Sound propagation and quantum-limited damping in an ultracold two-dimensional Fermi gas

Markus Bohlen

▶ **To cite this version:**

Markus Bohlen. Sound propagation and quantum-limited damping in an ultracold two-dimensional Fermi gas. Quantum Physics [quant-ph]. Université Paris sciences et lettres, 2020. English. NNT : 2020UPSLE068 . tel-03651235

HAL Id: tel-03651235
https://theses.hal.science/tel-03651235

Submitted on 25 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

PSL
UNIVERSITÉ PARIS

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée au Laboratoire de Neurosciences Cognitives et Computationnelles (LNC2), UMR 960, DEC, ENS, Inserm

Neural responses to heartbeats during subjective choices in humans

Soutenue par
Damiano Cristiano AZZALINI
Le 20 novembre 2020

Ecole doctorale n° 158
Cerveau, Cognition, Comportement

Spécialité
Sciences cognitives, neurosciences, psychologie

Composition du jury :

Mathias, PESSIGLIONE Dr., Université Pierre et Marie Curie	*Président*
Olaf, BLANKE Pr., Ecole Polytechnique Fédérale de Lausanne	*Rapporteur*
Benedetto, DE MARTINO Dr., University College London	*Rapporteur*
Frederike, PETZSCHNER Dr., ETH Zürich/Brown University	*Examinatrice*
Catherine, TALLON-BAUDRY Dr., Ecole Normale Supérieure – PSL University	*Directrice de thèse*

ENS | PSL

"*Can you feel my heart beat?*
Can you feel my heart beat?"

Nick Cave, Higgs Boson Blues

"Il faut être conscient pour choisir
et il faut choisir pour être conscient.
Choix et conscience sont une seule et même chose."

Jean-Paul Sartre, L'être et le néant, p. 506

Acknowledgements

I would like to start by thanking all the members of my jury for having found the time and the energy in such a stormy period.
Catherine, thank you for your committed supervision. You have taught me the strictest scientific rigor and the importance of the slightest detail. Thanks for being supportive when things did not follow the plans and comprehensive when I made some mistakes. After five years, I am still mesmerised by the way you question and understand the data.
Thanks to the ENP team: Laurent, Laure, Yvette and Laetitia, who made this PhD possible.
I would like to thank the MEG platform at the ICM that was fundamental to this research. Thank you, Nathalie, Denis, Laurent. Christophe, thank you for making the MEG recording sessions so amusing and for expanding my French vocabulary with all sort of idioms.
What would be the LNC without its unique roster of PIs? Etienne, Julie and Valentin. A special thanks goes to Stefano. I found in you a brilliant and curious mind with whom I could openly discuss about everything that matters.
Thanks to all the members of the LNC for the nice discussions about movies (Julie), economics (Basile), and confidence (Tarryn); for the photos and stories of exotic travels (Margaux), for exhibiting my ignorance in many domains with puzzling quiz questions (Sophie and Marine) and for providing over-detailed explications afterwards (Clémence). Thanks to the past members of the LNC, friends who welcomed me in the lab making it such a human and cheerful place to live in: Christina, Hannah, Emma, Marion and Marwa.
Thanks to the current and past members of my team: Janina, Juliette, Nicolai, Tahnée, Max, Ignacio, Stephen, Anne and Mariana. Thank you for your positive vibes (Ignacio), your merciless comments (Stephen), for your support, patience and your homemade figs (Anne). Thank you Tahnée for carefully reading the lengthy Introduction of my thesis.
Paris has been also the place where I met unique friends: Thibaud, Albert, Henri, Victor, Julien, Sarah and Simon. You made these five years so full of memories and good times. Rocco, you are somewhat apart. You taught me to play the guitar, pushed me higher on the climbing wall, and you have been such a present and remarkable friend.
The decision to undertake a PhD arose during my time in Rovereto, where I met a group of amazing persons that changed me profoundly: Evelyn, Marco, Eli, Ben, Jean and Vassiki. Scattered around the world, you made each continent become our new Jelly's.
This journey would have been much different without the friends that followed me from my hometown, Turin. Thank you, Francesco, Giorgio, Andrea and Claudio. Writing these lines made me realise how fortunate I am to have shared all these years with you.
My deepest gratitude is for Mariana. Thank you for giving me the courage of facing the hardest challenges, for giving me every day some reasons to smile. Thank you for the serenity and bliss that you brought into my life.
Finally, I would like to thank my parents, Gabriella and Fabio. I can only imagine what it feels like to have seen me going from the primary school of our neighbourhood to carrying out a PhD in Paris. Thanks for having given me the opportunities that led me here, for your tireless comprehension and trust. The person that I am is the result of your dedication.

I would like to dedicate this thesis to my grandmother, Nonna Angela, the strongest person that I have ever met.

Grazie.

Contents

ACKNOWLEDGEMENTS **5**

GENERAL INTRODUCTION **10**

CHAPTER I: PREFERENCE-BASED DECISIONS **12**

1. VALUE: DEFINITION AND MEASUREMENTS 12
   - *A. What is a preference?* *12*
   - *B. How to define values?* *13*
   - *C. How to measure value?* *14*
   - *Summary* *18*
2. HOW DO WE CHOOSE? COMPUTATIONAL MODELS OF DECISIONS 19
   - *A. The decision problem* *19*
   - *B. Argmax* *20*
   - *C. Softmax* *21*
   - *D. Signal detection theory* *21*
   - *E. Sequential Sampling Models* *22*
   - *F. The Neural Circuit Model* *24*
   - *Summary* *25*
3. THE NEURAL UNDERPINNINGS OF VALUE-BASED DECISIONS 27
   - *A. Value-related signals* *27*
   - *B. Brain regions encoding value* *29*
     1. Ventral Striatum 29
     2. Ventro-medial prefrontal cortex and posterior cingulate cortex 30
     3. Hippocampus 31
     4. Dorsal anterior cingulate cortex and anterior insula 32
   - *C. Temporal unfolding of the decision process: electrophysiological studies of value-based decisions* *34*
     1. Serial representation of option values 34
     2. Neural signatures of accumulation dynamics 35
   - *D. Which functional role for the multiple value-related regions?* *37*
   - *E. The debate around vmPFC* *38*
   - *Summary* *41*
   - *F. The neural mechanisms underlying choice variability* *42*
     1. Choice variability and random noise 42
     2. Contextual effects 43
     3. Choice-induced preference changes 46
     4. Internal factors 48
   - *Summary* *50*
   - *G. Conclusion* *51*

CHAPTER II: FROM EVALUATION TO SELF-REPRESENTATION. AN EMBODIED PERSPECTIVE OF SUBJECTIVE VALUE **52**

1. THE REPRESENTATION OF THE CURRENT STATE AS BUILDING-BLOCK OF EVALUATION 53
2. THE NEURAL REPRESENTATION OF CURRENT STATES 55
3. FROM HOMEOSTATIC TO COGNITIVE SELF-REPRESENTATIONS 56
   - *Summary* *57*
4. COGNITIVE COMPONENTS OF SELF-REPRESENTATION AND THEIR NEURAL UNDERPINNINGS 58
   - *A. Self-attribution* *58*
   - *B. Self-monitoring and self-beliefs* *59*
   - *C. Autobiographical memory, and future scenarios* *59*
   - *D. Spontaneous thoughts* *60*
   - *E. Self-representation in psychiatric disorders* *60*
   - *Summary* *64*

5 CONCLUSION 64

CHAPTER III: THE NEURAL MONITORING OF CARDIAC SIGNALS AS A GROUNDING MECHANISM FOR THE SELF **65**

1. ASCENDING BODILY SIGNALS 66

A. Changes in bodily states and affective experience *66*
B. Visceral signals: the heart and the gastro-intestinal tract *67*
C. Grounding the self in visceral signals *67*
D. Comparing different theoretical frameworks *68*
E. Implicit and explicit interoception *70*
Summary *72*
2. HEART-BRAIN COUPLING: PHYSIOLOGY AND NEURAL PATHWAYS 73
A. The heart and the cardiac cycle *73*
B. Ascending cardiac information *74*
C. Ascending neural pathways *75*
Summary *76*
3. NEURAL RESPONSES TO HEARTBEATS 78
A. HER: measures and confounds *78*
B. Cortical sources of HERs *81*
Summary *81*
C. Functional coupling between HERs and self-related cognition *83*
1. Inward oriented attention 83
2. Bodily self-awareness 84
3. Self-related thoughts 85
4. Affective experiences 87
5. Conscious perception 88
Summary *91*
4. CARDIAC CYCLE EFFECTS AND CARDIAC TIMINGS 92
A. Cardiac cycle effects *92*
Summary *93*
B. Cardiac synchronicity in multi-sensory integration *95*
Summary *96*
5. CONCLUSION AND RELEVANCE FOR MY PHD 97

CHAPTER IV: RESPONSES TO HEARTBEATS IN VENTROMEDIAL PREFRONTAL CORTEX CONTRIBUTE TO SUBJECTIVE PREFERENCE-BASED DECISIONS 99

Abstract *100*
Significance statement *101*
Introduction *102*
Materials and Methods *104*
Results *115*
Discussion *130*
Conflict of interest *133*
Acknowledgements *133*
References *134*

CHAPTER V: NEURAL RESPONSES TO HEARTBEATS PREDICT CHOICE CONSISTENCY BETWEEN INCOMMENSURABLE GOODS 139

Abstract *140*
Introduction *141*
Material and Methods *143*
Results *152*
References *163*
Supplementary materials *167*

CHAPTER VI: GENERAL DISCUSSION 169

1. SUMMARY OF THE EXPERIMENTAL RESULTS 169
2. COMPARING EXPERIMENTAL RESULTS 171
A. Cortical localisation *171*
1. Integration of HERs to local computations 171
2. The importance of experimental contrasts in HER investigation 172
B. Latencies of HER effect *174*
C. The relationship between HERs and choice consistency *174*
3. WHAT DO THE CURRENT RESULTS TELL US ABOUT HERS? 175
A. HERs reflect artefactual cardiac activity *175*
B. HERs differences reflect changes in cardiac inputs *175*

C. HERs as indexes of cortical states *176*
D. HERs are genuine neural responses to cardiac events *177*
4. ISOLATING SELF-REFLECTION EXPERIMENTALLY 177
5. HERS AND SELF-REFERENTIAL PROCESSING 179
A. HERs as the biological root of self-representation *179*
B. HERs as index of enhanced body-related representation *179*
C. HERs and self-relevant bodily aspects *181*
6. THE ROLE OF HERS IN PREFERENCE-BASED DECISIONS 182
7. OPEN QUESTIONS AND FURTHER DIRECTIONS 184
8. FINAL WORD 186

APPENDIX A: VISCERAL SIGNALS SHAPE BRAIN DYNAMICS AND COGNITION **187**

APPENDIX B: DOES STROKE VOLUME INFLUENCE HEARTBEAT EVOKED RESPONSES? **210**

Abstract *212*
Introduction *213*
Materials and Methods *215*
Results *221*
Discussion *225*
Conclusion *228*
References *229*
Figures *233*

REFERENCES **238**

General introduction

Do you prefer tea or coffee? **Preferences are inherently subjective**, my personal preference (coffee) is probably different from the one of the reader. In addition, **preferences are variable**: what I choose on one occasion may change in another context. The computational and neural mechanisms underpinning subjective value representation and comparison have been the object of the rapidly expanding field of neuroeconomics (*Chapter I*). However, the inherent subjective nature of evaluation, arising from the unique perspective of the decision-maker, has remained largely unaddressed.

The core idea of my PhD is that **the unique perspective from which an agent can ascribe values to the available options requires a form of self-reference**: only because I can represent a given context and some goals in relation to myself, I can evaluate the different options that are before me. Subjective value representation and self-referential processes may thus be functionally coupled, a hypothesis in line with experimental evidence showing the recruitment of common neural resources, for instance the ventro-medial prefrontal cortex (*Chapter II*).

How is self-reference implemented biologically? Various theoretical proposals argue that it may be grounded on the **neural monitoring of** bodily signals. Among these, ascending **visceral signals coming from the heart** and the stomach have been suggested to have a special status (*Chapter III*). Their continuous and autonomous activity, which is constantly monitored by the brain, provides a particularly relevant signal for tracking the integrity of the living organism as a whole. This evolutionary ancient form of neural self-monitoring **may thus underlie more abstract and cognitively complex forms of self-representation**. Indeed, a growing body of experimental evidence, which I had the opportunity to review (Azzalini et al., 2019b; cf. *Appendix A*), suggests that the neural monitoring of visceral signals, and neural responses to heartbeats more specifically, are linked to various cognitive processes that require some form of self-reference. In my PhD thesis, **I investigated how the neural monitoring of cardiac activity relates to the self-reflective process underlying preference-based decisions and how their relationship may account for subjective evaluation, choice variability and the comparison between goods of different kinds**.

In my first experiment (*Chapter IV*; Azzalini et al., 2020), I showed that neural responses to heartbeats, recorded using MEG in healthy human subjects, are differentially

recruited when decisions involve self-reference and when they do not. We operationalize these decisions contrasting subjective preferences and objective perceptual discrimination. In addition, I revealed that a portion of preference inconsistency, which is generally ascribed to unspecified neural noise, is functionally related to fluctuations in neural responses to heartbeats. These results thus suggest that the neural monitoring of cardiac signals contributes to preference-based decisions, possibly by enabling a precise self-representation, which in turn produce a more stable evaluation and hence more consistent choices.

Evaluating different options is particularly challenging when the nature of the alternatives is not the same. For example, during the weekend, you may want to choose whether to have an ice-cream outside or to stay at home reading a book. **How can goods that share almost no features be mapped onto a single value scale?** We reasoned that neural responses to heartbeats may provide a common reference signal that coordinates information pertaining to different sensory and conceptual spaces. This coordination will result in a common self-referential space in which items can be evaluated and compared. We test this hypothesis in a second MEG experiment (*Chapter V*). Participants indicated their preference either between two items belonging to the same category of goods, such as two food items, or between items belonging to two different categories, such as between a food item and a movie. Consistent with our hypothesis, we found that neural responses to heartbeats differed between the two types of decisions (same vs. different categories), presumably reflecting their differential recruitment needed to map incommensurable goods into a common space. Moreover, fluctuations in neural responses to heartbeats predicted choice precision.

Importantly, in all the experiments we run strict controls to ascertain that differences in neural responses to heartbeats could not be trivially explained by changes in cardiac parameters (an issue analyzed in more depth in a methodological paper, cf. *Appendix B*), arousal and ongoing brain states. Altogether, our results indicate that our effects stem from genuine neural differences in how the brain responds and integrates ascending cardiac signals in the decision process.

At the end of the manuscript (*Chapter VI*), I will discuss the consistency between experimental findings, put forth a putative role of neural monitoring of heartbeats in preference-based decisions and outline current controversies and open questions.

Chapter I: Preference-based decisions

1. Value: definition and measurements

A. What is a preference?

Preferences are often invoked during everyday discussions and expressing them represents a ubiquitous activity in our daily life. Despite this, a clear consensus of what preferences are is still missing. From a purely descriptive standpoint, preferences can be defined as a relation between options, such that when someone prefers *A* to *B*, she believes that *A* is more worthy or valuable than *B*. Where does this relation stem from? On the one hand, preferences can be thought as something that individuals intrinsically possess and express through their choice. This view considers preferences as hypothetical choices and choices as revealed preferences (Hansson and Grüne-Yanoff, 2018). On the other hand, preferences can be considered as the product of a construction taking place at the time of choice and not something that is retrieved (Lichtenstein and Slovic, 2006). While the former account captures well the fact that humans generally tend to be consistent in their choices, it faces some difficulties in explaining why – in numerous contexts – humans appear to reverse their preferences, leading to decisions that, from a normative standpoint, are irrational.

Both accounts, independently of their ontological stance, appeal to the notion of **value, which is the input quantity to some function that enables individuals to make decisions**. In pure economic choices, such as when deciding whether to win 10€ or 5€, value is clearly identifiable: it is the scalar associated with monetary prizes. In these contexts, individuals use these quantities to guide their decisions in order to select the highest amount. However, the mapping between options and value is not straightforward in many other contexts. These more complex decisions are the subject of value-based and preference-based decision-making research. For example, imagine that you have to choose between a small but secure amount of money (5€) or to bet on a random lottery that will give you either a larger amount of money (20€) with some probability (0.3) or nothing otherwise. Although normatively everyone should accept the bet (which has a higher expected value 6€ = 0.3*20€), the willingness to accept the bet greatly varies across individuals. This is because not all people evaluate the options and their characteristics in the same way. The role of subjective evaluation in guiding our choice is

even more apparent in other types of decision, which we face every day. Imagine that you are sitting in a café, and you want something to drink. Would you prefer to have tea or coffee? In this case, the value of each option is strongly dependent on your own taste and the mapping between these well-defined extrinsic quantities onto a single value is complicated. These examples show the difficulty characterising the study of preference-based decisions: ***value* is inherently *subjective* and depends on the unique *perspective* of the individual making the decision**.

Psychology and neuroeconomics have been putting a large amount of effort in developing experimental approaches to retrieve reliable measures of subjective values. Before detailing these approaches, it is useful to consider how value has been defined and what it is thought to reflect.

B. How to define values?

The rapidly expanding field of neuroeconomics has provided some useful operational definitions to distinguish different senses in which value is used. For example O'Doherty (O'Doherty, 2014) distinguishes four possible definitions of value:

1. The relative *attractiveness* of a stimulus, which is maximized at the moment of choice.
2. The *utility* of an option, which enables the ordering of subjective preferences over a set of options.
3. The *motivating properties* of an option, namely the amount of work or effort that an animal is willing to undertake to obtain that option.
4. The *hedonic experience*, i.e. the subjective pleasure, that the option provides.

These definitions reflect distinct ontological perspectives and each of them falls short in accounting for value-based decisions in which different facets of value come into conflict. Consider drug addiction. A drug addict is probably willing to exert a great amount of effort to obtain some drug, despite its consumption may provide no pleasure at all. Based on the third definition, we should ascribe a high value to the drug, however, if we consider value as the amount of pleasure it provides, drug should have little value, if any. In addition, *a posteriori* definitions of value, such as the one based on utility, appear to be at odds with its intuitive psychological nature. Intuitively, value should be something that pre-exists decisions and determines them: we choose something because its valuable, we do not estimate something as

valuable because we choose it. However, as I shall present later in this chapter, these two aspects most probably co-exist. The opposite perspective, that is defining value as a pure internal representation of the amount of pleasure an option provides, also faces problems. Consider people that undergo an extreme physical effort to finish a marathon. This activity can be hardly defined as pleasing; nonetheless, people choose to do it and persevere in doing it. This example suggests that pleasure cannot fully explain how subjective value is computed and a more complex definition that integrates other aspects is required. Although a comprehensive definition of value is still missing, all these definitions uncover some important dimensions of what value is and how it relates to decisions.

C. How to measure value?

At least four different methods are generally used to obtain an explicit or implicit measure of subjective values (**Figure 1.1**). These methods provide cardinal (subjective ratings, willingness to pay and effort tasks) and ordinal (choice task) measures of subjective values and despite their methodological and conceptual differences, they provide consistent results (Lopez-Persem et al., 2017).

1. *Subjective ratings*. This method consists in asking participants to rate some items based on 'how much they like them'. Subjective ratings have been used for different kinds of stimuli: visual (Lebreton et al., 2009), acoustic (Blood et al., 1999) and gustatory (Plassmann et al., 2008). In spite of its directness, this method suffers from two main limitations. First, it is not suited to investigate value-based decisions in animals, limiting the generalizability from one species to others. Second, subjective ratings are inconsequential: participants are not incentivized to precisely report their subjective values. This may produce noisier estimates of true subjective values as well as distorted evaluations.

2. *Willingness-to-pay (WTP)*. Becker-DeGroot-Marschack (BDM) auction task (Becker et al., 1964) provides a solution to incentivize participants to faithfully report subjective values. Participants are endowed with some amount of money and they have to say, for each presented item, how much they are willing to pay for it. After all the bids are placed, one of these bids is randomly selected. Then a random number, from uniform distribution spanning the possible bids participants could place, is drawn. If the random number is equal to or lower than the bid that the participant has placed for that item, she

must pay the amount of money equal to the randomly drawn number and will obtain the item. This procedure has been shown to incentivize a faithful report of the subjective values, as the optimal strategy for participants is to place a bid corresponding to the exact amount they are willing to spend for that given item. Since the random number determines the price that participants pay, there is no incentive in placing a lower bid, nor a higher one. Some studies force subjects to consume the item that they obtained through this procedure to make the task even more consequential (Plassmann et al., 2007).

3. *Effort tasks*. As presented in the section above, value can be partly defined based on its motivational aspect. In the Becker-DeGroot-Marschack task, motivation assumes an abstract (economic) dimension as it is indexed by the amount of money participants are willing to spend. In effort tasks, motivation is measured by subject's willingness to exert some effort to probabilistically obtain an option (Pessiglione et al., 2007; Lopez-Persem et al., 2017). This method provides a more implicit measure of subjective value associated with a given option.

4. *Choice task*. Another approach to determine subjective values is to directly infer them from participant's choices. This paradigm assumes that if an individual chooses *A* over *B*, it means that the value she assigns to *A* is larger than *B*. This operational definition of value is based on a basic axiom of utility theory, whereby it is possible to order each option in a choice menu (assign them a *utility*) so that single choices maximize utility, i.e. individuals select the option with the highest value (Morgenstern and Von Neumann, 1953).

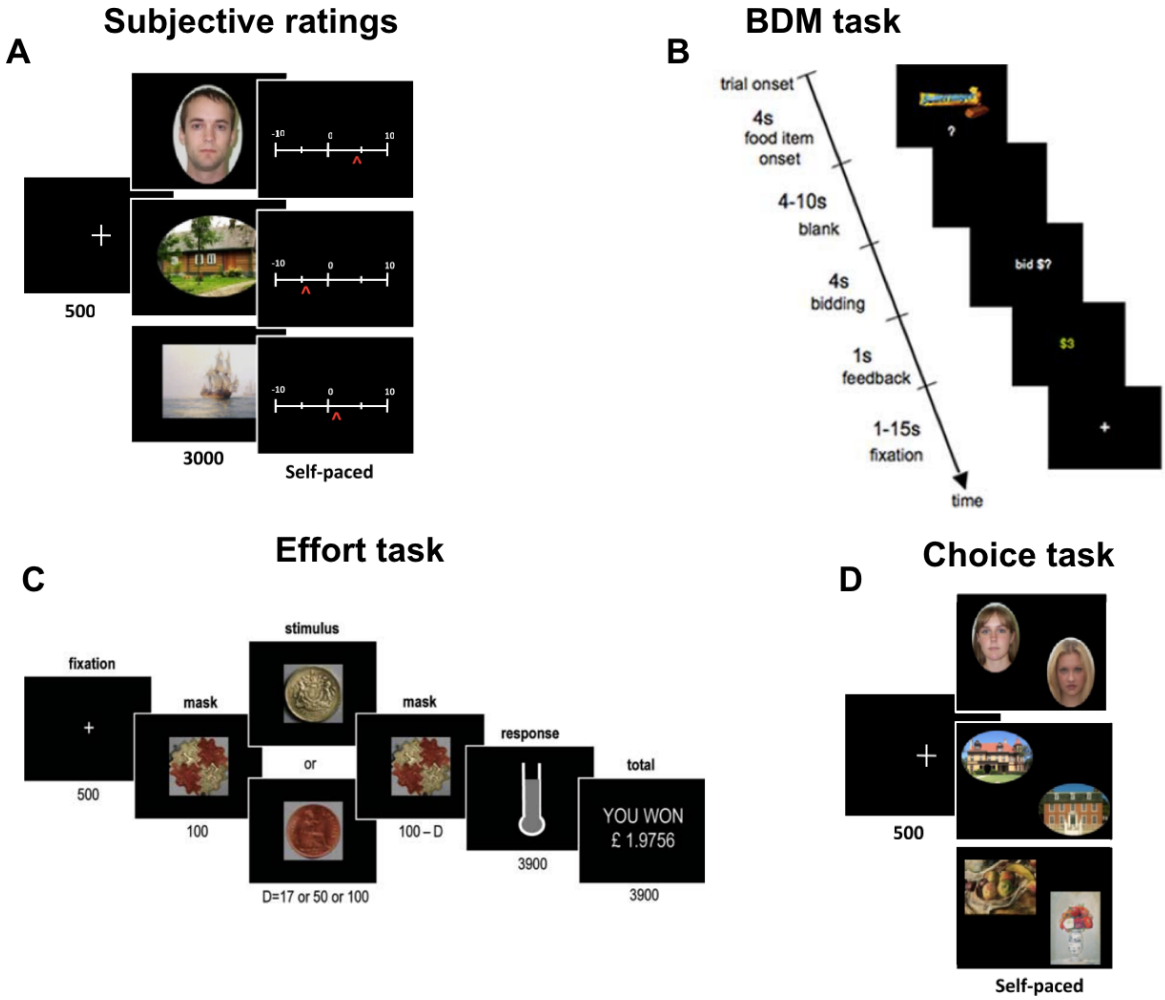


Figure 1.1 Different paradigms used to investigate subjective values. **A**, *Lebreton et al (2009)* asked subjects to rate on a 21-point scale how much they liked the different items (faces, houses and paintings) displayed on the screen. **B**, *Plassmann et al (2007)* used BDM task to assess how much participants were willing to pay to obtain the food item presented on the screen. **C**, *Pessiglione et al. (2007)* measured how willing participants were to squeeze a hand grip to obtain the reward on the screen (subliminally presented). **D**, *Lebreton et al (2009)* asked participants to choose the preferred item between the two presented. Task illustrations used in the figure are readapted from the cited papers.

A precise mapping between the experimental paradigms and the theoretical definitions of value is not straightforward. Whereas effort tasks nicely fit with a motivational definition of value, the other tasks seem to involve different value-related facets. For example, willingness-to-pay paradigms rely on both attractiveness and the motivation aspect of an item.

Experimentally, each method comes with some benefits and drawbacks. For example, willingness to pay, effort and choice tasks (for this latter, in the specific case in which one of the choices is randomly drawn and the selected option/outcome allotted to participants) all aim to incentivise 'true' subjective values, similar to those guiding everyday behaviour. Indeed, these approaches have proven successful to study consumer behaviour. However, they may fail to capture some other features of preferences. To understand why, imagine that a participant must choose between a chocolate bar and an apple. She picks the apple, despite her passion for chocolate. The reason behind her choice is that she ate too much chocolate the day before and she wants to pause consuming it for some days. Incentivized paradigms would infer that apple has a greater value for the participant, despite participant's 'true' preference for chocolate. Incentive paradigms may thus be more influenced by contextual factors (e.g., momentarily changes in states) than likeability ratings, which can reveal long-lasting preferences.

To conclude, **there is no unique solution to experimentally retrieve measures of subjective values**. The choice of the paradigm to be used ultimately rests on the aspect of value that the experimenter is more interested in and a combination of multiple approaches may reveal deeper insights otherwise inaccessible. As the objective of my PhD experiments was to investigate how spontaneous fluctuations in self-related processes influence decision coherence, I used subjective ratings to obtain more abstract and long-lasting subjective preferences and choice task to probe their consistency.

Summary

- Preferences can be defined as a relation between alternatives such as **when one prefers A to B, she thinks that A has more value than B**. In order to explain preference relation, it is fundamental to appeal to the notion of value.
- The way **values are** ascribed is **inherently subjective**: different individuals prefer different things. Despite this general agreement, a clear definition of subjective value, whether it is something that we possess and reveal or construct while choosing, is still a matter of debate.
- Psychology and, more recently, neuroeconomics have provided different **operational definitions of subjective value**. Definitions are based on **stimulus' attractiveness**, its **motivational properties**, the **pleasure** it provides or value is inferred from the **pattern of choices**.
- **How can subjective values be experimentally retrieved**? Four major experimental paradigms have been used: **subjective ratings**, **willingness to pay**, **effort tasks** and **forced decisions**. Although each paradigm has benefits and drawbacks, all paradigms provide coherent measures of subjective values.

2. How do we choose? Computational models of decisions

A. The decision problem

How are subjective values used to reach a decision? Rangel and colleagues (Rangel et al., 2008) described value-based decisions as a 5-steps process (**Figure 1.2**).

1. The first step is to outline the *decision problem* and it includes the *representation* of the agent current state (with her needs, conditions and so forth), the state of the environment (for example contextual factors, such as the presence of an external threat) and the available options/actions among which to choose.

2. During *evaluation*, values (either new or learnt) are associated with the available options in light of the current internal needs and external factors. It is worth noting that "anticipated value", a commonly employed term in the neuroeconomics literature to describe the attractiveness and motivation of choice alternatives, refers to this stage of the decision process and it is thought to guide later ones.

3. The agent must then compare the options and select the one with the highest value (*value comparison and action selection*). This stage represents what is intuitively thought of as the decision problem, that is, how the agent decides between alternatives. It is at this stage that other value-related concepts, such as decision value or value difference, refer to. These quantities are key factors regulating the predictions of some behavioural models such as drift diffusion models and psychometric functions.

4. Once a decision is made, the agent obtains and consumes the chosen option. Outcome value, yet another decision-related concept, refers to this phase and reflects the amount of reward or pleasure derived from outcome consumption (*outcome evaluation*).

5. Evaluation of the obtained outcome is then used to update the value of the options involved in the decision process (*learning phase*). This stage has been the key target of one of the most prolific domains of decision-making research: reinforcement learning.

As will become clear in *Chapters II* and *III*, the main contribution of my experimental work concerns step 1 and the influence that this step exerts on later stages, namely 2 and 3.

In the next sections, I will focus on the third step of the decision problem, value comparison and selection, presenting different computational models that have been proposed to formalise how these processes operate.

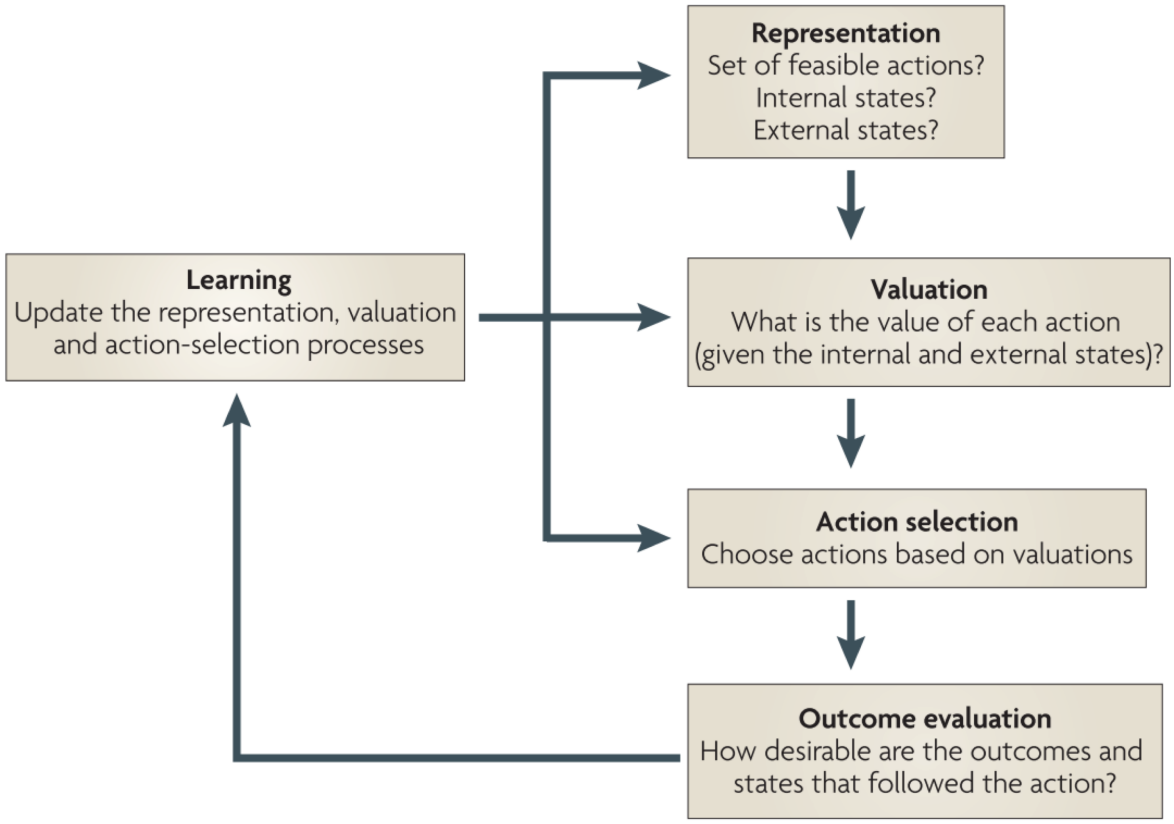


Figure 1.2 The 5 steps involved in value-based decisions. Figure from Rangel et al., 2008.

B. Argmax

According to a normative definition of how choice should operate in value-based decisions, individuals select the option that maximizes their utility, that is, they select the option with the highest value. More formally:

$$Choice = argmax(A, B)$$

Describing choice through an argmax function makes option selection fully deterministic. This entails that errors, corresponding to the selection of the suboptimal option, are impossible even for the smallest difference between A and B, provided that the representation of the value of each option is correct.

C. Softmax

In order to capture the inherent choice variability and suboptimality observed in animal's behaviour, we need to introduce some stochasticity in option selection or some uncertainty about the value representation. The first solution is achieved by replacing the *argmax* function with a *softmax* function (Luce, 1959a), allowing choices to be probabilistic (**Figure 1.3B**). This approach is generally modelled as an exponential function of the form

$$p(A) = \frac{1}{1 + exp((\alpha - DV)/\beta)}$$

Where p(*A*), the probability of choosing A, is determined by the value difference between the two options *A* and *B*, also known as decision variable (*DV*). In order to capture the individual specificity, two parameters α and β are generally fitted to single participant's data. The α parameter captures the bias that agents may have in selecting one option even when the two alternatives are equally valuable. The α parameter thus shifts the curve along the x-axis (**Figure 1.3B**). On the contrary, the β parameter determines the precision of the choice, by modifying the slope of the softmax function (larger β produces flatter curves, hence more stochastic choices). Slope changes principally affect choices between similar values (i.e. small *DVs*) by lowering the probability of selecting the option with the highest value. Importantly, α and β are independent parameters, a feature that enables to distinguish two sources of variability in choice.

D. Signal detection theory

A second way of accounting for choice stochasticity comes from signal detection theory (*SDT*), a widely used approach in perceptual decision-making (Tanner Jr and Swets, 1954). *SDT* assumes that decisions are based on evidence sampled from an internal representation and compared to some internal criterion (**Figure 1.3C**). If the evidence in favour of one alternative exceeds the criterion, that option is selected. As in the case of the *softmax* function, bias and sensitivity can be retrieved independently. Bias corresponds to a shift in the decision criterion, i.e. the amount of evidence needed to select the option. Sensitivity, the precision in choosing the option with the highest value, is constrained by the amount of overlap between the two distributions from which evidence is drawn. Therefore, when the available options are similar

in value, more errors are to be expected as the randomly drawn evidence may be mistakenly selected or rejected.

Although *softmax* and *SDT* give quantitative similar predictions, they suggest slightly different sources of decision stochasticity. On the one side, *softmax* function assumes that the decision maker correctly represents the true values associated with the options but her comparison process is corrupted by noise during the selection process. On the other hand, *SDT* highlights the probabilistic nature of values. These two possible interpretations mirror what I have pointed out in the section *What is a preference?*: whether preferences are something that we possess and express, possibly with some noise (*softmax*) or something fluid that is probabilistic in nature (*SDT*).

The *softmax* and *SDT* models are easy-to-implement and robust procedures providing a formalized description of the decision process. However, they suffer from two limitations. First, they only consider decision probabilities, leaving the temporal aspect of the decision process unspecified. As we know from our daily experience, difficult decisions take more time. Second, they offer a limited mechanistic insight into how the different components of the decision process can be implemented. A second class of models, which can be called dynamic, try to address these limitations by modelling jointly choice probabilities and reaction times, offering at the same time a more mechanistic implementation of this process.

E. Sequential Sampling Models

Sequential sampling models, of which the most diffused is probably the *drift diffusion model* (DDM; Ratcliff, 1978), represent decision as a dynamic evidence accumulation process – corrupted by some gaussian noise – which drifts toward a threshold, called boundary. Once a boundary is reached, an option is selected and an action takes place (**Figure 1.3D**). At every time step the amount of the accumulated evidence (corresponding to the position of the drifting particle) is determined by:

$$R_{(t+1)} = R_{(t)} + v * DV + \epsilon_{(t)}$$

Where *v*DV* represents the drift rate, i.e. the speed at which the evidence drifts towards a boundary. The drift rate is determined by *v*, a fitted parameter, which scales the influence of decision evidence (*DV*), representing the relative evidence in favour of one option. Their

interaction determines reaction times. Smaller decision evidence will produce longer decision times, as the drifting particle will be slower to reach the decision boundary. In addition, accumulation is corrupted at every time step by gaussian noise ($\varepsilon_{(t)}$). In most cases, this model is used to explain two-alternatives choice, and therefore uses two symmetrical boundaries, *a* and *-a*, which represent the two possible actions/options (*A* and *B*) that are selected once one of the boundaries is hit. For the same reason, *DV* can take either a positive or negative value, as it represents the relative evidence in favour of one of the two options. Another important component of this model is the starting point (*z*), expressed relative to the boundaries, which determines the relative position from where evidence starts to be accumulated. Finally, to account for the time needed for sensory processes (for example visual processing) and action execution to be performed, the drift diffusion model also comprises another parameter called non-decision time (*ndT*).

In this model, choice stochasticity may stem from several sources. First, it may depend on noise ($\varepsilon_{(t)}$), which can randomly push the particle beyond a threshold triggering an incorrect decision. Noise maximally exerts its influence when decision evidence (*DV*) is weak. This effect would partially map onto a shallower slope in the psychometric function and a low sensitivity in *SDT* framework. The effect of noise can be further amplified when the amount of evidence needed for decision is reduced, corresponding to a shrinkage in the distance between the starting point of evidence accumulation (*z*) and decision boundary (*a*). This reduction may be due to task requirements such as speeded responses obliging participants to 'set' lower decision boundaries. A similar shrinkage can be obtained by shifting the starting point towards one of the two boundaries. This operation corresponds to a bias towards one of the two options, similar to a shift in criterion in *softmax* and *SDT* models. Although both scenarios will produce faster responses, only the latter will predict higher choice probability for one option (the one closer to the decision bound) (for a thorough review of the characteristics of this model see (Ratcliff et al., 2016)).

Instead of accumulating the relative evidence in favour of either option, *race models* consider the evidence of each option to be accumulated in separate modules (Vickers, 1970). Sequential sampling models have been also modified to accommodate the effects of other cognitive factors. For example, attention has been proposed to modulate the drift rate according to which option is currently considered (Krajbich et al., 2010). Many other flavours and variants

of sequential sampling models exist (Wagenmakers et al., 2007; Brown and Heathcote, 2008), but they are outside the scope of this section.

F. The Neural Circuit Model

The drift diffusion model, in its original version, rests on the strong assumption that evidence integration is perfect, that is, the time constant with which evidence is maintained is infinite (for a version of the sequential sampling model allowing integration to be leaky, see (Usher and McClelland, 2001)). This assumption is challenged by the knowledge of electrophysiological properties of neurons, which have time constants of tens of milliseconds (Kandel et al., 2000). To integrate these biological constraints into the evidence accumulation process, Wang (Wang, 2002) proposed the *neural circuit model* (NCM – **Figure 1.3E**). This model comprises two pools of neurons, each of which is selective for one alternative, and it is characterized by plausible time constants and synaptic dynamics. The two pools have strong recurrent excitatory connections and compete through mutual inhibition. Noisy incoming evidence (the difference in inputs) that favours one of the alternatives is integrated over time through stochastic ramping activity. When the ramping activity of one of the pools reaches a predefined threshold, it triggers an all-or-none response in downstream brain regions enabling an overt behavioural response, while the activity in the other pool decays. An important consequence resulting from the attractor dynamics of this decision network is that early evidence has a larger influence on the decision process than a later one, as in the latter the network is already into one attractor mode with a stable pattern of activity. However, this prediction seems at odds with experimental data showing that late evidence can reverse the initial commitment to a choice (Resulaj et al., 2009). In NCM, noise arises from the intrinsic variability in neural responses in the recurrent neural dynamics, which represent at once evidence accumulation and the categorical choice (Wang, 2008).

The computational models described in this section offer multiple levels of analysis of the decision dynamics – from simpler psychological descriptions of value comparisons to more detailed neurobiological implementations. Although the ultimate goal will be a full-blown description of how decisions are biologically implemented, the sophistication of more advanced models of behaviour such as DDM and NCM comes with a cost. These models require large amounts of data and experimental manipulations that are able to isolate specific subcomponents

of the decision process. These conditions are not always easy to meet. In addition, consistent but small behavioural effects may become more difficult to capture as the number of parameters grows, as their effect may "spread" on different features of the model.

In my experimental work, I have thus decided to use a *softmax* function to characterise participants' behaviour. This simpler description of the decision process allowed me nonetheless to relate specific behavioural signatures to precise neurobiological events.

Summary

- How do subjective values guide preference-based decisions? Broadly, **the decision problem can be decomposed in five steps**: (a) representation, (b) evaluation, (c) comparison and selection, (d) outcome evaluation, (e) learning.
- The computational models of decision reviewed in this section, focus on how subjective values are compared and selected.
- Simple **models of choice** comprise: **argmax** and **softmax** functions, and **signal detection theory**. These models are robust descriptions of choice behaviour, but they lack a deeper mechanistic insight into the decision process.
- **Sequential sampling models** and **biophysically inspired models** of choice are able to **capture both choice and reaction times,** describing the unfolding of decision dynamics over time. This deeper mechanistic insight comes with a cost of more powerful experimental manipulation – able to isolate the influence of experimental factors on parameters model – and more complex fitting procedures.

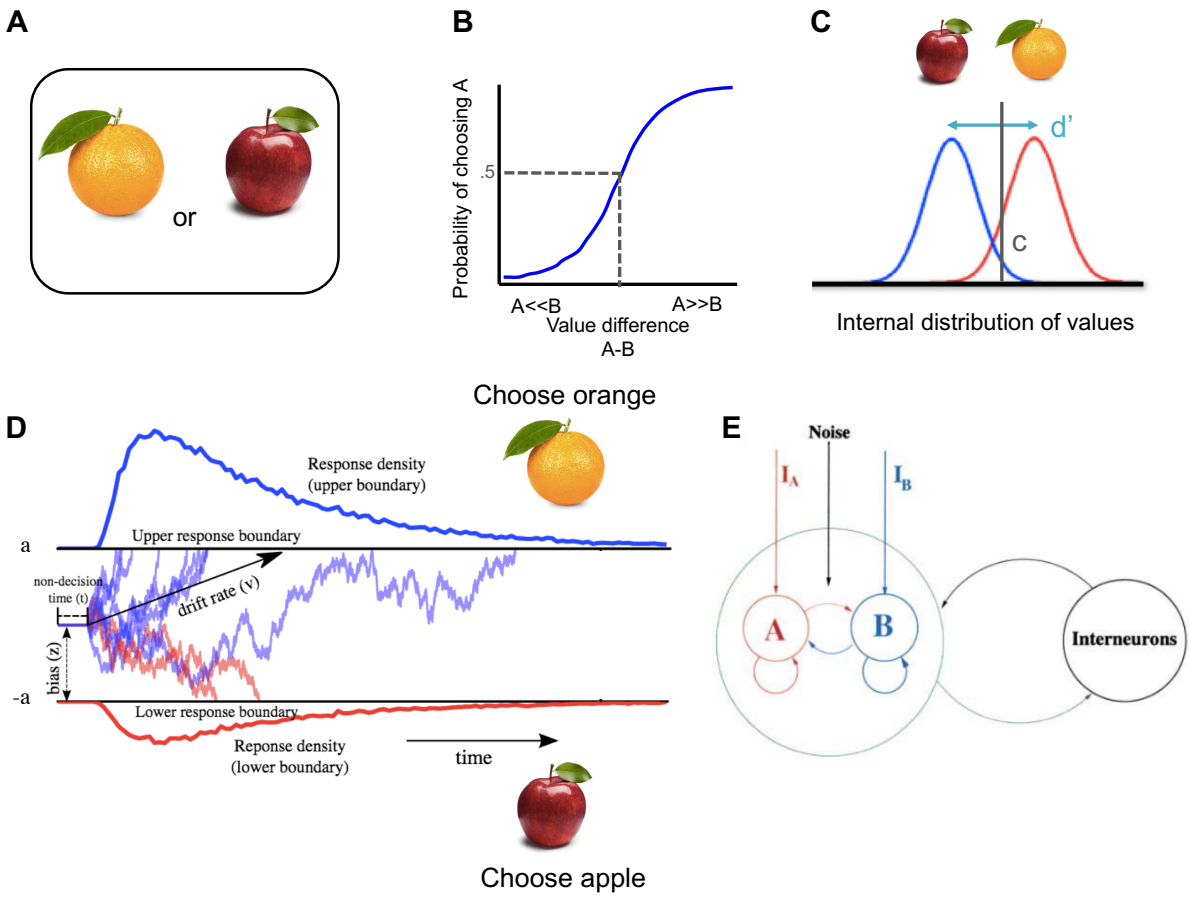


Figure 1.3. **Models of decision**. **A**, 2-alternatives forced choice example between 2 (food) items. **B**, Softmax function fitted to choices differing in the values of the two options. **C,** Signal detection framework applied to value-based decisions. Values are gaussian distributions from which a sample is drawn and compared to a criterion (c, dark grey) to decide. The distance between the two distributions determines the sensitivity (d', light blue). **D,** Schematic representation of DDM (Ratcliff, 1978), in which a decision is made when evidence is accumulated and hits one of the two symmetric thresholds (a/-a). DDM predicts response proportion as well as RTs. **E,** The neural circuit model (Wang, 2002) describes decisions as a mutual inhibition process between two pools of neurons selective for the value of the two alternatives.

3. The neural underpinnings of value-based decisions

In the previous section, I have described how values come into play during the decision process and how this latter can be formalised through a mathematical framework. On the other hand, understanding how values are represented in the brain and which brain regions contribute to the decision process can constrain modelling work and refine psychological frameworks of value-based decisions. In the current section, I will thus review the experimental findings about the neural underpinnings of value representation and value-based decisions. This body of evidence has highlighted the presence of a distributed network of brain regions associated with value representation, comprising the ventromedial prefrontal cortex (vmPFC), posterior cingulate cortex (PCC), ventral striatum (VS) and hippocampus (HC).

A. Value-related signals

To study how values are represented in the brain it is necessary to outline the different value-related quantities that have been used and how they map onto psychological and computational processes. As outlined above, the term value does not have a unique meaning and often describes various decision-related aspects. For example, value can be considered either as the evidence upon which decision is based, ensuing from the evaluation of the available options, or as the hedonic experience produced by the outcome receipt. Below, I will provide a brief overview of the value-related quantities most commonly used to investigate how value is represented in the brain.

1. **Stimulus value and goal value** (O'Doherty, 2014) refer to the value of the available alternatives and it is thought as the **input to the decision process.** In other words, it corresponds to its **prospective value**. The literature refers to this quantity also in terms of offer value (Padoa-Schioppa and Assad, 2006; Padoa-Schioppa, 2013).

2. **Outcome value** is a notion tightly connected to stimulus value and it reflects the **evaluation ensuing from outcome receipt** (O'Doherty, 2014). This quantity has been generally investigated by looking at changes in brain activity following the delivery of pleasant vs. unpleasant (or neutral stimuli). Alternatively, this quantity has been investigated in the context of devaluation, such as when a food item is no longer rewarding because the animal has been satiated (O'Doherty, 2004).

3. **Chosen** and **unchosen values** are two other quantities generally employed to study value-related brain signals (Padoa-Schioppa, 2013). These quantities are also known as **post-decision value signals** as they can only be determined *post-hoc*, once the decision is made. It is currently debated which role they play in the decision process and whether they are actually represented as such in the brain. Some authors suggest that neural correlates of chosen and unchosen values are a by-product of network's behaviour, thus artefacts of the analysis (Hunt and Hayden, 2017). On the other hand, chosen and unchosen values have been shown to have functional relevance, for example in context where decisions are taken successively and individuals are required to learn the value of each option by trial and error. Experimental evidence showed that the rate of learning for chosen and unchosen option values is different depending on the valence of the context and feedback information (Palminteri et al., 2017).

4. In models such as DDM and NCM, **value difference** corresponds to the **relative evidence** that is accumulated during decision and determines the difficulty of the choice. In turn, value difference influences the accuracy of the decision as well as its reaction time. Even though the influence of value difference on behaviour is straightforward, how the neural representation of this quantity supports the decision process is still debated.

5. Finally, the **sum of option values** reflects the overall attractiveness of the choice menu, also defined as **context/state value** (Palminteri et al., 2015). A robust behavioural effect of value context is that decisions between options with higher sum of values (two rewarding options) are faster than decisions between low-valued options (two punishments). The overall value of the available (or recently encountered) options is a central notion in accounts that propose normalized representations of value encoding (Rangel and Clithero, 2012). In addition, relative value coding can also explain the intriguing finding that humans tend to prefer options that yield small punishments to small rewards, when these values are learnt in loss and gain contexts respectively (Palminteri et al., 2015).

These quantities highlight distinct value-related aspects and they have been widely employed to investigate the neural underpinnings of value-based decisions. However, it is important to bear in mind that they differ with respect to the stage of the decision they refer to, the function they take up during decision as well as how they are experimentally defined. It

follows that when talking about value-related signals, one should specify which quantity a neural signal refers to, especially since the encoding of these different quantities appears in similar brain regions (Padoa-Schioppa and Assad, 2006; Bartra et al., 2013; Padoa-Schioppa, 2013).

Another important point to consider is the distinction between value signals and other confounding quantities. For example, brain signals representing values should be dissociated from those (linked, but different) underlying motor execution. This is normally achieved by decoupling motor requirements from the best option or by revealing stimulus-action mapping only after enough time has elapsed and a decision has been made so that value-related activity during stimulus presentation is not confounded with motor acts. Another confounding variable is saliency – the property by which valuable options draw more attention and possibly enhance cognitive and perceptual processes. Saliency can be decoupled from value since salient stimuli can be both highly appetitive or highly aversive. It follows that whereas values increase monotonically, saliency peaks at the extremes of the range of values. Using stimuli that span the full range of values, from aversive to appetitive, is thus important to dissociate these two dimensions (Rangel and Clithero, 2013).

In the following sections, I will provide a broad overview of the brain regions encoding value-related signals. I will present results from brain-imaging studies in monkeys and humans investigating value representation during both decision time and outcome receipt. These results point to the presence of a network of brain regions reliably involved in the representation of different value-related quantity across different stage of the decision process.

B. Brain regions encoding value

1. Ventral Striatum

Historically, animal studies have first used reinforcement learning paradigms to investigate the representation of values guiding choices. In one of these seminal studies, Schultz and colleagues (Schultz et al., 1992) found that activity in dopaminergic neurons in the ventral striatum (VS), a subcortical structure which is densely connected to many other cortical and subcortical brain regions, predicts the delivery of reward and is modulated by stimulus appetitive value. More recently, electrophysiological recordings in monkeys revealed that

ventral striatal activity is not only involved in reward-based learning, rather it encodes several value-related quantities (Strait et al., 2015). In this study, monkeys performed a two-alternative forced choice task between probabilistic rewards. Firing rates in VS during deliberation encode the value of the two options, their difference, and the value of the chosen option. After choice, VS neurons also encode the reward outcome (gain or loss) obtained from the selected option.

Value-related signals in VS were also observed for multiple types of stimuli in humans using fMRI. Namely, VS BOLD activity scales with the subjective value of delayed monetary rewards (Kable and Glimcher, 2007), the pleasantness of thermal somatosensory stimulation on the hand (Rolls et al., 2008) and the pleasantness of various categories of visual stimuli (faces, houses and paintings; Lebreton et al., 2009).

2. Ventro-medial prefrontal cortex and posterior cingulate cortex

Another brain region consistently associated with the computation of value is the ventromedial prefrontal cortex (vmPFC). BOLD activity in vmPFC has been shown to scale with the subjective value of the available options (indexed by pre-scan willingness to pay bids) during two-alternatives choice tasks (Chib et al., 2009), pleasantness ratings (Lebreton et al., 2009) and monetary bids (Plassmann et al., 2007). Interestingly, in two-alternatives forced choice tasks, the magnitude of vmPFC activity predicted participants' consistency in choice, i.e. their overall tendency to select the item they rated higher in a pre-scan session (Grueschow et al., 2015).

While vmPFC seems to encode value when this is relevant for deciding, the posterior cingulate cortex (PCC) appears to represent value even when this quantity is task-irrelevant. Using fMRI in healthy human subjects, Grueschow and colleagues (Grueschow et al., 2015) found that both vmPFC and PCC activities positively encoded subjective values of the displayed items when subjects were deciding whether to purchase them or not. However, only the activity in PCC encoded subjective value during a control perceptual task and predicted the amount of value-driven attentional capture that a given item exerted on the perceptual judgment. The dissociation between vmPFC and PCC contrasts with results obtained both with fMRI and intracranial recordings (Lebreton et al., 2009; Lopez-Persem et al., 2020), who found that an extended network of brain regions, including vmPFC, VS, hippocampus and posterior cingulate cortex (PCC), automatically encodes subjective value irrespective of task requirements (**Figure 1.4A**).

Whether involved in task-relevant computations or not, value encoding in VS and vmPFC is a consistent finding in the value-based decision-making literature, and it has been further corroborated by two meta-analysis involving 206 and 81 functional magnetic resonance imaging (fMRI) studies (Bartra et al., 2013; Clithero and Rangel, 2014). Both meta-analyses found value signals in these regions independent of the decision phase (i.e. both during deliberation and outcome receipt; **Figure 1.4B**) and type of rewards. Although vmPFC appears to encode subjective value of different categories of goods, an anterior-to-posterior gradient seems also to emerge within vmPFC (**Figure 1.4C**). Altogether, these results suggest that VS and vmPFC are involved in the representation of multiple value signals during different decision stages.

3. Hippocampus

The cognitive resources required in value-based decisions may vary depending on the type of tasks and rewards used in the study. Therefore, additional regions, beyond VS and vmPFC, may contribute to value-based decisions. Value-based decisions have been widely investigated using economic gambles where values can be 'easily' computed by combining the amount of reward (or loss) with probabilities. Conversely, preference-based decisions about food items or cultural objects are likely to tap onto a broader set of cognitive processes among which autobiographical memory is likely to play an important role. Imagine that you are asked if you prefer Miles Davis or Pearl Jam. Faced with this decision, you are likely to think back to your high-school days in which with your best friends you were listening to some rock music or you may think to the more recent past in which you have started to appreciate jazz. A recent study has provided direct evidence for the contribution of the hippocampus, a well-known memory-related region, to preference-based decisions (Bakkour et al., 2019). Hippocampal BOLD activity predicted reaction times of participants performing a 2AFC task between familiar snack items: longer reaction times were associated with higher activity in the hippocampus. Using the same experimental task in amnesic patients, the authors found that patients were generally more stochastic and slower than healthy controls in preference-based decisions, but not in perceptual choice. The results suggest that the hippocampus may provide necessary value-related evidence during deliberation, even though in this study it has not been found to encode value *per se*. In addition, it appears that hippocampal contribution is particularly relevant when decisions are difficult – that is when more internally generated evidence is needed to tip the balance.

Beyond decisions involving familiar stimuli, hippocampus appears to be also involved in the learning and generalization of new stimulus-values associations never encountered before (Wimmer and Shohamy, 2012). In this experiment, participants first performed a target detection task, while seeing pairs of pictures (of faces, scenes and bodies) and abstract images. In a second phase, while undergoing fMRI recordings, participants learnt associations between abstract images and rewards through classical conditioning. In the final decision phase, participants had to choose between two visual stimuli, either the directly conditioned abstract images or the pictures seen during detection task. Participants chose more frequently pictures that in the detection task were associated with abstract images later rewarded and the strength of this bias correlated with hippocampal activity during the conditioning phase. Although hippocampus has already been found to represent subjective value (Lebreton et al., 2009), its cognitive contribution was not directly tackled with a specific experimental paradigm. These results can thus may lead us to consider value as a more complex quantity wherein other cognitive components may participate.

4. Dorsal anterior cingulate cortex and anterior insula

Although their role is still highly controversial, two other brain regions are worth mentioning when presenting experimental evidence about the neural underpinnings of value representation. In their extensive meta-analysis, Bartra and colleagues (Bartra et al., 2013) found activity in dorsal anterior cingulate cortex (dACC) being negatively associated with subjective value, a signal that according to some authors possibly represents the value of alternative options (Kolling et al., 2012). Contrasting this view, another fMRI study revealed that dACC activity encodes the difficulty of the decision (Shenhav et al., 2014). These findings inspired two different perspectives on dACC function: value comparator (Rushworth et al., 2012) and cognitive controller (Shenhav et al., 2013). The interested reader may find it valuable to refer to a recent review (Heilbronner and Hayden, 2016).

The anterior portion of the insula has also been linked to value-related signals (Bartra et al., 2013). However, in numerous experimental paradigms, value is highly correlated with salience, a correlation that must be accounted for when investigating value signals (see the beginning of this section). Indeed one fMRI study found that when carefully controlling for these two related but distinct dimensions (Litt et al., 2011), BOLD activity in the insula

specifically reflected the salience of the presented food item, while only BOLD activity in vmPFC, rostral ACC and PCC scaled with the subjective value.

The brain-imaging studies presented in this section highlighted the existence of a distributed neural network, encompassing VS, vmPFC, hippocampus and PCC, which has been consistently related to multiple value-related quantities revealed by different experimental paradigms and a variety of stimuli. To understand whether these different value-related signals may serve distinct functions and how they coordinate to guide decision, it is useful to look at how value-related signals unfold over time.

