. (2010). My body in the brain: A neurocognitive model of body-ownership. *Neuropsychologia*, *48*(3), 703–712. https://doi.org/10.1016/j.neuropsychologia.2009.09.034

Tsakiris, M., & Haggard, P. (2005). Experimenting with the acting self. *Cognitive Neuropsychology*, *22*(3), 387–407. https://doi.org/10.1080/02643290442000158

Tsakiris, M., & Haggard, P. (2005b). The rubber hand illusion revisited: visuotactile integration and self-attribution. Journal of Experimental Psychology: Human Perception and Performance, 31(1), 80.

Tsakiris, M., Tajadura-Jiménez, A., & Costantini, M. (2011). Just a heartbeat away from one's body:Interoceptive sensitivity predicts malleability of body-representations.

Proceedings of the Royal Society B: Biological Sciences, *278*(1717), 2470–2476. https://doi.org/10.1098/rspb.2010.2547

Tsakiris, M., Tajadura-Jiménez, A., & Costantini, M. (2011). Just a heartbeat away from one’s body:Interoceptive sensitivity predicts malleability of body-representations. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1717), 2470–2476. https://doi.org/10.1098/rspb.2010.2547

Van Boeijen, I., & Saraiva, R. (2018). egalPsych: A tool for calculating calibration statistics in eyewitness research. Retrieved from https://github.com/IngerMathilde/legalPsych

Van der Wel, R. P., Sebanz, N., & Knoblich, G. (2015). A joint action perspective on embodiment. In *Conceptual and interactive embodiment: Foundations of embodied cognition* (Vol. 2, pp. 165–181). https://doi.org/10.4324/9781315751962

Veldhuijzen Van Zanten, J. J., De Boer, D., Harrison, L. K., Ring, C., Carroll, D., Willemsen, G., & De Geus, E. J. (2002). Competitiveness and hemodynamic reactions to competition. *Psychophysiology*, *39*(6), 759–766. https://doi.org/10.1111/1469-8986.3960759

Vilensky, J. A., Robertson, W., & Suarez-Quian, C. A. (2015). The Clinical Anatomy of the Cranial Nerves: The Nerves of" On Old Olympus Towering Top". John Wiley & Sons.

Voeten, C. (2019). buildmer: Stepwise Elimination and Term Reordering for Mixed-Effects Regression. Retrieved from https://github.com/cvoeten/buildmer/issues

Von Mohr, M., Finotti, G., Villani, V., & Tsakiris, M. (2019). Taking the pulse of social cognition: cardiac afferent activity and interoceptive accuracy modulate emotional egocentricity bias.

Vossel, S., Mathys, C., Daunizeau, J., Bauer, M., Driver, J., Friston, K. J., & Stephan, K. E. (2014). Spatial Attention, Precision, and Bayesian Inference: A Study of Saccadic Response Speed. *Cerebral Cortex*, *24*(6), 1436–1450. https://doi.org/10.1093/cercor/bhs418

Wallot, S. (2017). Recurrence Quantification Analysis of Processes and Products of Discourse: A Tutorial in R. *Discourse Processes, 00*(00), 1–24. https://doi.org/10.1080/0163853X.2017.1297921

Wallot, S., & Leonardi, G. (2018). Analyzing multivariate dynamics using cross-recurrence quantification analysis (CRQA), diagonal-cross-recurrence profiles (DCRP), and multidimensional recurrence quantification analysis (MdRQA) - A tutorial in R. *Frontiers in Psychology*, *9*(DEC), 2232. https://doi.org/10.3389/fpsyg.2018.02232

Wallot, S., Fusaroli, R., Tylén, K., & Jegindø, E.-M. (2013). Using complexity metrics with R-R intervals and BPM heart rate measures. *Frontiers in Physiology*, *4*, 211. https://doi.org/10.3389/fphys.2013.00211

Waters, S. F., West, T. V., & Mendes, W. B. (2014). Stress Contagion: Physiological Covariation Between Mothers and Infants. *Psychological Science*, *25*(4), 934–942. https://doi.org/10.1177/0956797613518352

Weisz, J., Bálazs, L., & ÁDám, G. (1988). The Influence of Self-Focused Attention on Heartbeat Perception. *Psychophysiology*, *25*(2), 193–199. https://doi.org/10.1111/j.1469-8986.1988.tb00987.x

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. Psychological bulletin, 88(3), 638.

Welsh, T. N., & McDougall, L. M. (2012). Negative priming in a joint selection task. *PLoS ONE*, *7*(8), e42963. https://doi.org/10.1371/journal.pone.0042963

Welsh, T. N., Elliott, D., Anson, J. G., Dhillon, V., Weeks, D. J., Lyons, J. L., & Chua, R. (2005). Does Joe influence Fred's action? Inhibition of return across different nervous systems. *Neuroscience Letters*, *385*(2), 99–104. https://doi.org/10.1016/j.neulet.2005.05.013

White, T. W., Holmes, D. S., & Bennett, D. H. (1977). Effects of instructions, biofeedback, and cognitive activities on heart rate control. *Journal of Experimental Psychology: Human Learning and Memory*, *3*(4), 477–484. https://doi.org/10.1037/0278-7393.3.4.477

Whitehead, W. E., Drescher, V. M., Heiman, P., & Blackwell, B. (1977). Relation of Heart Rate Control to Heartbeat Perception. *Biofeedback and Self-Regulation*, *2*(4). Retrieved from https://link.springer.com/content/pdf/10.1007{\%}2FBF00998623.pdf

Widen, S. C., & Russell, J. A. (2003). A closer look at preschoolers' freely produced labels for facial expressions. Developmental psychology, 39(1), 114.

Widen, S. C., & Russell, J. A. (2008). Children acquire emotion categories gradually. Cognitive development, 23(2), 291-312.

Williamson, D. A., & Blanchard, E. B. (1979a). Heart rate and blood pressure biofeedback - II. A review and integration of recent theoretical models. https://doi.org/10.1007/BF00998948

Williamson, D. A., & Blanchard, E. B. (1979b). Heart rate and blood pressure biofeedback - II. A review and integration of recent theoretical models. https://doi.org/10.1007/BF00998948

Wolf, W., Launay, J., & Dunbar, R. I. (2016). Joint attention, shared goals, and social bonding. *British Journal of Psychology*, *107*(2), 322–337. https://doi.org/10.1111/bjop.12144

Wolpert, D. M. (1997). in bilingual Computational approaches control. *Nature*, *1*(September).

Zamariola, G., Maurage, P., Luminet, O., & Corneille, O. (2018). Interoceptive accuracy scores from the heartbeat counting task are problematic: Evidence from simple bivariate correlations. *Biological Psychology*, *137*, 12–17. https://doi.org/10.1016/J.BIOPSYCHO.2018.06.006