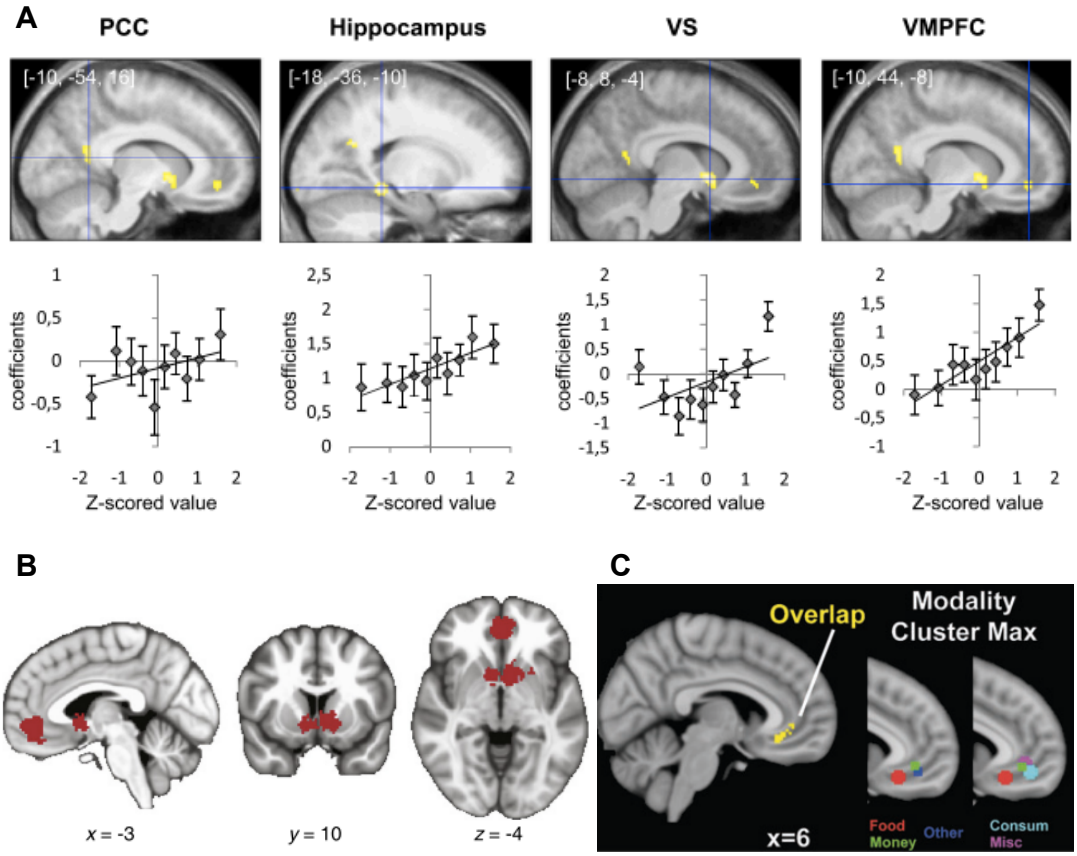


Figure 1.4. **Neural underpinnings of value representation**. **A**, Regions linearly encoding subjective value during rating (top panels) and correlation coefficients of the regression between regions activation and subjective values (bottom panel). Adapted from Lebreton et al., 2009. **B**, Meta-analysis (Bartra et al., 2013) showing brain regions (vmPFC and VS) positively scaling with value, independent of rewards type and decision phase. **C**, Coordinate-based meta-analysis (Clithero & Rangel, 2014) showing (left panel) vmPFC loci whose activity encodes value independent of reward type and vmPFC loci associated with the encoding reward-specific value (middle and right panels).

C. Temporal unfolding of the decision process: electrophysiological studies of value-based decisions

Electrophysiological techniques are particularly useful to access value representations rapidly evolving over time. In addition, these techniques have the advantage of directly recording neural activity, whereas fMRI can provide only an indirect measure of it through changes in BOLD. In this section, I will present studies looking at different types of neural activity such as local field potentials (LFPs), electroencephalography (EEG) and magnetoencephalography (MEG).

1. Serial representation of option values

fMRI studies, due to their poor temporal resolution, are poorly suited to investigate how the quick transition in considering one item or the other may affect the neural representation of value. On the contrary, electrophysiological techniques can access this information. For instance, chronically implanted electrodes in macaque's OFC (Rich and Wallis, 2016) revealed that the activity of these neurons (both from single neuron as well as LFP) encodes the value of the two presented options, but does so in an alternate fashion shifting from the representation of the value of the option the will be chosen to the value of the discarded one, starting around 150 milliseconds after option presentation (**Figure 1.5A**). In addition, the encoding of the chosen option's value predicted reaction times: the stronger the encoding the faster the response. Additional studies provide evidence that OFC and vmPFC neurons represent the values of both items as well as which one will be chosen. Yoo and Hayden (Yoo and Hayden, 2020) revealed that OFC and vmPFC neurons in macaques initially represent the values of two serially presented options. Interestingly, the activity in the same pools of neurons at a later phase of decisions is reorganised to enable the selection process. This functional reorganisation, supported by a change in the covariance structure between neurons, enables the same population of neurons to perform the evaluation and the selection processes at different moments. These findings are important as they may explain why neurons in OFC and vmPFC are often found to encode different value-related quantities such as the values of the offered goods, the value of the chosen option and the identity of the chosen good (Padoa-Schioppa and Assad, 2006). In addition, they indicate that evaluation and comparison process may occur serially.

2. *Neural signatures of accumulation dynamics*

In humans, invasive recordings are rare and temporal resolved neural dynamics are generally studied using electro-encephalography (EEG) and magneto-encephalography (MEG). Using temporal predictions issued from the DDM fit to value-based choices between snack items, Polania and colleagues (Polanía et al., 2014) revealed that EEG activity in 48-60 Hz frequency band at frontal and parietal sensors closely resembles an evidence accumulation process starting 1 second before response (**Figure 1.5B**). Interestingly, the coupling strength between frontal and parietal electrodes in this frequency range also predicted preference choice consistency.

Hunt and co-workers (Hunt et al., 2012) used the neural circuit model (Wang, 2002) to simulate the activity of a decision-making circuit performing a value-based decision task between risky prospects. They analysed network activity by regressing its frequency-decomposed activity against the sum and the difference in values of the available options. The analysis revealed that earlier in the trial, activity in the 10 Hz range encoded the sum of the available options, while later in the trial lower frequency activity (4-8 Hz range) encoded value difference. Next, the authors looked whether and which brain activity recorded with magnetoencephalography in human participants performing the same task behaved similar to model predictions. Two regions, vmPFC and IPS, transitioned from representing the overall value of the available alternatives to their difference in frequency ranges similar with those predicted by the biophysical model. Mutual inhibition therefore appears as a plausible model of value-based (and perceptual) decisions. Despite the similarity between model predictions and MEG activity in vmPFC and IPS, the reason why the sum and difference in values should be *explicitly* represented in neural signals remains to be elucidated.

Altogether, evidence from humans and non-human primates thus suggests a common involvement of prefrontal regions in value-based decision-making. A recent study (Hunt et al., 2015) directly compared the temporal profiles between human vmPFC and neural activity in OFC, DLPFC (dorso-lateral prefrontal cortex) and ACC in monkeys, finding remarkable similarities. All these regions encode the sum of option values early in the trial and their difference later on. In monkeys, only DLPFC activity was found to represent the value difference between the option early in the trial and the categorical decision reported by the animal later on, showing a temporal evolution from value comparison to action selection. This finding suggests that DLPFC may also perform both value representation and selection in the same way as OFC and vmPFC do (Yoo and Hayden, 2020). Finally, principal component

analysis (PCA) on LFP and MEG activity showed that the same second principal component – both in humans and monkeys – scales with chosen option's value during decision.

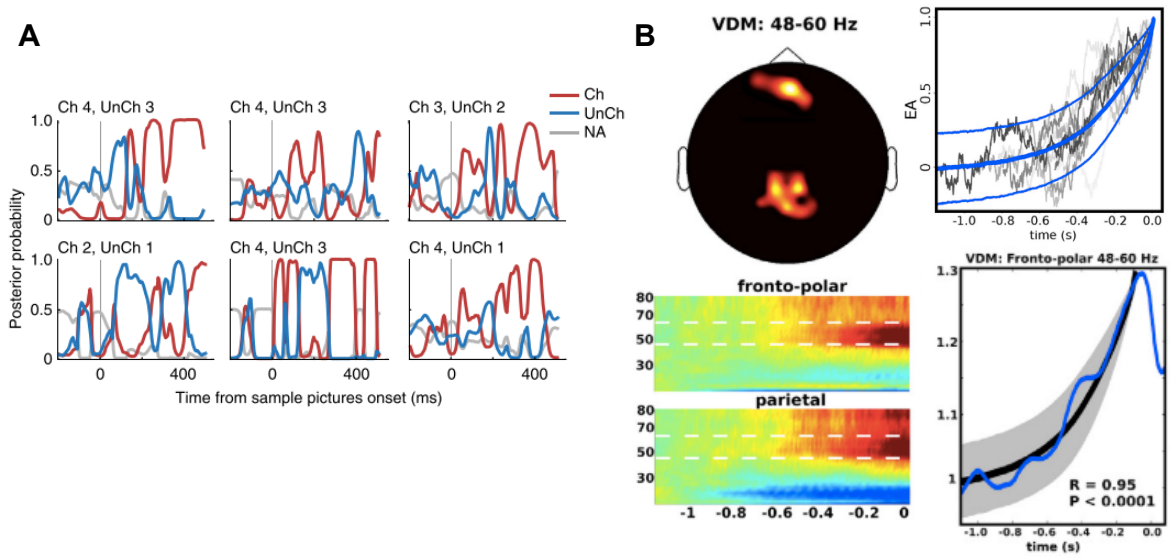


Figure 1.5. **Temporal unfolding of value-based decisions. A**, Six example trials of decoding accuracy of the chosen (Ch) and unchosen (UnCh) option values obtained from neural activity in monkey's OFC (Rich & Wallis, 2016). Decoding time-series clearly shows temporal alternation between encoding of the chosen and unchosen options. Figure from Rich & Wallis, 2016. **B**, Human EEG activity in 48-60 Hz frequency range in frontal and parietal electrode (left column) closely resembles the prediction of the drift diffusion model (top right panel) as visible in the averaged activity (across frequencies) of the fronto-polar cluster (bottom right panel) (Polanía et al., 2014).

D. Which functional role for the multiple value-related regions?

Despite the growing evidence showing the existence of an extended network of regions involved in value representation, it is not yet clear how these signals coordinate to support decision.

One possibility is that different brain regions may perform distinct and serial computations during the decision process. For example, vmPFC may first encode option values and then pass this information to other regions, e.g. dACC, that may in turn perform value comparison (Hare et al., 2011). This proposition seems at odds with recent findings showing that value comparison may happen already in vmPFC, through functional reorganisation (Yoo and Hayden, 2020). In addition, the value comparison process in vmPFC is compatible with the temporal profile of value encoding that shifts from one option value to the other (Rich and Wallis, 2016), it accounts for the encoding of value differences (Hunt et al., 2012) and why vmPFC damages impairs (some types of) decisions (Noonan et al., 2010). In this view, the role of other value-related brain regions is diminished as decision is already formed in vmPFC.

To reconcile these hypotheses, some authors have proposed that vmPFC performs value comparison, which are then passed on to dACC, which transforms value-based information into actions (Cai and Padoa-Schioppa, 2012). Others explain the involvement in vmPFC and dACC in terms of two distinct comparison processes (Rushworth et al., 2012). On the one hand, vmPFC performs specific value comparison guiding binary decisions, whereas dACC compares strategic values, namely whether to engage with a certain option or to search for better alternatives.

A more radical hypothesis claims that value is not explicitly represented in the brain as such, rather it emerges from parallel comparison processes that transform sensory information into action execution (Yoo and Hayden, 2018). According to this view, value-related regions do not perform any distinct decision-related computations, rather they are characterised by the types of inputs they receive.

To conclude, although a network of regions involved in value computation has been identified, a consensus about how the regions in this network coordinate has not yet been achieved.

E. The debate around vmPFC

The ventromedial prefrontal cortex has attracted a lot of attention in research on value-based decision. One reason is that vmPFC appears to represent the values of items as varied as food, money and non-consumable goods (Chib et al., 2009). This experimental evidence led to the hypothesis that vmPFC represents values in a 'common currency' (Levy and Glimcher, 2012), i.e. in an abstract format useful to compare different kinds of goods and enabling decision between them (Levy and Glimcher, 2011).

However, the debate is far from being settled. In fact, evidence suggests that vmPFC also maintains value representations that are dependent on stimulus type. For example, the metanalysis by Clithero and Rangel (Clithero and Rangel, 2014) reveals the presence of an posterior-to-anterior gradient of value representation in vmPFC going from more concrete goods (such as consumables and food) to more abstract ones (**Figure 1.4C**). More probably, vmPFC contains both types of value representations: category-dependent and category-independent. This view is in line with fMRI multivariate pattern analysis (MVPA) results showing that in the posterior and more ventral portion of vmPFC encoded the subjective value of food and trinkets, while its rostral part contained a category-independent value representation (McNamee et al., 2013). Recent evidence points that vmPFC and OFC, which are sometimes considered together (see last paragraph of the current section) have distinct functional roles: OFC may encode the value of specific attributes, while vmPFC may integrate value across attributes (Suzuki et al., 2017).

However, **the mechanisms supporting the transformation from a category-dependent to abstract values representation are currently unknown**. This transformation is particularly important when agents have to face decisions between goods of completely different natures, as these decisions require disparate attributes to be mapped into a unique representational space. For example, during the weekend you might decide whether to go out to have an ice-cream or stay in to finish reading your book. These alternatives are characterised by completely different features, yet we make these kinds of decisions every day. In my second experiment (*Chapter V*), I tested the hypothesis that the neural monitoring of cardiac inputs may provide a common neural signal which coordinates the mapping of distinct types of attributes into a unique space.

Findings from lesion studies provide evidence of a critical role of vmPFC in value-based decisions. Indeed, patients with vmPFC lesions are characterized by more stochastic choice

patterns in preference judgments than healthy controls (Fellows and Farah, 2007; Camille et al., 2011). The necessity of vmPFC for choice consistency is challenged by other lesion studies revealing that impairments in value-based decisions is present only for some type of real-world choice entailing social factors, such as choosing which person to marry, but not for other types of decision such as choosing which house to buy (Bowren et al., 2018). Rather than representing value *per se*, **vmPFC may be necessary to combine different sources of information and different attributes to establish the value of the available options**. Recent evidence supports this hypothesis showing that patients with vmPFC lesions are as consistent as healthy controls when choosing between options whose value can be predicted by one or two simple attributes. However, when each attribute is uninformative on its own and values must be inferred from a conjunction of attributes, patients reveal a striking impairment (Pelletier and Fellows, 2019). vmPFC thus appears to be necessary for the construction of values that merges stimulus attributes derived from mnemonic components and task structure (Vaidya and Fellows, 2020). This account resonates with other theories considering PFC as a key region for the appraisal of self-related processes combining several dimensions (Dixon et al., 2017).

The unique role of vmPFC as the endpoint wherein many sources of information are combined and compared for decision is challenged by theoretical accounts that embrace a parallel and distributed view of the decision process (Hunt and Hayden, 2017). Opposite to a modular and segregated perspective on decision, value representation is thought to emerge from mutual recurrent hierarchical competition processes. Two main arguments support this view: the dense anatomical and functional connectivity between regions thought to represent value, and the 'graceful degradation' of the network's capacity when some of its component are obliterated. As described above, although lesions on the medial face of PFC impair difficult decisions (Noonan et al., 2010) or those relying on configuration of attributes (Pelletier and Fellows, 2019), they leave simpler decisions unaffected, questioning whether value representation in this regions is necessary to guide decisions (Schoenbaum et al., 2011; Stalnaker et al., 2015).

Finally, it is important to note that authors tend to refer to vmPFC as a unique region, often neglecting the fact that vmPFC contains many sub-regions (**Figure 1.6**) distinguishable by their cytoarchitecture (Mackey and Petrides, 2014) and connectivity profiles (Neubert et al., 2015). The sub-regions mostly reported in value-based decisions are Broadman areas 10, 14, 25, 32 (Dixon et al., 2017). In addition, experimental evidence coming from different animal species (rodents, monkey and humans) and different neural recording techniques (single spike,

LFP, EEG, MEG and fMRI) are often pooled together to support the interpretation of value encoding in vmPFC. Although approaches based on different recording techniques and animal species are extremely useful to investigate value-related computations at multiple levels, some caveats should be kept in mind. First, the fact that across species task requirements may differ. Second, sub-regions recording sites may differ: the non-human primate literature preferentially records from OFC, while in humans the majority of the results comes from vmPFC. fMRI recordings from vmPFC are further complicated by artefacts affecting this region. In fact, the presence of air in the sinuses distorts the BOLD signal and reduces measurement sensitivity (Ojemann et al., 1997). Finally, generalising across recording techniques is not simple. Up to date, it is not clear how single-units, LFP and BOLD signals relate to one another (Wallis, 2012).

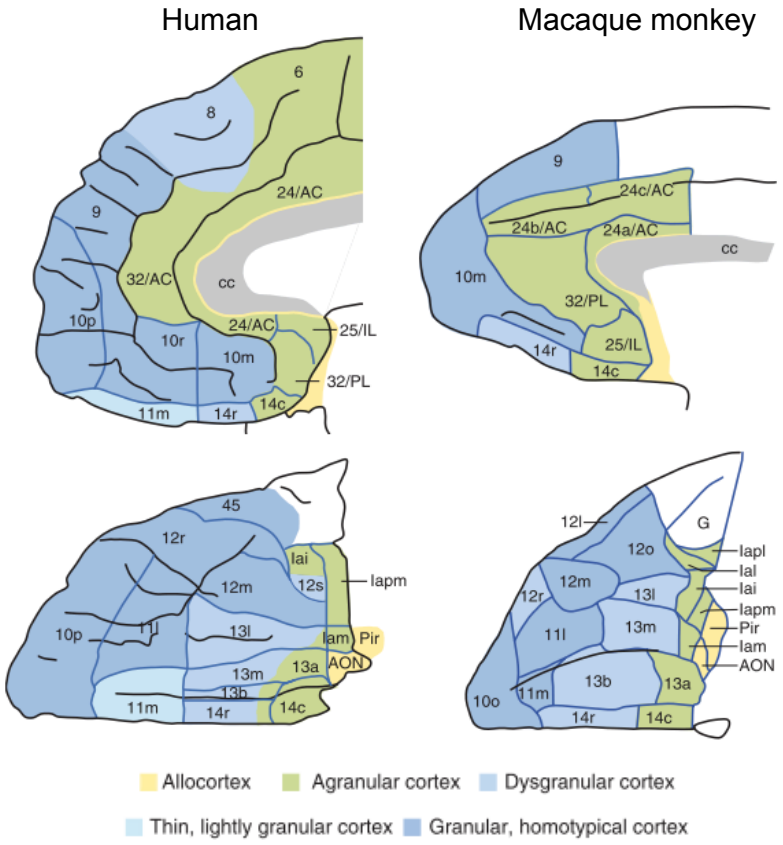


Figure 1.6. Cytoarchitectonic subdivisions of medial and lateral prefrontal cortex in humans (left) and macaque monkeys (right). Adapted from Wallis, 2012.

Summary

- Brain-imaging studies have been using **different value-related quantities** to investigate the neural representation of subjective values.
- Some of these **quantities** are thought to reflect the **input of the decision process**, such as the prospective value associated with a specific stimulus. Others reflect **post-decisional quantities**, such as the outcome value, chosen and unchosen option values. Finally, the difference between option values and sum of values reflect the **difficulty** and the particular **decision context** (gain or loss), respectively.
- Experimental evidence points to the presence of a **distributed network of regions** comprising the ventral striatum, the ventro-medial prefrontal cortex, hippocampus, anterior insula and dorsal anterior cingulate cortices encoding many value-related signals.
- Neurophysiological recordings further provide insight into the temporal evolution of these signals during decision. It has been found that vmPFC encodes shifts from encoding the value of the available options to the encoding of the one that is eventually chosen, suggesting that both evaluation and value comparison may take place there.
- Other accounts suggest that the multiple regions involved in value representation may contribute to distinct decision-related processes. For example, vmPFC may carry out the evaluation of the available options while dACC may perform comparison and action selection.
- The role of vmPFC in value-based decisions has been particularly debated. Some important points in this debate regard (a) whether vmPFC contains an abstract value representation common to different types of goods; (b) whether this region is necessary to guide value-based decisions; and (c) whether vmPFC sub-regions perform distinct value-related computations, as suggested by different cytoarchitectonic and connectivity profiles.

F. The neural mechanisms underlying choice variability

Preferences, and more generally value-based decisions, are stochastic in nature. Linking behavioural variability to its underlying neural mechanisms thus represents a valuable approach to understand the origin of choice stochasticity and further constrain theories about what values are and how they are used in decisions.

1. Choice variability and random noise

To some degree, choice variability can be explained by noisy computations. One class of theories proposes that choice variability in value-based decisions arises from noise in the selection process (Hey and Orme, 1994). Value-to-action transformation is thought to rest on functional connectivity between vmPFC, dACC and parietal regions (Hare et al., 2011), suggesting that a less efficient communication between these regions may lead to more stochastic decisions. In keeping with this hypothesis, Polania and colleagues (Polanía et al., 2014) found that phase coherence – an index of coupling between distant regions – between frontal and parietal EEG electrodes in 40-70 Hz frequency band predicted inter-individual variability in preference coherence for food items. In another study (Polanía et al., 2015), the authors perturbed this functional coupling with transcranial alternating current stimulation (tACS) to probe its necessity for coherent preference-based decisions. Subjects were indeed more stochastic in their choices when the stimulation impaired fronto-parietal connectivity (the two regions were stimulated using currents with opposite phases) with respect to the sham condition (i.e., no stimulation) and when stimulation over frontal and parietal regions was delivered with the same phase. Interestingly, anti-phase tACS stimulation appeared to affect trial-by-trial variability in drift rate, while leaving the overall quality of value representation (the mean drift rate) intact. However, the fact that tACS stimulation interacts with the strength of value-based evidence (tACS has greater influence on simpler decisions) is not easy to reconcile with the account proposing a simple value-to-action transformation.

Other theories – known as random utility theory – propose that variability may arise because preference relationships themselves are random (Loomes, 2005) or because the representation of values is stochastic (McFadden, 2001). This latter possibility has been recently linked to the intrinsic limitations in neural resources available for value representation (Polanía et al., 2019) and to the inherent noise in neural activity (Kurtz-David et al., 2019). Experimental findings suggest that variation in neural activity may indeed affect the neural

representation of value, leading to incoherent choices, as it has been observed in vmPFC (Grueschow et al., 2015), but also biasing evaluation, as observed by changes in subjective ratings produced by pre-stimulus activity (Abitbol et al., 2015; Lopez-Persem et al., 2020). On top of spontaneous and externally-induced changes in pre-stimulus activity, the results of my first experiment suggest that part of pre-stimulus activity in vmPFC is specifically linked to the neural monitoring of cardiac signals, a mechanism that putatively supports self-referential processes.

2. Contextual effects

Beyond random fluctuations, systematically inconsistent preferences have been observed in a number of well-defined scenarios. One of the most famous case is the *framing effect* described in the seminal work of Tversky and Kahneman (Tversky and Kahneman, 1981). The authors observed that depending on whether a choice is framed in terms of gain or loss, individuals reverse their preference. Subjects generally select riskier options in loss contexts and safer ones in gain contexts. The influence that the context exerts on value encoding in the brain was nicely captured by a fMRI study (De Martino et al., 2006; **Figure 1.7**). The authors found that the framing effect modulates amygdala activity, which displayed an inverse coding for the sure and uncertain gambles depending on the gain/loss context. Furthermore, inter-individual susceptibility to framing was predicted by activity in vmPFC, a region that – as suggested above – may play a role in integrating different sources of information into evaluation. Ventral striatum, prefrontal cortex and amygdala were also found to be related to the inter-individual difference towards loss aversion in other fMRI and lesion studies (Tom et al., 2007; De Martino et al., 2010). To conclude, preference inconsistency arising from reference-dependent evaluation, as postulated by prospect theory (Tversky and Kahneman, 1981), indeed appears to map on precise neural activity (De Martino et al., 2006; Palminteri et al., 2015).

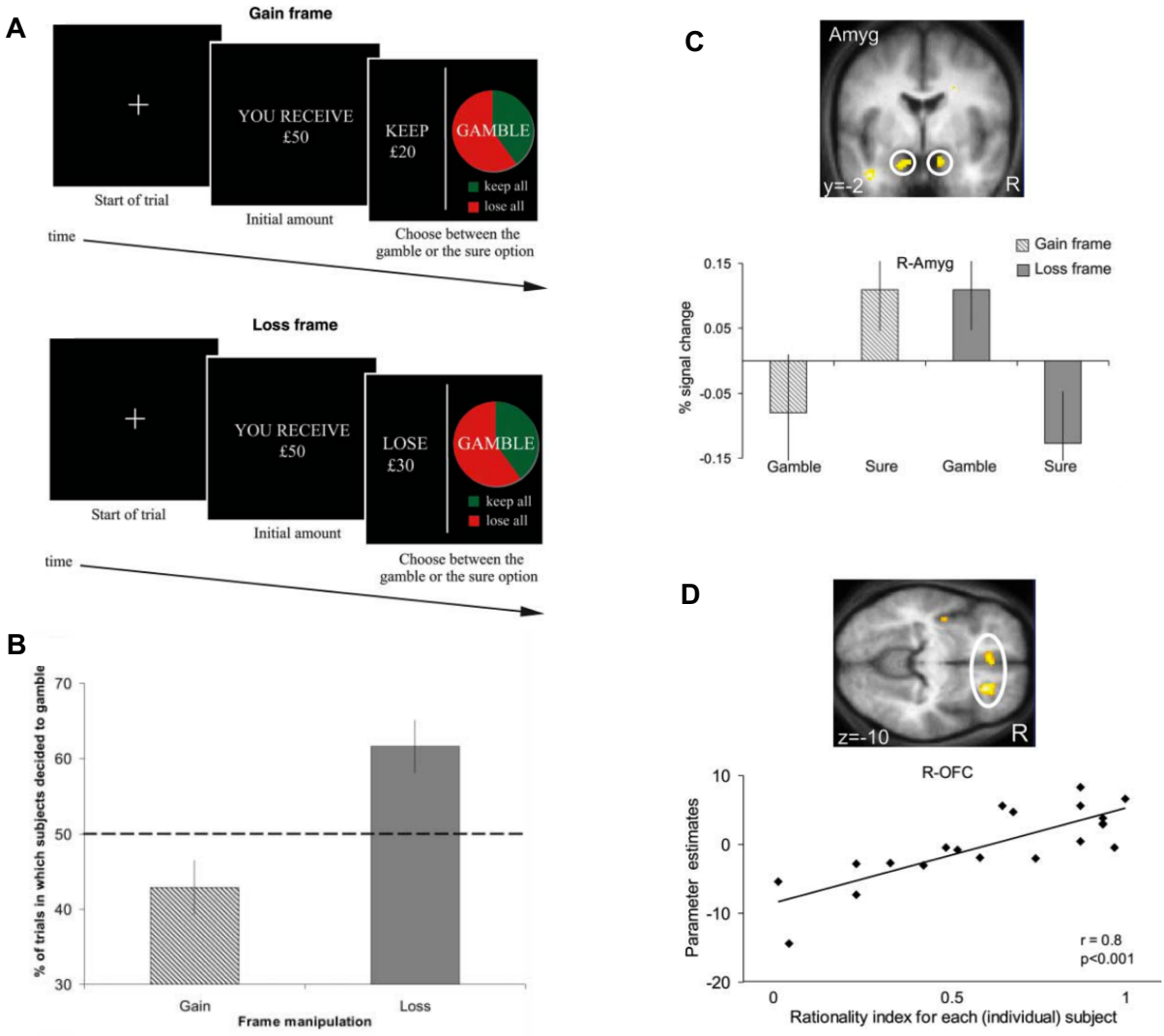


Figure 1.7 Neural underpinnings of framing effect. A, Trial design. Participants choose whether to accept a risky gamble or win/lose a sure amount of money. **B,** Behavioural effects of gamble framing: participants are more likely to gamble when the gamble is framed as a loss. **C,** Inverse coding in amygdala BOLD activity depending on whether subjects take the risky gamble or the sure option in gain or loss frame. **D,** R-OFC BOLD activity correlates with susceptibility to gamble framing. Figure readapted from De Martino et al., 2006.

The availability of additional information also affects evaluation. Plassman and colleagues (Plassmann et al., 2008) asked participants to blindly taste different wines and rate how much they like them. Importantly, before tasting, participants were told the price of each wine. The authors found that price significantly affected pleasantness ratings: wines associated with higher prices were judged more pleasant compared to the same wine presented as cheaper. The increase in pleasantness driven by higher price was mediated by higher activity in medial orbitofrontal cortex (mOFC). Subjective evaluation of a consumable goods thus appears to entail a combination of incoming sensory information, which the authors showed not to be modulated by price knowledge, and top-down influence of externally provided information that may function as a reference point from which evaluation ensues.

The number of options available for decision is another key factor influencing choice consistency. A core principle of choice theory – the normative description of how choices *should* be made – states that adding new options to a choice set should not change preference ordering, i.e. which option you select. However, this axiom, the Independence of Irrelevant Alternatives (IIA; Luce, 1959b), is constantly violated. Imagine that you are deciding between two restaurants that are equally valuable for you: one (*A*) prepares excellent food but has a thirty-minutes waiting list, the other one (*B*) serves regular food but it has a table immediately available. It has been observed that adding a third option, a *decoy* option, that is worse under all aspects with respect to B (serves junk food and there is a 15-minutes waiting list, i.e. it is asymmetrically dominated by B), shifts preferences towards *B*, hence it exerts an *attractive effect* (Huber et al., 1982). Instead, adding a third option that is more extreme than the other two (world-best restaurant for which you have to book weeks in advance) has been shown to increase the relative probability of choosing the one that sits in between the other two (restaurant *A*), an effect known as *compromise effect* (Simonson and Tversky, 1992). Interestingly, these contextual effects were observed in animals such as honeybees and birds (Shafir et al., 2002) (but also in unicellular amoebas! (Latty and Beekman, 2011)) suggesting that these patterns of behaviour may arise from biological constraints.

One such constraint may be imposed by the neural code employed to represent the properties of the environment. Divisive normalization, initially described in cat visual cortex (Heeger, 1992), states that the activity of a neuron selective for a given visual feature is normalized by the overall activity of the surrounding pool of neurons. This operation is useful to optimally bind neural activity within some physiological range. It has been proposed that similar forms of normalisation may occur in value-encoding neurons. For example, by scaling

the neural encoding of individual option value by the sum of values of the available options in a choice set (*divisive normalisation*), Webb and colleagues (Webb et al., 2020) could predict the intransitive patterns of preferences classically observed when the choice set is manipulated. These results ensue from the reduction in discriminability (i.e. difference in neural activity) for two values ($v1 > v2$) when a third value ($v3$), smaller than the other two, is added in a choice set. Note that divisive normalisation does not change the ranking of the available options. To account for a changes in option ranking observed in asymmetric dominance effects, another flavour of normalisation, *range normalization* (Soltani et al., 2012), has been proposed. According to this model, the weight that each attribute has on the evaluation is dependent on the range of presented attributes such that

$$w(i) = \frac{\Delta i}{\Delta \hat{\imath}}$$

where w is the weight of a given attribute, Δi is the range of values for a given attribute in an original choice set and $\Delta \hat{\imath}$ is the range of values for the same attribute after the addiction of a decoy. This implies that the more the decoy is similar with respect to some attribute, the more that attribute weights in the evaluation of each option. The change in attribute's weight can therefore modify the relative preferences participants have. As the authors suggest, different types of normalisation may come into play at different stages of the decision process, for example range normalisation during evaluation and divisive normalisation during action selection.

3. Choice-induced preference changes

The history of recent choices can also shape preferences. When individuals choose between equally preferred items, choosing an item generally increases its value (Brehm, 1956), making it more likely to be selected in the future (Vinckier et al., 2019). Choice-induce preference change is a robust phenomenon found in human adult, children and monkeys (Egan et al., 2010) and it has been thought to reflect the willingness of individuals to reduce the cognitive dissonance between their preferences and their actions (for a critique of the free choice paradigm see (Chen and Risen, 2010)). This interpretation received support from neuro-imaging evidence as well. Multiple fMRI studies (Sharot et al., 2009; Izuma et al., 2010) found that the neural activity in the striatum during a post-decisional rating for items that have been previously rejected was reduced (**Figure 1.8**). Interestingly, choice-induced preference change

seems to be present only for remembered item, indexed both by behavioural reports and neural activity in the hippocampus (Chammat et al., 2017). However, similar mechanisms accounting for preference change have been shown also for items that are not consciously perceived. A recent study (Ai et al., 2018) reported that cueing subjects during midday nap with the name of some snack items enhances the subjective value of these items, a process that appears to be mediated by increased EEG activity in theta and delta frequency range during sleep.

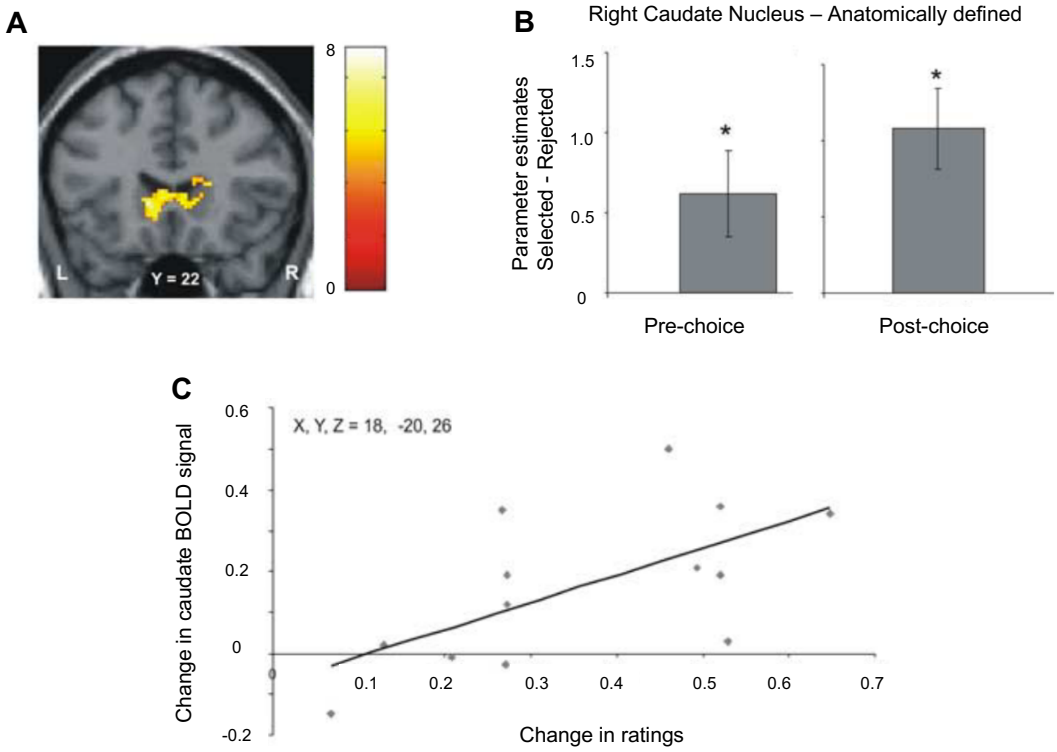


Figure 1.8 Caudate nucleus encodes choice-induce changes in ratings. **A**, Parametric encoding of ratings in caudate nucleus. **B**, BOLD activity in anatomically defined caudate nucleus encodes difference in selected-rejected item before (left) and after choice was made (right). **C**, Correlation between caudate nucleus activity and change in ratings after choice. Figure readapted from Sharot et al., 2009.

Knowing or not knowing *why* an item has been selected also affects its change in value. In an intriguing experiment, participants choose between two strings of symbols with no meaning (e.g., &%$#!*). After having selected one of the two, the true alternatives (travel destinations) were revealed in the same position of symbolic strings (Sharot et al., 2010). During a subsequent rating session, participants rated the item that they "blindly" selected higher.

A recent computational account for this phenomenon (Hornsby and Love, 2020) proposes that choosing an item specifically reinforces the values of its attributes resulting in an enhanced probability of choosing not only that item but similar ones in the future, thus moulding our preference space.

4. Internal factors

Choice variability may also stem from changes in the internal state of the decision maker. Imagine that you ate too much of your favourite food item on one occasion, say chocolate. During the next meal, given the choice between a chocolate cake and a fruit salad, you may choose the latter option, even though – usually – you value it less. This choice would not reflect a 'true' change in preference, rather the integration of the current physiological state in your decision. This behaviour is widely observed in animals, where appetitive responses to obtain food linearly decrease with the satiety level (Bouret and Richmond, 2010), a factor that modulates neural representation of options value in vmPFC (**Figure 1.9**).

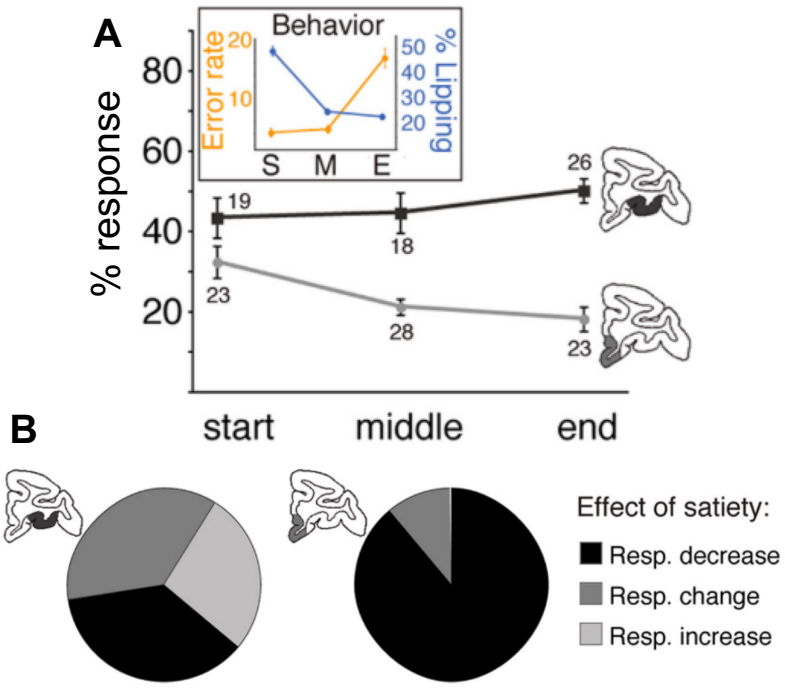


Figure 1.9 Modulation of behaviour and neural activity by satiety level in rhesus monkeys. A. Percentage of reward-selective neuron responses as a function of the elapsed time within a block in OFC (black line) and vmPFC (grey line). Inset shows the change in error rate and lipping behaviour as a function of the elapsed time. **B.** The chart pies show the change in neural selectivity for reward in OFC (left) and in vmPFC (right) following the administration of large bolus of water, indicating how vmPFC (but not OFC) neural encoding of reward is modulated by satiety. Figure readapted from Bouret and Richmond, 2010.

Mood is another factor that influences the evaluation process and affects choice. A recent study showed how positive mood increases the probability to engage in a motor precision challenge (Vinckier et al., 2018) by modifying the evaluation of choice prospects. The authors found that task-induced mood fluctuations affected baseline activity in anterior insula and vmPFC, which in turn predicted the probability of accepting the motor challenge task. Neural activity in these regions was specifically linked to the weights associated with potential gain and loss when evaluating the decision at hand. Although the influence of internal factors on choice is of outstanding importance, its study has received less attention than other "more cognitive" influence.

Summary

- Preferences have been often shown to be inconsistent. Where does variability come from?
- A portion of this **variability may be attributed to noise** or computational imprecisions intrinsic to the decision-making circuitry, and these factors can affect evaluation and option selection, leading to inconsistent and variable choices.
- **Extrinsic factors also produce preference inconsistency**. The way a decision is framed, additional information about the options or a change in the option menu have been consistently shown to change preferences.
- The recent history of one's choice, may also produce shifts in preference, as demonstrated by **choice-induced preference changes**.
- **Intrinsic fluctuations in the internal state of the individual**, for instance related to satiety or mood, can also influence evaluation, leading to transitory changes in preference.

G. Conclusion

Although preferences generally appear stable over time, a growing body of experimental evidence shows that many external and internal factors may transiently affect the evaluation process and hence choice. The effect that these factors exert on evaluation is related to a modulation of neural activity in value-related regions that possibly integrate internal and external sources of evidence.

The novel contribution of my PhD lies in the idea that preference-based decisions require the representation of the agents themselves and that an imprecise self-representation may lead to more variable and inconsistent choices. In the next chapter, I will present conceptual and experimental evidence suggesting that a simple representation of the organism is a necessary building block for evaluation. In addition, I will review brain-imaging evidence showing that the network supporting value-based decisions overlaps with brain regions involved in many cognitive processes supporting self-representation indicating that beyond co-localisation, these two processes may be functionally coupled. The functional link between self-representation and evaluation can thus account for preference stability over time: preferences are the reflection of decision-maker's identity.

Chapter II: From evaluation to self-representation. An embodied perspective of subjective value

The emerging field of neuroeconomics has extensively investigated the neural underpinnings of value and decisions and how fluctuations in value-related neural signals may produce variable and suboptimal behaviour. As described in *Chapter I*, neural and cognitive misrepresentations of stimulus value have been often suggested as sources of decision variability and suboptimality. However, values are not found in nature as such, rather they are inferred based on stimulus properties, external contexts and current internal needs and goals. This crucial decisional step is often subsumed under the term *evaluation* (Rangel et al., 2008). For example, the exact same hamburger may appear unattractive after a copious Sunday family lunch or succulent after an 8-hour fast. Although, changes in current states have been rightly highlighted as important factors of choice variability (Loewenstein, 1996), the ability of maintaining the representation of current states – *in itself* – has received far less attention. Think of an animal that is impaired in representing its current energetic needs, this reduced capacity will impact the way it evaluates food options, possibly driving aberrant decisions.

In this chapter, I will propose that a reliable representation of oneself is a necessary building block of the decision process. More specifically, I shall address how the capacity to **refer a set of available options to oneself**, in other words put them in relation to one's current state, needs and goals enables the evaluation of the available alternatives and influences choice coherence. To this aim, I will review some recent conceptual and computational modelling work showing how representing one's current state with respect to some goals is fundamental to ascribe values to available options and how this framework can account for some choice patterns observed in the literature. At the neural level, self-referential representations appear to be tracked in vmPFC, a region consistently involved in value-based decisions. The pivotal role of vmPFC in encoding more conceptual aspects of self-relevant information has been also consistently reported. Neural and conceptual evidence thus suggests that evaluation and self-representation are two cognitive processes tightly intertwined, sharing common neural resources. This consideration opens the possibility that some stochasticity in value-based decisions may stem from imprecise self-representation.

1. The representation of the current state as building-block of evaluation

Taking inspiration from previous homeostatic models of motivated behaviour (Hull, 1943; Keramati and Gutkin, 2014), Juechems and Summerfield (Juechems and Summerfield, 2019) propose that when an agent is choosing between two (or more) goods, she is seeking to restore her current imbalance among different needs (**Fig. 2.1A**). This operation is performed by reducing the distance between her current state and a given setpoint she wants to achieve, corresponding to some equilibrium across different needs. Imagine an animal that is severely thirsty and only a little hungry (**Fig. 2.1B**). It will ascribe more value to water than to a food item, as drinking will reduce thirst (the strongest need) bringing the animal closer to the desired setpoint. Once thirst is appeased, the value of water will be reduced and, if hunger is still present, the value of a food item increased. The value of an option thus emerges from a simple computation: the value of an alternative equals to the reduction in distance between the agent's current state and her goal state after the consumption of that option. This framework thus contrasts with standard choice theories, which consider decisions as a reward maximizing procedure among rewards that are already given.

Since decisions are dictated by the state and needs of the agent, which can be numerous, the value-space in which the decision occurs is itself multi-dimensional. This characteristic entails precise consequences for less trivial decisions. Consider now an animal that is equally thirsty and hungry and that has to decide between two options: a mixture of water and food (100 ml of water and 1 food pellet) or a larger amount of food (2 food pellets). Since the animal's goal is to reduce its distance from a setpoint in which it is neither hungry nor thirsty, it will prefer the option with a mixture of goods (**Fig. 2.1C**). In fact, choosing a larger amount of food would have moved it closer to a state in which it will be sated, but it will have worsened its thirst. Interestingly, this same mechanism can account for more complex patterns of economic decisions, as the ones observed in consumer's behaviour. In *Chapter I*, I presented what is known as the *compromise effect* (Simonson and Tversky, 1992): when an extreme option is added to a choice menu (42-inch TV that costs 1000 Euros) between two equivalent alternatives (30-inch TV at 500 Euros and 19-inch TV at 200 Euros), people tend to select the option that lies in the middle between the two extremes (30-inch TV at 500 Euros). This pattern of choice can be easily explained by the framework presented here. As the agent seeks to optimise across states and not rewards, she will prefer to have the largest TV possible spending

as little as she can. In a similar vein, the model can explain other counter-intuitive everyday behaviours. Despite the fact that money can be infinitely accumulated, humans prefer to stop working for money once a wealth setpoint is achieved. Cab drivers in New York, for example, appear to work until a pre-established daily earning goal is achieved, even if this means to give away a large gain due to favourable circumstances (such as increased consumer demands and higher fares (Camerer et al., 1997)). Cab drivers seem therefore to comply with the "stopping rule" prescribed by this homeostatic framework: instead of maximising reward, they maximise wellbeing, composed of a combination of money and free time.

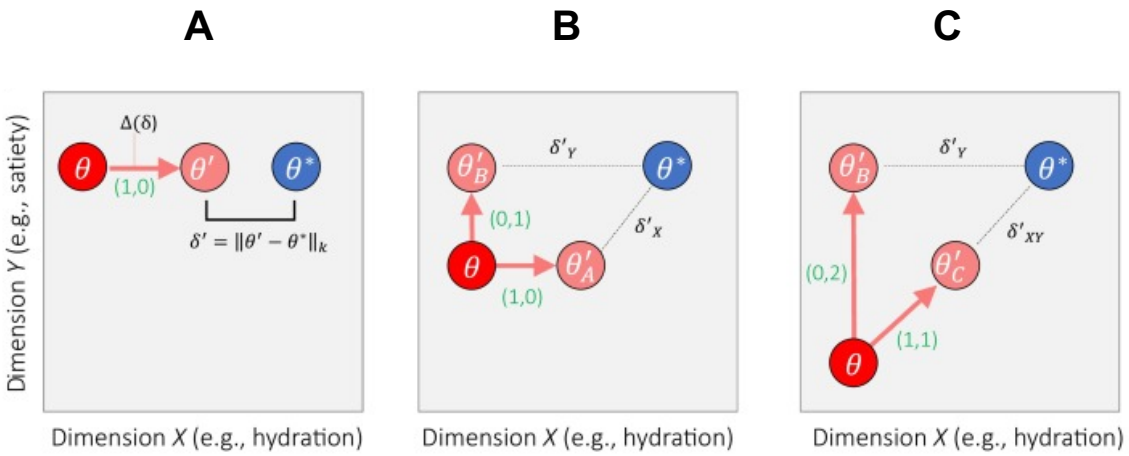


Figure 2.1. Value computation in a two-dimensional space. A, The value of a given option is computed as the distance reduction Δδ between agent's next state θ' (should that option been selected) and the setpoint θ* and the set point and her current state θ. The factor *k* indicates how much an agent prefers single asset (*k*<1), compounded bundle (*k*>1), or is indifferent to bundle composition (*k*=0). **B,** When the agent is more thirsty than hungry, an offer of 1 unity of water is more valuable than 1 unity of food. **C,** When the agent is as thirsty as hungry, a bundle of goods composed of 1 unity of water and 1 of food (1,1) will be more valuable than an offer of 2 unities of food (0,2), as the former option will approach her more to her desired setpoint θ*. Modified from Juechems and Summerfield, 2019.

2. The neural representation of current states

A model of decision that integrates the current state of an individual into the evaluation process seems therefore to account for a wide range of behavioural phenomena. How can this model be implemented biologically? Which neural substrates can support the tracking of the current (internal) state so that future behavioural policies can be adapted to the achievement of some internal goal? Juechems and co-workers used fMRI (Juechems et al., 2017) to investigate this matter, looking at how cumulative wealth influences decision policies. In this experiment, participants were asked to decide whether to accept a monetary gamble, in which they could either win or lose some money, or reject it incurring in a certain small loss. The authors found that as participants accumulate more and more money, they become more risk averse, and that BOLD activity in vmPFC tracked the latent accumulation of wealth. In opposition to the reward-maximising view of decisions, which predicts that risk-aversion should remain stable across time and hence that decision policy should not change, these results show that participants keep track of their accumulated level of wealth to avoid unnecessary loss and to ensure enough monetary resources.

Using a completely different paradigm, Schuck and colleagues (Schuck et al., 2016) revealed that vmPFC also represents the current position of an agent with respect to task-space, an extrinsically imposed goal, thus enabling adaptive behaviour. In this study, participants judged the age of visual stimuli composed of a superposition of a face and a house. The category (face/house) on which participants had to perform the age judgement was determined by what happened in the previous trial. A change in age of the same stimulus judged in the previous trial (young face$_{(t-1)}$ → old face$_{(t)}$) signalled participants they had to switch stimulus category on the next trial (at t+1 they should judge the age of houses). Using MVPA analysis, the authors revealed that BOLD activity in vmPFC encodes the latent information necessary to decide on which stimulus to perform the age judgement, that is the category and the age of the stimulus judged in the previous trial. Altogether, this body of experimental evidence suggests that vmPFC is involved in the representation of the agent's current state and her 'position' with respect to some goals.

3. From homeostatic to cognitive self-representations

Self-representation thus appears to be a fundamental step for evaluation: it allows an agent to confer a subjective value to the available options in their potential to bring the agent closer to a desired goal. To infer the reduction in distance that choosing an option produces, it is necessary to represent not only the long-term goal, but also one's current state, with all the factors that may contribute to it: physiological states, cognitive beliefs etc. In the first section, I described self-referential computations that exclusively concerned the physiological state of the individual (hunger, thirst). This simplification has been useful to portray the intrinsic relationship between evaluation and self-representation. However, self-representation also has important consequences in defining the goals an agent wants to achieve. Imagine that you are having dinner at your in-laws' and that after the main dish you are still hungry. Although in another context you would go for seconds, no other guest does and, as you believe that behaving politely at your in-laws' is more important than to be sated, you prefer not to ask. This shift in goals, and thus in the values ascribed to the different courses of action to achieve them, illustrates the fundamental role that self-representation has in guiding our behaviour. **Values only arise when they are in relation to an agent. In other words, the subject has to consider the available options from her own perspective and in relation to herself**.

This view thus encourages us to consider other origins of decision variability. On the one hand, choice inconsistency may arise from the inherent noise characterising neural activity that affects both option selection – thus generating 'lapses' – and value-related computations. Noisy computations have been already shown to affect perceptual decisions (Drugowitsch et al., 2016) and value-based learning (Findling et al., 2019) suggesting that they may be similarly present in preference-based decisions. In keeping with the idea presented in this chapter, a complementary source of **choice variability in value-based decisions may arise from an imprecise self-representation, which in turn affects evaluation**. A mechanism of this type may explain pathological disorders, which are characterised by impaired self-related processing.

In the remainder of *Chapter II*, I will focus on the cognitive dimensions associated with other forms of self-representation and show how all of them appear to recruit vmPFC. But how do homeostatic and more conceptual aspects of self-representations relate to each other? Although experimental data remain scarce, in *Chapter III*, I will review a recent stream of studies showing that diverse self-related cognitive processes are functionally coupled to the

neural monitoring of homeostatic signals suggesting that this mechanism can bridge multiple levels of self-representation.

Summary

- Values are not readily available in nature.
- The value of an option emerges from computing how much selecting that option will bring the agent closer to her desired goal.
- To perform this computation, a representation of one's current state is essential.
- This computation is a self-referential process: it is the capacity to relate the available options to oneself.
- Experimental results suggest that vmPFC can represent the current state of an agent.

4. Cognitive components of self-representation and their neural underpinnings

Self-representation is a complex cognitive process encompassing explicit and implicit components, and it has been often conceptualized as the study of the 'self as object of knowledge', in contrast to 'self as experiencing subject' (Christoff et al., 2011; D'Argembeau, 2013). Different experimental paradigms have been used to investigate the multiple dimensions associated with self-representation. Notably, all dimensions appear to commonly recruit vmPFC and suggest that an impaired self-representation, reflected by an abnormal activity in vmPFC, may underlie maladaptive behaviour in neurological and psychiatric disorders.

A. Self-attribution

One of the most widely used experimental paradigms investigating the neural substrates of explicit self-representation requires participants to judge whether some trait adjectives describe themselves or someone else. This type of paradigm consistently revealed the involvement of vmPFC and PCC (van der Meer et al., 2010). However, vmPFC does not seem to be uniquely involved in self-related judgement. A meta-analysis conducted on 107 fMRI studies revealed that vmPFC displays a gradient going from self-, more ventrally, to other-, more dorsally, related judgments (Denny et al., 2012). The spatial segregation between self- and other-related representations may be mediated by the similarity between oneself and others. In line with this possibility, fMRI results show that vmPFC activity appears to be more similar for self- and other-related judgments when participants judge personality traits of uniquely close individuals, such as best friends (Benoit et al., 2010).

Physical traits and mental states also contribute to the representation of oneself and judgements about these subjects appear to recruit vmPFC as well (Jenkins and Mitchell, 2011). Personality and physical traits are inherent characteristics that each individual has learnt to associate with herself over a long period of time. Do newly formed associations recruit the same neural resources? To address this question, some experimental paradigms require participants to associate abstract stimuli, such as geometric shapes or fractals, with themselves and compare these associations to those developed for someone else. People are generally faster and more accurate to categorise stimuli associated with themselves than stimuli

associated with others (Sui et al., 2012). This behavioural effect suggests that self-representation is flexible and that items that become part of it are differently processed, possibly prioritised. At the neural level, this effect is mediated by BOLD activity in vmPFC (Sui et al., 2013; Lockwood et al., 2018), suggesting that this region is more generally involved in the representation of information referred to oneself. The contribution of vmPFC in the organisation of self-relevant information is further confirmed by neuropsychological studies showing that patients with vmPFC and insular damage do not show the typical behavioural advantage observed for items associated with themselves (Sui et al., 2015).

B. Self-monitoring and self-beliefs

Self-representation also includes the ability of estimating the correctness of one's choices. This capacity has great importance to guide future behaviour, for example enabling to postpone a decision until more evidence is gathered or to change one's mind after an initial commitment (Resulaj et al., 2009; Desender et al., 2018). This cognitive ability is generally investigated by asking participants to overtly report their subjective confidence in a decision, a quantity that multiple meta-analyses showed to be encoded in vmPFC BOLD activity (Morales et al., 2018; Vaccaro and Fleming, 2018). The same region has been also shown to track trial-by-trial confidence levels, in absence of subjective overt reports, as predicted by latent computational variables (Lebreton et al., 2015; Bang and Fleming, 2018). The subjective feeling of confidence that accompanies individual decisions contributes to a general belief about one's ability in a given task (Rouault et al., 2019). Since vmPFC encodes self-relevant information, it is logical to hypothesise that this region is well suited to integrate local self-estimates about task performance into a more general self-belief, thus contributing to another important dimension of self-representation.

C. Autobiographical memory, and future scenarios

Personality traits and present concerns are not the only contents of self-representation. What happened in the past is another important factor that shape one's identity and can provide useful information to cope with novel situations. Memories of personal experiences, knowledge about one's personal life and the conceptual descriptions of one's identity are three other dimensions intimately linked with self-representation. Using activation likelihood estimates (ALE) to analyse 38 fMRI and PET studies, Martinelli and colleagues (Martinelli et

al., 2013) found that all three of these self-related dimensions commonly recruit vmPFC, while semantic and episodic memory preferentially engage PCC.

Although slightly more anteriorly, vmPFC is also involved when participants project themselves in future scenarios as compared to scenarios that they are not part of (Abraham et al., 2008).

This evidence suggests that vmPFC does not represent specific aspects of the self, rather it contributes to a construction of self-representation integrating multiple sources of information, such as semantic knowledge and episodic memory, enabling individuals to foresee the pleasantness of distant future scenarios (Benoit et al., 2014) and plan their behaviour accordingly.

D. Spontaneous thoughts

Up to now, I have reviewed studies investigating representations of the self in which participants are exogenously asked to focus on particular aspects of themselves or to imagine themselves in past and future scenarios. However, self-representation can occur spontaneously, during periods of rest in which self-related thoughts may naturally arise (Vanhaudenhuyse et al., 2010; D'Argembeau, 2013). This process can potentially account for why two regions of the default network (DN), vmPFC and PCC, are more active during periods of rest as compared to when participants are performing an experimental task (Buckner et al., 2008). Although the functional relation between self-representation and DN is still unresolved, some authors (Qin and Northoff, 2011) proposed that the interaction between spontaneous brain activity and stimulus-related activity may underlie the very notion of self, considered as a relational concept between internal and external components.

E. Self-representation in psychiatric disorders

Impaired self-representation mediated by vmPFC activity has been also proposed to underlie deficits in specific neuropsychiatric disorders. For example, up to 80% of schizophrenic patients are characterised by reduced insight about their clinical condition and mental processes (van der Meer et al., 2013). This impaired self-representation has been related to symptoms' severity and treatment compliance. Interestingly, BOLD activity in vmPFC during self-reflection has been shown to predict patients' 'cognitive insight' (Meer et al., 2012),

providing a putative link between neural activity, self-representation, and psychotic symptoms. Alzheimer's patients unaware of their condition also show reduced activity in vmPFC (Amanzio et al., 2011), suggesting that vmPFC may indeed be important for maintaining an accurate self-representation.

An impaired self-representation may produce distorted evaluation of what should be considered as self-relevant, triggering harmful behaviour akin to what is observed in drug addiction (for review, see Moeller and Goldstein, 2014). The imbalance in evaluating self-relevant stimuli is mirrored by vmPFC activity in addicted individuals: BOLD activity is enhanced for drug-related cues (Goldstein et al., 2009; Kühn and Gallinat, 2011) but weaker for other types of self-relevant stimuli not related to drugs, such as error feedback (Moeller et al., 2014).

Altogether, brain-imaging studies in healthy and clinical populations suggest that **vmPFC is a key region for maintaining a self-representation based on multiple types of information, including past experiences, conceptual knowledge about oneself and beliefs about one's ability**. These results thus generalise vmPFC function beyond more implicit and restricted forms of self-representation such as the encoding of an agent's cumulative wealth (Juechems et al., 2017) and relative position in task-space (Schuck et al., 2016). In addition, impairment of this representation may account for the severity of psychiatric symptoms as well as aberrant behaviours.

As noted before, the term vmPFC is often used in the literature to refer to many disparate sub-regions on the medial surface of PFC and the self-related literature reviewed here is not an exception. Despite this liberal usage of the term, many of the studies reviewed above point to a consistent overlap at least in Broadmann areas 10, 24 and 32 (**Fig 2.2**).

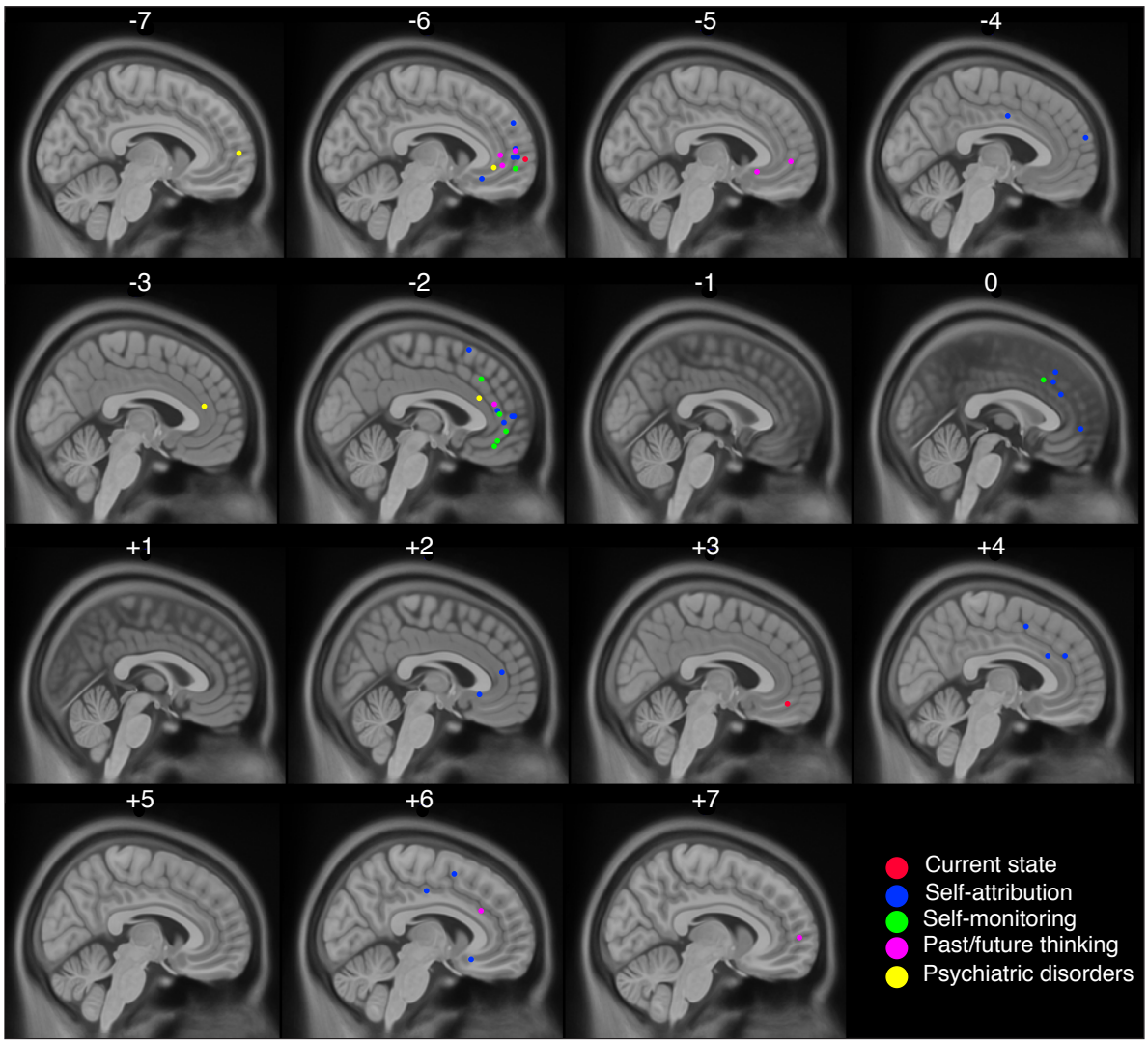


Figure 2.2. vmPFC activity associated with various aspects of self-representation in healthy and clinical population (yellow). Points represent peaks of activation drawn from the studies cited in this chapter (**Table 2.1**). Numbers above MRI sagittal slices indicate the x coordinate in MNI space.

Table 2.1. List of studies represented in **Figure 2.2**

Self-related aspect	*Study*	*Contrast/Parametric encoding*
Current state	Juechems et al. 2017	Parametric encoding accumulated evidence
	Schuck et al., 2016	Hidden states task-space
Self-attribution	Van der Meer et al., 2010	Self > baseline & self > other
	Denny et al., 2012	Self > baseline & self > other
	Benoit et al., 2010	Interaction (self > other)*similarity
	Jenkins and Mitchell, 2011	Self > other
	Sui et al., 2013	Self > other
	Lockwood et al., 2018	Self > other
Self-monitoring	Morales et al., 2018	Confidence classification
	Vaccaro and Fleming, 2018	Parametric encoding of confidence
	Lebreton et al., 2015	Parametric encoding of confidence (quadratic effect)
	Bang and Fleming, 2018	Expected confidence
Memory and future scenarios	Martinelli et al., 2013	Conjunction analysis: episodic memory, semantic memory and conceptual self
	Abraham et al., 2008	Personal > non-personal
	Benoit et al., 2014	Parametric encoding of pleasantness
Psychiatric disorders	van der Meer et al., 2013	Self > semantic associated with cognitive insight
	Amanzio et al., 2011	aware > unaware
	Kühn and Gallinat, 2011	ALE nicotine and alcohol craving
	Goldstein et al., 2009	Control > Cocaine addicted

Summary

- Self-representation involves multiple cognitive facets: attribution of personality and physical traits, beliefs about one's abilities, episodic and semantic knowledge about oneself.
- These multiple aspects of self-representation involve midline cortical structures, particularly PCC and vmPFC.
- Neurological and psychiatric disorders such as addiction, schizophrenia and Alzheimer's disease, characterized by impairment of self-representation, are associated with abnormal activity in vmPFC.

5 Conclusion

In this chapter, I have reviewed the conceptual and computational principles showing how **representing one's current state (e.g. the physiological state) is a fundamental step enabling evaluation.** Beyond homeostatic information, more conceptual aspects constitute the self-representation that may guide decisions. These conceptual aspects may be particularly relevant for preference-based decisions that dot not satisfy direct physiological needs, such as deciding whether to spend the vacations at the seaside or in the mountains. Interestingly, both types of self-related representation recruit vmPFC, a key region of the evaluation network.

To date, how homeostatic and conceptual self-related components relate to each other and how they contribute to value-based decisions has largely remained unexplored. In my PhD**, I propose that the neural monitoring of homeostatic signals, and in particular the cardiac ones**, **may provide a common self-specifying reference signal through which different types of self-relevant information are combined. The coordination of these diverse sources of information thus grounds the stable self-representation necessary to guide subjective evaluation. This hypothesis is developed in *Chapter III*.**

Chapter III: The neural monitoring of cardiac signals as a grounding mechanism for the self

A given form of self-representation is necessary to guide the subjective evaluation at play during value-based decisions. This necessity is evident for decisions involving primary needs: the value of a glass of water relative to a piece of bread depends on the current levels of thirst and hunger of the decision-maker. Value-based decisions that do not carry any direct consequence for the agent's survival also entail self-representation: choosing between going to the theatre or staying at home watching TV requires estimating which activity is more valuable given the unique subjective perspective and interests of the decision-maker. Despite the relevance of self-representation, its underlying mechanisms and its contribution to value-based decisions have been largely overlooked.

In the current chapter, I propose that **self-representation is rooted in a simple biological process**, **the neural monitoring of visceral signals, and more specifically cardiac ones**. This mechanism may provide **a common reference signal enabling** the mapping of diverse information into **a self-centred representational space generating the unique subjective perspective whence it is possible to evaluate the available options**.

To support this hypothesis, I shall review how different theories have related diverse aspects of subjective cognition to bodily signals and why ascending cardiac signals are particularly well suited to support a self-specifying mechanism. Next, I will detail how it is possible to experimentally investigate the neural monitoring of cardiac signals, where these signals arise from, and which information they are believed to reflect. Finally, I shall review recent experimental evidence showing the functional relationship between the neural monitoring of cardiac signals and different aspects of subjective cognition.

Before diving into these matters, I would like to draw the attention of the reader on two points. First, in this chapter, as for my PhD experimental work, I will specifically focus on the neural monitoring of **ascending cardiac signals**, and I will not consider the descending regulation of these signals by the brain. This is because self-representation, as we hypothesised, is based on a neural mechanism (i.e., the neural monitoring of cardiac signals) and it is linked to, but independent of, changes in cardiac activity triggered by descending influence of the

brain. Second, much of the conceptual scaffolding and literature exposed in this chapter capitalises on a recent review paper to which I contributed (Azzalini et al., 2019).

1. Ascending bodily signals

A. Changes in bodily states and affective experience

The idea that bodily signals are related to various aspects of subjective experience is certainly not novel. From the early days of Psychology, bodily signals have been recognised to be tightly intertwined to one of the most ubiquitous subjective experiences: emotions (James, 1890). Despite their intuitive relationship, the role of bodily signals in emotional experience has been long debated. For some authors (James, 1890), changes in bodily states triggered by external stimuli are the cause of emotional feelings: it is because the sight of a swift shadow moving in the dark made your heart pound that you experience fear. Conversely, other authors consider that it is the fact that you are experiencing an emotion that causes changes in the bodily states (Cannon, 1927). To date, this debate is still unresolved.

The central role of bodily signals in affective experiences has prompted Damasio to consider that the brain representation of bodily changes constitutes the most basic representation of the self, called the *material me* (Damasio, 2003a, 2003b). More complex components based on memory and language join this primordial construct to produce the extended notion of the self underlying the concepts of personhood and individuality. The idea that the integration of bodily states is an essential building block of subjective experiences has been more recently revisited by Craig (Craig, 2009). In his view, signals arising from the body are combined with environmental and cognitive information in a hierarchical fashion in the insula (through a posterior-to-anterior gradient). This integration would in turn produce a *global emotional moment* generating a meta-representation of the individual at a specific point in time that enables the existence of a sentient self.

However, not all bodily signals are equal. In the following section, I will describe why visceral signals coming from the heart and the gastro-intestinal tract enjoy a special status and why they may fulfil a broader role in subjective cognition, well beyond affective experiences.

B. Visceral signals: the heart and the gastro-intestinal tract

The heart and gastro-intestinal tract (GI) feature four main characteristics that set them apart from other bodily signals (Park and Tallon-Baudry, 2014; Tallon-Baudry et al., 2018; Azzalini et al., 2019):

1. **The monitoring of visceral signals is one of the core functions of the central nervous system.** Efficiently detecting deviations in the homeostatic balance concerning these signals allows the organism to rapidly find solutions to avoid critical states. The neural monitoring of these signals thus provides important vital information about the integrity of the organism as a whole.
2. **Both the heart and the GI tract are autonomous electrical pacemakers producing continuous activity**. The oscillatory electrical activities are created within these organs by pacemaker cells, in the sinoatrial nodes for the heart, and by the interstitial cells of Cajal along the entire GI tract. Descending influence from the central nervous system can modify the pace of the contractions, and it is generally a sign of health. However, these descending signals are not necessary for initiating the rhythmic activity of the heart and GI tract.
3. **The ascending cardiac and GI information** concerning the timing and the strength of contractions is **constantly relayed to a distributed network of regions involving cortical and subcortical structures**. A more detailed explanation of neural pathways and cortical hubs able to receive cardiac information will be provided below.
4. **Cardiac and GI oscillatory rhythms are compatible with the time-scale of cognition**. A cardiac cycle takes normally around 800ms, while the gastric one has a slower pace of 1 cycle every 20s.

C. Grounding the self in visceral signals

These characteristics may provide a general biological plausible mechanism accounting for two main features of *all* subjective experiences (hence not only – or primarily – emotions): **continuity** and **unity**.

The incessant activity of the heart and GI tract is continuously monitored by the brain making these signals particularly apt to provide a **biological backbone to subjective experience, which is characterised by temporal and spatial continuity**. Concerning self-

representation more specifically, these signals may underlie the perception of oneself as being one stable entity persisting in time and space. In line with this view, Park and Blanke (Park and Blanke, 2019a) proposed that visceral signals (encompassing ascending information from the heart, the stomach and respiration), along with other proprioceptive, vestibular and visual bodily inputs, underlie bodily self-consciousness. Bodily self-consciousness consists in the ownership and the identification of an individual with a body localised in time and space, and it has been proposed as a fundamental component of the *minimal phenomenal selfhood*. This simplest pre-reflective form of self-consciousness may also enable first person perspective, and thus subjective experiences, to arise (Blanke and Metzinger, 2009).

Visceral information appears to reach multiple cortical and subcortical regions. Park and Tallon-Baudry (Park and Tallon-Baudry, 2014) proposed that this **distributed neural representation** may act **as a self-specifying common reference signal, a biological anchoring point for all subjective experiences.** The rationale behind this proposal is that conscious subjective experiences arise from a combination of two elements. First, conscious experience needs to be *about something*, a content – e.g., the external world or the subject of the experience herself. Second, it needs the reference to an *agent*, an 'I'. The neural monitoring of visceral signals would instantiate the biological root of the latter component, the experiencing subject, the agent endowed with a unique first-person perspective (Tallon-Baudry et al., 2018). More recently, a more mechanistic account of how first-person perspective may arise has been proposed (Azzalini et al., 2019). The distributed neural monitoring of visceral signals may allow the production of **a common representational space through which the information expressed in sensory- and domain-specific formats can be coordinated and mapped onto**. This common self-specifying representational space may hence underpin the **unity of subjective experience (Figure 3.1A)**.

D. Comparing different theoretical frameworks

One useful way to differentiate the proposals presented above is to group them according to *what is*, with respect to bodily signals, that enables (given aspects of) subjective experience to arise. On the one side, James, Damasio (specifically in his somatic marker hypothesis, cf. Bechara and Damasio, 2005), and Craig base subjective experience on **changes in bodily signals (Figure 3.1B)**. This perspective has an intuitive appeal for emotional experiences, in which feelings are tightly linked to physical changes (for instance, the heart

beating faster when being fearful, the knot in the stomach when feeling anxious). However, the contribution of bodily signals to subjective experiences that do not necessarily entail bodily changes, for example the conscious experience of having seen a stimulus, is far less clear. The theoretical frameworks proposed by Blanke (Blanke and Metzinger, 2009; Park and Blanke, 2019a) and Tallon-Baudry (Park and Tallon-Baudry, 2014; Tallon-Baudry et al., 2018; Azzalini et al., 2019) offer a solution to this problem, by proposing that **the neural monitoring** of bodily and visceral signals ***per se*** is the condition for subjective experience to arise. As it shall become apparent later, my PhD work also falls under this second perspective.

Blanke (Blanke et al., 2015; Park and Blanke, 2019a) considers that **visceral and other bodily signals contribute to the production of the bodily self**, i.e. the identification of an individual with a body localised in time and space. In turn, this embodied self generates the first-person perspective that grounds subjective experiences (**Figure 3.1C**). Tallon-Baudry (Park and Tallon-Baudry, 2014; Tallon-Baudry et al., 2018; Azzalini et al., 2019) proposes that **first-person perspective is the direct product of the neural monitoring of visceral signals**, which in turn underlies different dimensions of subjective experience such as perceptual and emotional experiences as well as self-consciousness, of which bodily self can be thought as one aspect (**Figure 3.1D**).

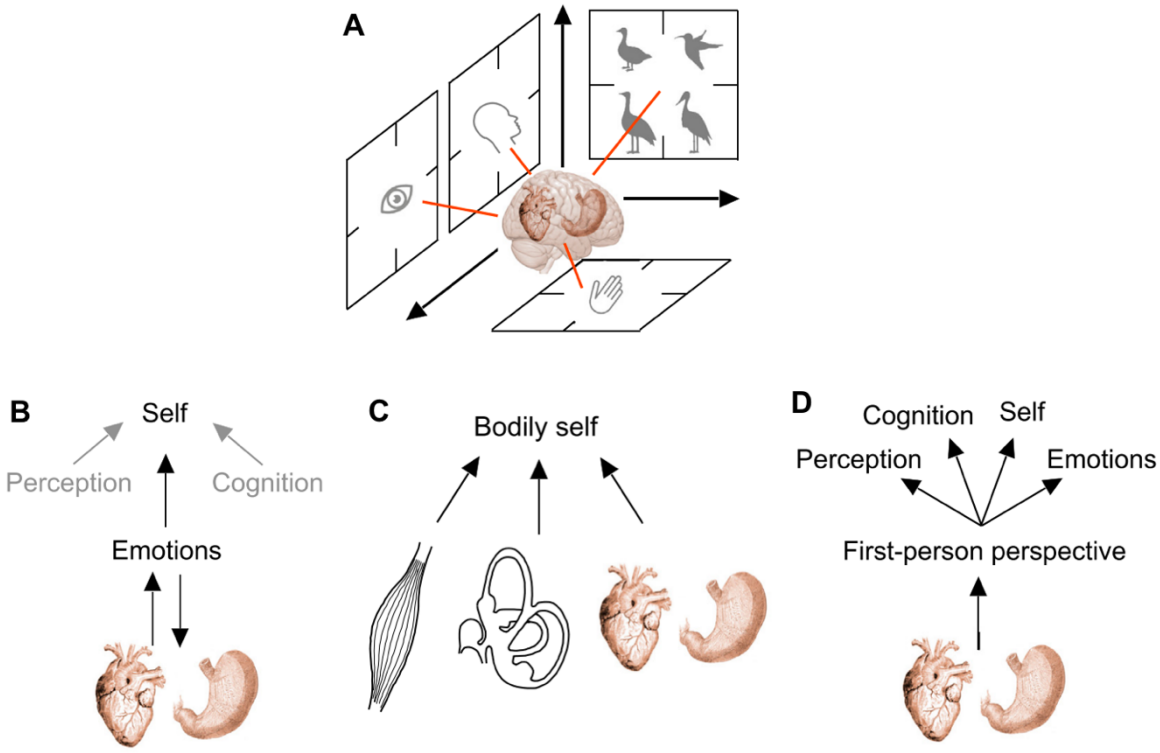


Figure 3.1 The functional roles of visceral signals. A, Visceral information acts as a common reference signal which allows the coordination of different perceptual and conceptual spaces. This neural mechanism potentially accounts for the unity of subjective experience. **B**, Theoretical frameworks such as those proposed by James, Damasio and Craig state that bodily and visceral signals primarily contribute to emotional experience, which in turn is the main substrate for the self. **C**, Other authors (Blanke et al., 2015; Park and Blanke, 2019a) consider visceral signals, along with other types of exteroceptive and proprioceptive signals, as a grounding mechanism for bodily self. **D**, Tallon-Baudry considers first-person perspective, grounded in the neural monitoring of visceral signals, as a necessary condition for all other aspects of subjective cognition to arise. Readapted from Azzalini et al., 2019.

E. Implicit and explicit interoception

The mechanisms by which the brain tracks visceral signals are generally subsumed under the term **interoception**. However, it is important to note two distinct meanings of this term. On the one side, interoception refers to the study of the **conscious experience of bodily signals**. Although most of the time we do not experience our heart beating, the rhythm of our breath, nor the activity of the stomach, occasionally salient changes in their activities enter conscious experience to trigger some actions. For instance, when we perceive the activity of the stomach because we are hungry. The conscious experience of this signal may prompt us to prepare a meal, to avoid starvation. In laboratory experiments, **explicit interoception** is

investigated by overtly asking participants to pay attention to some bodily signals, for example silently counting their heartbeats without taking their pulse (Critchley et al., 2004). These tasks assume that attention enhances the neural processing of cardiac signals, thus increasing the activity of interoceptive regions. However, there are rising concerns about the validity of such paradigms to uncover the neural bases of interoception (Brener and Ring, 2016; Ring and Brener, 2018; Zamariola et al., 2018). The major critiques about this task claim that participants use their previous knowledge to accomplish the task rather than using online interoceptive signals. In addition, the reliability of this measure to test interoceptive ability appears weak. These problems make it questionable that similar paradigms can provide a window onto the neural processes of interoception.

On the other hand, interoception may refer to the **automatic and unconscious monitoring** of visceral signals. This process can be studied by using brain-imaging techniques able to capture neural responses to physiological events, such as heartbeats, while subjects are resting or are performing unrelated cognitive tasks. The automatic and unconscious neural monitoring of visceral signals is the type of interoceptive process that I investigated in my PhD. However, before presenting how it is possible to record the neural monitoring of heartbeats and how this latter relates to different cognitive processes, I will review some basic properties of heart's physiology and on the transmission of cardiac signals to the brain.

Summary

- Bodily signals have been long considered important contributors to subjective experiences, ranging from affective experience to distinct aspects of selfhood.
- Visceral signals, coming from the heart and the gastro-intestinal tract, have a special status: they are autonomous pacemakers, whose ever-present activity, compatible with the time-scale of cognition, appears to reach a distributed network of regions.
- The neural monitoring of visceral signals provides a basic representation of an agent as living organism.
- The neural monitoring of visceral signals can provide an anchoring, self-specifying, signal to which subjective experiences (sensations, thoughts and decisions) can be referred.
- This binding mechanism may be grounded on the coordination of information expressed in sensory- and conceptual-specific formats through the neural monitoring of visceral information.
- The resulting common space may give raise to the unique perspective, the first-person perspective, inherent to all subjective experiences.
- The neural processing of visceral signals is often subsumed under the term of interoception. However, interoception can refer either to the conscious perception of one's visceral states or to the unconscious and automatic processing of these signals. The latter acceptation is the main interest of my PhD work.

2. Heart-brain coupling: physiology and neural pathways

A. The heart and the cardiac cycle

The heart generates its own intrinsic electrical rhythm, initiated by pacemaker cells in the sinoatrial node (SA), in the right atrium. These electrical impulses travel toward the atrioventricular node and, through the bundle of His, they trigger muscular contraction in the ventricles. Heart electrical activity can be recorded noninvasively by means of electrocardiography. The resulting electrocardiogram (ECG) is characterized by multiple waves corresponding to different phases of the cardiac cycle (**Figure 3.2A**). Atrial and ventricular contractions correspond to the P-wave and QRS complex, respectively, while the onset of ventricular relaxation is marked by T-wave. The length of a cardiac cycle, generally around 800 milliseconds, is often computed as the interval between two consecutive R-peaks (also known as inter-beat interval, or IBI). The cardiac cycle comprises two main phases: systole and diastole (**Figure 3.2B**). **During systole, the heart muscle actively contracts to eject blood into the arteries**. Conversely, **during diastole the heart relaxes**, allowing deoxygenated blood to flow in. During this second phase, roughly corresponding to the interval between the T-wave to the P-wave, heart's electrical activity is reduced. This time-window is particularly suited to study brain responses to heartbeats, as we shall see in the next sections.

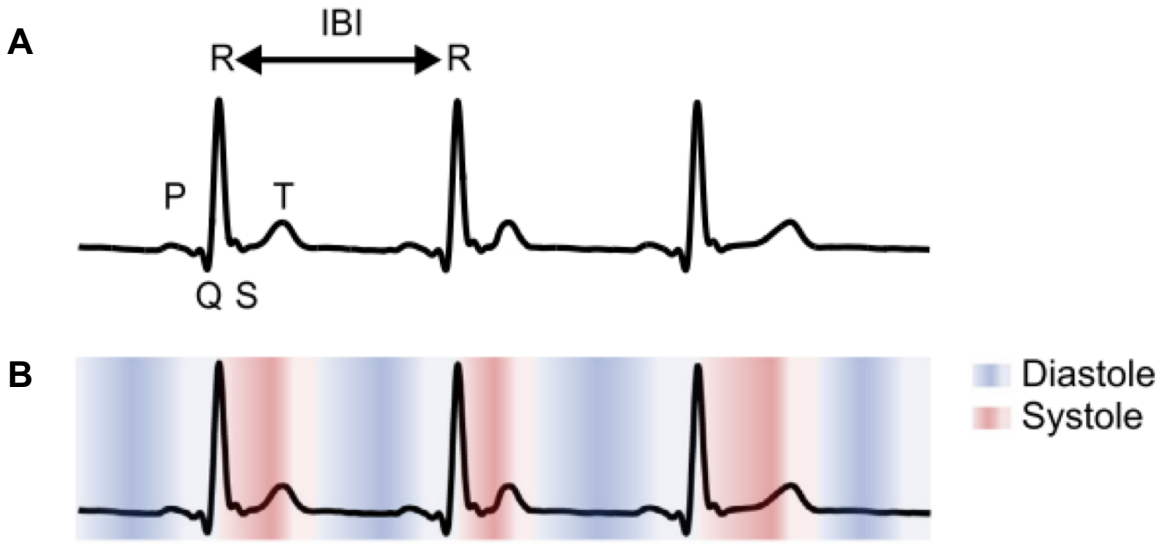


Figure 3.2 A, Stereotypical ECG trace. Letters indicate the names of the waves corresponding to cardiac events: atrial contraction (P-wave), ventricular contraction (QRS complex) and ventricular relaxation (T-wave). The interval between two consecutive cardiac cycles, the inter-beat interval (IBI), corresponds to the time between two consecutive R-peaks. **B**, Stereotypical ECG trace superposed with colours indicating cardiac phases: diastole (blue) and systole (red). Lighter colours indicate the blurred transition between what is generally considered systole and diastole. Figure from Azzalini et al., 2019.

B. Ascending cardiac information

The heart is an independent pacemaker that continues to beat even if its connections with the brain are resected. However, the two-way communication between heart and brain is essential for survival. At every heartbeat, different types of information about cardiac activity reach the brain. One type of cardiac-related information concerns the strength of the cardiac contractions sensed by mechanoreceptors in the walls of the heart and in the arteries (Bishop et al., 2011) which fire in response to a change in tissue distension. The distributed location of the mechanoreceptors (in the heart and in the arteries) implies that within the same cardiac phase, the brain receives information about mechanic distension of different portions of the circulatory system. This property is particularly relevant for studies that investigate the effect of specific phases of the cardiac cycle on stimulus processing (cf. section *Cardiac cycle effects*), as it is not easy to identify which cardiac information has already reached the brain. Other mechanoreceptors appear to fire at the occurrence of contractions themselves, thus simply

indicating the timing of heart's contraction, independently of their strength (Bishop et al., 2011).

Cardiac-related information may reach the brain through other neural pathways. For example, pulsatile activity has been shown to affect tactile receptors in human's finger pad (Macefield, 2003) and proprioceptive receptors in the muscle spindle of the leg (Birznieks et al., 2012).

Finally, arterioles' vascular tone has been also shown to be associate with a reduction of pyramidal cells' activity in brain slices of mice and rats (Kim et al., 2016), a phenomenon named vasculo-neuronal coupling.

C. Ascending neural pathways

Mechanic and somatosensory cardiac-related information are relayed to the brain via spinal and cranial nerves (vagal and glossopharyngeal nerves) through the nucleus of the solitary tract (NST) and parabrachial nucleus (PBN) located in the brain stem. These nuclei project to numerous subcortical structures such as locus coeruleus (LC), raphe nucleus (RN), hypothalamus, amygdala, striatum and the thalamus. Anatomical and functional studies suggest that many cortical regions receive cardiac information from these subcortical relays: primary and secondary somatosensory cortices (Amassian, 1951), insula (Cechetto and Saper, 1987), cingulate motor regions (Dum et al., 2009) as well as ventromedial prefrontal cortex (Vogt and Derbyshire, 2009). Although anatomical tracing is not available in humans, studies using cardiovascular adjustment to probe the neural underpinnings of visceral representation showed results that are compatible with functional and anatomical research in animals (Shoemaker et al., 2012; Thayer et al., 2012; Beissner et al., 2013).

This evidence thus indicates a rather distributed representation of cardiac signals, and of visceral signals more generally (**Figure 3.3**), a fact that is at odds with several theoretical proposals considering the insular cortex as the primary interoceptive region (Saper, 2002; Craig, 2009). For a specific discussion on the primary role of insula in interoception and critiques thereof, I will refer the interested reader to our recently published review article (*Appendix A*).

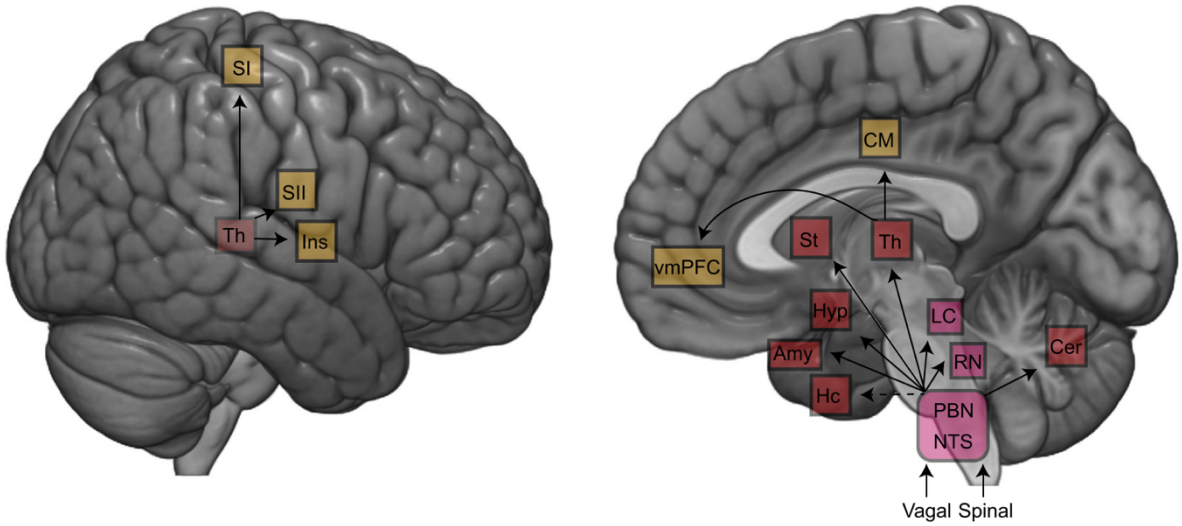


Figure 3.3 Ascending neural pathways. Vagal and spinal pathways convey afferent visceral signals to nucleus of the solitary tract (NTS) and parabrachial nucleus (PBN) in the brainstem (purple colours). From there, visceral information is relayed to locus coeruleus (LC) and raphae nucleus (RN), important neuromodulatory hubs. Visceral signals reach then multiple subcortical (red) and cortical (yellow) regions. Abbreviations: Hc, hippocampus; Amy, Amygdala; Hyp, hypothalamus; St, striatum; Th, thalamus; Cer, cerebellum; vmPFC, ventromedial prefrontal cortex; CM, cingulate motor region; SI and SII, primary and secondary somatosensory cortices; Ins, insula. Figure from Azzalini et al., 2019.

Summary

- Electrocardiography can be used to record heart electrical activity (ECG) whose typical features reflect various phases of the cardiac cycle.
- Within the cardiac cycle, two main phases can be distinguished: systole, corresponding to cardiac contraction, and diastole, corresponding to heart relaxation.
- The brain continuously receives various cardiac-related information such as the strength of mechanic distension of heart and blood vessel walls and the timing of cardiac contraction.
- Cardiac-related information can affect brain activity in other ways. For example, the pulsatile activity of blood vessel modulates somatosensory and proprioceptive peripheral receptors and appears to directly influence the activity of surrounding pyramidal cells (vasculo-neuronal coupling).
- Anatomical tracing and functional studies suggest a distributed representation of cardiac-related information targeting various cortical and subcortical regions: vmPFC, SI, SII, Insula and cingulate motor regions, thalamus, striatum, amygdala to name a few.

This information is relayed via vagal and spinal pathways through midbrain structures (NTS and PBN).

3. Neural responses to heartbeats

In the following sections, I will present how the neural monitoring of ascending cardiac information can be experimentally recorded and how this measure has been used to investigate the functional coupling between the neural monitoring of cardiac information and cognitive processes characterised by prominent subjective components.

A. HER: measures and confounds

The neural monitoring of heartbeats, indexed by heartbeat-evoked responses (HERs), can be measured by using non-invasive (M-EEG) and invasive (iEEG, ECoG) electrophysiological techniques. The process to obtain HERs is similar to what is generally done to investigate how the brain responds to externally presented stimuli, for example visual ones. Brain activity is recorded while multiple instances of a certain visual stimulus are presented to the participants. Next, the recorded brain activity is aligned to the timing of stimulus presentation and averaged across the multiple presentations. The resulting average is what is known as the transient evoked response to a given stimulus. To obtain HERs, the recorded neural activity is aligned to the occurrence of an internally generated stimulus, i.e. individual heartbeats (either to the R-peak or the T-wave) detected on the electrocardiogram (ECG) (**Figure 3.4**; Schandry et al., 1986).

The precise neural mechanisms underlying HER generation are not yet known. However, initial evidence suggests that transient responses to heartbeats may arise from the phase resetting of ongoing brain activity to the timing of heartbeats rather than an additive evoked mechanism (Park et al., 2017). In addition, differently from the neural responses observed in other sensory modalities (e.g. visual, auditory), the factors influencing the morphology of neural responses to heartbeats (such as their peaks and latencies) are also unknown. For example, studies reported differences in HER amplitude at multiple latencies ranging from 200 to 600 milliseconds after R-peak (Kern et al., 2013; Park and Blanke, 2019b), which may relate to the different sources of cardiac information and the neural pathways through which they are conveyed. The study of neural responses to heartbeats is further complicated by a number of caveats that should be kept in mind.

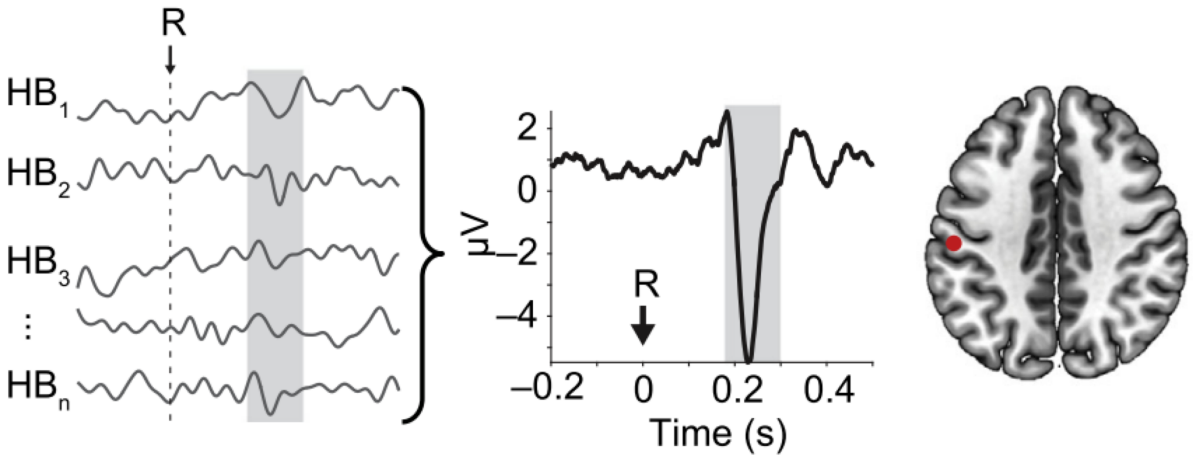


Figure 3.4 Neural responses to heartbeats. Multiple occurrences of neural responses to heartbeats from intracranial data aligned to the timing of the R-peak (left) and the resulting averaged neural response to heartbeats (middle panel). Data displayed in the figure were recorded from an electrode in the somatosensory cortex (right panel). Figure readapted from Azzalini et al., 2019.

First, heart electrical activity directly affects EEG and MEG sensors and creates what is known as the cardiac field artefact (CFA; Dirlich et al., 1997). To limit the influence of CFA on HERs, it is advisable to compute HERs by locking neural activity to the occurrence of the T-wave, as heart electrical activity is minimal in this period. In addition, since the artefactual component will still be present in the M-EEG data, it is recommended to apply independent component analysis correction (ICA)(Vigário et al., 2000).

As experimenters are often interested in how differences in heartbeat-evoked responses influence cognition, the first step is to ascertain that these differences do not simply reflect changes in cardiac inputs (cardiac electrical activity, rhythm or amount of blood ejected). A methodological study to which I contributed (*Appendix B*) sheds light onto the complex relation between cardiac parameters, ECG and MEG signals. We used impedance cardiography to measure the amount of blood ejected at every heartbeat from the left heart ventricle and found that this measure (henceforth stroke volume) could predict beat-by-beat HER amplitude, when MEG data were not corrected with ICA, but not when ICA correction was applied. In addition, differences in stroke volume were reflected in changes in inter-beat intervals and ECG activity. How should one then control that differences in HER amplitude do not simply reflect changes in cardiac parameters but changes in neural responses to constant signals? We suggest to (1) apply ICA correction to limit CFA and the influence of beat-by-beat stroke volume may have on HER amplitude and (2) to perform control analyses to ascertain that HER amplitude

differences are independent of inter-beat intervals and ECG activity. Although invasive recordings are less affected by heart electrical activity, volume distortion in the vessels produced by the flowing of blood may displace the implanted electrodes creating pulse-related amplitude differences (Kern et al., 2013; Park et al., 2017).

A second concern comes from the fact that changes in HERs may be artifactually produced by differences in the overall activity between the experimental conditions that are contrasted, which are sampled around heartbeats (**Figure 3.5**). An effective way to cope with this scenario is to compare the original difference in HER with the difference obtained from data in which the temporal relationship between the neural and cardiac activity has been broken. To this aim, one can simply re-compute synthetic (or surrogate) HERs based on the original portion of the data but in which the timings of heartbeats are randomly assigned. If the original HER difference is driven by differences in underlying neural activity, one should expect that randomly sampling the neural activity (that will not be locked to heartbeats anymore) will produce a difference as big as the original one.

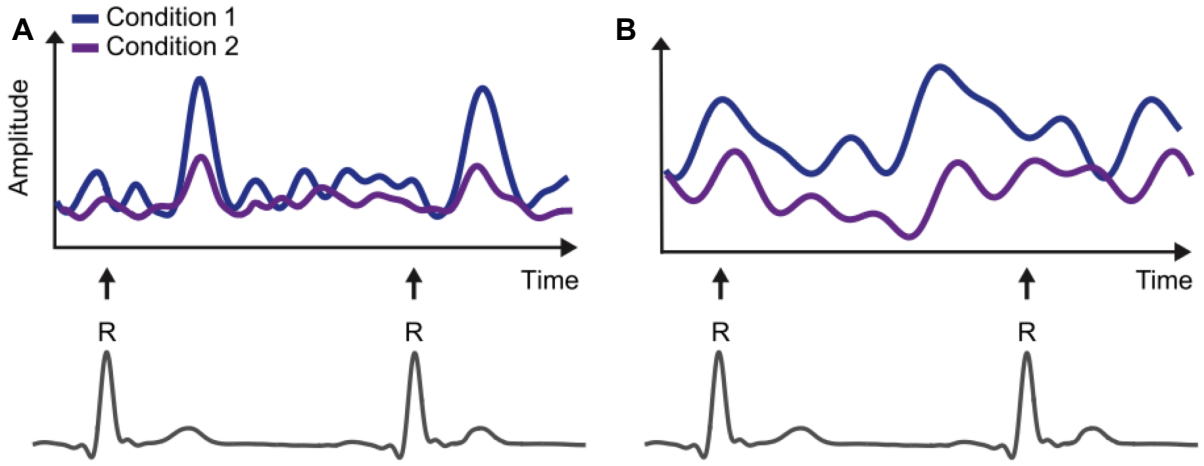


Figure 3.5 True and spurious differences in HERs. A, The difference between two conditions is time-locked to the heartbeat. **B**, Difference between two conditions is sustained and independent from neural responses to heartbeats. Figure from Azzalini et al., 2019.

Last, some experimental designs (Canales-Johnson et al., 2015; Sel et al., 2017) use stimulus presentation time-locked to heartbeats while also analysing HERs. Stimulus-evoked responses may thus create artefactual differences in HERs that should be explicitly addressed via additional controls.

Despite these technical difficulties, recent reviews (Azzalini et al., 2019; Park and Blanke, 2019b) show that the field is becoming increasingly more aware of these potential confounds and that careful controls are applied in the majority of the studies.

B. Cortical sources of HERs

As presented above, cardiac afferent signals comprise information from mechanoreceptors and somatic inputs. These different components may underlie the rather **distributed presence of HERs in different cortical regions**. Intracranially recorded neural responses to heartbeats have been observed in the **somatosensory cortices, insula, amygdala, frontal gyrus and vmPFC** (Kern et al., 2013; Canales-Johnson et al., 2015; Babo-Rebelo et al., 2016b; Park et al., 2017). In addition, cortical reconstruction of M-EEG recordings revealed neural responses to heartbeats in **posterior cingulate cortex, inferior parietal lobule** (IPL) **and vmPFC** (Park et al., 2014, 2016; Babo-Rebelo et al., 2016a, 2019).

Summary

- Heartbeat-evoked responses (HERs) recorded with invasive and non-invasive electrophysiological techniques provide a tool to investigate the functional coupling between neural monitoring of cardiac signals and cognition.
- The knowledge about which neural mechanisms generate HERs is still in its infancy. In addition, little is known about the cardiac parameters that may affect HERs amplitude (heart rate, ejected blood volume and heart's electrical activity).
- To ascertain that differences in neural responses to heartbeats reflect changes in how the brain processes an invariant signal rather than changes in the cardiac input, one must control for differences in cardiac parameters and in electrical cardiac activity.
- Researchers generally infer the functional role of HERs in different cognitive processes by testing whether HERs differ between two (or more) experimental conditions. However, one must control that HERs difference cannot be explained by changes in underlying brain activity unrelated to the occurrence of heartbeats. To this end, one has to compare HER difference against difference obtained when randomly sampling neural activity unrelated to heartbeats.

- Using both invasive and non-invasive electrophysiological recordings, neural responses to heartbeats have been found in somatosensory cortices (SI and SII), insula, amygdala, IPL, PCC and vmPFC.

C. Functional coupling between HERs and self-related cognition

A burgeoning wealth of experimental evidence indicates a functional association between HERs and distinct cognitive processes. This evidence includes the subjective experience of bodily self-awareness, emotional experiences, subjective conscious perception, and self-related thoughts. In the following sections, I will review these experimental findings and suggest that all these cognitive components can be cast in terms of different facets of a self-referential process, involving some form a self-representation. These results thus suggest that the neural monitoring of heartbeats can provide a biologically plausible self-specifying signal harnessed by all cognitive processes characterised by a subjective dimension.

1. Inward oriented attention

Being aware of one's current state comprises the integration of physiological signals, such as cardiac activity. To study how the brain processes interoceptive information, one possibility is to employ explicit interoceptive paradigms (see also section *Implicit and explicit interoception*). In a seminal study of HERs, Schandry and colleagues (Schandry et al., 1986) showed how orienting one's attention towards cardiac perception is related to a change in the cortical processing of these signals, as indexed by HER amplitude modulation. In this experiment, participants had to either silently count their heartbeats (without taking their pulse), a task normally referred to as heartbeat counting task, or to detect an auditory tone embedded in noise. The authors found that HER amplitude in EEG central electrodes was larger during the heartbeat counting task as compared to when attention was externally oriented. Further studies revealed a link between HER amplitude and inter-individual variability in the ability of correctly perceiving one's heartbeats. "Good" perceivers, characterised by lower discrepancy between actual and counted heartbeats, show larger heartbeat-evoked responses compared to "poor" perceivers (Pollatos and Schandry, 2004; Pollatos et al., 2005). Rather than being an immutable ability, heartbeat perception can be trained (Bornemann and Singer, 2017) and improvements appear to be reflected in changes in HER amplitude (Schandry and Weitkunat, 1990; Canales-Johnson et al., 2015). The beneficial effect of training on cardiac perception has been recently criticised (Ring et al., 2015) and thought to reflect improved knowledge of one's cardiac rhythm rather than enhanced interoceptive processing.

More recent research (Petzschner et al., 2019) has called into question the validity of previous experimental paradigms used to test whether larger HERs actually relate to an inward

attentional shift. In fact, attention manipulation in previous paradigms was often confounded with difference in task requirements and in external stimulation. For example, Schandry and colleagues (Schandry et al., 1986) presented auditory tones only in the exteroceptive condition, hampering the interpretability of the results. By asking participants to pay attention either to their heartbeats or to auditory white noise (continuously presented), Petzschner and co-workers (Petzschner et al., 2019) demonstrated that HER amplitude modulation due to inward oriented attention could be replicated when carefully controlling for experimental confounds.

Altogether, these results show that intentionally orienting attention towards one's cardiac activity can modulate the neural processing of heartbeats putatively reflecting the conscious experience of one's visceral state. **The conscious access to this information is crucial to evaluate one's current physiological state** and initiate actions to adjust possible homeostatic imbalance. HERs can thus provide an important neural index for the study of abnormal visceral processing characterising various psychiatric and neurological conditions (Paulus, 2007; Stephan et al., 2016; Petzschner et al., 2017; Khalsa et al., 2018).

2. Bodily self-awareness

The neural monitoring of interoceptive signals occurs even when they are not the focus of attention. It has been proposed that the cortical processing of cardiac information, combined with other bodily signals, may more generally contribute to the conscious awareness of one's own body (Blanke et al., 2015; Park and Blanke, 2019a). For example, Park and colleagues (Park et al., 2016) showed that HERs in posterior cingulate cortex relate to the strength of full-body illusion induced by visuo-tactile stimulation (**Figure 3.6A**). In this experiment, participants saw a video of their back being stroked through a virtual reality headset, while their own back was stroked. Crucially, the stroking pattern on their back was delivered either synchronously or asynchronously with respect to the visually presented video. When visuo-tactile stimulation was synchronous, participants reported identifying themselves more with the virtual avatar body, a condition also characterised by larger HERs. Furthermore, the amplitude of HERs within each experimental block predicted the strength with which participants identify themselves with the virtual avatar body (**Figure 3.6B**).

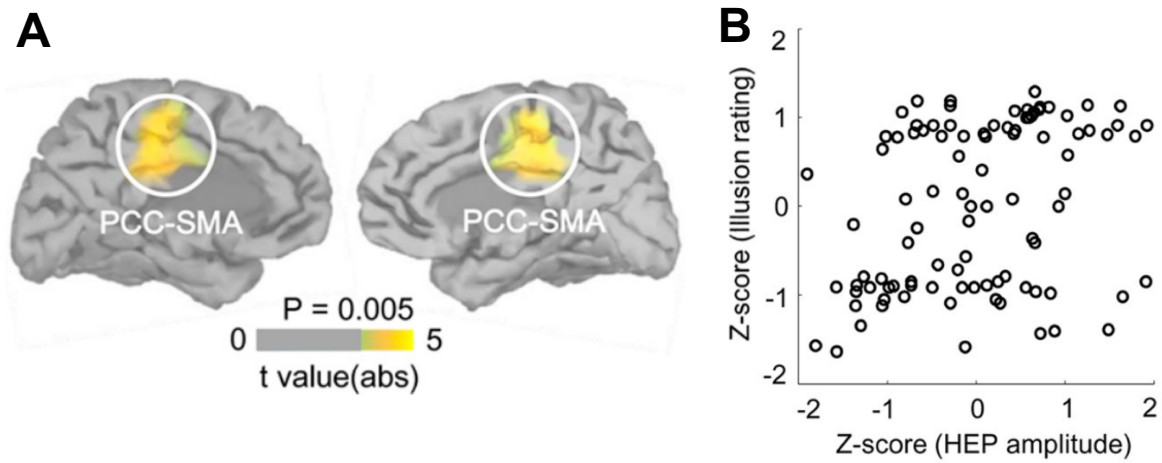


Figure 3.6 HERs scale with the strength of full-body illusion. A, HERs in PCC differ between synchronous and asynchronous visuo-tactile stimulation. **B,** The amplitude of HER scales with self-reported full-body illusion. Figure modified from Park et al., 2016.

Larger HERs were also observed when participants identified themselves with a morphed face containing facial features of both themselves and an unknown stranger, an effect called enfacement illusion (Sel et al., 2017). The interpretability of these latter results is complicated by the fact that the enfacement illusion was driven by facial stimuli presented in synchrony with heartbeats. The superposition of visual- and heart-evoked responses thus makes it difficult to disentangle if the observed difference is due to changes in HERs or stimulus-related activity (see also section *HER: Measures and confounds*). Nonetheless, both results suggest that **HERs may provide an internal signal supporting bodily self-consciousness** (Park and Blanke, 2019a), i.e. an explicit representation of oneself as being a body.

3. Self-related thoughts

Do HERs **provide a self-specifying signal even when no other bodily cues** (e.g. visual, somatosensory, proprioceptive etc.) **are available**, for example **during imagination**? Recent MEG data (Babo-Rebelo et al., 2019) support this hypothesis. The authors revealed that HER amplitude in anterior precuneus and posterior cingulate cortex distinguished when participants were imagining themselves vs. a friend during cued scenarios (**Figure 3.7A**). Importantly, differences in HERs could not be explained by differences in imagination ability, or arousal.

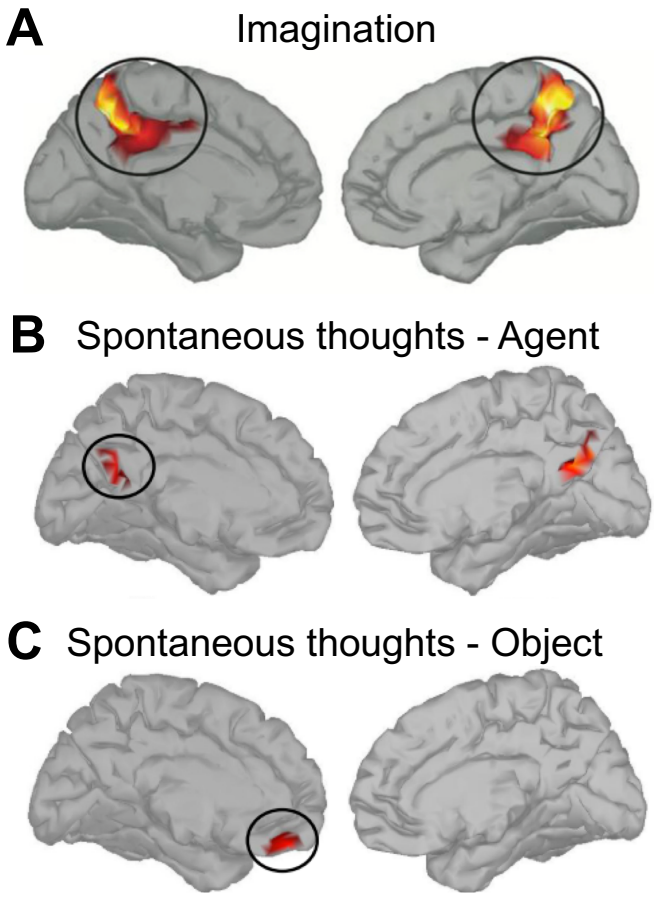


Figure 3.7 HERs encode the self during imagination and spontaneous thoughts. **A**, HERs in anterior precuneus and PCC distinguish self from other during imagined scenarios (Babo-Rebelo et al., 2019). **B-C**, HERs reflect self-relatedness of spontaneous thoughts along two dimensions: in PCC (**B**), HERs scale with the degree to which participants are the agents during the thoughts, while in vmPFC (**C**), HER amplitude reflects the degree to which participants are the object of the thoughts (Babo-Rebelo et al., 2016a; Babo-Rebelo et al., 2016b).

Self-related spontaneous thoughts were also associated with larger HERs. Both MEG (Babo-Rebelo et al., 2016a) and intracranial recordings (Babo-Rebelo et al., 2016b) revealed that the amplitude of HERs encode the degree to which participants' thoughts refer to themselves. In these experiments, participants were asked to let their mind wander for a variable period of time. At random moments, their thoughts were interrupted and participants were asked to report different features of the interrupted thought; namely, whether the thought was about themselves ('Me' dimension), whether they were the agent of an action ('I' dimension), the valence of the thought, and the time (past, present or future) in which the thought took place. Interestingly, HERs in different cortical regions captured two distinct dimensions of the self.

HER amplitude in PCC (and more weakly in anterior insula) encoded how much participants were the agent of the thought (i.e. those performing the action, **Figure 3.7B**), while the amplitude of HERs in vmPFC reflected the extent to which participants were the object of the thought (i.e. the degree to which they were reflecting about themselves, **Figure 3.7C**).

These results offer further evidence that **the neural monitoring of heartbeats** not only contributes to an explicit representation of one's own body, but **may provide relevant information to distinguish self-related mental processes from the ones referring to other people or about the external environment**.

4. Affective experiences

Emotions are probably one of the most intuitive types of subjective experience. They involve the affective evaluation of an external object (or event) from the unique perspective of a given individual. In other words, they are based on a self-referential process that meaningfully binds an object to the experiencing subject. As bodily states are thought to be primary contributors to emotions (LeDoux and Hofmann, 2018), it is reasonable to hypothesise a link between the neural monitoring of cardiac signals and emotional experiences.

In keeping with this hypothesis, Couto and colleagues (Couto et al., 2015) reported that HERs amplitude differed according to the emotional content (neutral, positive and negative videos) presented to the participants. Differences in HER amplitude were source-localised in right and left insular region. Despite the interesting findings, the lack of control for differences induced in cardiac activity by the different videos limits the interpretability of the results. HER differences between emotions may be a consequence of changes in cardiac parameters, rather than a change in the neural responses to an invariant cardiac information. However, a recent study (Kim et al., 2019) showed that HERs differences relate to the experience of specific emotions (i.e. sadness) without concurrent changes in cardiac activity.

Expectations about upcoming emotional stimuli also modulate HERs. Three studies showed a specific suppression of HERs for the expected repetition of angry faces, while HER enhancement was observed for the repetition of neutral, pained and sad faces (Marshall et al., 2017, 2018; Gentsch et al., 2018). The authors interpret these results in terms of top-down predictions about interoceptive states. More precisely, they argue that HERs are reduced in response to repeated negative (angry) emotional stimuli because changes in cardiac activity are correctly predicted. Although these studies control for differences in some cardiac parameters

(i.e. ECG amplitude), other confounds may remain. For example, different emotions are known to trigger changes in various other physiological parameters (e.g. skin conductance) and arousal (Lang and Davis, 2006), which have been also reported to affect heartbeat-evoked responses (Luft and Bhattacharya, 2015).

Another interesting possibility is that **HERs may contribute to subjective affective experience *per se***, rather than to the experience of a particular emotion. One study (Fukushima et al., 2011) contrasted neural responses to heartbeats when participants had to provide subjective judgements about the affective valence of facial expressions against judgments about facial symmetry. Whereas the former type of deliberation requires a subjective emotional appraisal, the latter rests on more objectively defined criteria, such as the relative distance of the eyes from the nose. Endorsing the idea that HERs support affective experiences *per se* rather than specific types thereof, the results reveal that the subjective emotional appraisal was characterised by larger (more negative) HERs responses as compared to judgments of facial symmetry.

5. Conscious perception

As for emotional experiences, conscious perception may also rest on a binding mechanism integrating the processing of an external stimulus to the unique perspective of the experiencing subject. This mechanism could explain why even when the neural processing of external stimuli is deep, they may fail to enter conscious perception (Tallon-Baudry et al., 2018). If HERs index the self-referential process needed to anchor the perceptual processing to an experiencing subject, one should expect changes in HERs to be functionally coupled to changes in conscious perception. In a seminal study, Park and co-workers (Park et al., 2014) found that spontaneous fluctuations in HER amplitude in bilateral vmPFC and right inferior parietal lobule (**Figure 3.8A**) *before* stimulus presentation predicted whether participants detected a faint visual grating presented at threshold. Interestingly, larger HERs were related to increased sensitivity to perceptual evidence and not to a shift in criterion (**Figure 3.8B**), indicating that larger HERs were functionally related to a qualitative change of conscious experience. Importantly, the authors performed several control analyses to exclude that the observed effects could be accounted by changes in arousal and visual processing.

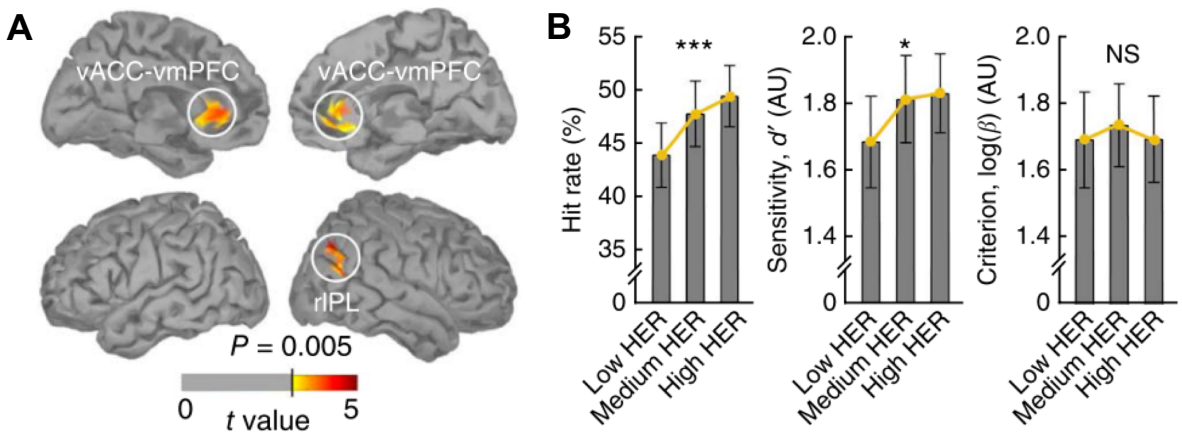


Figure 3.8 Fluctuations in HERs predict conscious visual perception. A, Cortical sources of HER difference between hits and misses before stimulus presentation. **B**, Larger HER amplitude corresponds to an increase in hit rate (left) mediated by a larger perceptual sensitivity (middle) but no change in criterion (right). Figure modified from Park et al., 2014.

The role of HERs in conscious perception has been recently extended to the somatosensory modality (Al et al., 2020). In this experiment, participants were presented with weak electrical stimulation either on the index or middle finger and they were asked to report whether stimulation was present and on which finger it was applied. Before stimulus presentation, and contrarily to what has been reported in vision, misses were characterised by larger HERs (**Figure 3.9A, B**). In addition, the change in detection performance was associated with a shift in decision criterion, with larger HERs featuring a more conservative criterion (**Figure 3.9D**). HER amplitude was also related to discrimination performance, with reduced discrimination for larger HERs (**Figure 3.9E**). These results seem at odds with better detection performance for larger HERs observed in the visual domain. However, comparing amplitude differences between EEG and MEG sensor-level activity is not straightforward. In addition, vagal and somatosensory afferent pathways may exert opposite effects on perception in different modalities. For example, larger HER responses reflecting somatosensory information may be detrimental to somatosensory stimulation presented at threshold.

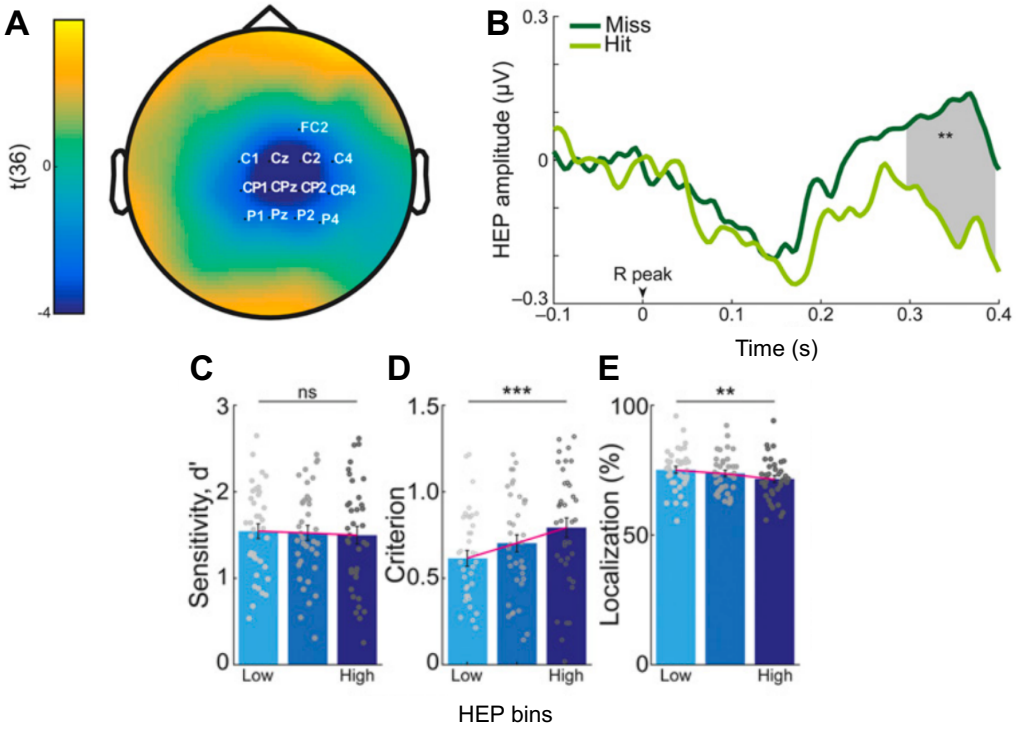


Figure 3.9 HERs predict somatosensory perception and discrimination. A-B, Topography and time-course of the HERs difference between hits and misses before stimulus presentation. Larger HERs were associated with a more conservative criterion **(D)** and no change in sensitivity **(C)**. **E,** Discrimination performance was also related to HERs amplitude, with larger HERs featuring reduced discrimination. Figure modified from Al et al., 2020.

Despite the differences between the two studies, these results show that **spontaneous fluctuations in the neural monitoring of heartbeats are linked to changes in the conscious perception of external stimuli**, **indicating that subjective experience integrates the information about the external world and interoceptive self-related processes**.

Altogether, these findings show that the neural monitoring of heartbeats is functionally involved in many cognitive processes that are characterised by a marked subjective component. These processes comprise explicit awareness of one's bodily signals, emotional experiences, conscious perception and self-related thoughts. In the final section of this chapter, I will attempt to reconcile these results in a common theoretical framework that considers the neural monitoring of heartbeats as a simple biological mechanism providing a self-specifying signal that grounds cognitive processes that involve self-reference.

Summary

- Fluctuations in the neural monitoring of ascending cardiac signals have been associated with various aspects of subjective cognition.
- HERs are enhanced when individuals voluntarily orient their attention towards cardiac sensation, a major interoceptive source of information for the monitoring of current physiological states.
- HERs scale with the strength of full-body illusions, indicating a role for the neural monitoring of cardiac signals in bodily self-awareness.
- HERs act as a neural marker that distinguishes self from other during imagination and spontaneous thoughts.
- Emotional appraisal is related to increased processing of cardiac inputs, a process which may be further modulated by the specific emotional content.
- Changes in visual and somatosensory conscious perception are predicted by fluctuations in HERs, before stimulus presentation.

4. Cardiac cycle effects and cardiac timings

For the sake of completeness, it should be mentioned that other experimental approaches have been used to investigate the role of ascending cardiac information in cognition and behaviour. These approaches are based on the hypothesis that information about the strength of cardiac contraction and vessel distension is conveyed to the brain at specific phases of the cardiac cycle, putatively at systole. It follows that if cardiac-related information is functionally linked to cognitive processes, one should observe effects on behaviour arising at specific timings of the cardiac cycle. Experimentally, this is obtained either by locking stimulus presentation at specific cardiac phases (to R-peak, T-wave or randomly **Figure 3.10**) or by sorting experimental events *post hoc* according to their timing within the cardiac cycle.

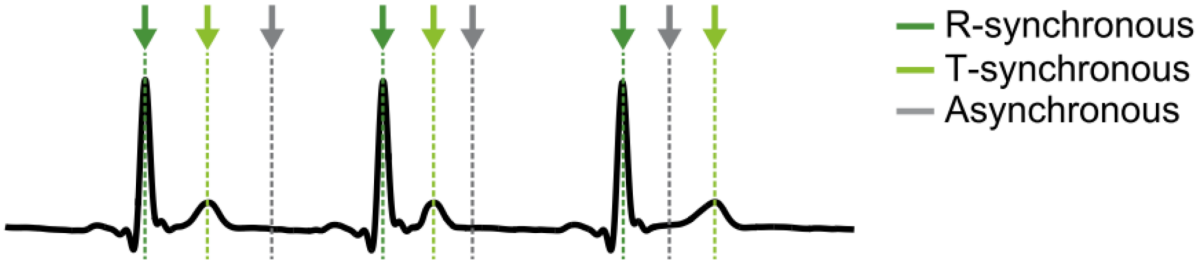


Figure 3.10. Cardiac-dependent stimulus presentation. One procedure to investigate the effect of cardiac phase on stimulus processing is to present stimuli either during systole (i.e. synchronously with T-wave; light green) or at diastole (i.e. synchronously with R-peak; dark green). For an example study using this procedure, see Garfinkel et al. 2014. Other studies investigating cardiac synchronicity present stimuli either locked to one cardiac event (such as R-peak, green lines) or randomly within the cardiac cycle (grey lines) (cf. Aspell et al., 2013). Figure from Azzalini et al., 2019.

A. Cardiac cycle effects

Investigating the influence of the phase of the cardiac cycle on cognition and behaviour was initially motivated by the **baroreceptor hypothesis**. Baroreceptors are mechanoreceptors in the aorta and carotid sinus that fire as a consequence of the pressure exerted on arterial tissue by the passing of blood. This information is crucial for the baroreflex, an automatic process occurring in the brain stem, which regulates blood pressure through the slowing of cardiac rhythm (via the parasympathetic system) and the constriction of the arterioles (via the sympathetic system). The activity of baroreceptors is thought to be maximal at systole, when the blood is ejected from the left ventricle into the arteries.

According to the original proposal (Lacey, 1967), baroreceptor activity at systole was hypothesised to inhibit ongoing cortical processing. However, experimental data only partially support this hypothesis. For example, some studies found that pain perception is reduced at systole (Dworkin et al., 1994; Wilkinson et al., 2013), while experimental results from other sensory modalities such as vision (Pramme et al., 2014, 2016) and somatosensory detection (Edwards et al., 2009) reveal the opposite effect. The effects of cardiac cycle on perception are far to be established, as other studies have failed to find any modulation by the cardiac phase, for instance in visual modality (Elliott and Graf, 1972; Park et al., 2014).

The presentation of stimuli locked to specific cardiac phases has been also used to study emotional processing (Garfinkel and Critchley, 2016). During systole, participants subjectively perceive disgusted (Gray et al., 2012) and fearful faces (Garfinkel et al., 2014) as more intense than when presented at diastole. Although, the influence of bodily signals on emotional processing has been long hypothesized (James, 1890; Cannon, 1927), the reason why afferent cardiac signals should exert their effect only on specific emotions (fear and disgust) and not others (happy, sad or neural) is not transparent. A possible explanation is that the prioritized processing of fearful stimuli at systole is a by-product of enhanced processing of threatening stimuli during aroused physiological states, compatible with the idea that baroreceptor firing at systole conveys information about cardiovascular arousal (Garfinkel and Critchley, 2016).

Altogether, this body of evidence does not allow a clear interpretation of cardiac cycle effects on perception and cognition, revealing that the influence of cardiac afferent signals on ongoing cortical processing is certainly more complex than initially proposed. Three reasons may account for this difficult pattern of results. Concerning baroreceptor influence exclusively, the distinct mechanoreceptors located in different portions of the cardiovascular system may fire at different moments of the cardiac cycle, thus reaching the brain at different latencies. Second, multiple cardiac signals (coming from baroreceptors and somatosensory pathways) can contribute to the observed effects. Finally, ascending cardiac signals are likely to influence activity of subcortical nuclei involved in neuromodulation.

Summary

- Congruent effects of cardiac phases on perception are far from being established.
- Both facilitating and inhibitory effects of stimuli presented at systole have been found in different perceptual modalities.

- However, the perception of fearful stimuli appears to be robustly enhanced during systole.

B. Cardiac synchronicity in multi-sensory integration

The temporal congruency between sensory modalities is a fundamental principle of **multisensory integration** (Stein and Stanford, 2008). A growing body of evidence has shown that synchronicity between exteroceptive and interoceptive signals, such as the cyclical cardiac activity [but other types of interoceptive signal may contribute as well, e.g. (Adler et al., 2014; Allard et al., 2017)], also plays an important role in bodily self-consciousness (Blanke et al., 2015; Park and Blanke, 2019a). A seminal study (Aspell et al., 2013) showed that when participants see the outline of their virtual body – projected some meters ahead of them through a virtual reality headset – illuminated in synchrony with their heart, they self-identify more with the virtual avatar as compared to asynchronous illumination (at 80-120% of their actual heartbeat frequency). More implicit behavioural indices of self-location, known as perceptual drift, also showed that participants judged their position, as well as their tactile perception, shifted towards the location of the avatar (**Figure 3.11**). Body parts, such as morphed faces containing personal and unfamiliar traits (Sel et al., 2017) and virtual hands (Suzuki et al., 2013), also engender stronger illusions of belonging when they are pulsating at the same frequency of participant's heart.

A possible account for this effect is that the pulsatile activity of the external object is more likely to be generated by one's own cardiac activity, and hence it is interpreted as a part of oneself. It follows that stimuli that are briefly presented in synchrony with heartbeats should be more difficult to detect, as they are more likely to be interpreted as self-generated. Experimental evidence for conscious visual perception is in line with this hypothesis. Visual stimuli presented at the frequency of participants' cardiac activity are perceived less accurately and they take longer to be consciously perceived (Salomon et al., 2016). Interestingly, this effect is not linked to the cardiac phase in which the visual stimulus falls, but only to the frequency of presentation, in contrast to the baroreceptor hypothesis.

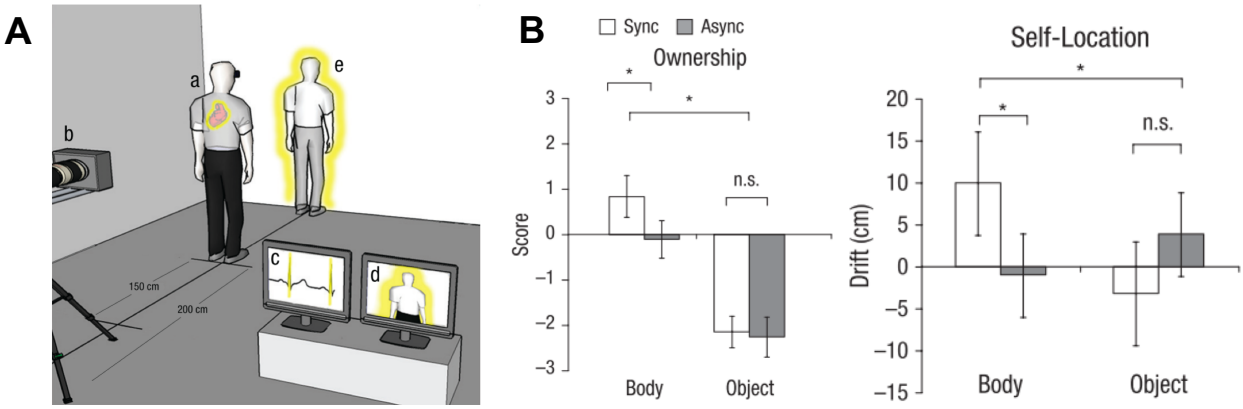


Figure 3.11. Body self-awareness is modulated by visuo-cardiac synchronicity. A, Paradigm from Aspell et al., 2013. Subjects saw a virtual image of their body through a virtual reality headset flashing either in synchrony with their heart or asynchronously. **B.** When the virtual body was flashing in synchrony with their heart, subjects identified themselves more with it (left) and they judge their position as shifted towards the virtual body (right). No effect was observed when the body silhouette was replaced with a rectangular object (right hand side of the graphs in **B**).

Summary

- The synchronicity between cardiac events and external stimuli enhances the self-identification with external stimuli (rubber hands, virtual avatar, morphed faces).
- This mechanism may account for less accurate perception of visual stimuli presented synchronously with the heart, which are considered as self-generated, thus suppressed.
- How these results fit with cardiac cycle effects is not clear, as they appear to be specifically linked to the cardiac frequency and not cardiac phase.

5. Conclusion and relevance for my PhD

The neural monitoring of ascending cardiac signals has been linked experimentally and theoretically to various aspects of subjective cognition. The type of cognitive processes required in all the aforementioned experiments require, to various degrees, a form of self-reference. This is evident in experimental paradigms in which participants are asked about the conscious experience of their own body, but it is also present when they have to explicitly report the content of their thoughts. Although implicit, self-reference is also at play in conscious perception, specifically when participants have to detect the presence of a near-threshold stimulus (Park and Tallon-Baudry, 2014; Tallon-Baudry et al., 2018). The self-referential process may be more prominent in judgements about absence (Mazor et al., 2020), as the current state of the subject provides important evidence to infer the likelihood of stimulus absence.

As noted in *Chapter II*, value-based decisions also entail a self-referential process that specifically enables the evaluation of the available options. In the first MEG experiment (*Chapter IV*), **I tested the hypothesis that spontaneous fluctuations in neural responses to heartbeats would account for a portion of the variability in preference-based decisions, putatively indicating a functional link between subjective evaluation and the precision of self-representation.**

A theoretical contribution of my PhD is a more mechanistic hypothesis of the role of the neural monitoring of heartbeats in subjective cognition. Namely, I propose that **the neural monitoring of cardiac signals may enable the coordination of different types of information expressed in sensory- or domain-specific formats. The reference to this distributed and ever-present signal may generate a common self-referential space in which the mapping of diverse information can be performed** (Azzalini et al., 2019). This hypothesis bears interesting consequences for value-based decisions. Some authors proposed that value may act as a common currency allowing the comparison of incommensurable goods (Levy and Glimcher, 2012). However, how different sensory and conceptual attributes are integrated into an abstract value is yet unknown. In the second MEG experiment (*Chapter V*), **I tested the prediction that neural responses to heartbeats are differentially recruited when comparing items belonging to different categories of goods vs. items of the same category. The former type of comparison specifically involves the transformation of**

sensory- and domain-specific information into a single comparable quantity that enables choice.

In the final chapter (*Chapter VI*), I will discuss the congruency between the two studies, their limitations, the implication for preference-based decisions and future directions.

Chapter IV: Responses to heartbeats in ventromedial prefrontal cortex contribute to subjective preference-based decisions

Responses to heartbeats in ventromedial prefrontal cortex contribute to subjective preference-based decisions

Damiano Azzalini*, Anne Buot, Stefano Palminteri & Catherine Tallon-Baudry*

Laboratoire de Neurosciences Cognitives et Computationnelles, Ecole Normale Supérieure, PSL University & Institut National de la Santé et de la Recherche Médicale, 75005, Paris, France

*Corresponding authors: damiano.azzalini@gmail.com, catherine.tallon-baudry@ens.psl.eu

Abstract

Forrest Gump or Matrix? Preference-based decisions are subjective and entail self-reflection. However, these self-related features are unaccounted for by known neural mechanisms of valuation and choice. Self-related processes have been linked to a basic interoceptive biological mechanism, the neural monitoring of heartbeats, in particular in ventromedial prefrontal cortex (vmPFC), a region also involved in value encoding. We thus hypothesized a functional coupling between the neural monitoring of heartbeats and the precision of value encoding in vmPFC. Human participants were presented with pairs of movie titles. They indicated either which movie they preferred, or performed a control objective visual discrimination that did not require self-reflection. Using magnetoencephalography, we measured heartbeat-evoked responses (HERs) before option presentation, and confirmed that HERs in vmPFC were larger when preparing to the subjective, self-related task. We retrieved the expected cortical value network during choice with time-resolved statistical modeling. Crucially, we show that larger HERs before option presentation are followed by stronger value encoding during choice in vmPFC. This effect is independent of the overall vmPFC baseline activity. The neural interaction between HERs and value encoding predicted preference-based choice consistency over time, accounting for both inter-individual differences and trial-to-trial fluctuations within individuals. Neither cardiac activity nor arousal fluctuations could account for any of the effects. HERs did not interact with the encoding of perceptual evidence in the discrimination task. Our results show that the self-reflection underlying preference-based

decisions involves the integration of HERs to subjective value encoding in vmPFC, and that this integration contributes to preference stability.

Significance statement

Deciding whether you prefer Forrest Gump or Matrix is based on subjective values, which only you, the decision-maker, can estimate and compare, by asking yourself. Yet, how self-reflection is biologically implemented and its contribution to subjective valuation are not known. We show that in ventromedial prefrontal cortex, the neural response to heartbeats, an interoceptive self-related process, influences the cortical representation of subjective value. The neural interaction between the cortical monitoring of heartbeats and value encoding predicts choice consistency, i.e. whether you consistently prefer Forrest Gump over Matrix over time. Our results pave the way for the quantification of self-related process in decision making and may shed new light on the relationship between maladaptive decisions and impaired interoception.

Introduction

Do you prefer Forrest Gump or Matrix? The decision is subjective: only *you* know which movie you like best. The subjective values used in preference-based decision making are internally generated, intrinsically private, and entail self-reflection. In other words, estimating a subjective value requires a reflection about how an item affects *you*. In contrast, the evidence required to decide which of the two words 'listen' and 'look' has more characters is publicly and objectively available to any reader of this article. While the neural underpinnings of valuation and choice have been well studied, the biological mechanism supporting the self-reflection intrinsic to subjective decisions remains unspecified. It might derive from the simplest biological implementation of self-reflection, i.e., the monitoring of one's current physiological state (Craig, 2002; Blanke and Metzinger, 2009; Damasio, 2010; Park and Tallon-Baudry, 2014; Azzalini et al., 2019) required to select the most appropriate behavior to restore homeostatic balance thus ensuring the integrity of the living organism. It follows that the organism needs to track its internal state to assign a value to a given option (Keramati and Gutkin, 2014; Juechems and Summerfield, 2019), and that an imprecise representation of the internal state may lead to suboptimal choice (Paulus, 2007).

The monitoring of current physiological state is notably indexed by the transient neural response automatically elicited by each heartbeat, also known as the heartbeat-evoked response (HER)(Montoya et al., 1993; Kern et al., 2013). HERs have been linked to subjective, self-related cognitive processes in ventromedial prefrontal cortex (vmPFC)(Park et al., 2014; Babo-Rebelo et al., 2016a, 2016b). A separate stream of studies repeatedly showed that vmPFC encodes subjective values (Lebreton et al., 2009; Bartra et al., 2013; Grueschow et al., 2015). We thus hypothesized that (a) HERs in vmPFC would signal the recruitment of self-reflective processes in preparation to a subjective decision, but absent when preparing to an objective decision and (b) that HER fluctuations would affect valuation in subjective preference-based decisions, but not in decisions based on objective evidence publicly available in the outside world, such as perceptual discriminations. We tested these hypotheses in a paradigm where participants performed either a subjective, preference-based choice, or a control, objective perceptual discrimination, between two visually presented movie titles (**Fig. 1**), while their neural and cardiac activity were measured with magnetoencephalography (MEG) and electrocardiography (ECG), respectively. Each trial began with an instruction period, with a symbol indicating which type of decision to perform, during which we measured HERs. Once options were displayed, participants selected the title of the movie they preferred in subjective

preference trials, and the title written with the highest contrast in objective perceptual discrimination trials. We found that (a) HERs during the instruction period were larger when preparing for preference-based decisions than for discrimination ones, and we demonstrated that (b) HER amplitude interacted with the neural encoding of subjective value in vmPFC during choice. The neural interaction between HER and value encoding was associated with more consistent subjective choices. This functional coupling was specific to subjective decisions: HER did not interact with the encoding of perceptual evidence in objective visual discrimination trials.

Materials and Methods

Participants

24 right-handed volunteers with normal or corrected-to-normal vision took part in the study after having given written informed consent. They received monetary compensation for their participation. The ethics committee CPP Ile de France III approved all experimental procedures. Three subjects were excluded from further analysis: one subject for too low overall performance (74%, 2 standard deviations below the mean = 87.6%), one subject was excluded for excessive number of artifacts (29.7% of trials, above 2 standard deviations, mean = 7%), one subject was excluded because the ICA correction of the cardiac artifact was not successful.

21 subjects were thus retained for all subsequent analyses (9 male; mean age: 23.57 ± 2.4 years; mean ± SD).

Tasks and procedure

Participants came on two consecutive days to the lab (mean elapsed time between the two sessions 22.28 ± 3.55 hours) to complete two experimental sessions. The first session was a likability rating on movies (behavior only), from which we drew the stimuli used in the second experimental session, during which brain activity was recorded with magnetoencephalography (MEG).

Rating session. We selected 540 popular movies form the Web (allocine.fr) whose title maximal length was 16 characters (spaces included). DVD covers and titles of the pre-selected movies were displayed one by one on a computer screen and subjects had to indicate whether they had previously watched the currently displayed movie by pressing a 'yes' or 'no' key on a computer keyboard, without any time constraint. Participants were then presented with the list of movie titles they had previously watched and asked to name the 2 movies they liked the most and the two they liked the least. Participants were explicitly instructed to use these 4 movies as reference points (the extremes of the rating scale) to rate all other movies. Last, the titles and the covers of the movies belonging to the list were displayed one by one at the center of the computer screen in random order. Participants assigned to each movie a likability rating by displacing (with arrow keys) a cursor on a 21-points Likert scale and validated their choice with an additional button press. Likability ratings were self-paced and the starting position of the cursor was randomized at every trial.

Stimuli. Experimental stimuli consisted of 256 pairs of written movie titles drawn from the list of movies that each participant had rated on the first day. Each movie title was characterized along two experimental dimensions: its likability rating (as provided by the participant) and its contrast. The mean contrast was obtained by averaging the luminance value (between 40 and 100; grey background at 190) randomly assigned to each character of the title. We manipulated trials difficulty by pairing movie titles so that the differences between the two items along the two dimensions (i.e., likability and contrast) were parametric and orthogonal. Additionally, we controlled that the sum of ratings and the sum of contrast within each difficulty level was independent of their difference and evenly distributed. Each pair of stimuli was presented twice in the experiment: one per decision type. A given movie title could appear in up to 10 different pairs. The position of the movie titles on the screen was pseudo-randomly assigned so that the position of the correct option (higher likeability rating or higher contrast) was fully counter-balanced.

Experimental task. On the second day, subjects performed a two-alternative forced-choice (2AFC) task while brain activity was recorded with MEG. At each trial, participants were instructed to perform one of the two decisions types on the pair of movie titles (**Fig. 1A**): either a preference decision, in which they had to indicate the item they liked the most, or a perceptual discrimination, in which they had to indicate the title written with the higher contrast. Each trial began with a fixation period of variable duration (uniformly distributed between 0.8 and 1.2 seconds in step of 0.05 s) indicated by a black fixation dot surrounded by a black ring (internal dot, 0.20° of visual angle; external black ring, 0.40° of visual angle), starting from which participants were required not to blink anymore. Next, the outer ring of the fixation turned either into a square or a diamond (0.40° and 0.56° visual angle, respectively) indicating which type of decision participants were to perform (preference-based or perceptual, counter-balanced across participants), for 1.5 seconds. Then, the outer shape turned again into a ring and two movie titles appeared above and below it (visual angle 1.09°). Options remained on screen until response was provided (via button press with the right hand) or until 3 seconds had elapsed. After response delivery, movie titles disappeared and the black fixation dot surrounded by the black circle remained on screen for 1 more second. The central dot turned green and stayed on screen for variable time (uniformly distributed between 2.5 and 3 seconds in step of 0.05 s), indicating participants that they were allowed to blink before the beginning of next trial. Each recording session consisted of 8 blocks of 64 trials each.

Prior to the recording session, participants familiarized themselves with the experimental task by carrying out 3 training blocks. The first 2 blocks (10 trials each) comprised trials of one type only, hence preceded by the same cue symbol. The last block contained interleaved trials (n=20), as in the actual recording. The movie pairs used during training were not presented again during the recording session.

Heartbeat counting task. After performing the 8 experimental blocks, we assessed participants' interoceptive abilities by asking them to count their heartbeats by focusing on their bodily sensations, while fixating the screen (Schandry, 1981). Subjects performed six blocks of different durations (30, 45, 60, 80, 100, 120 seconds) in randomized order. No feedback on performance was provided. Since the acquisition of our dataset, this widely used paradigm has been criticized in several respects (Ring et al., 2015; Desmedt et al., 2018; Zamariola et al., 2018).

Questionnaires. Once subjects left the MEG room, they filled 4 questionnaires in French: Beck's Depression Inventory (BDI) (Beck et al., 1961), Peter's et al. Delusions Inventory (PDI) (Peters et al., 2004), the Trait Anxiety Inventory (STAI) (Spielberger et al., 1983) and the Obsessive-Compulsive Inventory (OCI) (Foa et al., 2002).

Recordings

Neural activity was continuously recorded using a MEG system with 102 magnetometers and 204 planar gradiometers (Elekta Neuromag TRIUX, sampling rate 1000 Hz, online low-pass filter at 330 Hz). Cardiac activity was simultaneously recorded (BIOPAC Systems, Inc.; sampling frequency 1000 Hz; online filter 0.05-35 Hz). The electrocardiogram was obtained from 4 electrodes (2 placed in over the left and right clavicles, 2 over left and right supraspinatus muscles (Gray et al., 2007)) and referenced to another electrode on the left iliac region of the abdomen, corresponding to four vertical derivations. The four horizontal derivations were computed offline by subtracting the activity of two adjacent electrodes. Additionally, we measured beat-to-beat changes in cardiac impedance, to compute the beat-by-beat stroke volume (i.e. the volume of blood ejected by the heart at each heartbeat (Kubicek et al., 1970)). Impedance cardiography is a non-invasive technique based on the impedance changes in the thorax due to the changes in fluid volume (blood). A very low-intensity (400 μA rms) high frequency (12.5 kHz) electric current was injected via two source electrodes: the first one was placed on the left side of the neck and the second 30 cm below it (roughly on the sixth rib). Two other monitoring electrodes (placed 4 cm apart from the source ones: below the source

electrode on the neck and above the source electrode on the rib cage) measured the voltage across the tissue. To determine left ventricular ejection time, aortic valve activity was recorded by placing an a-magnetic homemade microphone (online band-pass filter 0.05-300 Hz) on the chest of the subject.

Pupil diameter and eye movements were tracked using an eye-tracker device (EyeLink 1000, SR Research) and 4 electrodes (2 electrodes placed on the left and right temples and 2 electrodes placed above and below participant's dominant eye).

Cardiac events and parameters

Cardiac events were detected on the right clavicle-left abdomen ECG derivation in all participants. We computed a template of the cardiac cycle, by averaging a subset of cardiac cycles, which was then convolved with the ECG time series. R-peaks were identified as peaks of the result of the convolution, normalized between 0 and 1, exceeding 0.6. All other cardiac waves were defined with respect R-peak. In particular, T-waves were identified as the maximum amplitude occurring within 420 milliseconds after the Q-wave. R-peak and T-wave automatic detection was visually verified for each participant.

Inter-beat intervals (IBIs) were defined for each phase of the trial as the intervals between two consecutive R-peaks. More specifically, we considered for 'fixation', 'instruction period' and 'response' phases the two R-peaks around their occurrence. IBIs during 'choice' were based on the two R-peaks preceding response delivery. Inter-beat variability was defined as the standard deviation across trials of IBIs in a given trial phase.

Stroke volume was computed according to the formula (Kubicek et al., 1970; Sherwood et al., 1990):

$$SV = \rho \times \left(\frac{L^2}{Zo^2}\right) \times LVET \times \frac{dZ}{dT_{(max)}}$$

where ρ is the resistivity of the blood (135 Ohms*cm) (Berntson et al., 2007), L^2 is the distance between the two source electrodes, Zo^2 is the base impedance, $LVET$ is the systolic left ventricular ejection time (in seconds), $dZ/dT_{(max)}$ is the largest impedance change during systole (Ohms/sec). Note that because we obtained stroke volume by injecting a current at 12.5 kHz, rather than the more typical frequency of 100 kHz, absolute stroke volumes are systematically underestimated, but relative values are preserved.

MEG data preprocessing

External noise was removed from the continuous data using MaxFilter algorithm. Continuous data were then high-pass filtered at 0.5 Hz (4^{th} order Butterworth filter). Trials (defined as epochs ranging from fixation period to 1 second after response) contaminated by muscle and movement artifacts were manually identified and discarded from further analyses (6% of trials on average, ranging from 0% to 15%).

Independent component analysis (ICA) (Delorme and Makeig, 2004), as implemented in FieldTrip Toolbox (Oostenveld et al., 2011) was used to attenuate the cardiac artifact on MEG data. ICA was computed on MEG data epoched ±200 ms around the R-peak of the ECG, in data segments that were free of artifacts, blinks and saccades above 3 degrees. The number of independent components to compute was set to be equal to the rank of the MEG data. Mean pairwise phase consistency (PPC) was estimated for each independent component (Vinck et al., 2010) with the right clavicle-left abdomen ECG derivation signal in the frequency band 0-25 Hz. Components (up to 3) that exceeded 3 standard deviations from mean PPC were then removed from the continuous data.

To correct for blinks, 2-seconds segments of data were used to estimate blink and eye-movement components. Mean PCC was then computed with respect to vertical EOG signal, and components exceeding mean PCC + 3 standard deviations were removed from continuous data. The procedure was iterated until no component was beyond 3 standard deviations or until 3 components in total were removed. Stereotypical blink components were manually selected in two participants as the automated procedure failed to identified them.

ICA-corrected data were then low-pass filtered at 25 Hz (6^{th} order Butterworth filter).

Trials selection

Trials had to meet the following criteria to be included in all subsequent analysis: no movement artifacts, sum of blinking periods less than 20% of total trial time, at least one T-peak during instruction period (cf. HERs section), and reaction time neither too short (at least 250 ms) nor too long. To identify exceedingly long RTs, we binned the trials of each task in 4 difficulty levels based on the difference of the two options (i.e., difference in ratings in preference-based choice and difference in contrast for the perceptual ones). Within each

difficulty level, for correct and error trials separately, we excluded the trials with reaction times exceeding the participant's mean RT + 2 standard deviations.

The average number of trials retained per participant was 421.67 ± 43.36 (mean ± SD).

Heartbeat-evoked responses

Heartbeat-evoked responses were computed on MEG data time-locked to T-wave occurring during the instruction period. T-waves had to be at minimum 400 ms distance from the subsequent R-peak. In order to avoid contamination by transient visual responses or by preparation to the subsequent visual presentation, we only retained T-waves that occurred at least 300 milliseconds after the onset of the instruction cue and 350 milliseconds before the onset of options presentation. If more than one T-wave occurred in this period, HERs for that trial were averaged. HERs were analyzed from T-wave + 50 ms to minimize contamination by the residual cardiac artifact (Dirlich et al., 1997) after ICA correction.

We verify that differences in HERs between the two types of decision were truly locked to heartbeats, and that a difference of similar magnitude could not arise by locking the data to any time point of the instruction period. To this end, we created surrogate timings for heartbeats (within the instruction period), to break the temporal relationship between neural data and heartbeats, and computed surrogate HERs. We created 500 surrogate heartbeat data set, by permuting the timings of the real T-wave between trials belonging to the same decision type (i.e., the timing of the T-wave at trial *i* was randomly assigned to trial *j*). We then searched for surrogate HER differences between trial types using a cluster-based permutation test (Maris and Oostenveld, 2007) (see below). For each of the 500 iterations, we retained the value of the largest cluster statistics (sum(t)) to estimate the distribution of the largest difference that could be obtained randomly sampling ongoing neural activity during the same instruction period. To assess statistical significance, we compared the cluster statistics from the original data against the distribution of surrogate statistics.

Nonparametric statistical testing of MEG data

HERs difference between preference-based and perceptual trials during instruction presentation was tested for statistical significance using cluster-based permutation two-tailed t-test (Maris and Oostenveld, 2007) as implemented in FieldTrip toolbox (Oostenveld et al., 2011), on magnetometer activity in the time-window 50-300 ms after T-wave. This method defines candidate clusters of neural activity based on spatio-temporal adjacency exceeding a statistical threshold ($p < 0.05$) for a given number of neighboring sensors (n=3). Each candidate cluster is assigned a cluster-level test statistics corresponding to the sum of *t* values of all samples belonging to the given cluster. The null distribution is obtained non-parametrically by randomly shuffling conditions labels 10,000 times, computing at each iteration the cluster statistics and saving the largest positive and negative *t* sum. Monte Carlo p value corresponds to the proportion of cluster statistics under the null distribution that exceed the original cluster-level test statistics. Because the largest chance values are retained to construct the null distribution, this method intrinsically corrects for multiple comparisons across time and space. Controls analyses involving the clustering procedure were performed with the same parameters.

The significance of beta time-series obtained from GLM analyses at the sensor level was obtained using cluster-based permutation two-tailed t-test against zero.

Bayes factor

We used Bayes factors (BF) to quantify the evidence in support of the null hypothesis (H_0 = no difference between 2 measures). To this aim, we computed the maximum log-likelihood of a gaussian model in favor of the alternative hypothesis and for the model favoring the null adjusting the effect size to correspond to a $p = 0.05$ for our sample size (n = 21 for all analyses except for pupil for which n = 16 and for 3 ECG derivations for which n=20). Finally, we computed Bayesian information criterion and the corresponding Bayes factor. As a summary indication, $BF < 0.33$ provides substantial evidence in favor of the null hypothesis, BF between 0.33 and 3 does not provide enough evidence for or against the null (Kass and Raftery, 1995).

For regression analyses, Bayes Factor was computed using the online calculator tool (http://pcl.missouri.edu/bf-reg) based on Liang and colleagues (Liang et al., 2008).

Generalized linear model on response-locked single trials

To analyze how task-related variables are encoded in neural activity during decision, we ran a generalized linear model (GLM) on baseline-corrected (-500 to -200 ms before instruction presentation) single trial MEG data time-locked to button press. We predicted z-scored MEG activity at each time-point and channel using task-relevant experimental variables. For preference-based decisions we modeled MEG activity as:

$$MEG_{t,c} = \beta_0 + \beta_{Chosen_{SV}} + \beta_{UnchosenSV} + \beta_{Button\ press}\ (GLM1a)$$

where *t* and *c* represent MEG activity at time-point *t* at channel *c*, β_0 is the intercept, $\beta_{ChosenSV}$ are the z-scored ratings of the chosen option, $\beta_{UnchosenSV}$ is the z-scored rating of the alternative unchosen option and $\beta_{Button\ press}$ is a categorical variable representing motor response (i.e. top or bottom).

Similarly, for perceptual decisions we used:

$$MEG_{t,c} = \beta_0 + \beta_{ChosenCtrs} + \beta_{UnchosenCtrs} + \beta_{Button\ press}\ (GLM1b)$$

where $\beta_{ChosenCtrs}$ and $\beta_{UnchosenCtrs}$ are the z-scored contrast of the chosen and unchosen option, respectively.

This procedure provided us with time series of beta values at each channel that could be tested against zero for significance using spatio-temporal clustering (Maris and Oostenveld, 2007). Once significant clusters encoding task-related variables were identified at the sensor level, we reconstructed the cortical sources corresponding to the sensor-level activity averaged within the significant time-window. We modeled source-reconstructed neural activity with the same GLMs to identify the cortical areas mostly contributing to the significant sensor-level effect.

Generalized linear model on posterior right vmPFC

To quantify the influence of HER in anterior r-vmPFC during instructions on subjective value encoding during choice, we modeled the activity of posterior r-vmPFC, encoding subjective value with the following GLM:

$$Posterior\ r\text{-}\ vmPFC\ =\ \beta_0 + \beta_{Chosen_{SV}}\ +\ \beta_{HER} + \beta_{HER*ChosenSV}\ (GLM2)$$

where, β_0 is the intercept, $\beta_{Chosen_{SV}}$ are the z-scored ratings of the chosen option, β_{HER} is the z-scored activity in the anterior right vmPFC cluster defined by comparing HERs in preference-based vs. perceptual choices and $\beta_{Chosen_{SV*HER}}$ is the interaction term obtained by multiplying the z-scored previous predictors.

To verify that the interaction between subjective value encoding and HER amplitude was specifically time-locked to heartbeats and not a general influence of baseline activity in anterior r-vmPFC, we ran an alternative model to explain posterior r-vmPFC activity:

$$Posterior\ r\text{-}\ vmPFC = \beta_0 + \beta_{Chosen_{SV}}\ +\ \beta_{BL\ vmPFC} + \beta_{BL\ vmPFC*ChosenSV}\ (GLM3)$$

where, β_0 is the intercept, $\beta_{Chosen_{SV}}$ are the z-scored ratings of the chosen option, $\beta_{BL\ vmPFC}$ is the z-scored activity in anterior r-vmPFC during instructions averaged across the whole instruction period, not time-locked to heartbeats and $\beta_{BL\ vmPFC*ChosenSV}$ is the interaction between the two preceding predictors. Note that regressors are not orthogonalized in any of the GLMs.

Anatomical MR acquisition and preprocessing

An anatomical T1 scan was acquired for each participant on a 3 Tesla Siemens TRIO (n = 2) or Siemens PRISMA (n = 20) or Siemens VERIO (n = 2). Cortical segmentation was obtained by using automated procedure as implemented in the FreeSurfer software package (Fischl et al., 2004). The results were visually inspected and used for minimum-norm estimation.

Source reconstruction

Cortical localization of neural activity was performed with BrainStorm toolbox (Tadel et al., 2011). After co-registration of individual anatomy and MEG sensors, 15,003 current dipoles were estimated using a linear inverse solution from time-series of magnetometers and planar gradiometers (weighted minimum-norm, SNR of 3, whitening PCA, depth weighting of 0.5) using overlapping-spheres head model. Current dipoles were constrained to be normally oriented to cortical surface, based on individual anatomy. Source activity was obtained by averaging sensor-level time-series in the time-windows showing significant effects (difference between HERs and beta values different from zero), spatially smoothed (FWHM 6 mm) and projected onto standard brain model (ICBM152_T1, 15,003 vertices). Note that sources in subcortical regions cannot be retrieved with the reconstruction method used here.

To assess which cortical areas contributed the most to the effects observed at the sensor-level, we ran parametric two-tailed t-test and reported all clusters of activity spatially extending more than 20 vertices with individual t-values corresponding to $p < 0.005$ (uncorrected for multiple comparisons). We reported the coordinates of vertices with the maximal t value and their anatomical labels according to AAL atlas (Tzourio-Mazoyer et al., 2002). For clusters falling into prefrontal cortices, we reported the corresponding areas according to the connectivity-based parcellation developed by Neubert and colleagues (Neubert et al., 2015).

Pupil data analysis

Pupil data that contained blinks (automatically detected with EyeLink software and extended before and after by 150 ms), saccades beyond 2 degrees and segments in which pupil size changed abruptly (signal temporal derivative exceeding 0.3, arbitrary unit) were linearly interpolated. All interpolated portions of the data that exceeded 1 second were removed from further analyses. Continuous pupil data from each experimental block were then band-pass filtered between 0.01 and 10 Hz (second order Butterworth) and z-scored. 16 subjects were retained for pupil analysis; 5 subjects were excluded due to too low quality of data. Pupil analysis was performed in two ways: 1) averaged pupil diameter in the same time period used for HER computation (i.e., 300 ms after instruction presentation until 350 ms before options display) and 2) averaged pupil diameter in the time-window spanning 1 second before button press until its execution.

Code accessibility

The custom code and the source data used for the main analyses of this paper can be accessed online at https://github.com/DamianoAzzalini/HER-preferences. Participants did not give any formal agreement to publicly share the MEG and physiological data, hence the raw data supporting the findings are available from the corresponding authors upon reasonable request.

Results

Behavioral results

Participants were asked to choose between two simultaneously presented movie titles according either to their subjective preferences or to the visual contrast of movie titles, as indicated by trial-by-trial instructions presented before the alternatives (**Fig. 1A**). Decision difficulty, operationalized as the difference between the two options (in the preference task: difference between likeability ratings measured one day before the MEG session [see Material and Methods], in the perceptual task: difference between contrasts), had the expected impact on behavior in both tasks. Both preference consistency and discrimination accuracy increased and reaction times decreased for easier decisions (**Fig. 1C**; preference task, one-way repeated measure ANOVA, main effect of difficulty: accuracy, $F_{(3,60)} = 99.25$, $p < 10^{-15}$; RT, $F_{(3,60)} = 41.14$, $p < 10^{-13}$; perceptual task, main effect of difficulty: accuracy, $F_{(3,60)} = 280.2$, $p < 10^{-15}$; RT, $F_{(3,60)} = 87.67$, $p < 10^{-15}$). Preference and perceptual decisions were matched in accuracy (two-way repeated measures ANOVA, main effect of task on accuracy, $F_{(1,20)} = 0.38$, $p = 0.55$, interaction task x difficulty, $F_{(3,60)} = 2.53$, $p = 0.07$), but preference-based decisions were generally slower (two-way repeated measures ANOVA, main effect of task, RT, $F_{(1,20)} = 57.64$, $p < 10^{-6}$) and reaction times decreased less rapidly for easier decisions (interaction task x difficulty, $F_{(3,60)} = 4.08$, $p = 0.01$). Participants only used task-relevant information (subjective value or objective contrast) to decide, since non-relevant information could not predict choice (**Fig. 1D**; **Table 1**).

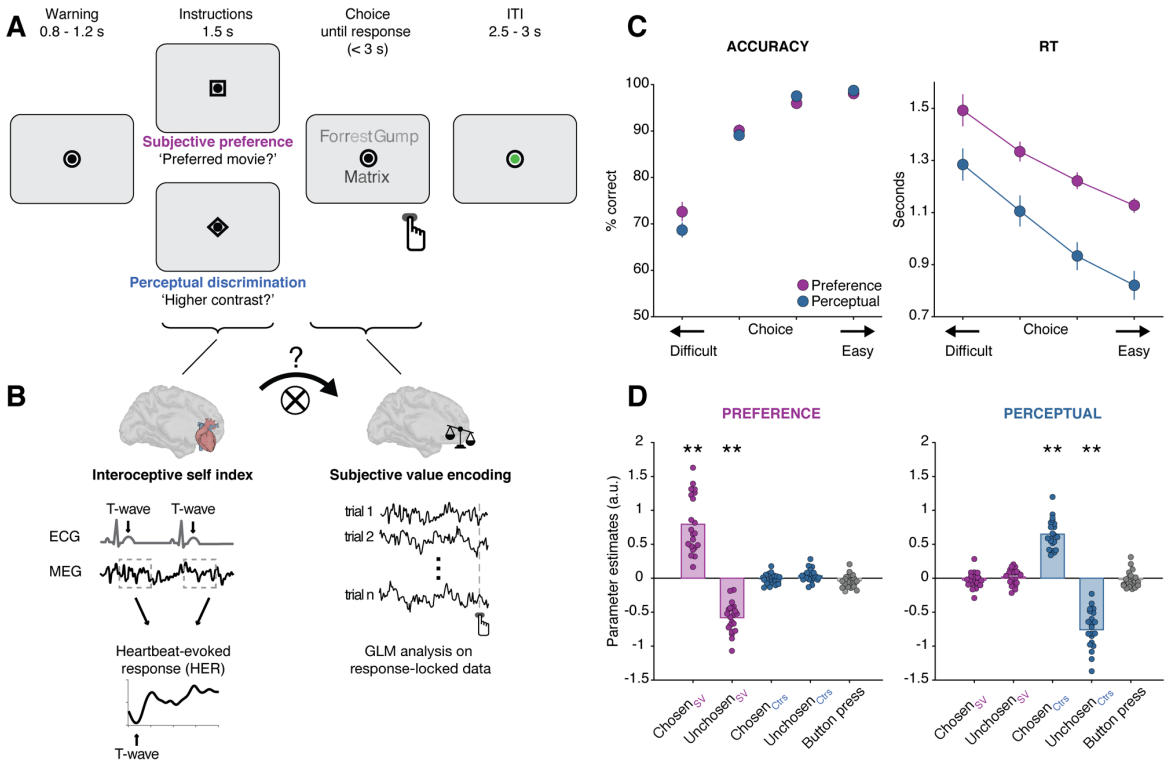


Figure 1. Experimental design and behavioral results. **A**, Trial time-course. After a fixation period of variable duration, a symbol (square or diamond) instructed participants on the type of decision to perform on the upcoming movie titles, either a subjective preference-based choice or an objective perceptual discrimination task. Decision type varied on a trial-by-trial basis. The two movie titles appeared above and below fixation and remained on screen until response or until 3 seconds had elapsed. **B**, Rationale for data analysis. Left, interoceptive self-related processes were indexed by heartbeat-evoked responses (HERs) computed during the instruction period, before option presentation, by averaging MEG activity locked to the T-waves of the electrocardiogram. Right, response-locked MEG activity during choice was modeled on a trial-by-trial basis with a GLM to isolate the spatio-temporal patterns of neural activity encoding value. The central question is whether HERs before option presentation and value encoding interact. **C**, Behavioral results. In both tasks, performance (choice consistency or discrimination accuracy) increased ($F_{(3,60)} = 99.25, p < 10^{-15}$) and response times decreased ($F_{(3,60)} = 41.14, p < 10^{-13}$) for easier choices (i.e. larger difference in subjective value for preference-based decisions, or difference in contrast for perceptual ones). **D**, Only task-relevant information significantly contributed to choice in both preference-based and perceptual decisions as estimated by logistic regression (two-tailed t-test against zero). Each dot represents one participant. **p<0.01.

Table1. Logistic regression against choice, for task-relevant and task-irrelevant stimulus information.

Regressor	Preference decisions two-tailed t-test against 0 (uncorrected p)	Perceptual decisions two-tailed t-test against 0 (uncorrected p)
$Chosen_{SV}$	$t_{(20)} = 8.59$, $p = 3.8*10^{-8}$	$t_{(20)} = -1.14$, $p = 0.27$
$Unchosen_{SV}$	$t_{(20)} = -11.95$, $p = 1.46*10^{-10}$	$t_{(20)} = 0.78$, $p = 0.45$
$Chosen_{Ctrs}$	$t_{(20)} = -0.50$, $p = 0.62$	$t_{(20)} = 13.65$, $p = 1.36*10^{-11}$
$Unchosen_{Ctrs}$	$t_{(20)} = 1.83$, $p = 0.08$	$t_{(20)} = -12.17$, $p = 1.05*10^{-10}$
Button Press	$t_{(20)} = -1.88$, $p = 0.07$	$t_{(20)} = -0.79$, $p = 0.44$

Neural responses to heartbeats are larger when preparing for preference-based decisions

HERs in vmPFC have been previously shown to index self-related processes (Park et al., 2014; Babo-Rebelo et al., 2016a, 2016b). We thus predicted that HERs would be larger when preparing for subjective, preference-based decisions relying on self-reflection than when preparing for objective, perceptual ones. Using a non-parametric clustering procedure which corrects for multiple comparisons across time and space (Maris and Oostenveld, 2007), we found that HERs during the instruction period, before option presentation, were indeed larger when participants prepared for preference-based decisions than for perceptual ones (**Fig. 2A**, **2B**; non-parametric clustering, 201-262 ms after T-wave, sum(t) = 1789, Monte Carlo cluster level p = 0.037). Averaging cluster activity separately in the two conditions results in an effect size Cohen's d of 1.28. The cortical regions that mostly contributed to this effect (**Fig. 2C**) were localized as expected in right and left anterior vmPFC (areas 11m and 14 bilaterally; cluster peak at MNI coordinates, [1 57 -21] and [-3 47 -6]; t-values, 4.68 and 3.76, respectively), but also in the right post-central complex ([32 -22 56]; t-value = 5.57) and right supramarginal gyrus ([41 -33 43]; t-value = 3.98).

The HER difference between subjective preference-based trials and objective perceptual discrimination trials was not accompanied by any difference in ECG activity (paired t-test on 4 ECG vertical derivations, all $p \geq 0.89$, all $BF \leq 0.24$; paired t-test on 4 ECG horizontal derivations, all $p \geq 0.34$, all $BF \leq 0.47$), in cardiac parameters (inter-beat intervals, inter-beat intervals variability, stroke volume) or arousal indices (alpha power and pupil diameter) measured during the instruction period (**Table 2**). The HER difference is thus neither due to differences in cardiac inputs nor to overall changes in brain state. Importantly, the HER difference was time-locked to heartbeats and thus did not reflect a baseline difference between conditions (Monte-Carlo p=0.026. see **Methods** for details).

Table 2. Cardiac parameters and arousal measures during instructions do not differ between preference-based and perceptual decisions.

Measure	Mean ± SEM preference	Mean ± SEM perceptual	Two-tailed paired t-test (uncorrected p)	Bayes Factor
Inter-beat interval (ms)	862.81 ± 28.10	863.49 ± 28.05	$t_{(20)}$= -0.55, p = 0.59	0.30 Substantial
Inter-beat variability (ms)	65.14 ± 5.74	65.24 ± 5.47	$t_{(20)}$ =-0.07, p = 0.94	0.24 Substantial
Stroke volume (ml)	39.97 ± 8.11	40.54 ± 8.74	$t_{(20)}$ = -0.85, p = 0.40	0.40 Inconclusive
Pupil diameter (a.u.)	-0.075 ± 0.05	-0.073 ± 0.05	$t_{(15)}$ = -0.08, p = 0.94	0.23 Substantial
Alpha power (fT^2/Hz)	182 ± 22	179 ± 22	$t_{(20)}$ = 1.17, p = 0.26	0.62 Inconclusive

The subjective value of the chosen option is encoded in medial prefrontal cortices in preference-based decisions

We then identified when and where subjective value was encoded during preference-based choice. First, we modeled single trial response-locked neural activity at the sensor level using a GLM (GLM1a, see **Methods**), using as regressors the subjective values of the chosen ($Chosen_{SV}$) and unchosen ($Unchosen_{SV}$) options, as well as the response button used. Neural activity over frontal sensors encoded the subjective value of the chosen option in two neighboring time-windows ($ß_{ChosenSV}$, first cluster: -580 to -370 ms before response, sum(t)=-7613, Monte Carlo p = 0.004. Second cluster: -336 to -197 ms before response, sum(t) = -4405, Monte Carlo p= 0.033; **Fig. 2D** and **Fig. 2E**). No cluster of neural activity significantly encoded the subjective value of the unchosen option. Motor preparation was encoded later in time in two posterior-parietal clusters of opposite polarities ($ß_{Button\ Press}$, negative cluster: -287 to -28 ms before response, sum(t) = -10918, Monte Carlo p = 0.003; positive cluster: -373 to -196 ms before response, sum(t) = 5848, Monte Carlo p = 0.02).

To identify the cortical regions contributing to the encoding of subjective value at sensor-level, we used the same model (GLM1a) to predict source-reconstructed activity averaged in the time-window identified at sensor level (-580 to -197 ms before response). The subjective value of the chosen option was encoded as expected in medial prefrontal regions (right posterior vmPFC area 32 and 24, cluster peak at MNI coordinates: [7 40 0]; t-value = 4.52); right dorso-medial prefrontal cortex dmPFC area 8m [5 30 40]; t-value = 3.73), as well as in bilateral occipital poles ([6 -77 11], [-1 -85 16]; t-values, 4.17 and 3.85, respectively) and mid-posterior left insula ([-34 -27 17]; t-value = 4.48) (**Fig. 2F**).

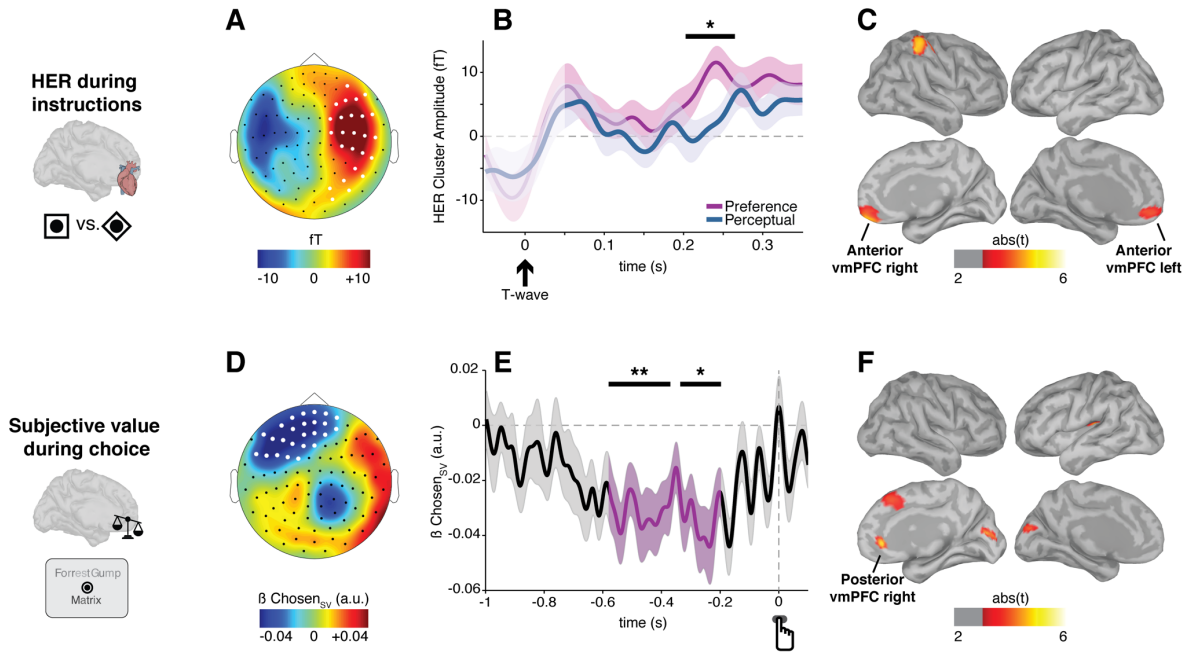


Figure 2. HERs and subjective value encoding. A, Topography of the significant HER difference between preference-based and perceptual decisions during the instruction period (201-262 ms after cardiac T-wave, cluster Monte Carlo p = 0.037). **B**, Time-course of HER (± SEM) for preference-based and perceptual decisions in the cluster highlighted in white in A. The portion of the signal (50 ms after T-wave) still potentially contaminated by the cardiac artifact appears in lighter color. The black bar indicates the significant time-window, as established by non-parametric clustering procedure. **C**, Brain regions mostly contributing to the HER difference between preference-based and perceptual decisions (at least 20 contiguous vertices with uncorrected $p < 0.005$). **D,** Topography of the significant encoding of the chosen option subjective value (-580 to -197 ms before motor response) during choice in preference-based trials. **E**, Time-course (± SEM) of the GLM parameter estimate for the chosen option subjective value in the cluster highlighted in white in D. Black bars indicate significant time-windows, as established by non-parametric clustering procedure. **F**, Brain regions mostly contributing to the encoding of the subjective value of the chosen option (at least 20 contiguous vertices with uncorrected $p < 0.005$). *p<0.05, **p<0.01.

HER amplitude during instruction interacts with subjective value encoding in right vmPFC during choice on a trial-by-trial basis

We thus show that two different sub-regions of vmPFC were involved at different moments in a trial: during the instruction period, HERs were larger when participants prepared for preference-based vs. perceptual decisions in left and right anterior vmPFC, and during the choice period, subjective value was encoded in right posterior vmPFC. We then addressed our main question (**Fig. 1B**): does the amplitude of neural responses to heartbeats during the instruction period interact with the encoding of subjective value during choice in vmPFC?

We tested whether subjective value encoding in right posterior vmPFC was affected by HER amplitude measured in either left or right anterior vmPFC in a two-by-two ANOVA with HER amplitude (high or low, median split across trials) and hemisphere as factors. The ANOVA revealed a significant interaction between HER amplitude and hemisphere (**Fig. 3A**; $F_{(1,40)}$=5.07, p=0.036; no main effect of HER amplitude $F_{(1,20)}$=2.69, p=0.12, no main effect of hemisphere $F_{(1,20)}$=0.19, p=0.67). This interaction corresponded to a significantly stronger subjective value encoding in trials where HERs in right vmPFC were larger during instructions (two-tailed paired t-test on $ß_{ChosenSV}$ in large vs. small HER values in right vmPFC, $t_{(20)}$ = 2.52, p = 0.02, Cohen's d = 0.55).

The influence of HER amplitude was specific to value encoding: the amplitude of visual responses evoked by option presentation was unrelated to HER amplitude (**Table 3**). HER amplitude is thus not a mere index of cortical responsiveness, interacting with any other brain response. HER amplitude in r-vmPFC did not vary with pupil diameter nor alpha power, neither during choice nor during value encoding (**Table 3**), indicating that HER fluctuations are not driven by an overall change in brain state. Last, to definitively rule out an influence of attention/arousal, we tested whether the strength of value encoding was modulated by fluctuations in alpha power or pupil diameter measured during instructions. We median-split trials based on either alpha power or pupil but found no difference in value encoding (alpha, paired t-test on $ß_{ChosenSV}$ $t_{(20)}$=0.19, p=0.85, BF=0.25; pupil, paired t-test $t_{(20)}$=-0.23, p=0.82, BF=0.24).

Table 3. Arousal states and physiological parameters do not differ between low and high HER amplitude in preference trials.

Measure	Low HER (Mean ± SEM)	High HER (Mean ± SEM)	Two-tailed t-test (uncorrected p)	Bayes Factor
Pupil diameter (a.u.) during option presentation, [-1 0] s	0.27 ± 0.07	0.24 ± 0.07	$t_{(15)} = 0.57$, $p = 0.58$	0.29 Substantial
Pupil diameter (a.u.) during value encoding [-0.580 -0.197] s	0.32 ± 0.08	0.29 ± 0.08	$t_{(15)} = 0.57$, $p = 0.58$	0.29 Substantial
Alpha power (fT^2/Hz) max 15 channels during option presentation, [-1 0] s	$1.24 \times 10^2 \pm 0.14 \times 10^2$	$1.25 \times 10^2 \pm 0.13 \times 10^2$	$t_{(20)} = -0.42$, $p = 0.68$	0.27 Substantial
Alpha power (fT^2/Hz) in the value encoding cluster [-0.580 -0.197] s	$0.56 \times 10^2 \pm 0.03 \times 10^2$	$0.56 \times 10^2 \pm 0.03 \times 10^2$	$t_{(20)} = -0.71$, $p = 0.48$	0.35 Inconclusive
RMS of visual response to options (fT) [0-250] ms	62.8 ± 2.43	62.5 ± 2.40	$t_{(20)} = 0.28$ $p = 0.78$	0.25 Substantial

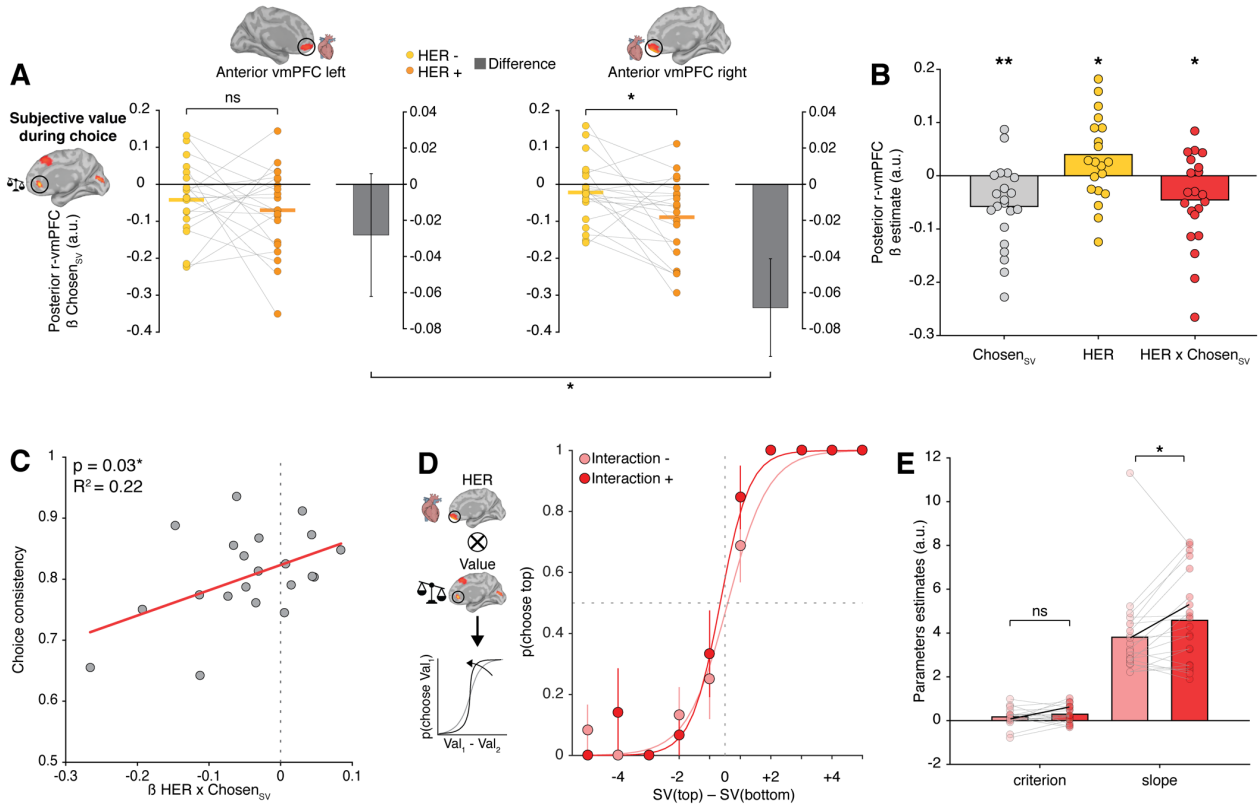


Figure 3. The interaction between HER and value encoding accounts for inter-individual variability in choice consistency and for intra-individual trial-by-trial fluctuations of choice precision. A, Parameter estimates for the encoding of the chosen option value in posterior r-vmPFC during choice, in trials where HERs during task preparation were small (yellow) or large (orange) in left anterior vmPFC (left), or in right anterior vmPFC (right). Each dot represents a participant, the horizontal line indicates the mean. The difference in value encoding between large and small HER (grey bars; error bars indicate SEM) was significant in right vmPFC (ANOVA on value encoding, significant interaction HER amplitude x hemisphere, $F_{(1,40)}$=5.07 p=0.036); two tailed paired t-test comparing encoding strength for trials with large or small HERs in right vmPFC, $t_{(20)}$=2.52 p=0.02) **B**, Activity in posterior r-vmPFC is by design explained by the chosen option subjective value, but it is also explained by HER amplitude in anterior r-vmPFC during instructions and by the interaction between HER and value encoding. **C,** Robust regression shows that the magnitude of the interaction between HER and value encoding positively predicts inter-individual variability in choice consistency **D,** Behavioral data (dots) and fitted psychometric function (lines) for one representative participant, in trials with a large or small interaction between HER and value encoding. Error bars represent SEM. **E**, Criterion and slope of the psychometric function in all participants, revealing a significantly steeper slope for trials with large interaction between HER and value-related vmPFC activity (p = 0.037). Decision criterion is unaffected. Black lines represent the parameters estimates of the participant displayed in D. *p<0.05, **p<0.01.

The interaction between HER amplitude and subjective value encoding in right vmPFC was further tested using a full parametric approach. Here (GLM2), we predicted the activity of posterior r-vmPFC during choice from the subjective value of the chosen option, the HER amplitude in anterior r-vmPFC during instruction period and the interaction between these two terms (**Fig. 3B**). Since the posterior vmPFC region of interest was defined based on its encoding of the chosen value, the parameter estimate for chosen value was, as expected, large ($ß_{ChosenSV}$ = -0.06 ± 0.02, two-tailed t-test against 0, $t_{(20)}$ = -3.37, p = 0.003). Activity in posterior vmPFC was also predicted by the amplitude of HERs occurring about 1.5 s earlier, during the instruction period, independently from the chosen value ($ß_{HER}$ = 0.04 ± 0.02, two-tailed t-test against 0, $t_{(20)}$= 2.13, p = 0.046), and importantly by the interaction between HERs and chosen value ($ß_{HER*ChosenSV}$ = -0.05 ± 0.02, two-tailed t-test against 0, $t_{(20)}$= -2.41, p = 0.025). Both the median-split analysis and the parametric model thus reveal a significant interaction between the amplitude of HERs during instruction period and the neural encoding of subjective value during choice.

We then verified that the effect on the neural encoding of subjective value was specific to HER amplitude, and not due to an overall baseline shift in anterior r-vmPFC during the instruction period. We ran an alternative model (GLM3) in which the activity in posterior r-vmPFC was predicted from the subjective value of the chosen option, the activity in anterior r-vmPFC averaged during the whole instruction period, i.e. activity not time-locked to heartbeats, and the interaction between the two terms. This analysis revealed that while the subjective value of the chosen option still significantly predicted the activity of posterior r-vmPFC ($ß_{ChosenSV}$ = -0.05 ± 0.02, two-tailed t-test against 0, $t_{(20)}$ = -3.27, p = 0.004), the other two terms did not (activity in anterior r-vmPFC averaged during instruction period: $ß_{BL\ vmPFC}$ = 0.006 ± 0.03, two-tailed t-test against 0, $t_{(20)}$ = 0.22, p = 0.83, BF = 0.25; interaction: $ß_{BL\ vmPFC*\ ChosenSV}$ -0.03 ± 0.02, two-tailed t-test against 0, $t_{(20)}$ = -1.55, p = 0.14, BF = 1.14). The encoding of subjective value is thus specifically modulated by HER amplitude in anterior r-vmPFC and not by an overall baseline shift unrelated to heartbeats in the same region.

The functional coupling between HERs and subjective value encoding was also region specific: HER amplitude in anterior r-vmPFC was unrelated to the strength of value encoding in any other value-related regions (two-tailed paired t-test on $ß_{ChosenSV}$ in large vs. small HER values in right dmPFC, $t_{(20)}$ = -0.89, p = 0.38, BF = 0.43; right occipital pole, $t_{(20)}$ = -0.86, p = 0.40, BF = 0.41; left occipital pole, $t_{(20)}$ = -1.60, p = 0.13, BF = 0.81; left posterior insula, $t_{(20)}$= 1.00, p = 0.33, BF = 0.49). Conversely, HERs outside anterior r-vmPFC did not significantly interact with value encoding in posterior r-vmPFC. Splitting trials based on the amplitude in

the two other cortical regions showing differential heartbeats-evoked responses (**Fig. 2C**) showed no significant modulation of value encoding in right posterior vmPFC (post-central complex: two-tailed paired t-test on $ß_{ChosenSV}$, $t_{(20)}$= -1.41, p = 0.17, BF = 0.90; right supramarginal gyrus: two-tailed paired t-test, $t_{(20)}$= -1.96, p = 0.06, BF = 2.41).

The interaction between HER and value encoding predicts choice consistency

To what extent does the interaction between HER and value encoding in vmPFC predict behavior? We first tested whether the interaction between HER and value encoding relates to inter-individual differences in choice consistency, i.e. whether participants selected the movie to which they had attributed the greatest likeability rating the day before. Given the overall high consistency in preference-based decisions, which may reduce our ability to detect significant relationships, we computed mean choice consistency using the top-50% most difficult trials (i.e. trials above median difficulty in each participant). We regressed the model parameter of the interaction between HER and value encoding ($ß_{HER*ChosenSV}$ obtained from GLM2) against mean choice consistency across participants. The larger the interaction between HER and value encoding, the more consistent were participants in their choices ($ß_{robust}$ = 0.41, robust regression R^2 = 0.22, $t_{(19)}$ = 2.29, p = 0.03; **Fig. 3C**). In other words, 22% of inter-individual difference in behavioral consistency is explained by the magnitude of the interaction between HER and value encoding.

The correlation between neural activity and behavior was specific to the interaction parameter: inter-individual differences in choice consistency could not be predicted from the model parameter estimate of HER ($ß_{HER}$ from GLM2; $ß_{robust}$ = 0.02, $R^2 = 4*10^{-4}$, $t_{(19)}$ = 0.09, p = 0.93, BF = 0.39), nor from the parameter estimate of value ($ß_{ChosenSV}$ from GLM2; $ß_{robust}$ = -0.19, R^2 = 0.04, $t_{(19)}$ = -0.88, p = 0.39, BF = 0.52). The interaction between HER and subjective value encoding did not co-vary significantly neither with the personality traits, assessed through self-reported questionnaires (robust regressions on BDI scores, $t_{(19)}$ = -0.17, p = 0.87, BF = 0.40; STAI scores, $t_{(19)}$ = 0.90, p = 0.38, BF = 0.52; OCI scores, $t_{(19)}$ = -1.23, p = 0.22, BF = 0.70; PDI scores, $t_{(19)}$ = -0.32, p = 0.75, BF = 0.60) nor interoceptive ability, assessed with the heartbeat counting task ($t_{(19)}$= -0.73,p = 0.48, BF = 0.48).

So far, results are based on parameter estimates computed across trials for a given participant. To assess how within-participant trial-by-trial fluctuations in behavior relates to the interaction between HERs and subjective value encoding, we computed the z-scored product of the HER amplitude in anterior r-vmPFC during the instruction period and the value-related

activity in posterior r-vmPFC during choice. We then median-split the trials according to this product and modeled participants' choices separately for trials with a small vs. large interaction (**Fig. 3D**). When the interaction was large, psychometric curves featured a steeper slope, corresponding to an increased choice precision (two-tailed paired t-test, $t_{(20)}$= -2.24, p = 0.037, Cohen's d = -0.49; after removal of the unique outlier with a slope estimate exceeding 3 SD from population mean, $t_{(19)}$ = -3.30, p = 0.003, Cohen's d= -0.74; **Fig. 3E**), while decision criterion was not affected (two-tailed paired t-test, $t_{(20)}$= -1.20, p = 0.25, BF = 0.64; after outlier removal $t_{(19)}$ =-0.96, p = 0.35, BF = 0.46; **Fig. 3E**).

To control for the specificity of the interaction, we estimated the psychometric function on trials median-split on HER amplitude alone but found no difference in choice precision (two-tailed paired t-test, $t_{(20)}$= 0.41, p = 0.69, BF = 0.27) nor in criterion (two-tailed paired t-test, $t_{(20)}$= 0.52, p = 0.61, BF = 0.29). Similarly, median-splitting trials on value-related posterior r-vmPFC activity alone revealed no difference in the psychometric curve (two-tailed paired t-test, slope, $t_{(20)}$= 0.21, p = 0.84, BF = 0.25; criterion: $t_{(20)}$= -0.57, p = 0.58, BF = 0.30). Therefore, our results indicate that trial-by-trial choice precision is specifically related to the interaction between HERs in anterior r-vmPFC and value-related activity in posterior r-vmPFC.

Altogether, these results indicate that the interaction between HER amplitude and subjective value encoding accounts both for within-subject inter-trial variability, and for inter-individual differences in preference-based choice consistency.

HER effects are specific to preference-based choices

Finally, we tested whether the effect of HER was specific to subjective value or whether it is a more general mechanism interacting with all types of decision-relevant evidence. To this aim, we analyzed perceptual discrimination trials using the same approach as for preference-based trials. First, we modeled the single trial response-locked MEG sensor-level data using a GLM (GLM1b) with the parameters accounting for choice in the perceptual task (i.e. contrast of the chosen option – $Chosen_{Ctrs}$ – and the contrast of the unchosen option – $Unchosen_{Ctrs}$), as well as response button. The non-parametric clustering procedure revealed the presence of a fronto-central cluster encoding the contrast of the chosen option ($ß_{ChosenCtrs}$, -257 to -25 ms before response, sum(t)= 8121, Monte Carlo p = 0.005; **Fig. 4A** and **B**). We also found two clusters of opposite polarities encoding the contrast of the unchosen option ($ß_{UnchosenCtrs,}$ positive cluster, -250 to -79 ms before response, sum(t) = 6182, Monte Carlo p = 0.008; negative cluster, -211 to -88 ms before response, sum(t) = -4127, Monte Carlo p = 0.04) and two clusters encoding motor

preparation ($ß_{\text{Button Press,}}$ positive cluster, -193 to 0 ms before response, sum(t) = 11103, Monte Carlo p = 0.0004; negative cluster, -222 to 0 ms before response, sum(t) = -10395, Monte Carlo p = 0.0004). The same model (GLM1b) applied to source-reconstructed activity averaged in the time-window identified at sensor level (-257 to -25 ms before response) revealed four cortical areas encoding the contrast of the chosen option (**Fig. 4C**): left midcingulate area (peak at MNI coordinates [-11 -27 43], t-value = 5.14), left superior frontal gyrus ([-15 7 71], t-value = 4.70) and bilateral inferior parietal lobule (right IPL, [-47 -55 53], t-value = 4.25; left IPL, [44 -42 49], t-value = 5.51). Finally, we median-split perceptual trials according to the amplitude of HERs in anterior r-vmPFC. The encoding strength of the contrast of the chosen option did not interact with heartbeat-evoked responses amplitude in any of the contrast-encoding regions (all $p \geq 0.26$, $BF \leq 0.62$; **Table 4**). The results thus indicate that HER amplitude in r-vmPFC is specifically linked to the cortical encoding of subjective value.

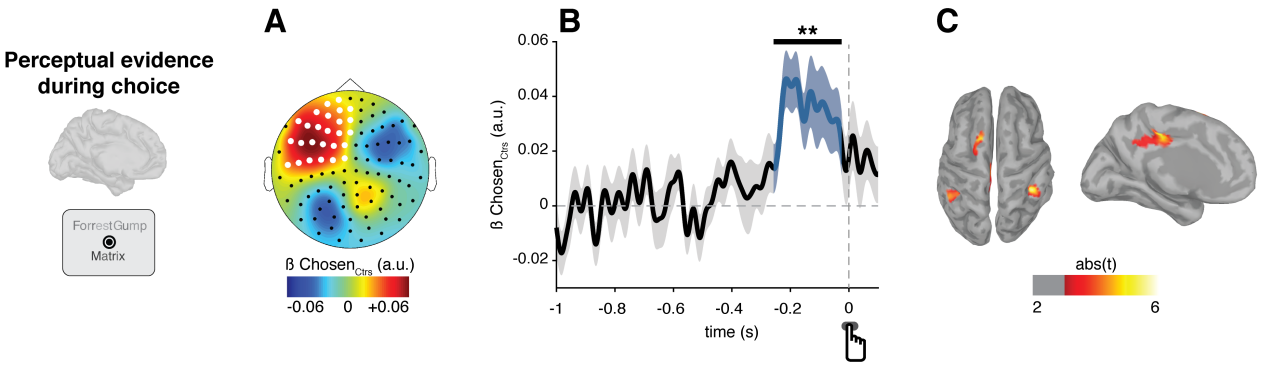


Figure 4. Neural encoding of perceptual evidence. A, Topography of the significant encoding of the chosen option contrast (-257 to -25 ms before motor response) during choice in objective visual discrimination trials. **B**, Time-course (± SEM) of the GLM parameter estimate for the chosen option contrast in the cluster highlighted in white in A. The black bar indicates the significant time-window, as established by non-parametric clustering procedure. **C**, Brain regions mostly contributing to the encoding of the contrast of the chosen option (at least 20 contiguous vertices with uncorrected $p < 0.005$). **p<0.01.

Table 4. Encoding strength for contrast during perceptual decisions does not depend on HER amplitude in anterior r-vmPFC.

Region	Two-tailed paired t-test HER - vs. HER +	Bayes Factor
L midcingulate area	$t_{(20)} = 1.17, p = 0.26$	0.62 Inconclusive
L superior frontal gyrus	$t_{(20)} = 0.90, p = 0.38$	0.43 Inconclusive
L inferior parietal lobule	$t_{(20)} = 0.11, p = 0.92$	0.24 Substantial
R inferior parietal lobule	$t_{(20)} = 0.79, p = 0.44$	0.38 Inconclusive

Discussion

We show that preparing for subjective preference-based decisions led to larger responses to heartbeats in vmPFC, as expected from previous studies relating self and HERs (Park et al., 2014; Babo-Rebelo et al., 2016a, 2016b), and in post-central gyrus, a region known to respond to heartbeats (Kern et al., 2013; Azzalini et al., 2019; Al et al., 2020). We further reveal that HERs before option presentation interact specifically with subjective value encoding during choice in vmPFC. The interaction between HERs and value encoding predicted preference-based choices: it was associated to more precise decisions at the single trial level, and predicted inter-individual variability in choice consistency over time. No interaction between HER and the encoding of perceptual evidence could be found in the control objective task. Neither the HER difference between the two tasks, nor the interaction between HER and value in the subjective preference task, could be trivially explained by changes in cardiac parameters (heart rate, heart rate variability, ECG, stroke volume) nor by changes in arousal state (pupil diameter, alpha power). Altogether, our results reveal how the self-reflection intrinsic to preference-based decisions involves the neural read-out of a physiological variable and its integration in the subjective valuation process.

In line with previous studies relating the self with HERs in vmPFC (Park et al., 2014; Babo-Rebelo et al., 2016a, 2016b), we find that HERs are more pronounced when preparing for the subjective preference task than when preparing for the objective discrimination task. A number of self-related processes might take place specifically when preparing for the subjective preference-based task, such as turning attention inward or pre-activating autobiographical memory circuits. Note that HERs cannot be influenced by processes triggered by the movie titles themselves, such as retrieving movie-specific information from memory, since options are not yet available in the instruction period where HERs are measured. HERs were not associated with any change in the neural encoding of objective perceptual evidence. This result might seem at odds with previously reported effects of HERs on sensory detection at threshold (Park et al., 2014; Al et al., 2020). However, as opposed to the perceptual discrimination task used here, sensory detection at threshold is intrinsically subjective, since participants are asked to introspect and report their fluctuating subjective experience in response to physically and objectively constant stimuli (Ress and Heeger, 2003; Campana and Tallon-Baudry, 2013). Last, there was no difference in cardiac parameters between the two tasks. Our results thus differ from previous observations in risky decisions, where changes in peripheral bodily signals predicted behavioral performance (Bechara et al., 1997). More generally, this indicates that

HER fluctuations relevant for valuation and behavior correspond to changes in the quality of the neural monitoring in cardiac inputs, rather than to changes in cardiac activity.

We successfully retrieved with MEG the cortical valuation network described in the fMRI literature (Levy and Glimcher, 2012; Bartra et al., 2013; Clithero and Rangel, 2014), including dmPFC and vmPFC, during the choice period. These findings are interesting *per se*, as data on the temporal course of value-based choices in the human prefrontal cortex remain scarce (Hunt et al., 2012, 2015; Polanía et al., 2014; Lopez-Persem et al., 2020). Here, we find that chosen value is robustly encoded in the valuation network from 600 ms to 200 ms before motor response, with a temporal (but not spatial) overlap with motor preparation. Note that we did not find a robust encoding of the unchosen value, in line with electrophysiological recordings in the monkey orbitofrontal cortex where the encoding of the chosen value dominates (Padoa-Schioppa and Assad, 2006; Strait et al., 2014; Hunt et al., 2015; Rich and Wallis, 2016). Chosen value encoding interacted with HERs specifically in vmPFC, a region shown by separate streams of studies to play a role in valuation (Fellows, 2006; Delgado et al., 2016; Vaidya et al., 2017) and self-related cognition (Qin and Northoff, 2011; Andrews-Hanna et al., 2014), and long suspected to play an integrative role in decision making (Vaidya and Fellows, 2020). In the objective discrimination task, relevant perceptual evidence was encoded, among other regions, in posterior parietal cortex, consistent with monkey electrophysiology literature (Shadlen and Newsome, 2001; Heekeren et al., 2008).

Because fluctuations in HERs occurred during task preparation, before option presentation, their influence on value encoding might generally pertain to interactions between ongoing activity (during task preparation) and stimulus-evoked activity (in response to option presentation). HERs constitute a very specific subset of such interactions: HERs interacted only with value, but not with visual evoked responses for instance, and only in vmPFC in the subjective task. Conversely, no interaction between overall ongoing activity in vmPFC and value encoding could be found. Besides, HERs modulated value encoding in a multiplicative manner. Our results thus differ from a previously reported vmPFC baseline-shift additive effect in pleasantness ratings (Abitbol et al., 2015; Lopez-Persem et al., 2020). Altogether, our results indicate that part of the unspecified 'neural noise' driving fluctuations in choice consistency (Padoa-Schioppa, 2013; Kurtz-David et al., 2019; Webb et al., 2019) comes from the interaction between interoceptive self-related processes, indexed by neural monitoring of cardiac signals, and the neural encoding of subjective value. A more detailed mechanistic account of how responses to heartbeat during task preparation influence the subjective valuation process taking

place about 1.5 second later remains to be established. Still, our results pave the way for the quantification of self-related processes in decision making, an aspect mostly absent from computational models of decision-making despite its relevance to understand maladaptive decisions in psychiatric disorders (Paulus, 2007; Moeller and Goldstein, 2014; Sui and Gu, 2017).

Decisions on primary goods such as food integrate information about internal bodily states to select options that preserve the integrity of the organism that needs to be fed and protected – the simplest notion of self. We show here that the subjective valuation of cultural goods, that relies on the same cortical valuation network as employed for primary goods (Chib et al., 2009; Lebreton et al., 2009; Levy and Glimcher, 2011; McNamee et al., 2013; Sescousse et al., 2013), has inherited a functional link with the central monitoring of physiological variables. Even when choosing between cultural goods that do not fulfill any immediate basic need, the neural monitoring of heartbeats supports self-reflection underlying evaluation, contributing to the precision of subjective decisions and fostering the stable expression of long-lasting preferences that define, at least in part, our identity.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

This work was supported by funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 670325, Advanced grant BRAVIUS) and a senior fellowship of the Canadian Institute For Advanced Research (CIFAR) program in Brain, Mind and Consciousness to CTB; by a doctoral fellowship from the Ecole des Neurosciences de Paris Ile de France to DA; and by ANR-17-EURE-0017.

The authors thank Clémence Alméras, Maximilien Chaumon and Christophe Gitton for assistance in data acquisition.

References

Abitbol R, Lebreton M, Hollard G, Richmond BJ, Bouret S, Pessiglione M (2015) Neural Mechanisms Underlying Contextual Dependency of Subjective Values: Converging Evidence from Monkeys and Humans. J Neurosci 35:2308–2320.

Al E, Iliopoulos F, Forschack N, Nierhaus T, Grund M, Motyka P, Gaebler M, Nikulin V V., Villringer A (2020) Heart-brain interactions shape somatosensory perception and evoked potentials. Proc Natl Acad Sci U S A 117:10575–10584.

Andrews-Hanna JR, Smallwood J, Spreng RN (2014) The default network and self-generated thought: Component processes, dynamic control, and clinical relevance. Ann N Y Acad Sci 1316:29–52.

Azzalini D, Rebollo I, Tallon-Baudry C (2019) Visceral Signals Shape Brain Dynamics and Cognition. Trends Cogn Sci 23:488–509.

Babo-Rebelo M, Richter CG, Tallon-Baudry C (2016a) Neural Responses to Heartbeats in the Default Network Encode the Self in Spontaneous Thoughts. J Neurosci 36:7829–7840.

Babo-Rebelo M, Wolpert N, Adam C, Hasboun D, Tallon-Baudry C (2016b) Is the cardiac monitoring function related to the self in both the default network and right anterior insula? Philos Trans R Soc B Biol Sci 371:20160004.

Bartra O, McGuire JT, Kable JW (2013) The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. Neuroimage 76:412–427.

Bechara A, Damasio H, Tranel D, Damasio AR (1997) Deciding Advantageously Before Knowing the Advantageous Strategy. Science (80-) 275:1293–1295.

Beck A, Ward CH, Mendelson M, Mock J, Erbaught J (1961) An Inventory for Measuring Depression. Arch Gen Psychiatry 4:561–571.

Berntson GG, Quigley KS, Lozano D (2007) Cardiovascular Psychophysiology. In: Handbook of Psychophysiology, 3rd ed. (Cacioppo JT, Tassinary LG, Berntson G, eds), pp 182–210. Cambridge University Press.

Blanke O, Metzinger T (2009) Full-body illusions and minimal phenomenal selfhood. Trends Cogn Sci 13:7–13.

Campana F, Tallon-Baudry C (2013) Anchoring visual subjective experience in a neural model: The coarse vividness hypothesis. Neuropsychologia 51:1050–1060.

Chib VS, Rangel A, Shimojo S, O'Doherty JP (2009) Evidence for a Common Representation of Decision Values for Dissimilar Goods in Human Ventromedial Prefrontal Cortex. J

Neurosci 29:12315–12320.

Clithero JA, Rangel A (2014) Informatic parcellation of the network involved in the computation of subjective value. Soc Cogn Affect Neurosci 9:1289–1302.

Craig AD (2002) How do you feel? Interoception: the sense of the physiological condition of the body. Nat Rev Neurosci 3:655–666.

Damasio AR (2010) Self Comes to Mind: Constructing the Conscious Brain. Pantheon Books.

Delgado MR, Beer JS, Fellows LK, Huettel SA, Platt ML, Quirk GJ, Schiller D (2016) Viewpoints: Dialogues on the functional role of the ventromedial prefrontal cortex. Nat Neurosci 19:1545–1552.

Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 134:9–21.

Desmedt O, Luminet O, Corneille O (2018) The heartbeat counting task largely involves non-interoceptive processes: Evidence from both the original and an adapted counting task. Biol Psychol 138:185–188.

Dirlich G, Vogl L, Plaschke M, Strian F (1997) Cardiac field effects on the EEG. Electroencephalogr Clin Neurophysiol 102:307–315.

Fellows LK (2006) Deciding how to decide: Ventromedial frontal lobe damage affects information acquisition in multi-attribute decision making. Brain 129:944–952.

Fischl B, Van Der Kouwe A, Destrieux C, Halgren E, Ségonne F, Salat DH, Busa E, Seidman LJ, Goldstein J, Kennedy D, Caviness V, Makris N, Rosen B, Dale AM (2004) Automatically Parcellating the Human Cerebral Cortex. Cereb Cortex 14:11–22.

Foa EB, Huppert JD, Leiberg S, Langner R, Kichic R, Hajcak G, Salkovskis PM (2002) The Obsessive-Compulsive Inventory: Development and validation of a short version. Psychol Assess 14:485–496.

Gray MA, Taggart P, Sutton PM, Groves D, Holdright DR, Bradbury D, Brull D, Critchley HD (2007) A cortical potential reflecting cardiac function. PNAS 104:6818–6823.

Grueschow M, Polania R, Hare TA, Ruff CC (2015) Automatic versus Choice-Dependent Value Representations in the Human Brain. Neuron 85:874–885.

Heekeren HR, Marrett S, Ungerleider LG (2008) The neural systems that mediate human perceptual decision making. Nat Rev Neurosci 9:467–479.

Hunt LT, Behrens TEJ, Hosokawa T, Wallis JD, Kennerley SW (2015) Capturing the temporal evolution of choice across prefrontal cortex. Elife 4:1–25.

Hunt LT, Kolling N, Soltani A, Woolrich MW, Rushworth MFS, Behrens TEJ (2012) Mechanisms underlying cortical activity during value-guided choice. Nat Neurosci

15:470–476.

Juechems K, Summerfield C (2019) Where does value come from? Trends Cogn Sci:836–850.

Kass RE, Raftery AE (1995) Bayes factors. J Am Stat Assoc 90:773–795.

Keramati M, Gutkin B (2014) Homeostatic reinforcement learning for integrating reward collection and physiological stability. Elife 3:1–26.

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.

Kubicek WG, Patterson RP, Witsoe DA (1970) Impedance Cardiography As a Noninvasive Method of Monitoring Cardiac Function and Other Parameters of the Cardiovascular System. Ann N Y Acad Sci 170:724–732.

Kurtz-David V, Persitz D, Webb R, Levy DJ (2019) The neural computation of inconsistent choice behavior. Nat Commun 10:1583.

Lebreton M, Jorge S, Michel V, Thirion B, Pessiglione M (2009) An Automatic Valuation System in the Human Brain: Evidence from Functional Neuroimaging. Neuron 64:431–439.

Levy DJ, Glimcher PW (2011) Comparing Apples and Oranges: Using Reward-Specific and Reward-General Subjective Value Representation in the Brain. J Neurosci 31:14693–14707.

Levy DJ, Glimcher PW (2012) The root of all value: A neural common currency for choice. Curr Opin Neurobiol 22:1027–1038.

Liang F, Paulo R, Molina G, Clyde MA, Berger JO (2008) Mixtures of g priors for Bayesian variable selection. J Am Stat Assoc 103:410–423.

Lopez-Persem A, Bastin J, Petton M, Abitbol R, Lehongre K, Adam C, Navarro V, Rheims S, Kahane P, Domenech P, Pessiglione M (2020) Four core properties of the human brain valuation system demonstrated in intracranial signals. Nat Neurosci 23:664–675.

Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. J Neurosci Methods 164:177–190.

McNamee D, Rangel A, O'Doherty JP (2013) Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex. Nat Neurosci 16:479–485.

Moeller SJ, Goldstein RZ (2014) Impaired self-awareness in human addiction : deficient attribution of personal relevance. Trends Cogn Sci 18:635–641.

Montoya P, Schandry R, Müller A (1993) Heartbeat evoked potentials (HEP): topography and

influence of cardiac awareness and focus of attention. Electroencephalogr Clin Neurophysiol Evoked Potentials 88:163–172.

Neubert F-X, Mars RB, Sallet J, Rushworth MFS (2015) Connectivity reveals relationship of brain areas for reward-guided learning and decision making in human and monkey frontal cortex. Proc Natl Acad Sci U S A 112:E2695-704.

Oostenveld R, Fries P, Maris E, Schoffelen JM (2011) FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput Intell Neurosci 2011:1–9.

Padoa-Schioppa C (2013) Neuronal origins of choice variability in economic decisions. Neuron 80:1322–1336.

Padoa-Schioppa C, Assad JA (2006) Neurons in the orbitofrontal cortex encode economic value. Nature 441:223–226.

Park H-D, Correia S, Ducorps A, Tallon-Baudry C (2014) Spontaneous fluctuations in neural responses to heartbeats predict visual detection. Nat Neurosci 17:612–618.

Park H-D, Tallon-Baudry C (2014) The neural subjective frame: from bodily signals to perceptual consciousness. Philos Trans R Soc Lond B Biol Sci 369:20130208.

Paulus MP (2007) Decision-Making Dysfunctions in Psychiatry—Altered Homeostatic Processing? Science (80-) 318:602–606.

Peters E, Joseph S, Day S, Garety P (2004) Measuring Delusional Ideation: The 21-Item Peters et aL Delusions Inventory (PDI). Schizophr Bull 30:1005–1022.

Polanía R, Krajbich I, Grueschow M, Ruff CC (2014) Neural Oscillations and Synchronization Differentially Support Evidence Accumulation in Perceptual and Value-Based Decision Making. Neuron 82:709–720.

Qin P, Northoff G (2011) How is our self related to midline regions and the default-mode network? Neuroimage 57:1221–1233.

Ress D, Heeger DJ (2003) Neuronal correlates of perception in early visual cortex. Nat Neurosci 6:414–420.

Rich EL, Wallis JD (2016) Decoding subjective decisions from orbitofrontal cortex. Nat Neurosci 19.

Ring C, Brener J, Knapp K, Mailloux J (2015) Effects of heartbeat feedback on beliefs about heart rate and heartbeat counting: A cautionary tale about interoceptive awareness. Biol Psychol 104:193–198.

Schandry R (1981) Heartbeat perception and emotional experience. Psychophysiology 18:483–488.

Sescousse G, Caldú X, Segura B, Dreher JC (2013) Processing of primary and secondary rewards: A quantitative meta-analysis and review of human functional neuroimaging studies. Neurosci Biobehav Rev 37:681–696.

Shadlen MN, Newsome WT (2001) Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. J Neurophysiol 86:1916–1936.

Sherwood A, Allen M, Fahrenberg J, Kelsey RM, Lovallo WR, van Doornen LJP (1990) Methodological Guidelines for Impedance Cardiography. Psychophysiology 27:1–23.

Spielberger CD, Gorsuch RL, Lushene R, Vagg PR, Jacobs GA (1983) Manual for the State-Trait Anxiety Inventory. In: Consulting Psychologists. Palo Alto, California.

Strait CE, Blanchard TC, Hayden BY (2014) Reward value comparison via mutual inhibition in ventromedial prefrontal cortex. Neuron 82:1357–1366.

Sui J, Gu X (2017) Self as Object: Emerging Trends in Self Research. Trends Neurosci 40:643–653.

Tadel F, Baillet S, Mosher JC, Pantazis D, Leahy RM (2011) Brainstorm: A user-friendly application for MEG/EEG analysis. Comput Intell Neurosci 2011:879716.

Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15:273–289.

Vaidya AR, Fellows LK (2020) Under construction: ventral and lateral frontal lobe contributions to value-based decision-making and learning. F1000Research 9:1–8.

Vaidya AR, Sefranek M, Fellows LK (2017) Ventromedial Frontal Lobe Damage Alters how Specific Attributes are Weighed in Subjective Valuation. Cereb Cortex 28:3857–3867.

Vinck M, van Wingerden M, Womelsdorf T, Fries P, Pennartz CMA (2010) The pairwise phase consistency: A bias-free measure of rhythmic neuronal synchronization. Neuroimage 51:112–122.

Webb R, Levy I, Lazzaro SC, Rutledge RB, Glimcher PW (2019) Neural random utility: Relating cardinal neural observables to stochastic choice behavior. J Neurosci Psychol Econ 12:45–72.

Zamariola G, Maurage P, Luminet O, Corneille O (2018) Interoceptive accuracy scores from the heartbeat counting task are problematic: Evidence from simple bivariate correlations. Biol Psychol 137:12–17.

Chapter V: Neural responses to heartbeats predict choice consistency between incommensurable goods

Neural responses to heartbeats predict choice consistency between incommensurable goods

Azzalini Damiano & Tallon-Baudry Catherine

Laboratoire de Neurosciences Cognitives et Computationnelles, Ecole Normale Supérieure, PSL University & Institut National de la Santé et de la Recherche Médicale, 75005, Paris, France

Abstract

Estimating the value of an item (e.g., an ice-cream) requires coordinating multiple attributes (its taste, smell and texture) into a single metric that guides choices. This coordination is accentuated in decisions between goods belonging to different categories (such as watching a movie vs. having an ice-cream outside) where very diverse, category-specific features have to be compared. How abstract value representations, found in ventromedial prefrontal cortex, arise from category-specific attributes, found in orbitofrontal cortex (OFC) remains unspecified. Neural responses to heartbeats have been shown to index self-related processes, including the computation of subjective values. Here, we test the more specific hypothesis that neural monitoring of heartbeats provides a format-independent signal enabling the coordination of distinct feature spaces. We recorded heartbeat-evoked responses (HERs) with magnetoencephalography (MEG) while 26 human participants performed a preference-based two-alternative force-choice task either between items belonging to the same category (e.g., movie vs. movie) or between items belonging to distinct categories (e.g., movie vs. food item). We reveal that, during decision, HERs in OFC and Insula are differentially recruited depending on decision type (same vs. different categories). Difference in HERs cannot be explained by changes in cardiac activity, in arousal, nor in behavioural differences between the two decision types. In addition, single trial HER amplitude fluctuations in OFC and Insula predict choice precision. Altogether, our results suggest that HERs enable the coordination of distinct feature spaces necessary for the comparison of otherwise incommensurable goods.

Introduction

Choosing between goods of different natures is an activity that we frequently perform in our daily life. For example, on a Sunday afternoon, you may ask yourself whether you would rather have an ice-cream outside or stay at home watching a movie. Making these decisions is subjectively simple, but the underlying neural implementation is not trivial. Goods of different natures are characterised by unique sensory and conceptual attributes (e.g., the taste, smell and texture of chocolate ice-cream vs. the emotional content and the visual aesthetics of a movie) that need to be mapped onto a common quantity, a subjective value, so that these dissimilar options can be compared. Some authors (Levy and Glimcher, 2011, 2012) propose that decisions between goods of different natures rely on an abstract neural representation of subjective values, which act as 'common currency'. Meta-analyses of fMRI studies indeed reveal that decisions between different categories of rewards appear to recruit common neural resources, such as ventromedial prefrontal cortex (vmPFC), ventral striatum (VS) and posterior cingulate cortex (PCC) (Levy and Glimcher, 2012; Bartra et al., 2013; Sescousse et al., 2013; Clithero and Rangel, 2014). Experimental results suggest the presence of a gradient in value representation going from category-specific value representation in central and lateral OFC (McNamee et al., 2013; Howard et al., 2015; Howard and Kahnt, 2017) to category-independent value representation in vmPFC (Chib et al., 2009; Fitzgerald et al., 2009; Levy and Glimcher, 2011; Gross et al., 2014). However, category-dependent and independent representations are not mutually exclusive, in fact vmPFC appears to encode both (McNamee et al., 2013; Zhang et al., 2017).

Despite the growing evidence of a common neural substrate for the representation of subjective value across categories, how different sensory and conceptual attributes can be mapped onto a unique value-space remains unknown. One possibility is that abstract, category-independent, subjective values found in vmPFC arise from the integration of disparate attributes from up-stream regions. In line with this view, studies show increased functional connectivity between vmPFC and temporal cortex during subjective ratings based on visuo-semantic information (Lim et al., 2013), as well as between vmPFC and OFC during food choice (Suzuki et al., 2017). However, the increased functional connectivity between areas does not specify the mechanism by which these attributes can be mapped onto a common space. In fact, subjectively evaluating an item is not a simple addition of its attributes, rather it relies on the weighted evaluation of the multifaceted subjective experience that the item produces (Vaidya et al., 2017).

The neural monitoring of heartbeats, a fundamental biological mechanism ensuring homeostasis, has been proposed to underlie the generation of first-person perspective, a key aspect of all subjective experiences (Park and Tallon-Baudry, 2014; Tallon-Baudry et al., 2018). Neural responses to heartbeats index self-relatedness in a number of experimental tasks (Babo-Rebelo et al., 2016a, 2016b, 2019; Park et al., 2016). Regarding subjective valuation, we have recently shown that the fluctuations in heartbeat-evoked responses (HERs) before option presentation contribute to preference-based decisions between cultural goods (i.e. movies) predicting the strength of neural encoding of subjective value and choice consistency (Azzalini et al., 2020). Beyond indexing subjective, self-related process, the neural monitoring of heartbeats would enable the coordination of information expressed in sensory- and conceptual-specific formats into a common space making it available to conscious self-reflective processes (Azzalini et al., 2019). The aim of this article is to experimentally test this hypothesis.

All subjective evaluations require some degree of coordination: a horror movie and a comedy are characterised by specific features. However, deciding between goods of the different categories require a more important coordination as their attributes are highly dissimilar, such as when comparing the plot of a movie to the taste of a chocolate ice-cream. We thus hypothesised that HERs would be differentially recruited depending on the amount of coordination required for subjective evaluation. Namely, HERs would be more involved in decisions between goods belonging to different categories require the coordination of more dissimilar feature spaces. To test this hypothesis, we recorded neural responses to heartbeats while 26 healthy human participants made preference-based decisions between goods belonging to either the same or different categories. We show that HERs in OFC and Insula differ depending on decision type. In addition, HER fluctuations in these two regions specifically predict changes in choice precision in decision between goods of different categories and same category, respectively. Altogether, these results show a functional contribution of HERs to subjective evaluation: they enable the coordination of attributes in distinct feature spaces that is particularly relevant for decisions between incommensurable goods.

Material and Methods

Participants

30 right-handed volunteers with normal or corrected-to-normal vision participated in the study after giving written informed consent. Participants received monetary compensation for their participation. All experimental procedures were approved by the ethics committee CPP Ile de France VI. Four subjects were excluded from analyses: one subjects for performance at chance level, one subject for excessive number of movement artefacts (20.76% of trials, 2 standard deviation above the mean 6,4%), two subjects for excessively fast cardiac rhythm (inter-beat interval: 0.66 and 0.53 second; mean ± SD: 0.83 ± 0.12 seconds). 26 subjects (13 female; mean age: 24.08 ± 2.46 years) were retained for subsequent analyses.

Task and procedure

Participants came to the lab for two experimental sessions (mean time elapsed between sessions: 34.1 ± 13 hours). In the first behavioural session, participants were asked to rate the likeability of individual items of different categories of goods. These ratings were subsequently used to build a participant-tailored two-alternative force choice (2AFC) design used in the second session, where magnetoencephalographic (MEG) data were acquired.

Rating session

Participants were presented with a list of 445 common items belonging to three different categories (activities, movies and food) and were asked to indicate the item that they liked the most and the least independently of the category. We explicitly instructed participants to use these items as reference points to rate all other items. Next, images and labels of each item were presented one by one at the centre of a computer screen in random order. Participants were asked to rate each item on a 21-point Likert scale using keyboard arrows and validate their choice with an additional button press. At each new item presentation, the cursor was randomly placed on the rating scale. If the presented item was unknown to the participants, they could skip that item (via keyboard button press) and move to the next one. No time constrain was imposed to participants during the rating phase.

Two-alternative forced-choice task design

We used likeability ratings from the first session to build 448 two-alternative force-choice trials. Each decision was either between items belonging to the same category of goods (e.g. chips vs. peanuts) or between goods of different categories (e.g. pizza vs. take a nap). We manipulated trial difficulty by pairing items with different subjective ratings, controlling that pairwise difference in value was independent of their sum. Participants performed 216 choices per condition (same/different category), with 72 trials for each possible 3 category combination (movie/movie, food/food and activity/activity in the same categories condition; activity/movie, activity/food and movie/food in the different categories condition). In the whole experiment, each item was presented a maximum of 6 times per experimental condition. The position on the screen of the most valuable item in each pair was fully counterbalanced so that the correct choice was presented an equal amount of times at the top and at the bottom of the screen. 16 pairs of stimuli (3.6% of trials; 8 pairs per condition) were assigned to catch trials (see below).

Experimental task

In the second session (**Figure 1A**), participants performed a two-alternative forced-choice task (2AFC) while sitting in the MEG. Each trial began with a fixation period of variable duration (uniformly distributed between 0.9 and 1.1 seconds in steps of 0.05 s) during which participants saw a black dot surrounded by a black ring (internal dot, 0.20° of visual angle; external black ring, 0.40° of visual angle) indicating that a new trial was about to start. Participants were instructed to refrain from blinking from that moment. Next, the outer ring of the fixation dot turned into either a square or diamond (0.40° and 0.56° visual angle, respectively; counterbalanced across participants) indicating which type of choice participants were about to face: they will be deciding either between items of a single category or between items of different categories. After a fixed delay of 1.5 seconds, the symbolic cue turned again into a ring and the names of two items were presented above and below the fixation dot. Participants had up to three seconds to indicate which of the two items they preferred using the index and the middle finger of their right hand. In standard trials (96.4%), the selected item remained on the screen for one second while the other disappeared. Next, participants were asked to express their level of confidence in the decision they just made on a 4-point scale within 3 seconds. Confidence levels were presented on the screen from lowest to highest (left to right) half of the trials and highest to lowest the other half (in random order). Finally, the central dot turned green and stayed on the screen for a variable duration of time (1.8 to 2.2

seconds in steps of 0.05 s) indicating that participants could blink and rest before a new trial began. The recording session consisted of 8 blocks of variable length (55 trials ± 1, depending on the number of catch trials).

In a small portion of trials (randomly ranging from 0 to 2 per block), participants were presented with pairs of stimuli that did not correspond to the condition cued by the geometric symbol (e.g., two items of different categories were displayed when the cue indicated that they would choose between two items of the same category). In this case, participants were required to report they had noticed the incongruence and press a button with their left index finger on another response pad. Participants received a positive (negative) feedback when they correctly (incorrectly) recognized these catch trials. The aim of catch trials was only to keep participants alert and engaged during the cue period, therefore we did not consider these trials in any further analyses.

Before the actual recording session, participants familiarised themselves with the experimental procedure by performing three training blocks. In each of the first two blocks (6 trials each), participants performed decisions of one type only (in the first block they always chose between items of the same category, in the second between items of different categories). Therefore, the cue symbol was the same within each block. In the third block (12 trials), choice types were randomly intermixed, similar to the recording session.

Second rating session

At the end of the 2AFC task, participants left the MEG chamber and performed a second rating session. The experimental procedure was the same as the one used in the first session, but only the items that participants already rated were used. As in the first session, participants carried out task without any time constrain.

Questionnaires

Finally, subjects filled six questionnaires in French: Beck's Depression Inventory (BDI) (Beck et al., 1961), Peter's et al. Delusions Inventory (PDI) (Peters et al., 2004), the Trait Anxiety Inventory (STAI) (Spielberger et al., 1983), the Obsessive-Compulsive Inventory (OCI) (Foa et al., 2002), the Dutch Eating Behaviour Questionnaire (DEBQ) (van Strien et al., 1986) and the Multidimensional Assessment Interoceptive Awareness (MAIA) (Mehling et al., 2012).

Recordings

Brain activity was recorded using a MEG system with 102 magnetometers and 204 planar gradiometers (Elekta Neuromag TRIUX, sampling rate 1000 Hz, online low-pass filter at 330 Hz). Heart electrical activity was simultaneously recorded with four electrodes (2 electrodes above the right and left clavicles and 2 over the left and right supraspinatus muscles (Gray et al., 2007); online filtered between 0.1-330 Hz, sampling frequency 1000 Hz) and referenced to another electrode on the left iliac region of the abdomen. This montage corresponds to four vertical derivations. Four horizontal derivations were obtained offline by subtracting the activity of two adjacent electrodes.

Eye blink muscular activity (vertical EOG) was recorded with two electrodes placed above and below participants' dominant eye. Pupil diameter and eye movements were recorded using an eye tracker device (EyeLink 1000, SR Research).

Cardiac events

For all participants, cardiac events were detected on the right-clavicle left abdomen electrocardiogram (ECG) vertical derivation. ECG time series were convolved with a cardiac cycle template, obtained from a clean portion of ECG data, and normalized between 0 and 1. R-peaks were identified as values exceeding 0.6 units. Timings with respect to the R-peak were used to identify all other cardiac waves. Specifically, the T-wave was identified as the maximum value occurring within 420 milliseconds after the Q-wave. The automatic labelling of R-peaks and T-waves was visually verified for each participant.

Inter-beat intervals (IBI) were computed as the difference between two consecutive R-peaks.

MEG data preprocessing

External noise was removed from MEG data applying MaxFilter algorithm (Elekta). Continuous data were high-pass filtered at 0.5 Hz (4^{th} order Butterworth filter). Muscle and movement artefacts occurring during the trial (i.e. from warning screen to the onset of inter-trial interval) were manually identified. All trials containing muscle or movement artefacts were excluded from further analyses (for participants included in the analysis, mean percentage of rejected trials 5 %, minimum 0 %, maximum 17%).

Independent component analysis (Delorme and Makeig, 2004), implemented in FieldTrip toolbox (Oostenveld et al., 2011), was applied to MEG data to reduce the influence cardiac artefact. Independent components were estimated on epochs of MEG data spanning 200 milliseconds before and after the R-peak. Only epochs that were free of any artefacts, eye blinks and saccades above 3 degrees were used for independent components estimation. We forced the maximal number of independent components to be equal to the rank of the MEG data. Mean pairwise phase consistency (PPC) was estimated between each independent component (Vinck et al., 2010) and the right clavicle-left abdomen ECG derivation signal in the frequency band 0-25 Hz. We removed from MEG continuous data all components (up to a maximum of three) that exceeded the mean PPC plus 3 standard deviations.

To correct for eye blink artefacts, we segmented clean MEG data into 2-seconds epochs and estimate independent components. Mean PCC was computed between independent components and vertical EOG. All independent components for which PPC exceeded mean + 3 standard deviations were removed. Stereotypical blink components were manually selected and removed from MEG data in six subjects, as automatized procedure failed to detect them.

ICA-corrected data were low-pass filtered at 25 Hz (6^{th} Butterworth filter).

Trials selection

All trials that were included in the analyses fulfil the following criteria. No muscular nor movement artefacts, the total amount of time affected by blink was less than 20% of the total trial time, response times were neither too short (< 250 ms) nor too long. To estimate long RTs, data in each experimental condition (choice between items of the same/different categories) were divided in six difficulty levels (corresponding to the difference in subjective values between the presented options). Next, mean RT for correct and incorrect trials separately was computed for each participant. All trials exceeding RT plus two standard deviations were considered too long and discarded. Catch trials, trials following catch trials, trials without any T-wave during decision phase were also removed from the analyses.

The mean number of trials retained for analyses was 332.62 ± 34.44 (mean ± SD, Min=250, Max=373).

Heartbeat-evoked responses

Heartbeat-evoked responses (HERs) were computed by locking MEG activity to the occurrences of T-peaks during the decision phase, i.e. when options were on the screen. In order to reduce the effect of transient visual responses to stimulus as well as those due to movement preparation, analyses were performed only on T-waves falling at least 300 milliseconds after options presentation and 350 milliseconds before button press. If multiple T-waves occurred during the decision phase, resulting heartbeat-evoked responses were averaged. To avoid contamination from residual cardiac artefactual activity (Dirlich et al., 1997), heartbeat-evoked responses were analyses from 50 ms after the T-wave.

We ensured that the difference in HERs between the two types of decision was not due to differences in ongoing brain activity unrelated to heartbeats. To this end, we broke the temporal alignment between MEG and ECG data by shuffling the timings of T-waves among trials of the same condition (e.g., the timings of the T-waves belonging to trial *i* were assigned to the trial *j* of the same condition). We then computed surrogate HERs using these shuffled timings for each condition separately. This procedure was repeated 500 times for each dataset (subject). For each permutation, we performed a clustering paired t-test and retained the largest sum(t) statistics for the positive and negative clusters. We thus obtained the distribution of the largest difference that could be obtained by randomly sampling neural activity in the choice period. To assess statistical significance, we computed the proportion of times in which the positive and negative surrogate statistics were at least as large as the ones obtained from the original HERs.

Nonparametric statistical testing of MEG data

We test for difference in HERs between decision types using cluster-based permutation two-tailed paired t-test (Maris and Oostenveld, 2007) implemented in the FieldTrip toolbox (Oostenveld et al., 2011). Sensor-level analyses were performed on magnetometer activity within a time-window spanning 50 to 300 milliseconds after the T-wave. This procedure identifies candidate clusters of neural activity based on spatio-temporal adjacency exceeding a statistical threshold ($p < 0.025$) for a given number of neighbouring sensors ($n = 3$). Each cluster is characterized by a cluster-level test statistic corresponding to the sum of all t values belonging to that cluster (sum(t)). Non-parametric null distribution of the test statistics is generated by randomly shuffling conditions labels 10,000 times and computing at each iteration the cluster statistics and saving the largest positive and negative sum(t). Monte Carlo p value corresponds

to the proportion of cluster statistics under the null distribution that exceed the original cluster-level test statistics. Note that because the largest chance values are retained to construct the null distribution, this method intrinsically corrects for multiple comparisons across time and space.

The same clustering procedure was used to assess the significance of the parameters estimates (two-tailed t-test against 0) obtained from the generalised linear model (see below) both at the sensor- and the source-level.

Generalised linear model

We used a generalised linear model (GLM) approach to control that differences in HERs could not been explained by differences in behaviour between the two decision types. To this aim, we modelled trial-by-trial z-scored HERs at each timepoint and channel as:

$$\boldsymbol{HER}_{(t,c)} = ß_0 + ß_1\boldsymbol{Decision\ type} + ß_2\boldsymbol{RT} + ß_3\boldsymbol{Consistency}$$

Where ***Decision type*** is a categorical variable indicating which type of decision was performed (+1 between goods of same category, -1 goods of different categories). ***RT*** is the z-scored log-transformed trial-by-trial reaction time. ***Consistency*** is a categorical variable indicating consistent or inconsistent choice (+1 and -1, respectively). Consistent choice was defined when participants chose the highest valued option, based on subjective ratings provided in the first session. The GLM procedure provided us with time series of parameter estimates (betas, ß) that we tested against zero for significance using the clustering procedure (Maris and Oostenveld, 2007). Parameter estimates time series were then reconstructed at the source-level.

MR acquisition and preprocessing

For all participants but one, we acquired a T1 MRI scan using a 3 Tesla Siemens VERIO. We used the automated procedure in FreeSurfer package (Fischl et al., 2004) to obtain the cortical segmentation from MRI. The results for each subject were visually inspected and used for minimum-norm estimation.

Source reconstruction

We reconstructed the cortical location of neural activity using Brainstorm toolbox (Tadel et al., 2011). To this end, we co-registered individual anatomy (for the subject without T1 MRI, we used the default anatomy provided by Brainstorm) and MEG sensors. Next, we estimated 15,002 current dipoles, constrained to be orthogonal to the cortical surface, using a linear inverse solution from the ensemble of MEG sensors (magnetometers and planar gradiometers; weighted minimum-norm, SNR = 3, whitening PCA, depth weighting = 0.5) with overlapping spheres head model.

We reconstructed beta time series at the source level for the entire HER time-window on participants' individual anatomy, then we spatially smoothed (FWHM 6 mm) the obtained time series and projected them onto a standard brain model (ICBM152_T1, 15,002 vertices). To reveal which cortical areas were responsible of the GLM effect observed at the sensor-level, we used the same non-parametric two-tailed t-test clustering procedure on source-reconstructed parameter estimates across time and space.

Median-split analysis on source-level HER amplitude was performed by reconstructing trial-by-trial HER activity in the entire HER time-window. Vertices and time-points belonging to the clusters identified through the non-parametric clustering procedure were then averaged to produce one HER amplitude per trial.

Pupillometry

Pupil data that contained blinks (automatically detected with EyeLink software and extended before and after by 150 ms), saccades beyond 2 degrees and segments in which pupil size changed abruptly (signal temporal derivative exceeding 0.3, arbitrary unit) were linearly interpolated. Continuous pupil data from each experimental block were then band-pass filtered between 0.01 and 10 Hz (second order Butterworth) and z-scored. Pupil data were baseline corrected using the average pupil diameter before cue onset (400-200 milliseconds before cue presentation). 25 subjects were retained for pupil analysis; one subject was excluded due to technical problems at the acquisition. Analyses were performed in two ways: (a) pupil signal was locked to option presentation and averaged from 0 to 1.4 seconds following the presentation, and (2) pupil signal was locked to response delivery and averaged in the time-window spanning 1 second before the response up to its occurrence.

Bayes Factor

We used Bayes Factor (BF) to quantify the evidence in support of the null hypothesis (i.e., no difference between two measures) in paired t-tests. We computed the maximum log-likelihood for a normally distributed model in favour of the alternative hypothesis and for the model favouring the null hypothesis, adjusting the reference effect to correspond to a p-value of 0.05 for our sample size (n = 26 for all analysis except for pupil analysis in which n = 25). We then compared the two models by computing Bayesian information criterion and the corresponding Bayes Factor.

Results

Behavioural results

Participants performed a two-alternative forced choice task in which they indicated their preferred option. Our experimental manipulation consisted in presenting participants either with two items belonging to the same category of good (food vs. food, movie vs. movie, activity vs. activity) or with two items belonging to different categories (e.g., food vs. movie, activity vs. food or activity vs. movie).

Participants consistently choose the item with the highest subjective value (retrieved from the rating session; mean choice consistency across decision types 71%; two-tailed t-test vs. chance, $t(25) = 17.41$, $p<10^{-16}$). Difference in subjective value (value difference, VD) between the two options significantly affected choice consistency (two-ways repeated measures ANOVA, main effect of VD on consistency, $F(5,125)=60.01$, $p < 10^{-6}$; **Figure 1B**) and reaction times (two-ways repeated measures ANOVA, main effect of VD on RT, $F(5,125)=7.11$, $p < 10^{-5}$; **Figure 1C**). When choosing between items belonging to the same category, participants were both more consistent and faster (two-ways repeated measures ANOVA, main effect of decision type on consistency, $F(1,25)=13.30$, $p = 0.001$; main effect of decision type on RT, $F(1,25)=26.23$, $p < 10^{-4}$). However, the difference in subjective value affected consistency and RT equally in both decision types (two-ways repeated measures ANOVA, no interaction effect of VD*decision type on performance, $F(5,125)=0.94$, $p = 0.46$; no interaction effect of VD*decision type on RT, $F(5,125)=1.87$, $p = 0.11$).

After response, subjects were prompted to report their subjective feeling of confidence in their decision. In both decision types, they were more confident when value difference between the items was large (two-ways repeated measures ANOVA, main effect of VD on confidence, $F(5,125)=22.87$, $p < 10^{-15}$; no interaction between VD and decision type, $F(5,125) = 0.75$, $p = 0.59$; **Figure 1D**). Participants were also overall more confident when choosing between items belonging to the same category (two-ways repeated measures ANOVA, main effect of decision type on confidence, $F(1,25)=15.50$, $p < 10^{-3}$).

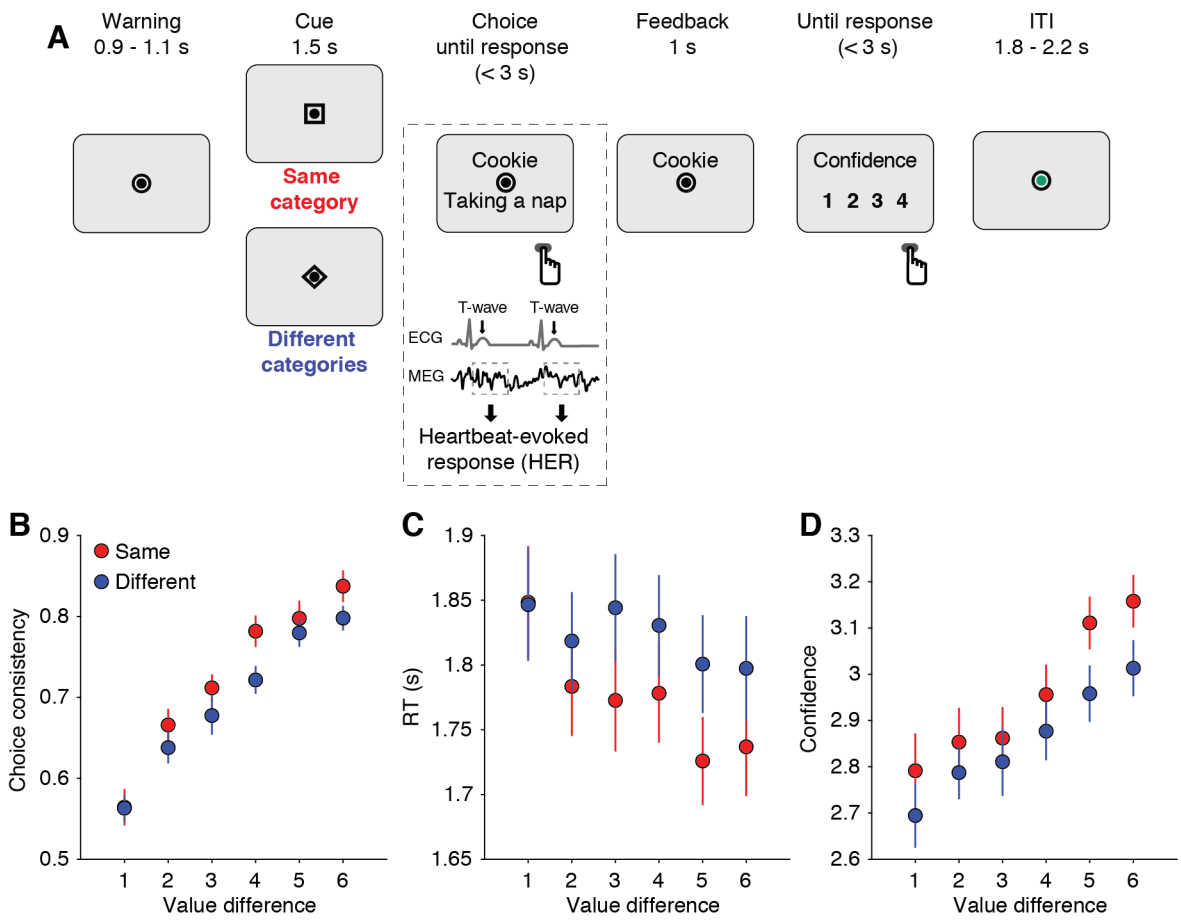


Figure 1. Task design and behavioural results. A, Each trial began with a warning period, indicated by a fixation dot, of variable duration (0.9 to 1.1 seconds) in which participants were asked to stop blinking and get ready. Next, a symbolic cue (square or diamond) appeared at the centre of the screen for a fixed delay (1.5 s) indicating which type of preference-based decision (either between items belonging to the same category or different ones) participants were about to perform. Participants had up to 3 seconds to choose their preferred item via button press. After decision, the selected option remained on the screen for 1 second. Finally, participants had to report the confidence in the decision they just made on a 4-point scale. **B**, Participants consistency (selection of the item with the highest value, as revealed in the preceding rating session) increased with the difference in value between the two options in both tasks ($F(5,125)=60.01$, $p < 10^{-6}$). Consistency was generally larger when deciding between options of the same category ($F(1,25)=13.30$, $p = 0.001$). **C**, Reaction times were overall faster for decision between items of the same category ($F(1,25)=26.23$, $p < 10^{-4}$), but the decrease in reaction times as difference in value between the alternatives increased was similar in both decision types ($F(5,125)=1.87$, $p = 0.11$). **D**, Subject were generally more confident when deciding between items belonging to the same category ($F(1,25)=15.50$, $p < 10^{-3}$). Confidence increases as the value difference between the alternative increased ($F(5,125)=22.87$, $p < 10^{-15}$) similarly in both decision types ($F(5,125) = 0.75$, $p = 0.59$).

Neural responses to heartbeats differ between decision types

We hypothesised that heartbeat-evoked responses would be differentially recruited when subjects decide between goods of the same category vs. goods of different ones. In the former type of decisions, items lie in the same feature space (for example, when choosing between two food items the features comprise the taste, textures, caloric content, …). Conversely, decisions between goods of different categories require a more considerable coordination by which category-specific attributes lying in distinct feature spaces must be aligned. Neural responses to heartbeats have been recently proposed to fulfil this coordination mechanism, by working as common reference signal for information spanning distinct spaces (Azzalini et al., 2019).

To test this hypothesis, we compared heartbeat-evoked responses during the two types of decisions. A non-parametric clustering procedure that corrects for multiple comparisons across space and time (Maris and Oostenveld, 2007), revealed that HERs differed depending on decision type in two distinct spatio-temporal clusters (positive cluster, 0.134 – 0.207 seconds after T-wave, sum(t) = 2038, Monte Carlo cluster-level p = 0.007; negative cluster, 0.203 – 0.267 seconds after T-wave, sum(t) = -1541, Monte Carlo cluster-level p = 0.02; **Figure 2A-B**). To ascertain that the observed difference in HER was time-locked to heartbeats and could not be explained by differences in underlying baseline activity between the two decision types, we performed the same analysis on "surrogate heartbeats" whereby the temporal relationship between neural and cardiac activity was broken (see Methods). The results show that the observed HER difference was indeed time-locked to heartbeats and could not be accounted by a difference in baseline activity (Monte Carlo p = 0.008).

As revealed by our behavioural analyses, reaction times and choice consistency differed between decision types. To ensure that these differences could not account for the observed HER effect, we modelled single trial HER time series using decision type, reaction time and consistency as regressors. Non-parametric clustering procedure on parameter estimates showed that decision type was encoded in two clusters of neural activity with topographies and time-courses very similar to those observed in the original HER contrast ($ß_{Decision\ type}$, positive cluster, 0.136 – 0.216 seconds after T-wave, sum(t) = 2663, Monte Carlo cluster-level p = 0.001; negative cluster, 0.203 – 0.273 seconds after T-wave, sum(t) = -2114, Monte Carlo cluster-level p = 0.004; **Figure 2C-D**). In addition, the GLM analysis shows that single trial HER amplitude was significantly predicted by reaction times in three spatio-temporal clusters (**Supplementary Table 1**), but not by choice consistency (3 candidate clusters, minimum

Monte Carlo p = 0.96). Altogether, these results show that HERs differ depending on decision type and that this difference cannot be explained by differences in ongoing brain activity unrelated to heartbeats nor by behavioural differences between the two tasks.

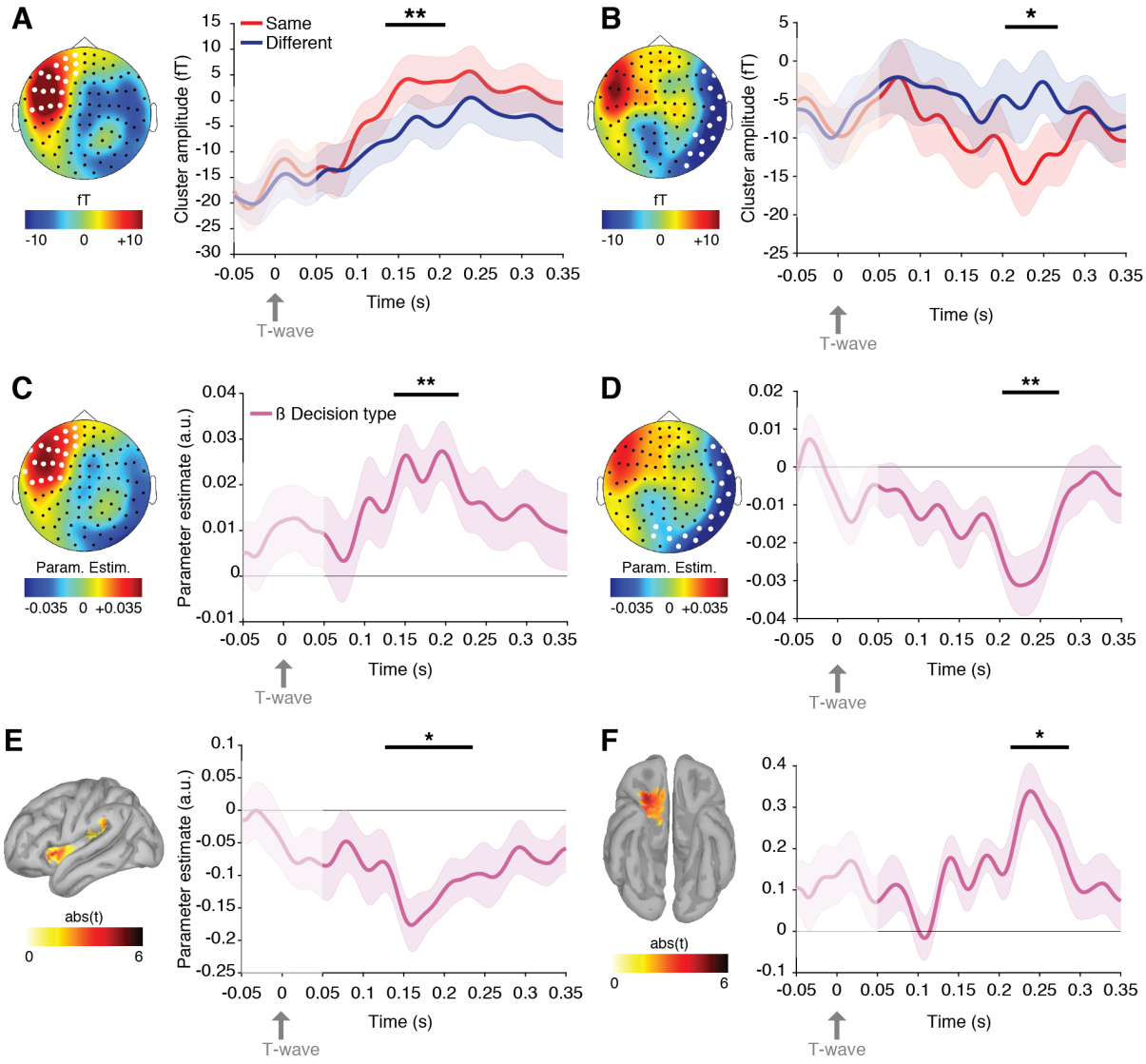


Figure 2. Heart-evoked responses amplitude depends on decision type. A, Topography (left) and time series (right, amplitude ± sem) of the significant positive HER cluster differing between decisions types (0.134 – 0.207 seconds after T-wave, Monte Carlo cluster-level p = 0.007). **B**, Same as A, but for the negative significant cluster (0.203 – 0.267 seconds after T-wave, Monte Carlo cluster-level p = 0.02). **C,** Topography (left) and time series (right, mean + sem) of the significant positive cluster encoding decision type as revealed by GLM approach (0.136 – 0.216 seconds after T-wave, Monte Carlo cluster-level p = 0.001;). **D**, Same as C, but for the significant negative cluster (0.203 – 0.273 seconds after T-wave, Monte Carlo cluster-level p = 0.004). **E-F**, T-maps of the cortical regions (left) and time series of the parameter estimate of the highlighted clusters (right, mean + sem) in which HERs significantly co-

vary with decision type (**E**, Left Insula, significant cluster 0.128 – 0.235 seconds after T-wave, Monte Carlo $p = 0.02$; **F**, right OFC, 0.214 – 0.286 seconds after T-wave, Monte Carlo cluster-level $p = 0.035$). Channels belonging to the significant clusters are highlighted in white in all topographic maps. Black bars on top of time series indicate the significant time-windows, as revealed by the clustering procedure. Shaded areas on time series indicate the 50 ms after T-wave that were not considered for statistical testing. *$p < 0.05$ ** $p < 0.01$

HER difference is of neural origin and cannot be accounted by changes in cardiac activity nor in arousal

The two decision types did not differ in heart electrical activity (t-test on mean ECG activity for vertical and horizontal ECG derivations, all $p \geq 0.45$, $BF \leq 0.39$; no significant clusters revealed by the non-parametric clustering t-test on the entire time-window; **Supplementary Table 2**) or in cardiac rhythm (two-tailed paired t-test on inter-beat intervals (IBI); IBI for two heartbeats around option presentation, $t(25)=0.8$, $p=0.45$, $BF = 0.37$; IBI for two heartbeats preceding response delivery, $t(25)=0.349$, $p=0.730$, $BF = 0.27$; IBI computed between the T-wave used for HER calculation and the following one, $t(25)=0.961$, $p=0.35$, $BF = 0.47$) indicating that HER difference is of neural origin and does not reflect changes in the cardiac input.

In addition, arousal measures during decision (pupil diameter, alpha power, stimulus-locked and response-locked) and brain responses to visual stimuli (visual-evoked potentials) did not differed between decision types (**Supplementary Table 3**). Altogether, these results reveal that HER differences reflect brain responses time-locked to heartbeats that cannot be accounted for by changes in cardiac inputs or by differences in overall brain states.

HER differences between decision types arise from right OFC and left Insula

To identify the cortical sources generating the HER effect observed at the sensor-level, we projected the time series of $ß_{Decision\ type}$ on individual brain anatomies and performed the non-parametrical clustering procedure with intrinsic correction for multiple comparisons against zero across time and space. Two clusters of HER activity encoded decision type (**Figure 2E-F**), one localised in left insular cortex (0.128 – 0.235 seconds after T-wave, abs(sum(t)) = 38379, Monte Carlo cluster-level p = 0.02; cluster peak at MNI coordinates [-43 18 -4], t-value = 6.18) and the other in right orbitofrontal cortex (0.214 – 0.286 seconds after T-wave, abs(sum(t)) = 34028, Monte Carlo cluster-level p = 0.035; cluster peak at MNI coordinates [20 21 -25], t-value = 6.22). The timings of these clusters are consistent with sensor-level analyses.

Trial-by-trial HER amplitude in OFC and Insula predicts choice consistency

If neural responses to heartbeats reflect the extent to which category-specific attributes are successfully mapped onto a common space, one would expect trial-by-trial fluctuations in HER amplitude to predict the precision of these type of choices. To test this hypothesis, we median-split trials in which participants chose between goods of different categories depending

on HER amplitude. Since the GLM at sensor-level showed that reaction times significantly predicted HERs, we partialled out from source-reconstructed HER amplitude the variance explained by reactions times before splitting the trials. This analysis revealed that trials with larger HER in right OFC are characterised by a steeper psychometric function as compared to trials with smaller HER (two-tailed paired t-test on slope, small vs. large HER: $t(25)=-2.21$, $p=0.04$; **Figure 3A-B**), but no difference in criterion ($t(25)=0.12$, $p=0.90$, $BF=0.25$). This effect was limited to OFC: the same analysis based on HER amplitude in the left Insula did not show any significant difference in slope (two-tailed paired t-test, small vs. large HER, $t(25)=-0.63$, $p=0.54$, $BF=0.33$) or criterion (two-tailed paired t-test, small vs. large HER, $t(25)=0.19$, $p=0.85$, $BF=0.25$; **Figure 3C-D**).

When performing the same analysis in trials where participants decided between goods belonging to the same category, we found the opposite result. Lower HER amplitude in the left Insula was associated with a steeper slope of the psychometric function (two-tailed paired t-test on slope, small vs. large HER, $t(25)=2.36$, $p=0.03$; two-tailed paired t-test on criterion, $t(25)=-0.37$, $p=0.72$, $BF = 0.27$; **Figure 3G-H**), while HER amplitude in right OFC revealed no effect (two-tailed paired t-test on slope, small vs. large HER, $t(25)=-1.26$, $p=0.22$, $BF = 0.71$; two-tailed paired t-test on criterion, $t(25)=0.16$, $p=0.87$, $BF = 0.25$; **Figure 3E-F**).

These results reveal that HER amplitude in right OFC and left Insula differently predicts single trial choice consistency depending on decision type. Whereas larger HERs in OFC predict more consistent choices in decisions involving goods belonging to different categories, more negative HERs in the left Insula predict more consistent choices for goods of the same nature.

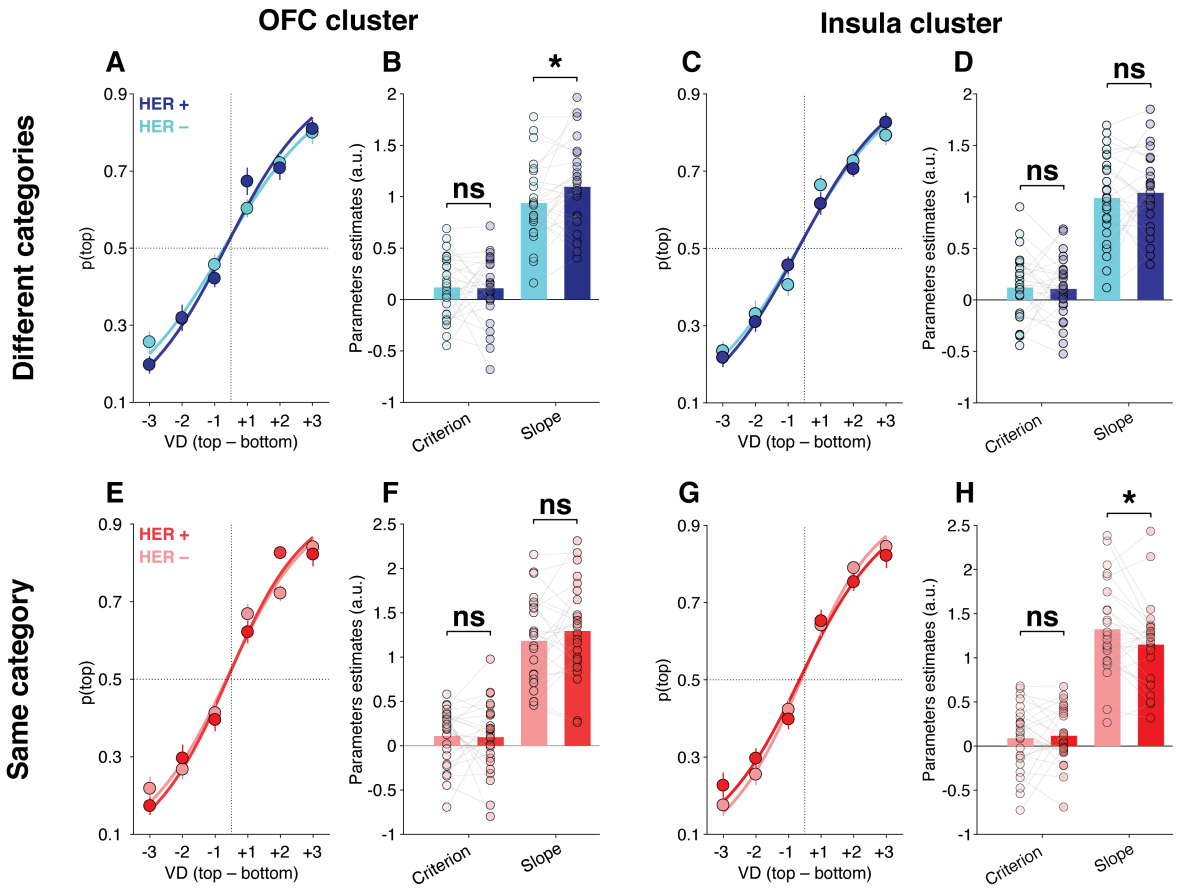


Figure 3. HER amplitude in right OFC and left Insula predicts choice precision in decision between items of different and same categories, respectively. A, Averaged psychometric functions (lines are psychometric fits, dots are data) for decisions between goods of different categories separately estimated for trials with large and small HER amplitude in right OFC. B, Criterion and slope average estimates (bars) for all participants (dots) for decisions between goods of different categories showing a steeper slope for trials with larger HER amplitude in OFC. C, Same as A, for trials median split based on HER amplitude in the left Insula. D, Same as B, showing no difference in psychometric estimates depending on HER amplitude. E, Averaged psychometric functions (lines are psychometric fits, dots are data) for decisions between goods of the same category separately estimated for trials with large and small HER amplitude in right OFC. F, Criterion and slope average estimates (bars) for all participants (dots) for decisions between goods of different categories show no difference depending on HER amplitude in OFC. G, Same as E but trials were median split according to HER amplitude in left Insula. H, Same as F, showing a steeper slope for trials with smaller HER amplitude in left Insula cluster. *p < 0.05

Discussion

Deciding between goods of different natures characterised by distinct attributes relies on mapping disparate sensory- and conceptual features onto a common space so that alternatives can be compared. Here, we provide evidence that neural responses to heartbeats contribute to this coordinating mechanism. We show that HER amplitudes differ depending on whether preference-based decisions are performed between goods of different categories or belonging to the same category. Our GLM analysis rules out the possibility that this effect arises from overall behavioural differences between the two decision types. In addition, we show that changes in cardiac activity, in arousal and in overall brain states cannot account for differences in HERs. Altogether, these results indicate that variations in HER amplitude reflect a genuine change in brain responses to cardiac inputs depending on decision type.

Source reconstruction revealed that the HER difference originates in the left anterior and posterior insular cortex, and in the right OFC. The posterior insular cortex is a known viscerosensory cortex (Cechetto and Saper, 1987), although probably not the only one (Azzalini et al., 2019). Previous intracranial electrophysiological recordings in humans showed the presence of neural responses to heartbeats in the insula (Canales-Johnson et al., 2015; Babo-Rebelo et al., 2016b; Park et al., 2017). HERs in OFC have seldom been reported (Canales-Johnson et al., 2015), with the majority of studies revealing neural responses to heartbeats in the ventromedial area of prefrontal cortex (Park et al., 2014; Babo-Rebelo et al., 2016a, 2016b; Azzalini et al., 2020). However, tracing studies in monkeys reveal that posterior (13m and 13l) and lateral (12m) OFC also receives visceral inputs from the ventromedial posterior nucleus of the thalamus (VMpC) via agranular and dysgranular insular areas (Carmichael and Price, 1995). The fact that HERs in OFC temporally lag those in the Insula would be in keeping with a projection from Insula to OFC. However, HERs in the insula influence within-category choices, whereas HERs in the OFC influence between-category choices, indicating functional independence. Further work is needed to assess whether activity in these regions reflects independent responses arising from separate thalamo-cortical pathways, or whether insular cortex directly feeds cardiac information to OFC.

Neural responses to heartbeats thus appear to reflect the extent to which the mapping of different attributes into a common space takes place. It follows that choice consistency in decisions between goods of different natures, that more significantly necessitate this coordination, should be related to trial-by-trial variations in HERs. The results provide evidence in favour of this hypothesis: single trial HER amplitude in OFC covaried with choice precision

in decisions between goods of different categories. Different populations of neurons in OFC encode the value of distinct attributes (Hosokawa et al., 2013; Suzuki et al., 2017) associated with specific reward identity (Howard et al., 2015). In addition, OFC has been proposed to act as a 'filter' enabling subjects to identify the information that is relevant for the task at hand (Wilson et al., 2014). Neural responses to heartbeats may thus bind the attributes relevant for value assignment into a unique space. These attributes may be subsequently passed on to other brain regions, such as vmPFC, for their explicit representation and comparison (Noonan et al., 2010; Walton et al., 2015). This view is compatible with proposals arguing that vmPFC encodes an abstract representation of subjective value (Levy and Glimcher, 2011, 2012; Suzuki et al., 2017), but it further specifies a neural mechanism through which the attributes encompassing different feature spaces can be held together enabling the neural representation that guides choice.

More unexpectedly, we find that fluctuations in HER amplitude in insular cortex predict the precision in choice consistency in decisions between goods belonging to the same category. Although these decisions also involve a coordination process, this is more limited as the attributes pertain to the same or very similar feature spaces (evaluating two food items is done on roughly the same attributes: the taste, the smell and the texture, for instance). The insula has been linked to various functional roles: the integration of interoceptive signals into ongoing cognitive processes (Craig, 2009) such as self-identification and self-location (Blanke et al., 2015), saliency encoding (Litt et al., 2011) and the switching between internal and external attention (Menon and Uddin, 2010). All these processes are present to some extent in the decision elicited by our paradigm. Therefore, a conclusive interpretation of the role of the HERs in the Insula during preference-based decision between goods of the same category is currently not possible.

We previously observed (Azzalini et al., 2020) that HERs fluctuations in vmPFC, not in Insula, predicted choice consistency in a preference judgement involving items of the same category, namely movies. A number of reasons can account for this difference. First, the vmPFC effect was revealed by contrasting within-category preference-based decisions against a perceptual discrimination task. Second, this contrast was performed during the instruction period – hence before option presentation – while here HERs are analysed during decision. Finally, in the current study, we included goods related to bodily aspects (food and activity). An interesting possibility is that the relationship between HER fluctuations in the Insula and choice precision for decisions between goods of the same category is driven by stimuli that are

directly related to bodily aspects. Should this be the case, we should observe a stronger relation between HER fluctuations in the Insula and choice precision for decisions involving bodily-related aspects, such as food items and (potentially) activities, and smaller change for decisions between movies.

The coordination of distinct feature spaces is crucial when deciding between goods of different natures, which share few – if any – attributes. Multiple studies have shown the existence of an abstract, category-independent, neural representation of subjective value in vmPFC, which can enable the comparison of goods of different natures. Nonetheless, the mechanism by which category-specific attributes are mapped onto a common value scale remains unspecified. Our results provide initial evidence that the neural monitoring of heartbeats may function as a common reference signal enabling the mapping of domain- and conceptual-specific attributes onto a common self-referential space from which a single metric – subjective value – can arise. This initial evidence opens new leads to the understanding of how the unity of subjective experience can arise from information expressed in distinct representational formats.

References

Azzalini D, Buot A, Palminteri S, Tallon-Baudry C (2020) Responses to heartbeats in ventromedial prefrontal cortex contribute to subjective preference-based decisions. bioRxiv:776047.

Azzalini D, Rebollo I, Tallon-Baudry C (2019) Visceral Signals Shape Brain Dynamics and Cognition. Trends Cogn Sci 23:488–509.

Babo-Rebelo M, Buot A, Tallon-Baudry C (2019) Neural responses to heartbeats distinguish self from other during imagination. Neuroimage 191:10–20.

Babo-Rebelo M, Richter CG, Tallon-Baudry C (2016a) Neural Responses to Heartbeats in the Default Network Encode the Self in Spontaneous Thoughts. J Neurosci 36:7829–7840.

Babo-Rebelo M, Wolpert N, Adam C, Hasboun D, Tallon-Baudry C (2016b) Is the cardiac monitoring function related to the self in both the default network and right anterior insula? Philos Trans R Soc B Biol Sci 371:20160004.

Bartra O, McGuire JT, Kable JW (2013) The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. Neuroimage 76:412–427.

Beck A, Ward CH, Mendelson M, Mock J, Erbaught J (1961) An Inventory for Measuring Depression. Arch Gen Psychiatry 4:561–571.

Blanke O, Slater M, Serino A (2015) Behavioral, Neural, and Computational Principles of Bodily Self-Consciousness. Neuron 88:145–166.

Canales-Johnson A, Silva C, Huepe D, Rivera-Rei Á, Noreika V, Del Carmen Garcia M, Silva W, Ciraolo C, Vaucheret E, Sedeño L, Couto B, Kargieman L, Baglivo F, Sigman M, Chennu S, Ibáñez A, Rodríguez E, Bekinschtein TA (2015) Auditory feedback differentially modulates behavioral and neural markers of objective and subjective performance when tapping to your heartbeat. Cereb Cortex 25:4490–4503.

Carmichael ST, Price JL (1995) Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. J Comp Neurol 363:642–664.

Cechetto DF, Saper CB (1987) Evidence for a viscerotopic sensory representation in the cortex and thalamus in the rat. J Comp Neurol 262:27–45.

Chib VS, Rangel A, Shimojo S, O'Doherty JP (2009) Evidence for a Common Representation of Decision Values for Dissimilar Goods in Human Ventromedial Prefrontal Cortex. J Neurosci 29:12315–12320.

Clithero JA, Rangel A (2014) Informatic parcellation of the network involved in the

computation of subjective value. Soc Cogn Affect Neurosci 9:1289–1302.

Craig AD (2009) How do you feel — now? The anterior insula and human awareness. Nat Rev Neurosci 10:59–70.

Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 134:9–21.

Dirlich G, Vogl L, Plaschke M, Strian F (1997) Cardiac field effects on the EEG. Electroencephalogr Clin Neurophysiol 102:307–315.

Fischl B, Van Der Kouwe A, Destrieux C, Halgren E, Ségonne F, Salat DH, Busa E, Seidman LJ, Goldstein J, Kennedy D, Caviness V, Makris N, Rosen B, Dale AM (2004) Automatically Parcellating the Human Cerebral Cortex. Cereb Cortex 14:11–22.

Fitzgerald THB, Seymour B, Dolan RJ (2009) The Role of Human Orbitofrontal Cortex in Value Comparison for Incommensurable Objects.

Foa EB, Huppert JD, Leiberg S, Langner R, Kichic R, Hajcak G, Salkovskis PM (2002) The Obsessive-Compulsive Inventory: Development and validation of a short version. Psychol Assess 14:485–496.

Gray MA, Taggart P, Sutton PM, Groves D, Holdright DR, Bradbury D, Brull D, Critchley HD (2007) A cortical potential reflecting cardiac function. PNAS 104:6818–6823.

Gross J, Woelbert E, Zimmermann J, Okamoto-Barth S, Riedl A, Goebel R (2014) Value Signals in the Prefrontal Cortex Predict Individual Preferences across Reward Categories. J Neurosci 34:7580–7586.

Hosokawa T, Kennerley SW, Sloan J, Wallis JD (2013) Single-neuron mechanisms underlying cost-benefit analysis in frontal cortex. J Neurosci 33:17385–17397.

Howard JD, Gottfried JA, Tobler PN, Kahnt T (2015) Identity-specific coding of future rewards in the human orbitofrontal cortex. Proc Natl Acad Sci U S A 112:5195–5200.

Howard JD, Kahnt T (2017) Identity-specific reward representations in orbitofrontal cortex are modulated by selective devaluation. J Neurosci 37:2627–2638.

Levy DJ, Glimcher PW (2011) Comparing Apples and Oranges: Using Reward-Specific and Reward-General Subjective Value Representation in the Brain. J Neurosci 31:14693–14707.

Levy DJ, Glimcher PW (2012) The root of all value: A neural common currency for choice. Curr Opin Neurobiol 22:1027–1038.

Lim S-L, O'Doherty JP, Rangel A (2013) Stimulus Value Signals in Ventromedial PFC Reflect the Integration of Attribute Value Signals Computed in Fusiform Gyrus and Posterior Superior Temporal Gyrus. J Neurosci 33:8729–8741.

Litt A, Plassmann H, Shiv B, Rangel A (2011) Dissociating valuation and saliency signals during decision-making. Cereb Cortex 21:95–102.

Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. J Neurosci Methods 164:177–190.

McNamee D, Rangel A, O'Doherty JP (2013) Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex. Nat Neurosci 16:479–485.

Mehling WE, Price C, Daubenmier JJ, Acree M, Bartmess E, Stewart A (2012) The Multidimensional Assessment of Interoceptive Awareness (MAIA) Tsakiris M, ed. PLoS One 7:e48230.

Menon V, Uddin LQ (2010) Saliency, switching, attention and control: a network model of insula function. Brain Struct Funct 214:655–667.

Noonan MP, Walton ME, Behrens TEJ, Sallet J, Buckley MJ, Rushworth MFS (2010) Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. Proc Natl Acad Sci U S A 107:20547–20552.

Oostenveld R, Fries P, Maris E, Schoffelen JM (2011) FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput Intell Neurosci 2011:1–9.

Park H-D, Bernasconi F, Bello-Ruiz J, Pfeiffer C, Salomon R, Blanke O (2016) Transient Modulations of Neural Responses to Heartbeats Covary with Bodily Self-Consciousness. J Neurosci 36:8453–8460.

Park H-D, Bernasconi F, Salomon R, Tallon-Baudry C, Spinelli L, Seeck M, Schaller K, Blanke O (2017) Neural Sources and Underlying Mechanisms of Neural Responses to Heartbeats, and their Role in Bodily Self-consciousness: An Intracranial EEG Study. Cereb Cortex 28:1–14.

Park H-D, Correia S, Ducorps A, Tallon-Baudry C (2014) Spontaneous fluctuations in neural responses to heartbeats predict visual detection. Nat Neurosci 17:612–618.

Park H-D, Tallon-Baudry C (2014) The neural subjective frame: from bodily signals to perceptual consciousness. Philos Trans R Soc Lond B Biol Sci 369:20130208.

Peters E, Joseph S, Day S, Garety P (2004) Measuring Delusional Ideation: The 21-Item Peters et aL Delusions Inventory (PDI). Schizophr Bull 30:1005–1022.

Sescousse G, Caldú X, Segura B, Dreher JC (2013) Processing of primary and secondary rewards: A quantitative meta-analysis and review of human functional neuroimaging studies. Neurosci Biobehav Rev 37:681–696.

Spielberger CD, Gorsuch RL, Lushene R, Vagg PR, Jacobs GA (1983) Manual for the State-

Trait Anxiety Inventory. In: Consulting Psychologists. Palo Alto, California.

Suzuki S, Cross L, O'Doherty JP (2017) Elucidating the underlying components of food valuation in the human orbitofrontal cortex. Nat Neurosci 20:1780–1786.

Tadel F, Baillet S, Mosher JC, Pantazis D, Leahy RM (2011) Brainstorm: A user-friendly application for MEG/EEG analysis. Comput Intell Neurosci 2011:879716.

Tallon-Baudry C, Campana F, Park H-D, Babo-Rebelo M (2018) The neural monitoring of visceral inputs, rather than attention, accounts for first-person perspective in conscious vision. Cortex 102:139–149.

Vaidya AR, Sefranek M, Fellows LK (2017) Ventromedial Frontal Lobe Damage Alters how Specific Attributes are Weighed in Subjective Valuation. Cereb Cortex 28:3857–3867.

van Strien T, Frijters JER, Bergers GPA, Defares PB (1986) The Dutch Eating Behavior Questionnaire (DEBQ) for assessment of restrained, emotional, and external eating behavior. Int J Eat Disord 5:295–315.

Vinck M, van Wingerden M, Womelsdorf T, Fries P, Pennartz CMA (2010) The pairwise phase consistency: A bias-free measure of rhythmic neuronal synchronization. Neuroimage 51:112–122.

Walton ME, Chau BKH, Kennerley SW (2015) Prioritising the relevant information for learning and decision making within orbital and ventromedial prefrontal cortex. Curr Opin Behav Sci 1:78–85.

Wilson RC, Takahashi YK, Schoenbaum G, Niv Y (2014) Orbitofrontal cortex as a cognitive map of task space. Neuron 81:267–279.

Zhang Z, Fanning J, Ehrlich DB, Chen W, Lee D, Levy I (2017) Distributed neural representation of saliency controlled value and category during anticipation of rewards and punishments. Nat Commun 8.

Supplementary materials

Supplementary Table 1. Spatio-temporal clusters of heartbeat evoked responses predicted by reaction times.

Parameter	Cluster	Latency after T-wave (sec)	Sum(t)	Monte Carlo p
$ß_{RT}$	1st positive	0.05 – 0.204	4335	p = 0.004
$ß_{RT}$	1st negative	0.075 – 0.224	-4654	p = 0.004
$ß_{RT}$	2nd negative	0.05 – 0.163	-2484	p = 0.02

Supplementary Table 2. ECG activity does not differ between decision types.

ECG derivation	Same category ECG activity mean ± SEM (μV)	Different categories ECG activity mean ± SEM (μV)	two-tailed paired t-test	Bayes Factor	Clustering on the entire time-window (0.05-0.3 s)
Vertical 1	67.8 ± 4.5	67.9 ± 4.5	t(25)=-0.31, p = 0.76	0.26	No candidate clusters
Vertical 2	37.5 ± 4.6	37.6 ± 4.7	t(25)=-0.42, p = 0.68	0.28	No candidate clusters
Vertical 3	61.9 ± 4.3	62.0 ± 4.4	t(25)=-0.10, p = 0.92	0.25	No candidate clusters
Vertical 4	40.8 ± 4.2	40.9 ± 4.25	t(25)=-0.31, p = 0.76	0.26	No candidate clusters
Horizontal 1	30.4 ± 2.0	30.4 ± 2.1	t(25)=0.03, p = 0.97	0.24	No candidate clusters
Horizontal 2	-33.5 ± 1.4	-33.4 ± 1.4	t(25)=-0.33, p = 0.75	0.26	1 candidate cluster (Monte Carlo p=0.21, sum(t)=-19.82, latency [0.108-0.115] s)
Horizontal 3	-21.1 ± 1.3	-21.1 ± 1.3	t(25)=-0.78, p = 0.45	0.38	1 candidate cluster (Monte Carlo p=0.34, sum(t)= -4.90, latency [0.186-0.187] s)
Horizontal 4	-59.0 ± 0.9	-59.5 ± 0.9	t(25)=0.80, p = 0.43	0.39	No candidate clusters

Supplementary Table 3. Average arousal measures during option presentations do not covary with decision types.

Measure	Time - window	Same category Mean ± SEM	Different categories Mean ± SEM	two-tailed paired t-test	Bayes Factor
Alpha power (fT^2/Hz) 15 channel maximal alpha suppression	Stimulus locked 0 – 1.4 s	-69 ± 11.3	-67.7 ± 11.4	t(25)=-1.16, p=0.26	0.61
Alpha power (fT^2/Hz) Channels positive HER cluster	Stimulus locked 0 – 1.4 s	-26.6 ± 5.1	-26.7 ± 4.8	t(25)=0.13, p=0.90	0.25
Alpha power (fT^2/Hz) Channels negative HER cluster	Stimulus locked 0 – 1.4 s	-45.0 ± 8.9	-43.9 ± 8.5	t(25)=-0.96, p=0.35	0.47
Alpha power (fT^2/Hz) 15 channel maximal alpha suppression	Response-locked -1 – 0 sec	-82.7 ± 13.4	-82.4 ± 13.6	t(25)=-0.16, p=0.88	0.25
Alpha power (fT^2/Hz) Channels positive HER cluster	Response-locked -1 – 0 sec	-34.8 ± 6.4	-35.7 ± 6.3	t(25)=1.01, p=0.32	0.50
Alpha power (fT^2/Hz) Channels negative HER cluster	Response-locked -1 – 0 sec	-51.6 ± 9.9	-51.0 ± 9.6	t(25)=-0.52, p=0.61	0.30
RMS fT	Stimulus-locked 0 – .500 s	69.2 ± 2.7	69.3 ± 2.7	t(25)=-0.11, p=0.91	0.25
VEP (fT) Channels positive HER cluster	Stimulus-locked 0 – .500 s	-14.1 ± 5.3	-11.9 ± 5.1	t(25)=-0.90, p=0.37	0.44
VEP (fT) Channels negative HER cluster	Stimulus-locked 0 – .500 s	5.2 ± 4.1	3.9 ± 4.6	t(25)=0.80, p=0.43	0.39
Pupil diameter (a.u.)	Stimulus-locked 0 – 1.4 s	0.37 +-0.08	0.34 +-0.07	t(24)=0.94, p=0.35	0.46
Pupil diameter (a.u.)	Response-locked -1 – 0 sec	0.66 +-0.08	0.67 +-0.08	t(24)=-0.39, p=0.7	0.27

Chapter VI: General discussion

1. Summary of the experimental results

The goal of my PhD was to investigate the functional role of the neural monitoring of heartbeats in subjective preference-based decisions.

The results from the first experiment reveal that the neural responses to heartbeats may support the self-reflective process specific to preference-based decisions (**Figure 4.1**, first row). We show that HERs in bilateral vmPFC and somatosensory cortex distinguish subjective preference-based decisions from objective perceptual discrimination, during decision preparation. In addition, HERs are functionally relevant for subjective evaluation: single trial HER fluctuations in vmPFC – before option presentation – predict the strength of subjective value encoding during preference-based decisions, with larger HERs predicting enhanced encoding of subjective value. The neural interaction occurring in vmPFC between HERs and value encoding also accounts for between- and within-subject variability in preference-based choice consistency. We also show that HERs specifically contribute to preference-based decisions, as HERs do not affect the encoding of perceptual evidence during perceptual discrimination trials.

In the second experiment, I investigated the mechanisms by which neural responses to heartbeats might enable subjective evaluation. We hypothesised that HERs support the mapping of different feature spaces onto a common dimension, enabling the integration of distinct attributes into a common value space. Multiple features determine the value of an item: for example, the value of an ice-cream depends on its taste, texture, caloric content etc. These features are very similar for items belonging to the same category, for instance an ice-cream and a cookie, thus requiring a smaller coordination between feature spaces. On the contrary, decisions between goods of different natures, for instance between having an ice-cream outside or staying home reading a book, require a much larger coordination, as the category-dependent attributes belong to very distinct feature spaces. HERs may provide a format-independent signal through which these spaces can be aligned. We experimentally addressed this hypothesis by testing whether HERs are differentially recruited depending on whether subjects were choosing between items belonging to the same category or different ones. We found that HERs in OFC and in the Insula depend on decision type and that HER fluctuations in OFC predict trial-by-

trial choice precision in decisions involving goods of different categories. Our whole brain analysis further reveals that HER fluctuations in the Insula predict choice precision between items belonging to the same category, a result we did not anticipate.

In both experiments, all the above-mentioned effects cannot be explained by differences in cardiac inputs, in overall brain states or in arousal between the conditions.

In what follows, I will compare the results obtained in the two experiments, highlighting how some of the discrepancies may be partly accounted for by the differences in the experimental tasks and the specific ways in which data have been analysed. I will then discuss what these results tell us about what HERs may reflect, their role in self-referential processing and in preference-based decision-making. Finally, I shall conclude by outlining some open questions.

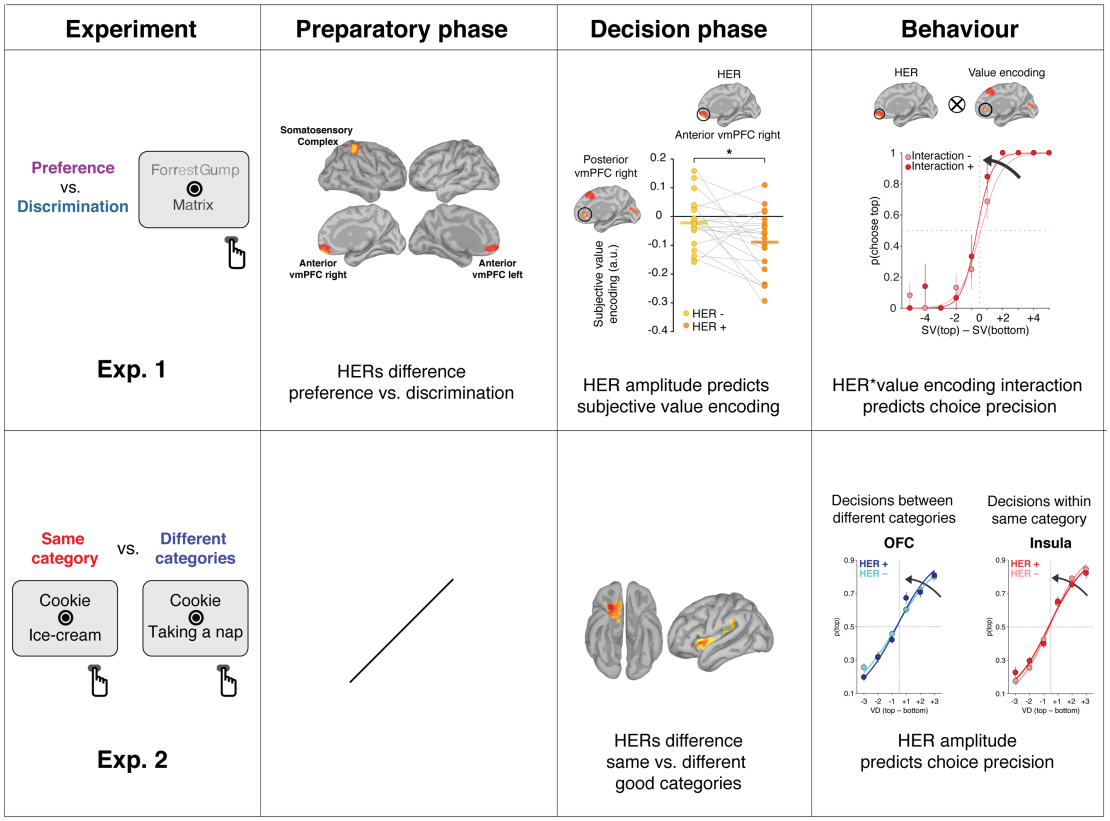


Figure 4.1. Summary of the experimental results. The rows in the table refer to the results from the first and second study, respectively. Each column corresponds to a different phase of the trial: preparatory phase (hence before option presentation), decision phase and behavioural response.

2. Comparing experimental results

A. Cortical localisation

Neural responses to heartbeats are located in different regions in the two experiments: bilateral (medial) vmPFC (Broadman areas 11m and 14) and right somatosensory cortex in the first experiment; right (ventral) OFC (BA 47) and left Insula in the second one.

Although different between the studies, these localisations are consistent with the existent literature. Previous studies using invasive and non-invasive electrophysiological recordings also found neural responses to heartbeats in vmPFC (Park et al., 2014; Babo-Rebelo et al., 2016a, 2016b), somatosensory cortex (Kern et al., 2013; Azzalini et al., 2019) and Insula (Babo-Rebelo et al., 2016b; Park et al., 2016, 2017). The presence of HERs in the OFC are only anecdotally reported (Canales-Johnson et al., 2015), but they are in line with tracing studies in monkeys showing visceral inputs to OFC (Carmichael and Price, 1995). OFC connectivity profiles between humans and monkeys are similar (Neubert et al., 2015), suggesting that visceral inputs may indeed reach human OFC as well.

Although cortical localisations retrieved with MEG are less precise than those accessible with fMRI or implanted electrodes, MEG spatial resolution is sufficient to reveal cortical locations as distinct as (medial) vmPFC and (ventral) OFC. The reasons why the two experiments may reveal two distinct sub-regions of prefrontal cortex involved in preference-based decisions is discussed in the following sections.

1. Integration of HERs to local computations

Anatomical and functional evidence suggest that neural responses to heartbeats are not predominantly found in a single region, rather they appear to be distributed across various cortical areas (Azzalini et al., 2019; Park and Blanke, 2019b). One possibility is that HER differences emerge in cortical regions involved in the experimental task, since this signal is integrated into local ongoing computations. Although this observation is based on the few studies that performed cortical reconstructions of the HER effects, some results are in line with this hypothesis. For instance, the vmPFC has been widely associated with self-reflective processes elicited by different experimental paradigms (D'Argembeau, 2013; Vaccaro and

Fleming, 2018) and HER differences have been reported in this region using tasks explicitly requiring self-reflection (Babo-Rebelo et al., 2016a, 2016b). HER differences were also observed in anterior precuneus, a region putatively involved in egocentric spatial transformations (Bicanski and Burgess, 2018), while participants were imagining themselves or others performing actions during cued scenarios (Babo-Rebelo et al., 2019), a cognitive operation that is likely to involve similar spatial transformations.

Similarly, in the first study, HERs may be integrated into ongoing self-reflective processes required to estimate the value of the available item, which are thought to occur in medial vmPFC (Roy et al., 2012; D'Argembeau, 2013). On the contrary, in the second study, HERs may support the coordination of the multiple attributes encoded in ventral OFC (McNamee et al., 2013; Suzuki et al., 2017; Zhang et al., 2017) necessary to compare goods of different natures.

Note that, although this hypothesis has the advantage of accounting for why HERs are observed in regions that are not necessarily associated with interoceptive processes, it cannot explain all results. In fact, we also observed HERs in somatosensory cortex and Insula, regions that are less often associated with preference-based decisions or self-reflective processes. As I shall discuss below (cf. *section 4.B*), this apparent contradiction can be explained by the fact that even abstract self-reflective processes may involve a certain degree of representation of bodily components.

2. The importance of experimental contrasts in HER investigation

Given the cyclical nature of the heartbeats, HERs cannot be identified by contrasting neural signals following the heartbeat against a baseline (e.g., a portion of neural data before the heartbeat has occurred), as this latter will also contain (a portion of) HERs. In addition, HERs, at least when measured non-invasively, are not characterised by any stereotyped waveforms, such as those of visual-evoked responses for instance, which feature clear peaks at precise latencies. It follows that one cannot select *a priori* portions of neural signals following the heartbeats and contrast these signals across conditions. To identify HERs and investigate their contribution to distinct cognitive processes, it is thus necessary to contrast HERs during two (or more) experimental conditions. Importantly, the kinds of experimental condition that are contrasted potentially affect where and when HER differences are observed, as different contrasts aim to isolate distinct cognitive processes to which HERs possibly contribute. This is

a possible reason that explains why in my two studies we observe HER differences at different cortical locations. Note that this argument is compatible with the idea the localisation of HERs partly depends on its integration into local computations.

The contrasts used in my two studies differ under multiple respects. In the first study, I contrasted subjective preference-based decisions vs. objective perceptual discriminations. In the second study, the contrast of interest involved in both cases subjective preference-based decisions that only differed in the type of goods compared. A second difference between the experiments concerns the phases of the trials in which HERs were analysed: during task preparation, *before* options were presented, in the first experiment, and *during* the decision process in the second one. The use of different contrasts across experiments is justified by the cognitive processes that we wanted to pin down. In the first study, we aim to isolate the self-reflective cognitive process needed to subjectively evaluate internal evidence by pitting it against the preparation to evaluate objective, externally available evidence. In the second study, we investigated how HERs can contribute to the comparison process of goods of different categories. As I shall detail below (cf. *section 5*), the seemingly different contributions of HERs in preference-based decisions may be accommodated by their general function in providing a biologically stable self-referential signal supporting self-reflection.

To facilitate the comparison between the two studies, in the second experiment we had introduced a cueing period, similar to the one used in first study. However, it seems that participants did not take advantage of the information delivered by the cue, and probably did not deliberately reflect on whether the upcoming choice was between items of the same or different categories. Indeed, when we asked participants whether the cue corresponded to the preference judgement required (same/different categories) in a few catch trials interspersed in the experiment, their performance was rather low (~60% on average), indicating that the meaning of the symbolic cue was only superficially processed. In any case, the hypothesised role of HERs in coordinating distinct feature spaces is more likely to occur when the information to coordinate is available, i.e. when decision alternatives are revealed. In keeping with this hypothesis, exploratory analyses during the cue period did not reveal any difference in HERs between decision types.

B. Latencies of HER effect

The latencies at which HER differences are observed also slightly differ between the two experiments. In the first one, the effect occurred 201-260 ms after the T-wave, while in the second study, the two clusters of neural activity spanned a larger time-window: from 128 to 286 ms after the T-wave. As noted in *Chapter III*, HER differences have been reported in a fairly wide range of latencies from 200 to 650 ms after the R-peak (Kern et al., 2013) , i.e. ~50-400 ms post T-wave, compatible with our results.

Differences in latencies between the two experiments may originate from the specific cortical regions in which HERs were observed. For instance, the Insula may receive cardiac information before other frontal regions, e.g. vmPFC and OFC, as suggested by tracing studies in monkeys (Carmichael and Price, 1995) and our own results. An alternative hypothesis is that cardiac information from different sources (baroreceptor in the carotid or heart wall, somatosensory signals) travels at different velocities depending on the neural pathways involved (Azzalini et al., 2019; Park and Blanke, 2019b) thus reaching cortical regions at different latencies.

C. The relationship between HERs and choice consistency

The relationship between HER fluctuations and choice consistency is also different between the experiments. In the first study, the relation with behaviour appears to be mediated by the interaction in vmPFC between HERs and the neural encoding of subjective value. In the second experiment, HER amplitude fluctuations in OFC and Insula are directly related to choice consistency. This apparent discrepancy may rest on the different times at which HERs have been analysed: before option presentation in the first experiment and during decision in the second, corresponding to putatively different influences of HERs in the decision process. HERs fluctuations in vmPFC before option presentation are likely to be related to differences in the self-reflective process underlying subjective evaluation. Therefore, it is not surprising that the relationship with choice consistency is indirect and mediated by the influence of HERs onto the product of self-reflection, i.e. subjective value encoding. On the contrary, in the second experiment, HERs may reflect the extent to which the mapping of category-specific attributes onto a common space successfully occurs, thus revealing a direct functional relationship with behaviour.

3. What do the current results tell us about HERs?

Although changes in HERs have been widely reported (for review cf. Azzalini et al., 2019; Park and Blanke, 2019a), what they reflect is still debated. In the following sections, I will present different alternatives and conclude how the results from my PhD provide evidence that how the brain responds to cardiac inputs is linked to the ongoing cognitive processes.

A. HERs reflect artefactual cardiac activity

One possibility is that HERs simply reflect artefactual activity. In M-EEG recordings, HERs would thus be the residuals of the cardiac field artefact (Dirlich et al., 1997), i.e. the electrical activity generated in the heart and directly sensed by M-EEG sensors on the scalp. Conversely, in invasive electrophysiological recordings, HERs may reflect the pulsatile activity in blood vessel in the proximity of recording sites (Kern et al., 2013). Little experimental evidence supports this hypothesis. In fact, most recent experiments performed with M-EEG perform ICA cardiac correction and control that HERs do not co-vary with heart electrical activity. These controls indicate that HERs are unlikely to merely reflect cardiac artefacts. Invasive recordings provide further support against this hypothesis by revealing that HERs are genuine neural responses distinct and distinguishable from pulsatility and other cardiac-cycle artefacts (Kern et al., 2013). Additional evidence against this hypothesis is provided by studies finding consistent results using different brain-imaging techniques for the same experimental paradigms (Babo-Rebelo et al., 2016a, 2016b; Park et al., 2016, 2017). There is little reason to think that different types of artefacts affect diverse recording techniques in the same way (despite some authors propose that this may be the case; cf. Dirlich et al., 1998) and in the same cortical regions.

B. HERs differences reflect changes in cardiac inputs

Alternatively, HERs may be true neural responses whose fluctuations may reflect changes in cardiac inputs rather than fluctuations in neural responses to an invariant cardiac stimulus. This concern is legitimate and receives some experimental support from between-

subjects studies (Schandry and Montoya, 1996; Gray et al., 2007). Our own work (*Appendix B*) on within-subject fluctuations of the volume of blood ejected at every heartbeat (i.e., stroke volume) and HERs does not favour this interpretation. Indeed, we show that a relationship between stroke volume and HERs exists, only when these latter are computed on raw MEG signal. However, after MEG data have been corrected for cardiac artefacts with ICA, HERs do not covary with stroke volume any longer. Besides, stroke volume fluctuations could be detected in heart rate with a higher sensitivity than in MEG data. To exclude that HER fluctuations reflect changes in cardiac input, it is thus necessary to explicitly control that HER differences are not accompanied by concurrent changes in cardiac activity (e.g. electrical activity, inter-beat interval). In all the studies reported in this PhD (*Chapters IV, V, Appendix B*), we explicitly controlled for this possibility providing quantitative measures (through Bayesian statistics) in favour of the absence of differences in cardiac parameters.

C. HERs as indexes of cortical states

A third possibility is that differences in HERs are a mere neural epiphenomenon – they would trivially reflect underlying differences in cortical states, such as difference in cortical excitability, associated with different experimental conditions. A consequence of this hypothesis is that cortical states should also drive different responses to other types of stimuli. Let us imagine that HERs in visual regions are enhanced because of a more excitable cortical state, visual responses to external stimuli are likely to be enhanced in a similar way. It must be acknowledged that it is not straightforward to control for this possibility: HERs occur in cortical regions for which response selectivity is not necessarily established. For example, it is not easy to find a stimulus that is known to be specifically processed in the Insula. However, controlling that HER differences are not accompanied by differences in neural markers of cortical excitability, such as alpha power, provides some evidence against this hypothesis. Indeed, previous results showed that differences in alpha power, induced by arousal manipulation, can affect HER amplitude (Luft and Bhattacharya, 2015). General arousal states induced by cognitive components can be also retrieved from more indirect measures, such as pupil diameter (Joshi and Gold, 2020). In both preference-based experimental tasks, we controlled that HER differences could not be explained by changes in cortical and general arousal states.

In addition, should HER differences be generated by dissimilar cortical states, one would expect this difference to be sustained over time, thus being reflected in baseline differences between the experimental conditions. Controlling that ongoing brain activity, unrelated to the occurrence of heartbeats, does not produce differences similar to those observed for HERs thus provides convincing evidence against this hypothesis.

Altogether, the control analyses performed in my PhD experiments make the possibility that the observed HER effects simply reflect difference in brain states unlikely.

D. HERs are genuine neural responses to cardiac events

My experimental results thus support and provide quantitative measures (Bayes Factors) in favour of the hypothesis that heartbeat-evoked responses are genuine neural responses to cardiac inputs, whose fluctuations reflect the differential processing of cardiac signals depending on the type of ongoing cognitive processes.

First, the cortical locations of HER differences retrieved from MEG data in my experiments are in accordance with previous results that used intracranial recordings (Canales-Johnson et al., 2015; Babo-Rebelo et al., 2016b; Park et al., 2017). This suggests that HER differences are unlikely to result from cardiac-related artefacts as different recording techniques are affected by different types of artefacts. In addition, we exclude that our differences arise from changes in cardiac rhythms and cardiac parameters such as the blood volume ejected at each heartbeat. Finally, we also rule out that these effects can be driven by changes in arousal states, as cortical (alpha power) and peripheral (pupil diameter) indices of arousal do not covary with HER fluctuations.

Below, I will discuss how we experimentally isolate the process we believe it is reflected in HER changes and other possible alternatives.

4. Isolating self-reflection experimentally

An important contribution of my PhD work, specifically addressed in the first experiment, concerns the study of **how the preparation to preference-based decision influences its later stages**. Most of the decision-making literature has focussed on the neural

processing underlying decision formation. However, as I pointed out in the *Chapter I* (*section 2.A*), a fundamental step is the self-reflective process by which an agent represents her current state.

To isolate this cognitive component, (a) we temporally separated the preparation process from the following ones, involving evaluation and decision; and (b) we contrasted the preparation to a subjective preference-based decision against the preparation to an objective visual discrimination task. It has to be noted that some types of perceptual decisions also involve a self-reflective component similar to the one at play in preference-based decisions. For instance, in near-threshold detection tasks, and more specifically for absence judgments (Fleming, 2020; Mazor et al., 2020), tracking one's current state is particularly useful to infer the likelihood of a stimulus being present. To isolate the self-reflecting aspect involved in the preference-based decisions we thus used a supra-threshold visual discrimination task as control condition.

In both experiments, we used naturalistic stimuli, i.e. stimuli whose value dimensions could not be manipulated *ad hoc* by the experimenter. In other words, the value of each item was entirely subjective. We explicitly avoid using probabilistic monetary rewards, as the evaluation of these stimuli does not rely on a self-reflective processing – the value of each option could be objectively estimated through mathematical operations. This does not exclude other possible paradigms that allow the experimenter to directly manipulate option value while eliciting strong self-reflection (cf. *Section 6*).

5. HERs and self-referential processing

Therefore, what do changes in HERs reflect?

A. HERs as the biological root of self-representation

Information about cardiac activity appears to reach multiple subcortical and cortical structures (Critchley and Harrison, 2013; Azzalini et al., 2019; Park and Blanke, 2019b). This feature is not surprising as the monitoring of cardiac information ensures individual's survival by enabling automatic and voluntary adjustments that keep the organism within an optimal homeostatic range. A portion of HER fluctuations may thus be related to this homeostatic regulation, as suggested by HER changes that depend on differences in cardiac parameters (Gray et al., 2007; Buot et al, *in prep – Appendix B*). However, our data show that HERs also depend on the ongoing cognitive process even in the absence of changes in cardiac inputs. **The simple neural representation of cardiac information, independent of its changes, may signal the biological integrity of the individual – hence its existence as living subject – providing the simplest form of self-representation** (Park and Tallon-Baudry, 2014; Tallon-Baudry et al., 2018; Azzalini et al., 2020). In my PhD, I have provided some experimental evidence indicating that **more abstract self-referential processes, such as the self-reflection enabling subjective evaluation, may be based on this basic biological signal**. In this view, **HER fluctuations reflect the extent of the self-referential process needed to accomplish the cognitive operation probed by the experimental task**. An important aspect of this mechanism is that, although it supports conscious cognitive representations of oneself, it does not need to be conscious itself. We propose that its recruitment is automatic and occurs at the sub-personal level. In other words, it is not more accessible to the subject than her neural responses in V1 that enable her to see the outside world.

B. HERs as index of enhanced body-related representation

An alternative interpretation of these experimental results is that the changes in HERs do not indicate any special status of the cardiac information for self-reference, rather, they reflect the fact that all self-referential processes necessarily entail an increased attention to one's body. The differential recruitment of HERs is thus a consequence of the self-referential processes rather than a condition thereof. This hypothesis seems at odds with experimental results. For example, it is not straightforward how thinking about one's own body during a visual detection experiment at threshold could increase the subjective consciousness of a faint

visual stimulus (Park et al., 2014). All other things being equal, one would expect that diverting the attention from the external stimulation towards one's bodily sensations would be detrimental to performance. Although this may be true, we cannot exclude, at least for the case of visual detection, that increased HERs observed *before* stimulus presentation index a more precise representation (independent of conscious attention) of one's own cardiac information, i.e. a bodily aspect, that can be leveraged as disambiguating cue *during* decision to infer stimulus presence. Imagine that a visual stimulus that is perfectly processed by the brain, and hence should enter into consciousness, induces a more important cardiac deceleration. In trials with larger HERs, one would have a more precise representation of bodily information that can be used as additional evidence to infer stimulus presence. In this view, HER differences will still be due to changes in the neural process of invariant cardiac inputs (the cardiac inputs eliciting HERs are stable). However, future cardiac changes induced by external stimulation would provide disambiguating cues that can be optimally integrated in the decision process only if their representation is precise enough (i.e., when HERs are large). Park and colleagues (Park et al., 2014) report that cardiac activity does not differ between hits and misses during the time-window of HER analysis, but it does so after response delivery. It will be interesting to know whether cardiac difference is present during decision time and whether this possibility can account for their results. This hypothesis resonates with the idea that external and internal information must be integrated to give raise to conscious awareness (Salomon et al., 2016).

The same hypothesis applied to my studies entails that the precision of subjective evaluation (and of the ensuing choice) would derive from the precise readout of cardiac changes induced by the stimulus during deliberation. Changes in cardiac activity are slower than the alternating evaluation of the alternatives, which through attention, appears to affect deliberation (Krajbich et al., 2010). It follows that the present experiments are not suited to test whether the strength of the neural encoding of subjective value depends on the precise representation of cardiac changes following stimulus presentation. A 2AFC task where options are sequentially presented can prove a more suitable tool to address this question in the context of preference-based decisions.

If self-referential processes require an increased representation of one's body, one would expect the neural responses to other bodily signals such as respiration, gastric activity and even proprioception to be enhanced akin to HERs. To my knowledge, this possibility has

just started to be addressed (Park and Blanke, 2019a) and may provide promising leads for the characterisation of the functional role of bodily inputs in self-referential processing.

C. HERs and self-relevant bodily aspects

A related, yet distinct, line of reasoning may propose that self-referential processes entailing bodily aspects are more affected by fluctuations in HERs than those involving more conceptual/abstract dimensions. According to this hypothesis, one may expect that, in my second experiment, HER fluctuations for decisions involving the comparisons between food items (a primary bodily need) would relate to stronger changes in choice precision as compared to decisions that have no direct physiological consequences, such as decisions between two movies. An effect of this type may be expected in the Insula, a region thought to integrate current physiological states into ongoing cognitive processes (Craig, 2009), and in which we found a functional relation between HER fluctuations and behaviour. Future analyses will specifically address this possibility to characterise the contribution of HERs in distinct brain regions to preference-based decisions.

6. The role of HERs in preference-based decisions

In my PhD, I have investigated two core features of value-based preference decisions: their subjective nature and the remarkable ability of choosing between goods of different natures. **The main contribution of my work is showing how both aspects may be grounded in the same underlying phenomenon: the implementation of self-referential processing through neural responses to heartbeats**.

The subjective nature of preference-based decisions has been long accepted (Bernoulli, 1954), but it has not been considered as an object of investigation *per se*, leaving the biological characterisation of subjectivity unaddressed. A possible reason for this lack of interest is the belief that there is nothing to be explained, as the subjective aspect inherent to preference-based decisions is the natural product of the neural activity going on in the brain of each of us. However, neurological (e.g., dementia (Strohminger and Nichols, 2015)) and psychiatric conditions (e.g., schizophrenia (Sass and Parnas, 2003), depersonalisation (Sierra and David, 2011)) reveal that a stable subjective perspective, a stable identity, is not something that can be taken for granted, rather it needs some biological account. Moreover, a stable representation of oneself, accessible through self-reflection, is a necessary condition for coherent subjective evaluation: ascribing a subjective meaning to a stimulus can only arise in relation to a given individual that has the capacity to accurately represent her current needs and goals (Paulus, 2007; Moeller and Goldstein, 2014; Juechems and Summerfield, 2019).

In the decision-making literature, homeostatic processes have long been considered a pivotal component only in relation to choices concerning primary needs, such as food items (Rangel, 2013). Indeed, accurately tracking one's state is of essence to establish which is the most suitable course of action to maintain physiological variables within a given range (Hull, 1943; Berridge, 2004). The automatic monitoring of basic physiological parameters can thus be considered as the simplest form of self-reflection. Subjective evaluation of secondary rewards also requires a self-reflective process. For example, knowing that I highly esteem movies that address contemporary societal issues will drive me to choose a Ken Loach's movie instead of a Hollywood comedy. The neural machinery underlying this more abstract type of self-referential process may have evolved from the monitoring of physiological variables underlying homeostatic mechanisms. In other words, the monitoring of basic physiological variables may act as the substrate for more abstract representations – concerning the same living organism – to be built. Among the many variables that signal the existence of a living entity,

the neural monitoring of cardiac information, in virtue of its continuous and automatic processing, are particularly suitable.

The results of my first experiment provide evidence in support of this hypothesis. We found that HERs differ when subjects need to prepare to reflect upon themselves to subjectively evaluate options as compared to evaluate visual properties of external stimuli. In addition, HER fluctuations in vmPFC (during decisions preparation) not only predict the strength of subjective value encoding (during decision) but also predict the intra- and inter-individual choice variability in subjective preference-based trials. We interpret the changes in the neural representation of subjective value and decision coherence as the result of a less precise subjective evaluation arising from the reduced precision in self-representation, indexed by smaller HER amplitude.

Another novel contribution of my PhD is the idea that the reference to a living organism that grants the existence of subjective evaluation also provides the key to understand how incommensurable goods can be compared. The comparison of goods of different natures, characterised by disparate sensory and conceptual attributes, requires the mapping of these features onto a single space. In turn, this coordination enables an agent to put the ensemble of these features in relation to herself, enabling evaluation and comparison. I hypothesise that this coordination is supported by neural responses to heartbeats. Although preliminary, the results from the second experiment provide evidence in this direction, as we observed a differential recruitment of HERs depending on the amount of coordination needed for good comparison: decisions between goods of different categories require more coordination between feature spaces than decisions between items characterised by attributes lying in a similar attribute space.

Altogether, the experimental results specify a biological mechanism underlying the self-referential processing involved in subjective preference-based decisions. The identification of this precise neural mechanism may prove particularly useful to quantify how changes in this self-referential processing are functionally linked to erratic decisions in psychiatric conditions such as depression and addiction. These pathologies are often characterised by abnormal interoceptive processes (Khalsa et al., 2018) indicating an interesting convergence between value-based decisions and homeostatic monitoring (Paulus, 2007). I regard this as one of the most interesting directions that this line of research can pursue.

7. Open questions and further directions

The results from my PhD indicate that HER fluctuations are associated with changes in subjective value encoding and choice precision. This in turns might suggest that fluctuations in **HERs may relate to changes in the accumulation of subjective evidence**. **Further work may address this question using dynamic models of the decision process**. To leverage the power of this approach, subjective values that are directly controllable by the experimenter combined with a task eliciting strong self-reflection will be highly beneficial. One possibility is to use forced-choice task paradigms in which individuals have to maintain their wealth (characterised by a combination of different assets) within a given range and contrast these decisions with ones on stimulus perceptual properties. This procedure will allow the experimenter to directly manipulate and control value-related parameters entering the decision process. In addition, to correctly perform the task, participants must monitor their current wealth state – i.e. self-reflect. The paradigm will thus optimally combine a precise estimation of subjective values with self-reflection. An example of this paradigm can be found in a recent article by Juechems and colleagues (Juechems et al., 2019).

The relationship between HERs and subjective evidence generated by self-reflective processes opens interesting questions about the role of HERs in metacognition. We started to tackle this question in my second experiment, but only tangentially. Several studies have shown how subjective evaluation and confidence estimation are intrinsically related (De Martino et al., 2013; Lebreton et al., 2015; Polanía et al., 2019; Lopez-Persem et al., 2020). Therefore, we reasoned that neural responses to heartbeats may relate to both first-order preference-based decisions, and second-order judgements of confidence in preference-based choices. In my second experiment, participants had to rate their confidence in each trial. It should be noted that given the constraints imposed on trial duration (keeping the trial duration below 12 seconds), the time-window to analyse HERs during confidence estimation might be too short. Nonetheless, future exploratory analyses may investigate whether HERs during the decision phase relates to subjective confidence.

The results from my second experiment also raise interesting questions about the **generality of neural responses to heartbeats in coordinating distinct information**, beyond preference-based decisions. Recent results already showed that HERs may contribute to the integration of visual and tactile information for bodily self-consciousness (Park et al., 2016, 2017). Future studies may extend these results by testing whether **HERs underlie other types**

of multi-sensory integration possibly revealing their role in the creation of a unitary subjective experience characterised by multiple distinct dimensions.

Another pivotal question concerns whether **HERs are causally relevant for self-reflective processes**, as current results remain correlational. Manipulating neural responses to heartbeats in humans is challenging: (1) the consequences of specifically silencing HERs are unknown and may be dangerous; (2) Transcranial magnetic stimulation (TMS) is poorly suited to reach the medial (vmPFC, PCC) and deep cortical structures (Insula), wherein HERs are often observed; and, (3) the effects of other techniques – such as tACS and tDCS – are still hotly debated (Vöröslakos et al., 2018). In this respect, animal studies are particularly promising, as they would enable specific causal manipulation of brain activity. However, HERs have not been yet investigated in animals and it is currently unknown whether they may fulfil similar cognitive functions. In addition, the dissimilarities in visceral pathways across different animal species (Azzalini et al., 2019) may hinder the extent to which these results can be directly translated to humans. Shall we accept that we will never able the functional role of HERs causally? Despite inter-species differences, I believe that some progress can be made by using animal models. A recent theoretical framework (Birch et al., 2020) addresses how various aspects of conscious experience, among which self-awareness, can be probed in animals. **Invasive techniques paired with clever experimental designs can thus be used to probe the causal relevance of HERs for self-related cognitive operations in animals**. For instance, it is likely that avoiding a devaluated option may require a more precise representation of the animal's physiological state as compared to choosing between an appetitive and a neural option that have not been devaluated. This is because cached values, developed through longstanding associations established via experience, cannot be used to infer an up-to-date value of the item that has been devaluated. If HERs are related to self-reflective aspects also in animals, perturbing HERs would have a larger impact on decisions involving devaluated options that non-devaluated ones.

When causally manipulating HERs is not possible, mechanistic models of how HERs contribute to self-referential processes are probably the most useful tool to advance our understanding. By developing such models, the experimenter may simulate the effects that should be observed when perturbing HERs. In turn, these simulations can be used to formulate precise predictions in experimental paradigms. Our work is a first step in this

direction, as it maps HERs fluctuations onto precise neural and behavioural components of decision.

8. Final word

My PhD work contributes to the idea that cognition is not only the product of the electrical activity of an isolated brain, rather it needs to be considered as arising from the complex interplay between the brain and the body. I demonstrated this by showing that more conceptual forms of self-representations may be grounded on a much more corporal version of ourselves: a living organism with a beating heart.

Appendix A: Visceral Signals Shape Brain Dynamics and Cognition

Trends in Cognitive Sciences

CellPress
REVIEWS

Feature Review

Visceral Signals Shape Brain Dynamics and Cognition

Damiano Azzalini,[1,2] Ignacio Rebollo,[1,2] and Catherine Tallon-Baudry[1,*]

Most research in cognitive neuroscience explores how external stimuli are processed by the brain. However, the brain also receives input from the internal body. We discuss here how the heart and gastrointestinal (GI) tract intrinsically generate their own electrical activity, thereby continuously sending information to the brain. These ongoing ascending signals actively shape brain dynamics at rest, complementing canonical resting-state networks (RSNs). Cardiac signals also influence the processing of external sensory information and the production of spontaneous, internal cognition. These findings are discussed in relation to interpretative frameworks regarding the functional role of visceral inputs. This active field of research offers a unique opportunity to draw new theories blurring the border between cognition, emotion, and consciousness, as well as between mind and body.

Highlights

Visceral signals automatically contribute to shaping spontaneous large-scale brain dynamics, even when not consciously experienced. To what extent ascending visceral signals can be consciously perceived remains a matter of debate.

Neural and behavioral responses to external stimuli can be influenced by temporal contingencies between heartbeats and stimuli.

Beyond temporal contingencies, the transient neural response evoked by each heartbeat plays a role in cognitive functions that are usually studied separately, such as emotion, self-related cognition, and also subjective visual perception, calling for new theoretical perspectives.

Neural responses to heartbeats may be involved in the generation of the unified viewpoint of consciousness.

A Paradigm Shift in Cognitive Neuroscience

'The brain is bombarded by information from the environment.' This sentence, in one form or another, acts as an introduction to many textbooks and articles about perception and cognition (e.g., [1]). Although the internal environment is sometimes mentioned, the vast majority of the experimental work pertains to the processing of external stimuli. This reflects the dominant paradigm in cognitive neuroscience, where an agent collects information from the external environment via the senses and then reacts to this environment by producing actions (Figure 1A). By studying the brain, cognitive neuroscience has been highly successful at providing mechanistic explanations of behavior and at revealing the existence of hidden rules and variables, such as for instance the reward prediction error [2]. In the following we review extensions of this paradigm that include the interplay between the brain and the internal bodily environment (Figure 1B) to account for the neural implementation of cognition. More specifically, we focus on the heart and the **gastrointestinal (GI) tract** (see Glossary), and refer the reader to [3] for a comprehensive review of other types of somatic influences, including notably humoral and immune factors.

There are several reasons why cardiac and GI inputs might influence brain dynamics and cognition. The first distinctive feature of these two organs is that they generate their own intrinsic oscillatory electrical activity, even when disconnected from the brain, from the first weeks of gestation until death. We will detail later the intrinsic properties of the heart and the GI tract – but consider an isolated heart in a vat, waiting to be transplanted: it is beating. The heart thus differs from respiration, for instance [4]. A respiratory system disconnected from the brain is inert, because breathing commands are generated in the central nervous system (CNS). By contrast, both the heart and the GI tract can be considered as intrinsic oscillators that continuously send information to the brain. Visceral signals can thus be considered as stimuli that influence spontaneous brain dynamics. Second, cardiac and GI signals are characterized by intrinsic time constants in the sub-second (heart) to 20 s range (stomach), compatible with the timescale of cognition. Importantly, these signals reach not only brainstem nuclei but also many subcortical and cortical regions involved in cognitive tasks. Last but not least, monitoring internal visceral variables is a

[1]Laboratoire de Neurosciences Cognitives et Computationnelles, Ecole Normale Supérieure, Université PSL (Paris Sciences et Lettres), and Institut National de la Santé et de la Recherche Médicale (INSERM), Paris, France
[2]Equal contributions

*Correspondence:
catherine.tallon-baudry@ens.fr
(C. Tallon-Baudry).

© 2019 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Check for updates

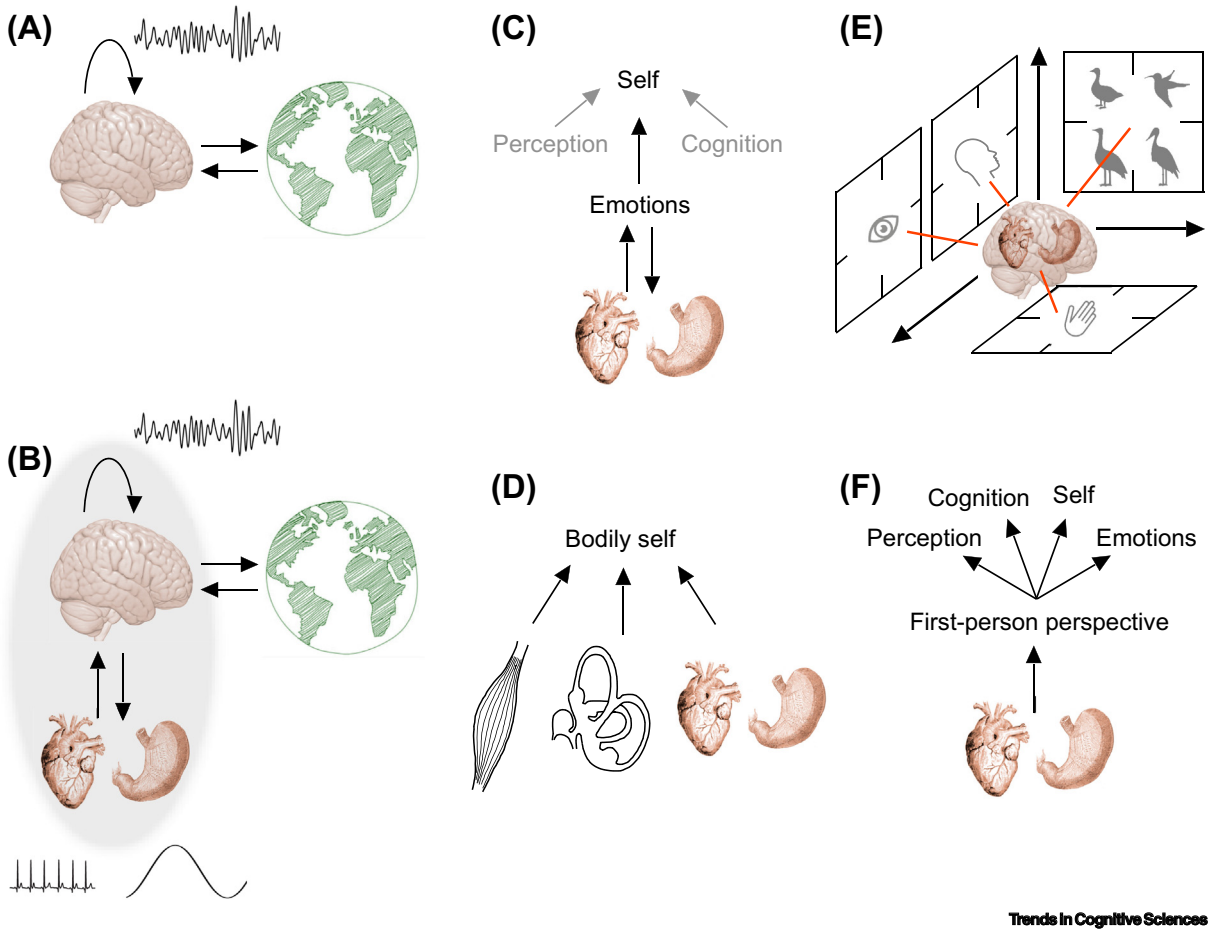


Figure 1. Brain–Viscera Interactions. (A) Classical paradigm of brain–environment interactions. (B) Extended paradigm that includes brain–viscera interactions, and more specifically the heart and gastrointestinal tract, which both intrinsically generate their own electrical activity and potentially contribute to brain dynamics by continuously and automatically stimulating the brain. (C) According to some theories [18,23–25], changes in visceral states could contribute to emotions and ultimately the self. (D) Another view [27] holds that the bodily self arises from the multisensory integration of proprioceptive, vestibular, and also visceral ascending signals. (E) A third proposal [28,29] is that neural monitoring of ascending visceral inputs is used to coordinate the different frames of reference in which sensory, motor, and cognitive information is organized, to (F) generate the first-person perspective inherent to conscious perception, cognition, self, and emotions. Bird icons generated by Freepik (www.flaticon.com).

core function of the CNS, and the drive behind many life-sustaining behaviors that are tightly related to high-level sensory and cognitive functions. For instance the basic behavior of searching for food is motivated by internal states such as hunger, and requires visual search, action planning, navigation, and learning and memory.

Given all these features, cardiac and GI signals stand out as a continuous source of information that may influence both brain dynamics and high-level cognition. However, experimental research has so far mostly focused on descending influences from brain to heart, analyzing in particular how pain, stress, physical effort, and emotion affect heart rate or **heart-rate variability** (HRV). In this article we will focus on ascending pathways from viscera to brain, their impact on resting-state brain dynamics, and their influence not only on emotions but also on perception and cognition.

The potential role of visceral signals in emotions, perception, and cognition is thought to originate from a common low-level function, homeostatic/allostatic regulation, that would have been harnessed by evolution to develop other, higher-level processes [5]. Homeostasis, a concept

identified by Claude Bernard in the 19th century, refers to the automatic adjustment of physiological variables within a given range. More relevant for cognition, allostasis [5–7] is a process that also aims at physiological stability, but that does so by activating physiological systems in anticipation to bodily needs, as for instance when preparing the body for physical effort. The best-known example of an allostatic regulation is the production of saliva by Pavlov's dog in response to a sound that the dog has learned to associate with food presentation. Other types of allostatic regulation result in adjustments of the cardiovascular system, for example in situations involving novelty, threat, and stress, or physical effort, with potential long-term consequences on health [8]. Allostatic regulations have been most often studied from the perspective of a descending control from brain to body, for instance a change in heart rate induced by a preparatory cue, and related to **arousal**. However, ascending inputs are required to inform the brain about current bodily states, as stated for instance in the recent reformulation of allostatic regulations in the predictive coding framework [9–13], where the brain constantly updates an internal model of the likely causes of inputs [14,15].

We begin this article by mapping the theoretical landscape on the putative functional roles of visceral inputs beyond allostatic regulations, and then present the basic physiology of cardiac and gastric signals and ascending anatomical pathways. The remainder of the review is devoted to experimental studies showing the impact of cardiac and gastric signals on spontaneous brain activity, and the different ways cardiac signals influence perception, emotion, cognition, and consciousness.

Mapping the Theoretical Landscape

What could be the functional roles of visceral inputs? Different proposals have been formulated, tapping onto domains ranging from emotions, to bodily self, to first-person perspective.

Emotions

Bodily signals, and in particular visceral signals, have long been associated with emotions. We tend to spontaneously associate different emotions with different combinations of bodily regions in a stereotypical manner, as recently shown in a set of elegant behavioral experiments [16,17]. However, the causal role of visceral signals in the generation of emotions is controversial. To briefly summarize a century-old debate, some [18,19] hold that subjective emotional feelings are consequences of bodily changes that are automatically triggered by emotional stimuli, whereas for others [20] it is the emotion that drives bodily changes. The various models of affective appraisal [21,22] leave open the possibility that changes in bodily states are included, among many other components, in affective responses. It has been suggested [23,24] that visceral signals, and more generally the sense of the physiological state of the body, are integrated with sensory and cognitive inputs in the insula to form a 'global emotional moment'. In another proposal [25], the self derives from 'primordial feelings' that are related to actual or mentally simulated changes in the body. The self here would be directly grounded in the subjective experience of emotions and their associated visceral states, and further enriched by perception and cognition (Figure 1C). Note that all those proposals link the emotional experience to 'changes' (actual, simulated or predicted) in bodily states and/or arousal.

Multisensory Integration and Bodily Self

Another line of proposals refers to the notion of the bodily self, viewed as a multisensory construct (Figure 1D) where tactile, **proprioceptive**, vestibular, and visceral signals are combined to define the sense of body location in space and the sense of body ownership [9,26]. The relative weight given to visceral inputs, compared with other sources of bodily information, varies from moderate, taking into account existing experimental evidence [26], to potentially important, in theoretical proposals linking active inference, body ownership and visceral signals [9,27]. Because the

Glossary

Arousal: physiological and/or self-reported fluctuations in alertness and responsiveness to external stimuli. Arousal depends on the cholinergic, serotonergic, noradrenergic, and dopaminergic neuromodulatory systems.
Atria: the two upper chambers of the heart. The right atrium fills with deoxygenated blood from the venous system (atrial diastole) and contracts to eject blood into the right ventricle (atrial systole). The left atrium fills with oxygenated blood from pulmonary veins and ejects it into the left ventricle.
Baroreceptors: specific types of mechanoreceptor, present in the heart, carotid, and aorta, that are sensitive to stretch and hence can detect blood pressure changes.
Body maps: spatially organized neural representations of body parts (limbs, hand, face, etc.) that are elicited by information from somatosensory, motor, or visual modalities.
Cranial nerve: the ensemble of ascending and descending fibers that contact the brainstem directly without traveling through the spinal cord.
Default network: a network of areas with temporally correlated fluctuations of activity that are more active at rest than during many, although not all, perceptual and cognitive tasks.
Diastole: the period in which all heart chambers relax and fill with blood. This period begins at the end of the T-wave and stops with the P-wave. It can be further divided into atrial and ventricular diastoles. Atrial diastole begins before ventricular diastole, at the R peak.
Electrocardiogram (ECG): noninvasive recording of the cardiac electrical activity obtained from cutaneous electrodes, typically placed on the right upper chest and left abdomen or leg.
Electrogastrogram (EGG): noninvasive recording of the gastric basal rhythm obtained from cutaneous electrodes placed on the left abdomen.
Gastrointestinal (GI) tract: the organ ensemble involved in food intake and digestion. It includes the mouth, esophagus, stomach, and intestine.
Glossopharyngeal: refers to the ninth cranial nerve. The glossopharyngeal nerve conveys information from the baroreceptors of the carotid sinus.
Heart-rate variability (HRV): fluctuations in heart rate, mostly driven by descending influences from the CNS

temporal coincidence between inputs is a key feature of multisensory integration, this hypothesis predicts that the relative timings of visceral and external signals should be relevant.

First-Person Perspective

Our own position is that the neural monitoring of cardiac and GI inputs plays a key role in the generation of first-person perspective by creating a subject-centered frame of reference [28,29]. First-person perspective refers to the unique bodily-centered viewpoint that we have on the environment [30,31] or on our own mental life. It pertains to the unity of consciousness and provides the intrinsic 'mineness' of conscious experience [32]. The existence of first-person perspective is often taken for granted (e.g., [33]), probably because, if there is a perspective at all, then it is mine because it happens in my brain [29]. Nevertheless, a great deal of information processing in the brain is unconscious and thus, by definition, is devoid of first-person perspective – a mechanism explaining how first-person perspective is generated thus becomes necessary.

How a unique viewpoint can emerge is not yet known. Information is expressed in the brain in sensory and cognitive maps, each of which has its own frame of reference, or coordinate system – gaze-centered, head-centered, etc. [34–36]. The expression of information in different frames of reference is also a general feature of cognitive maps [37–40]. Despite the multiplicity of coordinate systems in which information is encoded, the conscious experience of the external world and of inner mental life feels unified, as if experienced from a single point of view. The existence of a common input to different maps could facilitate the alignment and coordination of different frames of reference by providing a common reference point. Because visceral signals are constantly present and distributed across a large number of cortical and subcortical areas, they could provide such an input (Figure 1E) and hence help in binding those maps together into the integrated viewpoint of first-person perspective. In this view, visceral signals would play a role not only in emotions and bodily self but also more generally in conscious perception and cognition (Figure 1F).

Common Features and Differences between Theoretical Perspectives

The different perspectives discussed here share common features, and are related to one another. The first-person perspective account can encompass the proposals related to both emotions and bodily self because both are necessarily experienced from the first-person perspective. In addition, the viewpoint from which we experience the world originates from body location, an important component of the bodily self [30]. Finally, the sense of body ownership based on multisensory integration of proprioceptive and vestibular signals probably requires the coordination of different frames of reference, as hypothesized in the first-person perspective account.

However, there are also differences. In the first-person perspective view, and to a lesser extent in the multisensory integration account, visceral state does not matter *per se* (e.g., whether the heart is beating rapidly or slowly) provided that the viscera send ascending signals. By contrast, emotion-related theories postulate that visceral states directly contribute to emotional feelings. Furthermore, the three proposals target three different aspects of mental life: emotions, bodily self, and subjective experience. Finally, the different accounts use different levels of descriptions and propose different mechanisms: convergence in the insula, multisensory integration, or generation of a subject-centered frame of reference. Before reviewing experimental evidence on the functional role of visceral inputs, it is worth reviewing the known facts – and the many unknowns – about the generation of visceral signals, and their transduction and transmission up to their central representation.

The Heart and the GI Tract as Electrical Pacemakers

Both the heart and the GI tract are endowed with an intrinsic nervous system [41–43] and generate their own electrical rhythm. The heart contains pacemaker cells that initiate cardiac contractions, and that operate autonomously, enabling the heart to beat even when isolated from the

on the heart. The spectral analysis of HRV is often divided into low-frequency (0.04–0.15 Hz) and high-frequency HRV (0.15–0.4 Hz) domains.

Interoception: this term originally designated the mechanisms involved in the perception of visceral signals. It is presently often used in a much wider sense to describe the sense of the physiological state of the body.

Proprioception: the perception of the position of body parts, and of changes during movement, based on information from muscles and tendons.

Spinal: the ensemble of ascending and descending fibers traveling through the spinal cord.

Systole: often refers to ventricular systole, corresponding to ventricular contraction. It begins at the Q peak of the ECG and terminates at the end of the T wave. During ventricular systole, blood is ejected into the aorta and pulmonary trunk, and arterial blood pressure reaches its maximum. Atrial systole is the contraction of atria, which begins at the P-wave and ends at the R-peak of the ECG.

Vagal: refers to the vagus nerve, or tenth cranial nerve. The vagus senses and controls the heart, lungs, and GI tract. It also innervates the aortic arch.

Ventricles: the two lower chambers of the heart. During ventricular systole, the left ventricle ejects the oxygenated blood into the aorta, and the right ventricle ejects deoxygenated blood into pulmonary arteries.

Viscerotopic organization: spatially organized neural representation of visceral organs, where each organ maps onto a specific neuronal patch.

rest of the organism. In normal conditions, the rate of cardiac contractions is regulated by descending influences from the CNS, and by less well described local influences of the intrinsic cardiac nervous system. As a result, the heart rate of healthy participants is highly irregular. This irregularity is tightly related to the integrity of the CNS, and any sign that the heart is becoming too regular is a clinically relevant neurological warning sign. Recent studies showed that HRV provides information about the state of the CNS, and can be efficiently used not only to predict epileptic seizure onset [44] or seizure development [45] but also to detect residual signs of consciousness in noncommunicating patients [46,47]. However, we will set aside descending influences from brain to heart because this review focuses on ascending pathways – from heart to brain. How is the brain informed about the occurrence and strength of a cardiac contraction? This question pertains to **interoception** – the perception of visceral signals by the brain – as initially proposed by Sherrington, who coined the term 'interoceptor' [48,49]. The term interoception has later been extended to the sense of the physiological state of the body [24], including for instance information such as temperature.

Cardiac contractions, and resulting pulses of blood in vessels, are transduced into neural signals by numerous mechanoreceptors in the heart itself, both in the **atria** and in the **ventricles**, as well as in blood vessels (aorta, carotid artery, pulmonary and coronary vessels) [50–52]. Most cardiac receptors fire transiently at a given moment in the cardiac cycle, but latencies can vary considerably depending on receptor type and location [51]. Mechanoreceptors often signal the strength of cardiac contraction and blood pressure, but some cardiac receptors instead indicate the occurrence of a contraction, irrespective of cardiac contraction parameters [51]. In addition to cardiac and blood vessel receptors, the occurrence of cardiac contractions is also signaled through other, less well known pathways. The pulse modulates the neural activity of tactile [53] and proprioceptive [54,55] receptors, showing that information about cardiac activity is already present at the most peripheral level in the somatosensory modality. Another intriguing possibility is direct vasculoneuronal coupling in the CNS [56], an effect recently observed in rodent slices [57], where a change in pressure in the vessels modulates neural firing.

The GI tract also generates its own rhythm. The GI tract is lined with a specific cell type, the interstitial cells of Cajal, that continuously generate slow pacemaker currents [58,59]. During digestion, the gastric rhythm sets the pace of gastric contractions, but the rhythm itself is generated at all times, even outside digestion [60], and even when the stomach is disconnected from the CNS [61]. Interstitial cells of Cajal form direct synapse-like connections with afferent sensory neurons [62], in a pattern established during embryonic development [63], providing a pathway for ascending information from stomach to brain. In addition, the stomach contains multiples mechanoreceptors, some with multiple receptive fields distributed throughout the stomach [64], that are active when the stomach is distended by food ingestion or when the stomach contracts. Last, the chemoreceptor-like cells lining the gut epithelium were until recently thought to act on nerves only indirectly, through the slow action of hormones. However, it was recently discovered that those receptors can directly synapse with sensory **vagal** neurons [65].

The Central Representation of Visceral Organs

Anatomofunctional studies in animals (Box 1 for details and references) suggest a distributed central representation of visceral organs, in which visceral inputs target numerous cortical structures as well as major neuromodulation nuclei (Figure 2). Cortical targets include primary and secondary somatosensory cortices, insula (Box 2), ventromedial prefrontal cortex (vmPFC), and cingulate motor regions. Some puzzling anatomical projections are not represented in Figure 2, such as projections from brainstem nuclei that relay visceral information to the lateral geniculate nucleus [66,67]. These projections are massive, to the point that, based only on connectivity, one could conclude that the function of the (visual) lateral geniculate nucleus is to

Box 1. The Numerous Targets of Visceral Inputs

Signals arising from the heart and GI tract are relayed, through **spinal** or **cranial nerve** (mostly vagal but also **glossopharyngeal**) pathways, to brainstem nuclei (nucleus tractus solitarius, NTS; and parabrachial nucleus, PBN). Ascending sensory fibers represent 80% of the vagus nerve [157]. Visceral and somatosensory inputs converge on the same neurons [158] at different stages (spinal cord, NTS). As described in Figure 2, the brainstem nuclei directly influence both serotonergic (dorsal Raphe nucleus) and noradrenergic (locus coeruleus) pathways, and act as relays for direct thalamocortical pathways. From the thalamus numerous cortical areas receive visceral inputs: primary and secondary somatosensory cortex [159,160], insula [161], ventromedial prefrontal cortex [162], cingulate motor regions [163]. The hippocampus also receives inputs from the NTS via a multisynaptic pathway [164]. In addition, the NTS and PBN reach various subcortical structures, the hypothalamus, the cerebellum, the amygdala, and the striatum, structures that in turn project to several other cortical areas. The relative balance between spinal and vagal inputs, which differ in term of fiber velocity and activation threshold, is not yet known with certainty at the cortical level [163,165]. Furthermore, the core regions of this visceral network are functionally and anatomically connected [166]. Anatomy thus suggests a relatively distributed system, with numerous main cortical targets (also Box 2).

Viscerotopic organization is present in the brainstem relay nuclei [167] as well as in the thalamus and insula [161] (Box 2). Surprisingly, whether S1 shows a viscerotopic organization, and whether and how it merges with the somatotopic representation of the body, has not been extensively studied [160]. Only scarce information is available: bladder distension activates neurons in S1 that also respond to stimulation of the hand [168], and vagus nerve stimulation suggests a GI tract representation close to the mouth region [169].

Finally, there are numerous interspecies differences. For instance, the ratio between the different types of cardiac mechanoreceptors can vary by a factor of 1–10 between cats, dogs, and monkeys [51], and the direct projections from PBN to insula and to ventromedial prefrontal cortex in rats [170] seem to be absent in monkeys [171]. The overall picture of anatomical pathways presented in this article is a composite view derived from studies in different animal species that differ between themselves and that probably also differ from humans.

relay visceral information to the cortex [68]. In addition, because the monitoring of visceral inputs is a life-sustaining function, one would expect visceral pathways to be evolutionary ancient and preserved throughout evolution. However, there are many differences between mammalian species (see Box 1 for examples).

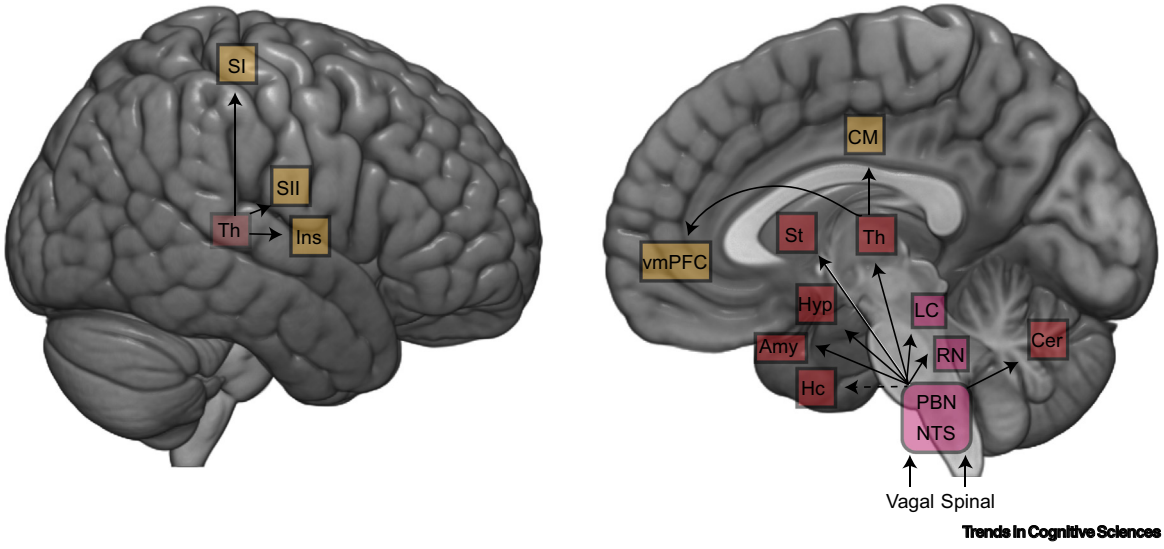


Figure 2. Central Targets of Visceral Signals. Visceral inputs reach the brain via vagal and spinal pathways which target NTS and PBN brainstem relay nuclei (purple). These in turn project to noradrenergic (LC) and serotonergic (RN) nuclei in the brainstem, and to subcortical (red) and cortical (yellow) regions. Note that this schematic representation does not highlight differences between species. Abbreviations: Amy, amygdala; Cer, cerebellum; CM, cingulate motor regions; Hc, hippocampus; Hyp, hypothalamus; Ins, insula; LC, locus coeruleus; NTS, nucleus of the solitary tract; PBN, parabrachial nucleus; RN, raphe nucleus; SI, primary somatosensory; SII, secondary somatosensory; St, striatum; Th, thalamus; vmPFC, ventromedial prefrontal cortex.

Box 2. The Insula in Interoception and Beyond

The insula is more and more often presented as 'the' primary visceral cortex in the literature on interoception. Although there is no doubt that the insula is involved in visceral processing, whether it represents the main entry point of visceral signals, or only one of the possible entry points (Figure 2), can be debated.

The insula has been labeled 'the' primary visceral cortex [172] because it contains a viscerotopic map [161]. However, studies from the 1950s show that primary and secondary somatosensory cortices also represent visceral information [159], and proposed to rename them somatovisceral cortices [173]. The viscerotopic organization of primary somatosensory cortex has scarcely been explored since then, and whether other cortical areas, for instance prefrontal or cingulate regions, are viscerotopically organized is not known.

The insula has been highlighted in experiments on explicit interoception. A pioneering study [81] revealed that attention to heartbeats increases activity in a distributed network, but the anterior insula was the only region where activity correlated with performance. Although a meta-analysis confirmed the involvement of the anterior insula [174], the necessity of the insula for explicit interoception has been questioned. A patient with insular lesions could perform well at interoceptive tasks provided that cardiac-related inputs mediated by the skin were available ([175], but see also [176]). In addition, functionally relevant HERs were observed predominantly outside the insula in several experiments [111,128,130,131,135] (Figure 5).

The insula is a very large region, of which viscerosensory and visceromotor representations occupy only a portion – in the mid- to posterior insula. Establishing the functional and structural organization of the insula with higher precision is a very active field of research [177–180]. The insula is also one of the most frequently activated regions across numerous imaging experiments [181,182]. Figure I highlights the overlap in the anterior insula of activations related to seemingly very disparate functional terms. The apparent convergence of very different functions in the insula might be due to the limited spatial resolution of standard fMRI. However, there might also be unexpected crosstalk between functions that are not usually considered together. In the cognitive neuroscience literature, the insula is more often associated with saliency [183] than with interoception. Because novelty detection alters heart rate in a systematic manner [46,113], the saliency-related function of the insula might be tightly intertwined with its visceral role. The insula could also be systematically informing value-based choices about the internal state of the body [184]. Conversely, insular activation in interoceptive tasks might be related to general cognitive functions that are not specific to interoception, such as saliency.

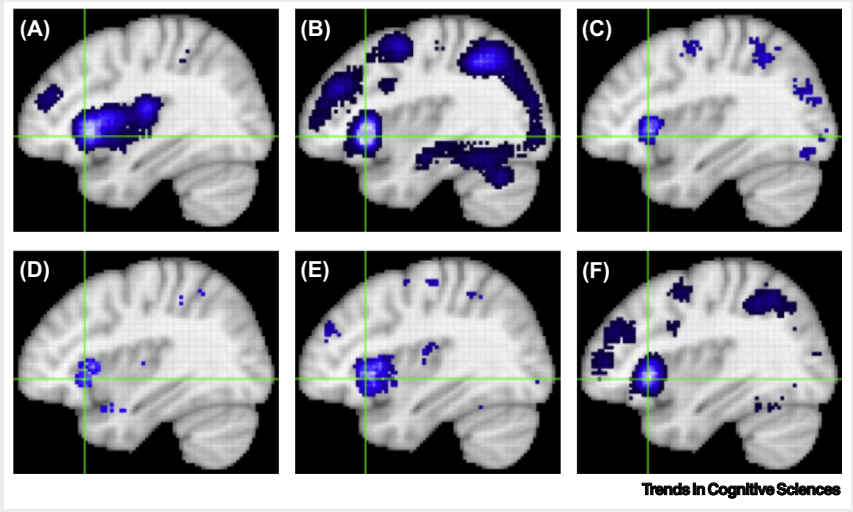


Figure I. Meta-Analytic Activation of the Insula. (A) Pain (516 studies), (B) memory (2744), (C) selective attention (146), (D) heart rate (72), (E) autonomic (117), (F) cognitive control (598), as retrieved from NeuroSynth [181], December 2018, using the uniformity test and default settings. The green crosshair corresponds to Montreal Neurological Institute (MNI) coordinates (36, 22, –4) in all panels.

Neither anatomical tracing from visceral organs to brain nor well-controlled purely ascending nerve stimulations are available in humans. Nonetheless, brain imaging studies during cardiovascular adjustments [69–72] point to the involvement of the vmPFC, insula, and hippocampal formation, as well as a dorsal anterior cingulate cortex (ACC) region compatible with cingulate motor regions. Note that in these studies it is difficult to distinguish between

ascending and descending signals. In addition, the primary somatosensory cortex displays electrophysiological responses to heartbeats [73], as will be detailed later. Other experimental paradigms based on mechanical distension of various parts of the GI tract [74], and/or on patients with GI disorders [75,76], reveal the involvement of somatomotor regions, insula, vmPFC, and mid-cingulate, as well as disconcerting modulations of activity in occipital regions [77,78]. Although some of those studies were careful to control for pain (e.g., [77,79]), the experimental set-up, with the insertion of inflatable balloons, is at best uncomfortable and stressful, and in any case nonecological. Noninvasive induction of gastric distention can be experimentally obtained by asking participants to drink moderate to large amounts of water [80], but it has not yet been used to map gastric distension with brain imaging. Nevertheless, both cardiovascular adjustments and artificial stimulation of the GI tract reveal in humans the activation of a network compatible with the anatomical pathways described in animals.

Another way to probe the visceral network is to ask participants to explicitly pay attention to visceral signals. This approach has been mostly used with heartbeats (Box 3), and shows that attention to heartbeats increases brain activity in a distributed network compatible with the known anatomy of visceral ascending pathways (somatosensory cortex, insula, mid-cingulate, and pre-motor regions), and decreases brain activity in occipital regions [81].

Dynamics of Viscera–Brain Coupling during Resting-State

Spontaneous brain activity is spatially and temporally structured, and offers a window into the basic functional organization of the brain. Spontaneous brain activity is typically studied in participants at rest, without any task or experimental stimulation, and is thus sometimes called 'intrinsic'. However, because ascending cardiac and GI signals are continuously generated, and reach numerous subcortical and cortical targets, these visceral

Box 3. Measuring Explicit Cardiac Interoception

In addition to being continuously and automatically monitored by the brain, visceral signals can sometimes become the object of conscious perception, resulting in explicit interoception. Sensitivity to internal signals in everyday life, as assessed through self-report questionnaires [185,186], appears to be altered in psychiatric disorders [3,187,188], notably in anxiety and depression [189], as well as in autism spectrum disorder [190].

Beyond questionnaires, the accuracy of cardiac interoception can be more objectively measured in experimental tasks. One paradigm that has been widely used is the heartbeat counting task [191], which consists of mentally tracking heartbeats without relying on external cues. It was recently pointed out that this paradigm has several limitations. The accuracy score negatively correlates with heart rate and is driven by heartbeat underestimation [192], it is biased by participants' belief about their heart rate [193,194], and counting heartbeats might rely more on somatosensory than on interoceptive signals [175,195]. To overcome some of these limitations, other paradigms use simultaneity judgments between heartbeats and either visual or auditory stimuli. A method of constant stimuli [196,197] was developed that involves presenting auditory stimuli at multiple latencies with respect to the R-peaks. Cardiac interoceptive accuracy is derived from participants' consistency in simultaneity judgments, and thus accommodates interindividual differences in the latency at which heartbeats may be perceived. Given the important methodological differences between the two paradigms, it is not surprising that the concordance in the results is only partial [195,198,199].

Whether and how self-report questionnaires relate to objective measures of interoceptive accuracy is not clear. For example, meditation practice, which encourages paying attention to bodily signals, increases participants' subjective confidence about interoceptive judgments, but not their objective accuracy [200–202].

To conclude, this currently very active field of research should soon propose more robust and standardized paradigms to measure explicit interoception. This should in turn help to quantify not only the relevance of cardiac interoception in psychiatric disorders but also interindividual differences in objective performance, subjective experience, and confidence judgments.

signals could act as internal stimuli that contribute to the organization of spontaneous brain activity.

Viscera–Brain Coupling Reveals New RSNs

In humans, spontaneous brain activity has been often studied with fMRI, and this revealed the existence of RSNs [82]. An RSN comprises a set of anatomically distinct regions with correlated slow fluctuations of brain activity. This approach has been extended to correlated fluctuations between brain and HRV (Figure 3A) ([70,71] for review). The HRV network (Figure 3B) is relatively extended and includes notably cingulate regions, insula, and hippocampal formation, and also the precuneus and motor cortex. Functional connectivity also fluctuates with heart rate [83,84] and pulse rate [85]. Because HRV is largely driven by the brain, these findings are thought to reflect descending influences from brain to heart. Note that it remains an experimental challenge to distinguish heart- and pulse-related physiological noise from neural signals that covary with cardiovascular changes [86–88].

The gastric rhythm is another interesting candidate for the study of brain–viscera coupling because its frequency falls in the same frequency range as RSNs. The gastric rhythm can be recorded noninvasively by placing cutaneous electrodes on the abdomen (Figure 3C) to record the **electrogastrogram** (EGG) [89,90]. In humans, the EGG spectral signature is sharp, with a peak at around 0.05 Hz corresponding to one cycle every 20 s (Figure 3D). Gastric–brain phase coupling at rest [91] reveals a network (Figure 3E) comprising several primary targets of visceral inputs, such as the primary and secondary somatosensory cortices, and mid-cingulate areas. This network unfolds over time, and a temporally structured sequence of neural events is locked to the gastric cycle. Within the gastric network, several regions are known to contain **body maps**: these include somatosensory regions and also cingulate motor regions that contain body maps activated by movement, as well as the extrastriate body area that is activated by viewing bodies and body parts [92]. The fact that a common gastric input reaches these different body maps might facilitate the coordination of the different frames of reference employed in touch, action, and vision, but this hypothesis remains to be experimentally tested. Nevertheless, current results indicate that the stomach, and probably more generally the entire GI tract [93], is coupled with the brain at rest in a novel, delayed-connectivity RSN, cutting across canonical RSNs.

Viscera–Brain Coupling and Spontaneous Brain Rhythms

Brain rhythms, as revealed by electrophysiological recordings, are another ubiquitous feature of neural activity. The dominant rhythm at rest is the alpha rhythm, ~10 Hz, but brain rhythms take place at numerous frequencies. An important organizing principle of rhythmic brain activity is phase–amplitude coupling [94], where the phase of the lower-frequency oscillation constrains the amplitude of the high-frequency oscillation (Figure 3F). Gastric–brain phase–amplitude coupling, as measured with electroencephalogram (EEG) and magnetoencephalography (MEG) [95], occurs selectively in the alpha frequency range, explaining 8% of alpha variance, and takes place in the parieto-occipital sulcus and right anterior insula (Figure 3E). Directionality analysis indicates that information flows from stomach to brain, supporting the view of a causal role of the gastric rhythm over alpha rhythm power fluctuations. Interestingly, although the stomach seems to contribute to the regulation of alpha amplitude, another study suggests that alpha peak frequency is related to heart rate [96], with a positive correlation between heart rate and occipital alpha peak frequency during wakefulness. During deep sleep, that is characterized by large-amplitude slow oscillations reflecting alternating states of hyperpolarization and depolarization of thalamocortical networks, heartbeats were more likely to occur at the beginning of the depolarization phase.

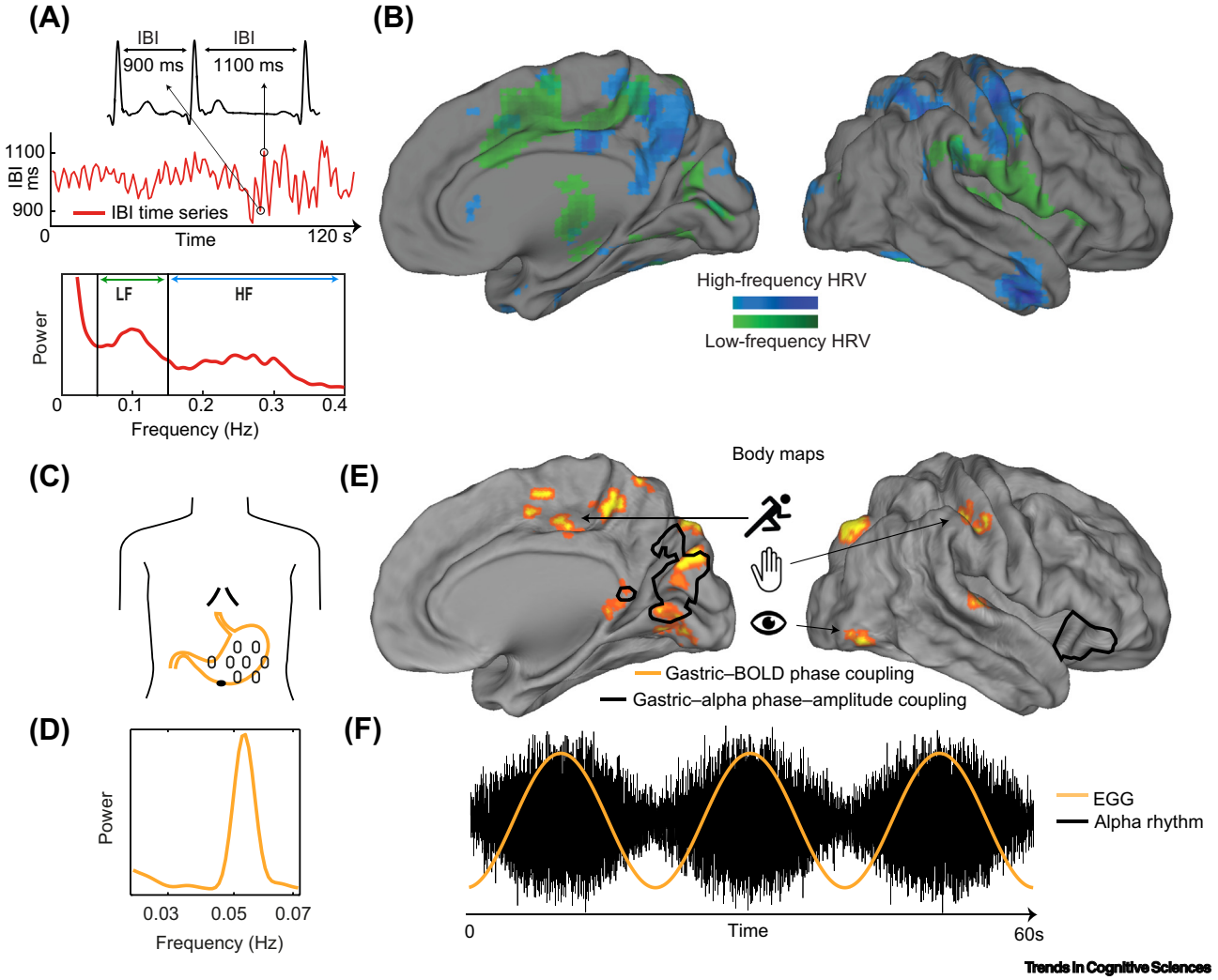


Figure 3. Heart–Brain and Gastric–Brain Coupling during Resting-State. (A) (Top) Interbeat intervals (IBI), between three consecutive heartbeats. (Middle) Two minute IBI time-series, or heart-rate variability (HRV). (Bottom) Power spectrum of HRV divided into low-frequency (0.04–0.15 Hz) and high-frequency (0.15–0.4 Hz) domains that reflect different mechanisms of autonomic control. (B) An example of the HRV network at rest [91]. (C) Electrode montage used to record the electrogastrogram (EGG). (D) The spectral signature of the gastric rhythm peaks at ~0.05 Hz. (E) Brain regions displaying delayed functional connectivity with the gastric rhythm (orange) [91]. This gastric network contains body maps involved with touch, vision, or action. The black contours correspond to brain regions where the amplitude of the alpha rhythm is constrained by gastric phase [95]. (F) Illustration (synthetic data) of phase–amplitude coupling between the phase of the gastric rhythm and the amplitude of the alpha rhythm, with large-amplitude alpha oscillations appearing at a specific moment of the slower gastric rhythm. Abbreviation: BOLD, blood oxygen level-dependent signal.

Viscera–Brain Coupling, Arousal Modulation, and Sleep Induction

The coupling between cardiovascular, GI, tract and brain activity has consequences for arousal and transition to sleep. Early observations in animals showed how artificially stimulating the mechanoreceptors of the carotid sinus induced immobility and then sleep in dogs and monkeys [97]. In cats, the mechanical distension of the carotid sinus, which causes **baroreceptor** activation, induces progressive reductions in cortical activity and the appearance of slow waves [98]. These results suggested that blood pressure 'is just as important a factor in maintaining the waking state, as the continuous inflow to the brain of proprioceptive and exteroceptive stimuli' [98].

These cardiovascular-related changes in cortical activity and behavior are mediated via modulation of noradrenergic activity in the locus coeruleus [99,100]. The GI tract also seems to be involved in the regulation of arousal. The stimulation of the GI tract induces sleep onset, electroencephalogram (EEG) synchronization, and an increase in slow-wave sleep duration [101], but whether this is mediated by inhibition of locus coeruleus activity is unknown. More generally, whether and how the spontaneous streams of heart and GI tract activity modulate serotonin and noradrenaline release is an important open question. The link between baroreceptor activation and sleep led to the so-called baroreceptor hypothesis, fully developed later, stating that under normal physiological conditions the transient activation of baroreceptors at each cardiac cycle leads to a transient dampening in cortical processing.

The relationship between cardiac/GI activity and sleep seems to go beyond sleep induction. During slow-wave sleep, the electrical activity of the duodenum is coupled with neuronal spiking activity in the cat visual cortex [102]. Furthermore, brain responses to heartbeats in humans (Box 4) decrease during sleep [96].

Viscera–Brain Coupling during Resting State: Conclusions

Taking into account the constant stream of visceral inputs stimulating the brain opens new avenues for understanding the nature and organization of spontaneous brain activity. This new source of variance of brain activity has not yet been included in biologically plausible models of spontaneous brain activity but might prove to be an important feature to further refine large-scale models of brain dynamics [103]. Similarly, brain–viscera interactions might constitute an interesting lead to understand the nature of 'background noise' in spontaneous neural firing [104], at least in the brain regions of the cardiac and gastric networks.

The Different Ways That Cardiac Signals Influence Perception and Cognition

Both the cardiac and gastric signals constrain brain activity at rest, but do they play a role in perception and cognition? Evidence for a link between the gastric rhythm and cognition is scarce (but see [80,105,106]), but the contribution of cardiac signals to cognition has been more extensively studied. Experimental studies on cardiac contribution to cognition can be grouped into three categories: neural responses to heartbeats, temporal contingencies between external stimuli and heartbeats, and heart-rate changes.

Before describing the rationale and methods in these three different approaches, it is worth reminding the reader of some basic facts about cardiac activity [107]. The different waves of the **electrocardiogram (ECG)** are labeled with conventional names (Figure 4A) corresponding to atrial (P wave) and ventricular (QRS complex) contractions, followed by ventricular relaxation (T wave). The part of the cardiac cycle where the heart muscles are active, contracting and ejecting blood, is called **systole**. It is followed by a phase in which the heart is relaxed, filling in with blood and where the ECG is more quiescent, called **diastole** (Figure 4B).

Neural responses to heartbeats (heartbeat evoked responses, HERs) are transient brain responses time-locked to heartbeats obtained by using noninvasive electrophysiological recordings such as magnetoencephalography and electroencephalography (M/EEG) as well as intracranial recordings (iEEG and electrocorticography, ECoG) (Figure 4D–G and Box 4). HERs are computed in the same way as standard evoked responses, by aligning and averaging brain activity to the occurrence of heartbeats [108]. HERs are considered to be responses evoked by the preceding heartbeat rather than being anticipatory of the next heartbeat. Indeed, because the intervals between successive heartbeats are variable, typically between 600 and 1300 ms in a healthy subject, an anticipatory command would be locked to the next heartbeat rather than to the preceding one. HERs are most likely related

Box 4. HERs – Artefacts, Pitfalls, and Controls

Averaging electrophysiological data locked to heartbeats is the first step of HER analysis (Figure 4). The heartbeat here is considered to be an internal stimulus that generates a neural evoked response – as an external visual or auditory stimulus would. However, the analysis of HERs presents some peculiarities, as compared with external stimuli. First, because cardiac activity is cyclical, there is no 'baseline activity': what happens before a heartbeat includes the response to the preceding heartbeat and/or a preparatory command for the current heartbeat ([73] for examples). Fortunately, because intervals between heartbeats are very irregular, an event precisely time-locked to a given heartbeat is related to this heartbeat rather than to the preceding or following heartbeat. Second, both MEG and EEG sensors detect the electrical activity of the heart (Figure 4D), as captured in the electrocardiogram. The cardiac artefact [203,204] (Figure 4E) can be attenuated, but not completely suppressed, using independent component analysis correction [205] (Figure 4F). Because the cardiac artefact is much reduced when the heart is electrically 'silent', during diastole [203], it is safer to restrict the analysis of HERs to this cardiac phase. Third, blood flow can generate electrode motion and/or changes in impedance, creating a pulse-artifact that varies from prominent in intracranial EEG data [73] to negligible in MEG data [204].

HERs are generally investigated by contrasting two conditions (for exceptions see [73,96,206–208]) to limit the contribution of the cardiac artefact. It still remains mandatory to verify that differences in HERs between two conditions, or between two groups of participants, do not arise from differences in the electrocardiogram itself or in cardiac parameters such as heart rate or HRV [109,110].

Analyzing brain responses to heartbeats when external stimuli are simultaneously presented raises additional challenges because the neural responses temporally overlap. When external stimuli are presented at a specific phase of the cardiac cycle, additional evidence must be provided to support the claim that differences in HERs are not due to differences in stimulus-evoked responses [156], and conversely that differences in response to external stimuli are not driven by differences in HER [209,210]. In the presence of slowly developing differences in ongoing neural activity, such as when subjects prepare to respond, HERs can exhibit nonspecific differences (Figure IB) that would be also present if the neural data were probed at any timepoint rather than being specifically locked to heartbeats, calling for additional controls [111,130].

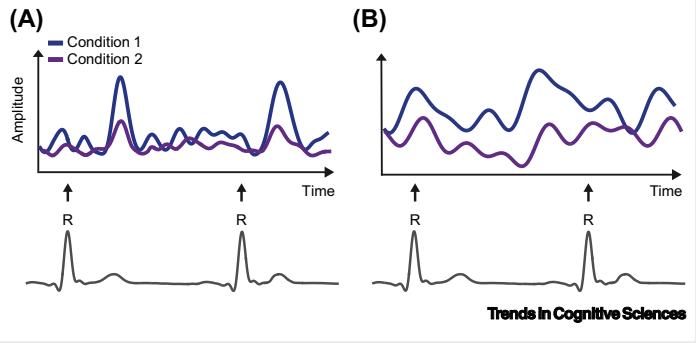


Figure I. True (A) versus Spurious (B) Differences in HERs between Conditions. (A) A transient neural activity (top), locked to the heartbeat (bottom), differs between condition 1 and condition 2. (B) The difference in neural data between condition 1 and condition 2 is sustained and is not locked to heartbeats.

to the activity of the different types of mechanoreceptors responding to cardiac contractions. However, the relative contributions of cardiac versus blood vessel mechanoreceptors, and of the pulse-sensing somatosensory and proprioceptive receptors, is not yet known. Because the various types of mechanoreceptors respond at different phases of the cardiac cycle, response latencies are difficult to predict. Last but not least, cardiac parameters, such as the volume of blood ejected at each heartbeat and the heart rate, can influence HER amplitude [109,110]. Knowledge about the physiological basis of HERs thus so far remains very limited compared with the precise description of evoked responses in other sensory modalities, where evoked responses can be used to track the cascade of activation from periphery to brainstem and cortex, and where variations in response amplitude and latency with input parameters are well documented. Links between HERs and cognition are reviewed in detail the next section.

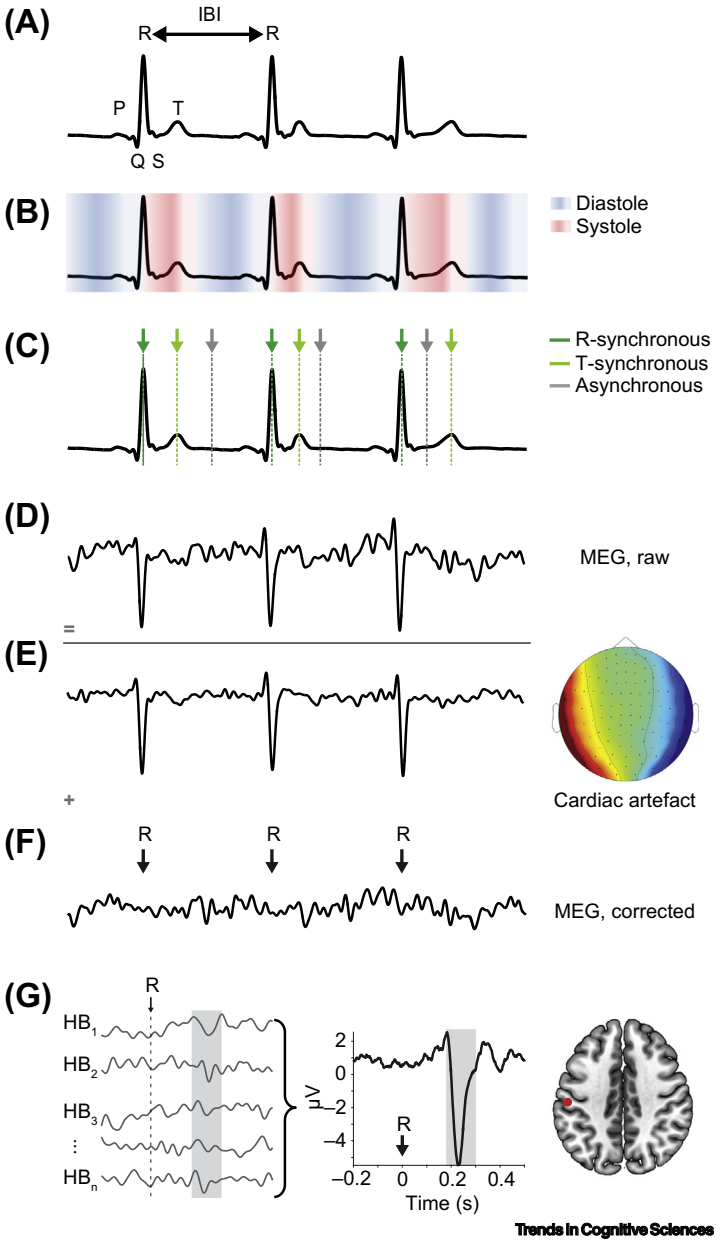


Figure 4. Electrocardiogram (ECG) and Heartbeat Evoked Responses (HERs). (A) Example of an ECG. Letters indicate the main components of a cardiac cycle. IBI is the inter-beat interval, computed between R-peaks. (B) Systole (red), when the heart contracts and ejects blood, and diastole (blue), when the heart relaxes and refills, superimposed on a standard ECG. Stimuli presented during systole and diastole are sometimes processed differently. (C) The fate of stimuli presented synchronously with respect to heartbeats (locked to R peak or T wave for instance) can be different from the fate of stimuli presented asynchronously. (D) Raw magnetoencephalographic (MEG) signal contaminated by the cardiac field artefact. (E) The cardiac field artefact (left, timecourse; right stereotypical topography) is extracted by independent component analysis (ICA). (F) MEG data after removal of the cardiac artefact ICA component. (G) The HER from intracranial data in the human somatosensory cortex. (Left) Several occurrences of responses to single heartbeats (HB); (middle) average HER; (right) electrode location.

A second way to probe the interaction between the neural processing of cardiac signals and the neural processing of an external stimulus is to test whether the occurrence of an external sensory stimulus at a specific moment in the cardiac cycle affects how this stimulus is processed (Figure 4C). For instance, a visual stimulus presented simultaneously with the R-peak might be processed differently than the same visual stimulus presented later in the cardiac cycle, or than a stimulus presented at random moments. There are several such examples, that we present later.

The third approach used to study brain–heart interplay is to measure heart-rate changes following a perceptual, cognitive, emotional, or motor event. Such changes are relatively systematic, even in nonemotional, nonstressing experimental paradigms such as detecting visual gratings [111] or auditory stimuli [46]. Because we focus here on ascending pathways from heart and GI tract to brain, rather than on the descending influences from the brain that lead to heart-rate changes, we do not describe those findings but instead refer the reader to [112–115]. Note that stimulus-driven heart-rate changes are sometimes called the cardiac response (or even the evoked cardiac response) [116], but they should not be confused with the HERs.

Heartbeat Evoked Responses

HERs index the central monitoring of cardiac inputs and can thus be used to experimentally probe the different theoretical hypotheses about the functional role of visceral inputs in emotions, bodily self, and first-person perspective (Figure 1). However, studies on HER historically adopted a simpler stance, testing whether paying attention to cardiac signals might enhance their neural processing.

Heartbeat Evoked Responses and Interoceptive Attention

The first studies on HERs were based on explicit interoceptive tasks (Box 3) in which participants were asked to mentally count their heartbeats [108,117,118]. Those pioneering studies, confirmed by more recent work [119,120], show that paying attention to heartbeats modulates the amplitude of HERs, reflecting a trade-off between internally oriented and externally oriented attention [119]. It has been further proposed that the strength of the HER modulation is related to how accurate participants are at detecting their heartbeats [108,117,118], but measuring explicit interoceptive accuracy has since proved to be a more difficult challenge than expected (Box 3). More recently, it was observed in infants that HERs covary with a new implicit measure of interoception, namely the extent to which infants prefer looking at a stimulus presented asynchronously with heartbeats [121].

Heartbeat Evoked Responses and Emotions

HERs vary with emotions while viewing mood-inducing video clips [122] or while judging the emotional expression of faces [123]. Based on the hypothesis that specific emotions might lead to precise interoceptive predictions [11], the expected repetition of an emotional facial expression leads to a reduction (for angry faces) or increase (for sad or pain faces) in HER amplitude [124,125]. Because emotions are accompanied by changes in bodily parameters, with a linear relationship between heart rate and valence, and between skin conductance and arousal [126], the emotion-related modulations of HERs might reflect either neural responses to changes in bodily state or the cortical mechanisms driving bodily changes [127]. It thus remains to be determined whether HERs drive emotional feelings or are consequences of emotional processing.

Heartbeat Evoked Responses, Bodily Self, and Self-Related Cognition

The experienced bodily self depends on neural responses to heartbeats [128]. In a paradigm using virtual reality to induce full-body illusions [128], the amplitude of responses to heartbeats in mid- to posterior cingulate cortex and supplementary motor area covaried with the degree to which participants identified themselves with an avatar and felt their bodily location displaced towards it (Figure 5A–C). Although this experiment was designed to probe the multisensory integration of HERs in the construction of bodily self, the connection with first-person perspective

becomes apparent when considering that, in normal situations, the spatial location of our body is also where we identify ourselves to be and the perspective from which we experience the world [30]. The correspondence between body location and first-person perspective is dissolved in patients with derealization/depersonalization disorders. These patients experience the world as unreal and/or experience themselves from outside their body, and notably they do not show the expected increase in HERs when paying attention to their heartbeats [129].

HERs also index the self-relatedness of spontaneous thoughts [130,131] (Figure 5D–F). Self-relatedness was broken down here into two components [132,133]: one related to the explicit self, when one thinks about oneself, as opposed to someone or something else ('Me' scale), and the other to the self as an agent, when one says 'I act/want/feel' as opposed to 'It's raining' or 'She is always late'. Participants rated their spontaneous thoughts along these two continuous dimensions. The scores on 'Me' scale varied with the amplitude of HERs in ventromedial prefrontal cortex, and the scores on the 'I' scale varied with the amplitude of HERs in posterior cingulate cortex. These two midline structures are known core areas for the self [134] and belong to the **default network**. In this experiment HERs did not vary with the emotional valence of spontaneous thoughts, showing that HERs can covary with aspects of cognition that are not related to emotions. Another study revealed that HERs in medial regions partially overlapping the default network distinguish between self and other [135].

Heartbeat Evoked Responses and Visual Perception

Experiments relating HERs and the different dimensions of the self require explicitly thinking about oneself, and thus might entail a component related to internally directed attention to bodily signals, which is known to increase HERs [108,117–120]. An account that explains HER amplitude solely based on internally directed attention could be ruled out in an experiment on vision at threshold. Indeed, internally directed attention should be detrimental when trying to detect a weak visual input. However, experimental data show that larger responses to heartbeats facilitate the detection of a visual stimulus at threshold [111].

In the experiment depicted in Figure 5G–I, a visual stimulus was presented at low contrast such that participants could detect it in only half of the trials. Whether or not the stimulus was consciously perceived (and thus was reported) depended on the amplitude of prestimulus responses to heartbeats in the ventromedial prefrontal cortex and right inferior parietal lobule. The HER effect was unrelated to changes in arousal or cortical excitability. Interestingly, HERs seemed to behave as sensory evidence, with larger HERs increasing perceptual sensitivity, but did not alter decision criterion. The results are in line with the predictions of the first-person perspective account. When the coordination of different frames of reference (as indexed by HERs) is weak, the visual stimulus fails to be consciously experienced [28,29].

Heartbeat Evoked Responses: Conclusions

The studies reviewed above show that HERs appear to be functionally relevant not only for emotion but also for basic visual perception and (nonemotional) self-related cognition. The effects were observed in various regions, and at different latencies. The reason for this variety of results is not yet known with certainty, but there are several likely candidates. First, there are multiple mechanisms of transduction of cardiac-related information, with in particular different cardiac mechanoreceptors firing at different phases of the cardiac cycle [51]. Second, the different anatomical pathways involve both fast myelinated fibers and slow unmyelinated fibers, as well as varying numbers of synaptic relays. The combination of those factors is likely to contribute to the emergence of effects at different latencies. Importantly, some of the above-mentioned studies included several physiological measures to verify that the HER effects are not due to changes in cardiac physiological parameters nor to unspecific modulations in arousal state (Box 4).

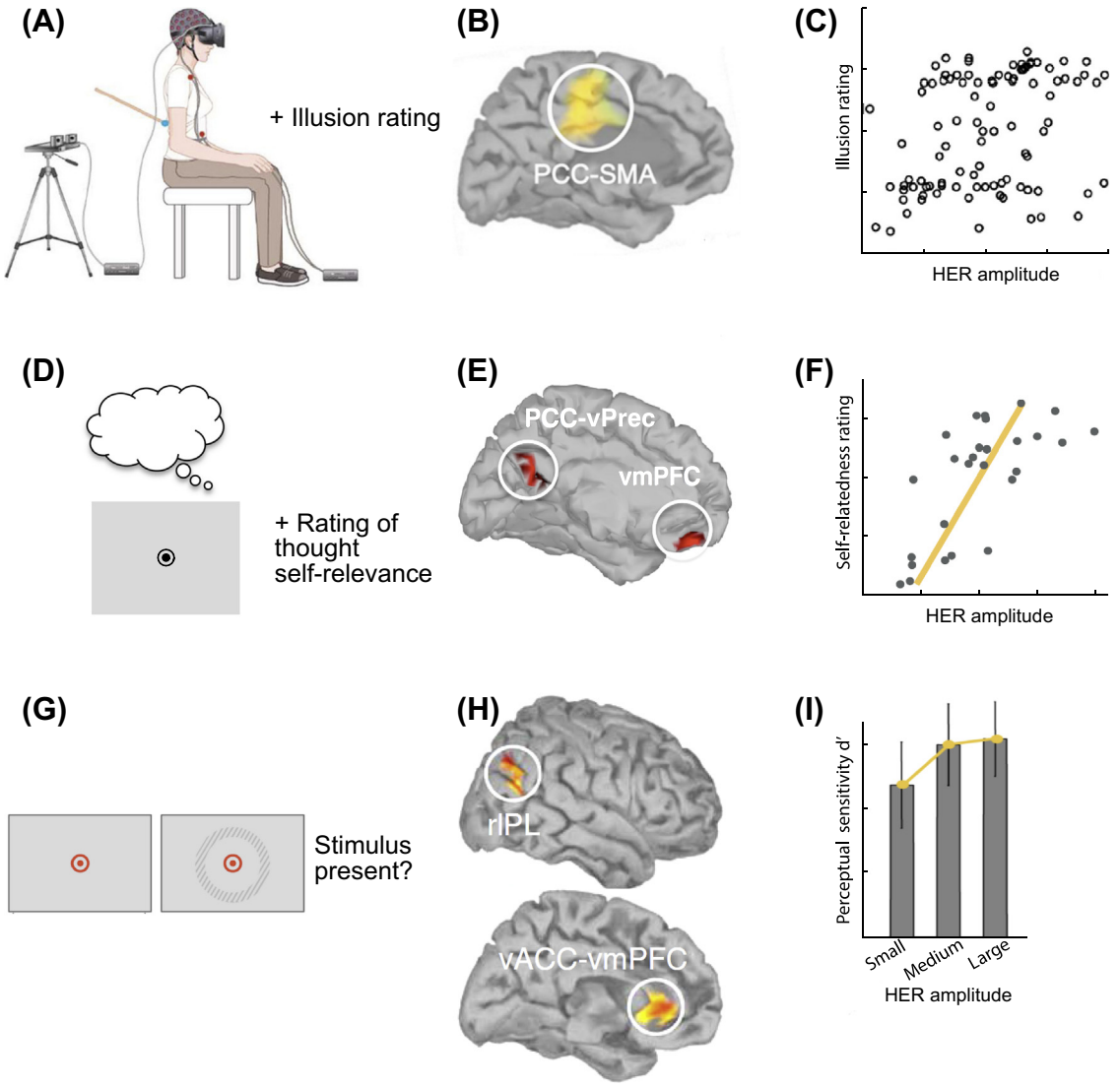


Figure 5. Heartbeat Evoked Responses (HERs) in Perception and Cognition. (A–C) HERs and bodily self [128]. (A) Participants rate their experience of a full-body illusion that depends on whether they are stroked on their back synchronously or asynchronously with the avatar they see in virtual reality. (B) HERs computed during illusion induction differ between synchronous and asynchronous conditions in posterior cingulate cortex (PCC) and supplementary motor area (SMA). (C) HER amplitude covaries with participants' ratings of illusion strength. (D–F) HERs and self in spontaneous thoughts [130,131]. (D) Participants mind-wandered while fixating, were occasionally interrupted, and were tasked to rate their thoughts according to whether they were thinking about themselves ('Me' scale) or engaged as an agent in the thought ('I' scale). (E) HERs computed during the rated thought vary with 'Me' ratings in ventromedial prefrontal cortex (vmPFC) and with 'I' ratings in posterior cingulate cortex (PCC) and ventral precuneus (vPrec). (F) Confirmation of the findings in intracranial recordings in vmPFC, with a correlation between single-trial HERs and 'Me' ratings. Another study revealed that HERs distinguish between self and other [135]. (G–I) HERs predict visual perception at threshold [111]. (G) Participants were required to report the presence or absence of a faint annulus presented around a central fixation point. (H) HERs before stimulus onset distinguished between hits and misses in right inferior parietal lobule (rIPL) and ventral anterior cingulate/ventromedial prefrontal cortex (vACC-vmPFC). (I) Larger HER amplitude before stimulus onset predicted higher perceptual sensitivity.

Temporal Contingencies between External Stimuli and Heartbeats

Another way to investigate the influence of the heart on perception and cognition is to look at how external stimuli are processed depending on their temporal relationship with

heartbeats. Two distinct lines of research have been pursued, one comparing stimuli presented early (systole) versus late (diastole) in the cardiac cycle, in a framework emphasizing bodily arousal and emotions, and the other looking at the effects of temporal synchronization between cardiac contractions and external stimuli, in the framework of multisensory integration.

The Baroreceptor Hypothesis: Systole versus Diastole

The baroreceptor hypothesis [136] inspired a long line of experimental research in humans. As summarized in [137], this hypothesis states that the activation of baroreceptors in the aorta and carotid, that is maximal at systole (when ejected blood distends artery walls), has an inhibitory effect on the CNS [138]. This hypothesis is rooted in the observations that artificial baroreceptor activation can induce sleep, as presented earlier in this review, and predicts that the processing of external sensory information would be dampened at systole. This hypothesis has been experimentally tested by comparing responses with external stimuli presented either at systole or diastole, or by comparing responses with external stimuli presented at systole versus stimuli presented randomly at any point of the cardiac cycle (Figure 4B,C).

Early experiments showed that applying neck suction during systole, a technique that artificially boosts baroreceptors activation, reduces pain sensation [139]. More recent pain studies report congruent results, with reduced pain sensation during normal baroreceptor activation without neck suction [140], and a less-pronounced startle reflex to acoustic stimuli at systole [141]. Although pain seems to be reduced at systole, in line with the baroreceptor hypothesis, other experiments find facilitatory effects for stimuli presented during systole in somatosensory detection [137], visual discrimination [142], and visual search [143] tasks, but not all results are consistent. In the visual modality, some older studies [144] as well as more recent experiments [111] did not observe any modulation of visual detection as a function of the cardiac cycle, and another study [145] found reduced visual awareness and less accuracy for stimuli presented during systole. Interestingly, motor control also seems to be dependent on the phase of the cardiac cycle. Systole is characterized by more frequent microsaccades [146], improved inhibitory motor control [147], and more frequent self-initiated movement [148].

The phasic nature of baroreceptor activity has been also investigated with emotional stimuli ([149] for recent review). The rationale for this line of research is that the physiological state of the body contributes to the experience of emotion: cardiovascular arousal – larger at systole – should therefore influence how emotional stimuli are processed. In agreement with this hypothesis, fearful faces are judged to be of greater emotional intensity and are more easily detected during systole [150]. Emotion perception can sometimes be maladaptive, as for instance in racial stereotypes where threat tends to be associated with Black people. This negative bias is enhanced during systole [151]. In addition to fear and threat, disgust perception is enhanced during late systole [152]. The available evidence thus suggests that fear, threat, and disgust perception are enhanced during systole. Processing emotional stimuli differently at systole and diastole may not convey a specific functional advantage *per se*, but in a state of high bodily arousal, when the heart is beating rapidly and diastole is shortened, the influence of baroreceptors on the brain would help in prioritizing threatening stimuli.

The results of different studies investigating the modulation of sensory and cognitive processing by cardiac phases are difficult to accommodate within a single account, and numerous questions remain to be answered. Why does baroreceptor activation sometimes facilitate stimulus processing, as for somatosensory detection or threat, and sometimes inhibit it, as for pain? Should all results be interpreted as reflecting baroreceptor activation, or does pulse-related information

mediated by tactile [53] and proprioceptive [54,55] receptors also play a role, with potentially a direct influence on muscle activity?

Multisensory Integration: Synchronous versus Asynchronous Conditions

Multisensory integration depends on temporal contingencies between stimuli, as repeatedly shown for exteroceptive stimuli. Assuming that the same rule holds true for the integration of interoceptive and exteroceptive stimuli, several studies investigated whether external events occurring synchronously (most frequently around the time of R-peak) have a different fate than external events occurring at random phases of the cardiac cycle (Figure 4C). According to the hypothesis that bodily ownership depends on multisensory integration (Figure 1D) [9,26], participants experienced enhanced body ownership of a virtual avatar hand [153] and a virtual avatar body [154,155] when these were illuminated in synchrony with their heartbeats. Participants also identified a face as their own more easily when the face was visually pulsating synchronously with their heartbeats [156].

Concluding Remarks and Future Perspectives

Despite the vital roles of the heart and GI tract, there are surprising gaps in our knowledge of brain–viscera interactions, including basic questions regarding transduction mechanisms and anatomical pathways or visceral representations in somatotopically organized cortices. Nevertheless, the recent findings reviewed here show that visceral signals are coupled with, and sometimes drive, spontaneous brain dynamics, indicating that the viscera and the brain form a single complex system. This view could open new leads to understanding the nature of spontaneous brain activity: some of the 'background noise' in spontaneous neural firing might be related to relevant brain–viscera interactions [104], and taking into account visceral inputs might prove useful in revisiting the origin and dynamics of RSNs (see Outstanding Questions). Moreover, the importance of visceral signals extends beyond emotions to include emotionless, basic visual perception, self-location, and spontaneous self-related cognition. An interpretative framework that could encompass numerous, but not all, experimental findings is that visceral inputs are useful to coordinate information expressed in different frames of reference, hence generating the unified viewpoint of the first-person perspective. Whether this hypothesis proves to hold true or not, this active field of research offers a unique opportunity to develop new theories that could blur the border between cognition and emotion, between consciousness and cognition, and between mind and body.

Outstanding Questions

How are the viscera represented? Is there a visceral homunculus, and if yes, how does it map onto the somatomotor homunculus? Is there a primary viscerosensory cortex, as for exteroceptive modalities, or is the visceral sensory system relatively distributed, as is the vestibular system for instance?

Are cardiac and GI tract inputs combined in the brain, and if yes how? Are the heart and GI tract related to different aspects of cognition, with different timescales?

What is the functional role of the link between occipital cortex activity and visceral signals?

To what extent will taking into account the constant stream of visceral inputs reaching the brain modify current views on spontaneous, resting-state brain activity? Can mechanistic models of spontaneous brain activity, either at whole-brain or unit level, be improved by including brain coupling with the heart and GI tract? Is the coordination between classical RSNs mediated by visceral inputs?

Could interventions on the heart or GI tract be informative? Several types of interventions can be considered, from invasive GI tract stimulations with inflatable balloons to patients with cardiac pacemakers, to study brain responses to visceral inputs and/or temporal contingencies between visceral and external inputs, or to design more reliable procedures for explicit interoception. It remains to be determined whether these artificial, nonecological stimulations tap onto the same mechanisms as physiological stimuli.

Different theoretical perspectives propose a functional role for ascending signals from the heart and GI tract. Can they be combined into a single account that encompasses emotions, arousal, self, and conscious perception and cognition, or are these different concepts separated by hard borders that cannot and should not be crossed?

Acknowledgments

C.T.B. is supported by the European Research Council (ERC) under the EU Horizon 2020 Research and Innovation Program (grant agreement 670325, Advanced grant BRAVIUS), a senior fellowship from the Canadian Institute for Advanced Research (CIFAR) program in Brain, Mind, and Consciousness, as well as by the Agence Nationale de la Recherche (ANR-17-EURE-0017). I.R. is supported by grants from Domaine d'Intérêt Majeur (DIM) Cerveau et Pensée and the Fondation Bettencourt-Schueller. D.A. is supported by a grant from the Ecole des Neurosciences de Paris Ile de France.

References

1. Kandel, E.R. *et al.* (2013) Nerve cells, neural circuitry, and behavior. In *Principles of Neural Science* (5th edn) (Kandel, E.R. *et al.*, eds), pp. 21–38, McGraw-Hill Professional
2. Schultz, W. *et al.* (1997) A neural substrate of prediction and reward. *Science* 275, 1593–1599
3. Critchley, H.D. and Harrison, N.A. (2013) Visceral influences on brain and behavior. *Neuron* 77, 624–638
4. Tort, A.B.L. *et al.* (2018) Respiration-entrained brain rhythms are global but often overlooked. *Trends Neurosci.* 41, 186–197
5. Smith, R. *et al.* (2017) The hierarchical basis of neurovisceral integration. *Neurosci. Biobehav. Rev.* 75, 274–296
6. Sterling, P. and Eyer, J. (1988) Allostasis: a new paradigm to explain arousal pathology. In *Handbook of Life Stress, Cognition and Health* (Fisher, S. and Reason, J., eds), pp. 629–649, John Wiley & Sons
7. Ramsay, D.S. and Woods, S.C. (2014) Clarifying the roles of homeostasis and allostasis in physiological regulation. *Psychol. Rev.* 121, 225–247
8. McEwen, B.S. and Seeman, T. (1999) Protective and damaging effects of mediators of stress. Elaborating and testing the concepts of allostasis and allostatic load. *Ann. N. Y. Acad. Sci.* 896, 30–47
9. Seth, A.K. (2013) Interoceptive inference, emotion, and the embodied self. *Trends Cogn. Sci.* 17, 565–573
10. Pezzulo, G. *et al.* (2015) Active Inference, homeostatic regulation and adaptive behavioural control. *Prog. Neurobiol.* 134, 17–35
11. Barrett, L.F. and Simmons, W.K. (2015) Interoceptive predictions in the brain. *Nat. Rev. Neurosci.* 16, 419–429
12. Marshall, A.C. *et al.* (2018) The interaction between interoceptive and action states within a framework of predictive coding. *Front. Psychol.* 9, 180

13. Allen, M. and Friston, K.J. (2018) From cognitivism to autopoiesis: towards a computational framework for the embodied mind. *Synthese* 195, 2459–2482
14. Rao, R.P.N. and Ballard, D.H. (1997) Dynamic model of visual recognition predicts neural response properties in the visual cortex. *Neural Comput.* 9, 721–763
15. Friston, K. (2010) The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138
16. Nummenmaa, L. *et al.* (2014) Bodily maps of emotions. *Proc. Natl. Acad. Sci. U. S. A.* 111, 646–651
17. Nummenmaa, L. *et al.* (2018) Maps of subjective feelings. *Proc. Natl. Acad. Sci. U. S. A.* 115, 9198–9203
18. James, W. (1890) *The Principles of Psychology*, Heny Holt
19. Damasio, A.R. (1996) The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 351, 1413–1420
20. Cannon, W.B. (1927) The James–Lange theory of emotions: a critical examination and an alternative theory. *Am. J. Psychol.* 39, 106–124
21. Schachter, S. and Singer, J.E. (1962) Cognitive, social, and physiological determinants of emotional state. *Psychol. Rev.* 69, 379–399
22. Scherer, K.R. *et al.* (2001) *Appraisal Processes in Emotion: Theory, Methods, Research*, Oxford University Press
23. Craig, A.D. (2009) How do you feel – now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* 10, 59–70
24. Craig, A.D. (2002) How do you feel? Interoception: the sense of the physiological condition of the body. *Nat. Rev. Neurosci.* 3, 655–666
25. Damasio, A. (2010) *Self Comes to Mind: Constructing the Conscious Brain*, Pantheon Books
26. Blanke, O. *et al.* (2015) Behavioral, neural, and computational principles of bodily self-consciousness. *Neuron* 88, 145–166
27. Seth, A.K. and Tsakiris, M. (2018) Being a beast machine: the somatic basis of selfhood. *Trends Cogn. Sci.* 22, 969–981
28. Park, H.D. and Tallon-Baudry, C. (2014) The neural subjective frame: from bodily signals to perceptual consciousness. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 369, 20130208
29. Tallon-Baudry, C. *et al.* (2018) The neural monitoring of visceral inputs, rather than attention, accounts for first-person perspective in conscious vision. *Cortex* 102, 139–149
30. Blanke, O. and Metzinger, T. (2009) Full-body illusions and minimal phenomenal selfhood. *Trends Cogn. Sci.* 13, 7–13
31. Vogeley, K. and Fink, G.R. (2003) Neural correlates of the first-person-perspective. *Trends Cogn. Sci.* 7, 38–42
32. Zahavi, D. (2005) *Subjectivity and Selfhood: Investigating the First-Person Perspective*, MIT Press
33. Saleem, A.B. *et al.* (2018) Coherent encoding of subjective spatial position in visual cortex and hippocampus. *Nature* 562, 124–127
34. Snyder, L.H. *et al.* (1998) Separate body- and world-referenced representations of visual space in parietal cortex. *Nature* 394, 887–891
35. Bernier, P.M. and Grafton, S.T. (2010) Human posterior parietal cortex flexibly determines reference frames for reaching based on sensory context. *Neuron* 68, 776–788
36. Chen, X. *et al.* (2018) Flexible egocentric and allocentric representations of heading signals in parietal cortex. *Proc. Natl. Acad. Sci. U. S. A.* 115, E3305–E3312
37. Vann, S.D. *et al.* (2009) What does the retrosplenial cortex do? *Nat. Rev. Neurosci.* 10, 792–802
38. Behrens, T.E.J. *et al.* (2018) What Is a cognitive map? Organizing knowledge for flexible behavior. *Neuron* 100, 490–509
39. Wang, C. *et al.* (2018) Egocentric coding of external items in the lateral entorhinal cortex. *Science* 362, 945–949
40. Bellmund, J.L.S. *et al.* (2018) Navigating cognition: spatial codes for human thinking. *Science* 362, eaat6766
41. Gershon, M.D. (1999) The enteric nervous system: a second brain. *Hosp. Pract.* 34, 31–52
42. Furness, J.B. (2012) The enteric nervous system and neurogastroenterology. *Nat. Rev. Gastroenterol. Hepatol.* 9, 286–294
43. Armour, J.A. *et al.* (1997) Gross and microscopic anatomy of the human intrinsic cardiac nervous system. *Anat. Rec.* 247, 289–298
44. Fujiwara, K. *et al.* (2016) Epileptic seizure prediction based on multivariate statistical process control of heart rate variability features. *IEEE Trans. Biomed. Eng.* 63, 1321–1332
45. Bahari, F. *et al.* (2018) A brain–heart biomarker for epileptogenesis. *J. Neurosci.* 38, 8473–8483
46. Raimondo, F. *et al.* (2017) Brain–heart interactions reveal consciousness in noncommunicating patients. *Ann. Neurol.* 82, 578–591
47. Riganello, F. *et al.* (2018) A heartbeat away from consciousness: heart rate variability entropy can discriminate disorders of consciousness and is correlated with resting-state fMRI brain connectivity of the central autonomic network. *Front. Neurol.* 9, 769
48. Sherrington, C.S. (1906) *The Integrative Action of the Nervous System*, Yale University Press
49. Ceunen, E. *et al.* (2016) On the origin of interoception. *Front. Psychol.* 7, 743
50. Janig, W. (1996) Neurobiology of visceral afferent neurons: neuroanatomy, functions, organ regulations and sensations. *Biol. Psychol.* 42, 29–51
51. Bishop, V.S. *et al.* (1983) Cardiac mechanoreceptors. In *Handbook of Physiology, Section 2: The Cardiovascular System* (Shephard, J.T. and Abbound, F.M., eds), pp. 497–555, Waverly Press
52. Zeng, W.Z. *et al.* (2018) PIEZOs mediate neuronal sensing of blood pressure and the baroreceptor reflex. *Science* 362, 464–467
53. Macefield, V.G. (2003) Cardiovascular and respiratory modulation of tactile afferents in the human finger pad. *Exp. Physiol.* 88, 617–625
54. Birznieks, I. *et al.* (2012) Modulation of human muscle spindle discharge by arterial pulsations - functional effects and consequences. *PLoS One* 7, e35091
55. Ford, T.W. and Kirkwood, P.A. (2018) Cardiac modulation of alpha motoneuron discharges. *J. Neurophysiol.* 119, 1723–1730
56. Iadecola, C. (2017) The neurovascular unit coming of age: a journey through neurovascular coupling in health and disease. *Neuron* 96, 17–42
57. Kim, K.J. *et al.* (2016) Vasculo-neuronal coupling: retrograde vascular communication to brain neurons. *J. Neurosci.* 36, 12624–12639
58. Sanders, K.M. *et al.* (2006) Interstitial cells of Cajal as pacemakers in the gastrointestinal tract. *Annu. Rev. Physiol.* 68, 307–343
59. Huizinga, J.D. and Chen, J.H. (2014) Interstitial cells of Cajal: update on basic and clinical science. *Curr. Gastroenterol. Rep.* 16, 363
60. Bozler, E. (1945) The action potentials of the stomach. *Am. J. Physiol.* 144, 693–700
61. Suzuki, N. *et al.* (1986) Boundary cells between longitudinal and circular layers: essential for electrical slow waves in cat intestine. *Am. J. Phys.* 250, G287–G294
62. Powley, T.L. and Phillips, R.J. (2011) Vagal intramuscular array afferents form complexes with interstitial cells of Cajal in gastrointestinal smooth muscle: analogues of muscle spindle organs? *Neuroscience* 186, 188–200
63. Hepworth, K.L. *et al.* (2015) Vagal fibers form associations with interstitial cells of Cajal during fetal development. *Anat. Rec. (Hoboken)* 298, 1780–1785
64. Berthoud, H.R. *et al.* (2001) Vagal and spinal mechanosensors in the rat stomach and colon have multiple receptive fields. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 280, R1371–R1381
65. Kaelberer, M.M. *et al.* (2018) A gut–brain neural circuit for nutrient sensory transduction. *Science* 361, eaat5236
66. Erisir, A. *et al.* (1997) Relative numbers of cortical and brainstem inputs to the lateral geniculate nucleus. *Proc. Natl. Acad. Sci. U. S. A.* 94, 1517–1520
67. Erisir, A. *et al.* (1997) Immunocytochemistry and distribution of parabrachial terminals in the lateral geniculate nucleus of the cat: a comparison with corticogeniculate terminals. *J. Comp. Neurol.* 377, 535–549
68. Guillery, R.W. and Sherman, S.M. (2002) Thalamic relay functions and their role in corticocortical communication: generalizations from the visual system. *Neuron* 33, 163–175

69. Shoemaker, J.K. *et al.* (2012) Cortical circuitry associated with reflex cardiovascular control in humans: does the cortical autonomic network 'speak' or 'listen' during cardiovascular arousal. *Anat. Rec. (Hoboken)* 295, 1375–1384
70. Thayer, J.F. *et al.* (2012) A meta-analysis of heart rate variability and neuroimaging studies: implications for heart rate variability as a marker of stress and health. *Neurosci. Biobehav. Rev.* 36, 747–756
71. Beissner, F. *et al.* (2013) The autonomic brain: an activation likelihood estimation meta-analysis for central processing of autonomic function. *J. Neurosci.* 33, 10503–10511
72. Gianaros, P.J. and Wager, T.D. (2015) Brain–body pathways linking psychological stress and physical health. *Curr. Dir. Psychol. Sci.* 24, 313–321
73. Kern, M. *et al.* (2013) Heart cycle-related effects on event-related potentials, spectral power changes, and connectivity patterns in the human ECoG. *Neuroimage* 81C, 178–190
74. Derbyshire, S.W. (2003) A systematic review of neuroimaging data during visceral stimulation. *Am. J. Gastroenterol.* 98, 12–20
75. Lee, I.S. *et al.* (2016) Functional neuroimaging studies in functional dyspepsia patients: a systematic review. *Neurogastroenterol. Motil.* 28, 793–805
76. Kano, M. *et al.* (2018) Understanding neurogastroenterology from neuroimaging perspective: a comprehensive review of functional and structural brain imaging in functional gastrointestinal disorders. *J. Neurogastroenterol. Motil.* 24, 512–527
77. Ladabaum, U. *et al.* (2001) Gastric distention correlates with activation of multiple cortical and subcortical regions. *Gastroenterology* 120, 369–376
78. van Oudenhove, L. *et al.* (2009) Cortical deactivations during gastric fundus distension in health: visceral pain-specific response or attenuation of 'default mode' brain function? A $H_2^{15}O$-PET study. *Neurogastroenterol. Motil.* 21, 259–271
79. Hobday, D.I. *et al.* (2001) A study of the cortical processing of ano-rectal sensation using functional MRI. *Brain* 124, 361–368
80. Schulz, A. *et al.* (2017) Gastric modulation of startle eye blink. *Biol. Psychol.* 127, 25–33
81. Critchley, H.D. *et al.* (2004) Neural systems supporting interoceptive awareness. *Nat. Neurosci.* 7, 189–195
82. Fox, M.D. and Raichle, M.E. (2007) Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat. Rev. Neurosci.* 8, 700–711
83. Chang, C. *et al.* (2013) Association between heart rate variability and fluctuations in resting-state functional connectivity. *Neuroimage* 68, 93–104
84. Nikolaou, F. *et al.* (2016) Spontaneous physiological variability modulates dynamic functional connectivity in resting-state functional magnetic resonance imaging. *Philos. Transact. A Math. Phys. Eng. Sci.* 374, 20150183
85. Shokri-Kojori, E. *et al.* (2018) An autonomic network: synchrony between slow rhythms of pulse and brain resting state is associated with personality and emotions. *Cereb. Cortex* 28, 3356–3371
86. Glover, G.H. *et al.* (2000) Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magn. Reson. Med.* 44, 162–167
87. Birn, R.M. (2012) The role of physiological noise in resting-state functional connectivity. *Neuroimage* 62, 864–870
88. Chang, C. *et al.* (2016) Brain–heart interactions: challenges and opportunities with functional magnetic resonance imaging at ultra-high field. *Philos. Transact. A Math. Phys. Eng. Sci.* 374, 20150188
89. Koch, K.L. and Stern, R.M. (2004) *Handbook of Electrogastrography*, Oxford University Press
90. Yin, J. and Chen, J.D. (2013) Electrogastrography: methodology, validation and applications. *J. Neurogastroenterol. Motil.* 19, 5–17
91. Rebollo, I. *et al.* (2018) Stomach-brain synchrony reveals a novel, delayed-connectivity resting-state network in humans. *Elife* 7, e33321
92. Orlov, T. *et al.* (2010) Topographic representation of the human body in the occipitotemporal cortex. *Neuron* 68, 586–600
93. Hashimoto, T. *et al.* (2015) Neural correlates of electrointestinography: insular activity modulated by signals recorded from the abdominal surface. *Neuroscience* 289, 1–8
94. Bragin, A. *et al.* (1995) Gamma (40–100 Hz) oscillation in the hippocampus of the behaving rat. *J. Neurosci.* 15, 47–60
95. Richter, C.G. *et al.* (2017) Phase–amplitude coupling at the organism level: the amplitude of spontaneous alpha rhythm fluctuations varies with the phase of the infra-slow gastric basal rhythm. *NeuroImage* 146, 951–958
96. Lechinger, J. *et al.* (2015) Heartbeat-related EEG amplitude and phase modulations from wakefulness to deep sleep: interactions with sleep spindles and slow oscillations. *Psychophysiology* 52, 1441–1450
97. Koch, E. (1932) Die Irradiation der pressoreceptorischen Kreislaufreflexe. *Klin. Wochenschr.* 11, 225–227
98. Bonvallet, M. *et al.* (1954) Tonus sympathique et activité électrique corticale. *Electroencephalogr. Clin. Neurophysiol.* 6, 119–144
99. Persson, B. and Svensson, T.H. (1981) Control of behaviour and brain noradrenaline neurons by peripheral blood volume receptors. *J. Neural Transm.* 52, 73–82
100. Elam, M. *et al.* (1984) Regulation of locus coeruleus neurons and splanchnic, sympathetic nerves by cardiovascular afferents. *Brain Res.* 290, 281–287
101. Kukorelli, T. and Juhasz, G. (1977) Sleep induced by intestinal stimulation in cats. *Physiol. Behav.* 19, 355–358
102. Pigarev, I.N. *et al.* (2013) Cortical visual areas process intestinal information during slow-wave sleep. *Neurogastroenterol. Motil.* 25, 268–275
103. Ponce-Alvarez, A. *et al.* (2015) Resting-state temporal synchronization networks emerge from connectivity topology and heterogeneity. *PLoS Comput. Biol.* 11, e1004100
104. Kim, K. *et al.* (2019) Resting-state neural firing rate is linked to cardiac cycle duration in the human cingulate and parahippocampal cortices. *J. Neurosci.* Published online March 6, 2019. https://doi.org/10.1523/JNEUROSCI.2291-18.2019
105. Vianna, E.P. and Tranel, D. (2006) Gastric myoelectrical activity as an index of emotional arousal. *Int. J. Psychophysiol.* 61, 70–76
106. Harrison, N.A. *et al.* (2010) The embodiment of emotional feelings in the brain. *J. Neurosci.* 30, 12878–12884
107. Berntson, G.G. *et al.* (2007) Cardiovascular psychophysiology. In *Handbook of Psychophysiology* (3rd edn) (Cacioppo, J.T. *et al.,* eds), pp. 182–210, Cambridge University Press
108. Montoya, P. *et al.* (1993) Heartbeat evoked potentials (HEP): topography and influence of cardiac awareness and focus of attention. *Electroencephalogr. Clin. Neurophysiol.* 88, 163–172
109. Schandry, R. and Montoya, P. (1996) Event-related brain potentials and the processing of cardiac activity. *Biol. Psychol.* 42, 75–85
110. Gray, M.A. *et al.* (2007) A cortical potential reflecting cardiac function. *Proc. Natl. Acad. Sci. U. S. A.* 104, 6818–6823
111. Park, H.D. *et al.* (2014) Spontaneous fluctuations in neural responses to heartbeats predict visual detection. *Nat. Neurosci.* 17, 612–618
112. Somsen, R.J. *et al.* (2004) The cardiac cycle time effect revisited: temporal dynamics of the central–vagal modulation of heart rate in human reaction time tasks. *Psychophysiology* 41, 941–953
113. Bradley, M.M. (2009) Natural selective attention: orienting and emotion. *Psychophysiology* 46, 1–11
114. Thayer, J.F. *et al.* (2009) Heart rate variability, prefrontal neural function, and cognitive performance: the neurovisceral integration perspective on self-regulation, adaptation, and health. *Ann. Behav. Med.* 37, 141–153
115. Allen, M. *et al.* (2016) Unexpected arousal modulates the influence of sensory noise on confidence. *Elife* 5, e18103
116. Lawrence, C.A. and Barry, R.J. (2010) Cognitive processing effects on auditory event-related potentials and the evoked cardiac response. *Int. J. Psychophysiol.* 78, 100–106
117. Schandry, R. *et al.* (1986) From the heart to the brain: a study of heartbeat contingent scalp potentials. *Int. J. Neurosci.* 30, 261–275

118. Pollatos, O. and Schandry, R. (2004) Accuracy of heartbeat perception is reflected in the amplitude of the heartbeat-evoked brain potential. *Psychophysiology* 41, 476–482
119. Villena-Gonzalez, M. *et al.* (2017) Attending to the heart is associated with posterior alpha band increase and a reduction in sensitivity to concurrent visual stimuli. *Psychophysiology* 54, 1483–1497
120. Petzschner, F.H. *et al.* (2019) Focus of attention modulates the heartbeat evoked potential. *Neuroimage* 180, 595–606
121. Maister, L. *et al.* (2017) Neurobehavioral evidence of interoceptive sensitivity in early infancy. *Elife* 6, e25318
122. Couto, B. *et al.* (2015) Heart evoked potential triggers brain responses to natural affective scenes: a preliminary study. *Auton. Neurosci.* 193, 132–137
123. Fukushima, H. *et al.* (2011) Association between interoception and empathy: evidence from heart-beat evoked brain potential. *Int. J. Psychophysiol.* 79, 259–265
124. Gentsch, A. *et al.* (2019) Affective interoceptive inference: evidence from heart-beat evoked brain potentials. *Hum. Brain Mapp.* 40, 20–33
125. Marshall, A.C. *et al.* (2018) Cardiac interoceptive learning is modulated by emotional valence perceived from facial expressions. *Soc. Cogn. Affect. Neurosci.* 13, 677–686
126. Lang, P.J. and Davis, M. (2006) Emotion, motivation, and the brain: reflex foundations in animal and human research. *Prog. Brain Res.* 156, 3–29
127. Luft, C.D. and Bhattacharya, J. (2015) Aroused with heart: modulation of heartbeat evoked potential by arousal induction and its oscillatory correlates. *Sci. Rep.* 5, 15717
128. Park, H.D. *et al.* (2016) Transient modulations of neural responses to heartbeats covary with bodily self-consciousness. *J. Neurosci.* 36, 8453–8460
129. Schulz, A. *et al.* (2015) Altered patterns of heartbeat-evoked potentials in depersonalization/derealization disorder: neurophysiological evidence for impaired cortical representation of bodily signals. *Psychosom. Med.* 77, 506–516
130. Babo-Rebelo, M. *et al.* (2016) Neural responses to heartbeats in the default network encode the self in spontaneous thoughts. *J. Neurosci.* 36, 7829–7840
131. Babo-Rebelo, M. *et al.* (2016) Is the cardiac monitoring function related to the self in both the default network and right anterior insula? *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 371
132. Legrand, D. and Ruby, P. (2009) What is self-specific? Theoretical investigation and critical review of neuroimaging results. *Psychol. Rev.* 116, 252–282
133. Christoff, K. *et al.* (2011) Specifying the self for cognitive neuroscience. *Trends Cogn. Sci.* 15, 104–112
134. Qin, P. and Northoff, G. (2011) How is our self related to midline regions and the default-mode network? *Neuroimage* 57, 1221–1233
135. Babo-Rebelo, M. *et al.* (2019) Neural responses to heartbeats distinguish self from other during imagination. *Neuroimage* 191, 10–20
136. Lacey, J.I. (1967) Somatic response patterning and stress: some revisions of activaction theory. In *Psychological stress: Issues in Research* (Appley, M.H. and Trumbull, R., eds), pp. 14–42, Appleton-Century-Crofts
137. Edwards, L. *et al.* (2009) Sensory detection thresholds are modulated across the cardiac cycle: evidence that cutaneous sensibility is greatest for systolic stimulation. *Psychophysiology* 46, 252–256
138. Rau, H. *et al.* (1993) Baroreceptor stimulation alters cortical activity. *Psychophysiology* 30, 322–325
139. Dworkin, B.R. *et al.* (1994) Central effects of baroreceptor activation in humans: attenuation of skeletal reflexes and pain perception. *Proc. Natl. Acad. Sci. U. S. A.* 91, 6329–6333
140. Wilkinson, M. *et al.* (2013) Electrocutaneous pain thresholds are higher during systole than diastole. *Biol. Psychol.* 94, 71–73
141. Schulz, A. *et al.* (2009) Cardiac modulation of startle eye blink. *Psychophysiology* 46, 234–240
142. Pramme, L. *et al.* (2014) Cardiac cycle time effects on mask inhibition. *Biol. Psychol.* 100, 115–121
143. Pramme, L. *et al.* (2016) Cardiac cycle time effects on selection efficiency in vision. *Psychophysiology* 53, 1702–1711
144. Elliott, R. and Graf, V. (1972) Visual sensitivity as a function of phase of cardiac cycle. *Psychophysiology* 9, 357–361
145. Salomon, R. *et al.* (2016) The insula mediates access to awareness of visual stimuli presented synchronously to the heartbeat. *J. Neurosci.* 36, 5115–5127
146. Ohl, S. *et al.* (2016) Microsaccades are coupled to heartbeat. *J. Neurosci.* 36, 1237–1241
147. Rae, C.L. *et al.* (2018) Response inhibition on the stop signal task improves during cardiac contraction. *Sci. Rep.* 8, 9136
148. Kunzendorf, S. *et al.* (2019) Active information sampling varies across the cardiac cycle. *Psychophysiology* 56, e13322
149. Garfinkel, S.N. and Critchley, H.D. (2016) Threat and the body: how the heart supports fear processing. *Trends Cogn. Sci.* 20, 34–46
150. Garfinkel, S.N. *et al.* (2014) Fear from the heart: sensitivity to fear stimuli depends on individual heartbeats. *J. Neurosci.* 34, 6573–6582
151. Azevedo, R.T. *et al.* (2017) Cardiac afferent activity modulates the expression of racial stereotypes. *Nat. Commun.* 8, 13854
152. Gray, M.A. *et al.* (2012) Emotional appraisal is influenced by cardiac afferent information. *Emotion* 12, 180–191
153. Suzuki, K. *et al.* (2013) Multisensory integration across exteroceptive and interoceptive domains modulates self-experience in the rubber-hand illusion. *Neuropsychologia* 51, 2909–2917
154. Aspell, J.E. *et al.* (2013) Turning body and self inside out: visualized heartbeats alter bodily self-consciousness and tactile perception. *Psychol. Sci.* 24, 2445–2453
155. Heydrich, L. *et al.* (2018) Cardio-visual full body illusion alters bodily self-consciousness and tactile processing in somatosensory cortex. *Sci. Rep.* 8, 9230
156. Sel, A. *et al.* (2017) Heartfelt self: cardio-visual integration affects self-face recognition and interoceptive cortical processing. *Cereb. Cortex* 27, 5144–5155
157. Agostoni, E. *et al.* (1957) Functional and histological studies of the vagus nerve and its branches to the heart, lungs and abdominal viscera in the cat. *J. Physiol.* 135, 182–205
158. Cervero, F. and Tattersall, J.E. (1987) Somatic and visceral inputs to the thoracic spinal cord of the cat: marginal zone (lamina I) of the dorsal horn. *J. Physiol.* 388, 383–395
159. Amassian, V.E. (1951) Cortical representation of visceral afferents. *J. Neurophysiol.* 14, 433–444
160. Downman, C.B. (1951) Cerebral destination of splanchnic afferent impulses. *J. Physiol.* 113, 434–441
161. Cechetto, D.F. and Saper, C.B. (1987) Evidence for a viscerotopic sensory representation in the cortex and thalamus in the rat. *J. Comp. Neurol.* 262, 27–45
162. Vogt, B.A. and Derbyshire, S.W.G. (2009) Visceral circuits and cingulate-mediated autonomic functions. In *Cingulate Neurobiology and Disease* (Vogt, B.A., ed.), pp. 220–235, Oxford University Press
163. Dum, R.P. *et al.* (2009) The spinothalamic system targets motor and sensory areas in the cerebral cortex of monkeys. *J. Neurosci.* 29, 14223–14235
164. Castle, M. *et al.* (2005) Autonomic brainstem nuclei are linked to the hippocampus. *Neuroscience* 134, 657–669
165. Willis Jr., W.D. *et al.* (2002) A critical review of the role of the proposed VMpo nucleus in pain. *J. Pain* 3, 79–94
166. Kleckner, I.R. *et al.* (2017) Evidence for a large-scale brain system supporting allostasis and interoception in humans. *Nat. Hum. Behav.* 1, 0069
167. Altschuler, S.M. *et al.* (1989) Viscerotopic representation of the upper alimentary tract in the rat: sensory ganglia and nuclei of the solitary and spinal trigeminal tracts. *J. Comp. Neurol.* 283, 248–268
168. Bruggemann, J. *et al.* (1997) Viscero-somatic neurons in the primary somatosensory cortex (SI) of the squirrel monkey. *Brain Res.* 756, 297–300
169. Ito, S. (2002) Visceral region in the rat primary somatosensory cortex identified by vagal evoked potential. *J. Comp. Neurol.* 444, 10–24
170. Shipley, M.T. and Sanders, M.S. (1982) Special senses are really special: evidence for a reciprocal, bilateral pathway between insular cortex and nucleus parabrachialis. *Brain Res. Bull.* 8, 493–501

171. Pritchard, T.C. *et al.* (2000) Projections of the parabrachial nucleus in the old world monkey. *Exp. Neurol.* 165, 101–117
172. Saper, C.B. (2002) The central autonomic nervous system: conscious visceral perception and autonomic pattern generation. *Annu. Rev. Neurosci.* 25, 433–469
173. Ruch, T.C. *et al.* (1952) Topographical and functional determinats of cortical localization patterns. *Res. Publ. Assoc. Res. Nerv. Ment. Dis.* 30, 403–429
174. Schulz, S.M. (2016) Neural correlates of heart-focused interoception: a functional magnetic resonance imaging meta-analysis. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 371
175. Khalsa, S.S. *et al.* (2009) The pathways of interoceptive awareness. *Nat. Neurosci.* 12, 1494–1496
176. Ronchi, R. *et al.* (2015) Right insular damage decreases heartbeat awareness and alters cardio-visual effects on bodily self-consciousness. *Neuropsychologia* 70, 11–20
177. Deen, B. *et al.* (2011) Three systems of insular functional connectivity identified with cluster analysis. *Cereb. Cortex* 21, 1498–1506
178. Chang, L.J. *et al.* (2013) Decoding the role of the insula in human cognition: functional parcellation and large-scale reverse inference. *Cereb. Cortex* 23, 739–749
179. Evrard, H.C. (2018) von Economo and fork neurons in the monkey insula, implications for evolution of cognition. *Curr. Opin. Behav. Sci.* 21, 182–190
180. Nomi, J.S. *et al.* (2018) Structural connections of functionally defined human insular subdivisions. *Cereb. Cortex* 28, 3445–3456
181. Yarkoni, T. *et al.* (2011) Large-scale automated synthesis of human functional neuroimaging data. *Nat. Methods* 8, 665–670
182. Behrens, T.E. *et al.* (2013) What is the most interesting part of the brain? *Trends Cogn. Sci.* 17, 2–4
183. Menon, V. and Uddin, L.Q. (2010) Saliency, switching, attention and control: a network model of insula function. *Brain Struct. Funct.* 214, 655–667
184. Gu, X. and FitzGerald, T.H. (2014) Interoceptive inference: homeostasis and decision-making. *Trends Cogn. Sci.* 18, 269–270
185. Mandler, G. *et al.* (1958) Autonomic feedback: the perception of autonomic activity. *J. Abnorm. Psychol.* 56, 367–373
186. Mehling, W.E. *et al.* (2012) The multidimensional assessment of interoceptive awareness (MAIA). *PLoS One* 7, e48230
187. Khalsa, S.S. *et al.* (2018) Interoception and mental health: a roadmap. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 3, 501–513
188. Quadt, L. *et al.* (2018) The neurobiology of interoception in health and disease. *Ann. N. Y. Acad. Sci.* 1428, 112–128
189. Paulus, M.P. and Stein, M.B. (2010) Interoception in anxiety and depression. *Brain Struct. Funct.* 214, 451–463
190. Garfinkel, S.N. *et al.* (2016) Discrepancies between dimensions of interoception in autism: implications for emotion and anxiety. *Biol. Psychol.* 114, 117–126
191. Schandry, R. (1981) Heart beat perception and emotional experience. *Psychophysiology* 18, 483–488
192. Zamariola, G. *et al.* (2018) Interoceptive accuracy scores from the heartbeat counting task are problematic: evidence from simple bivariate correlations. *Biol. Psychol.* 137, 12–17
193. Ring, C. *et al.* (2015) Effects of heartbeat feedback on beliefs about heart rate and heartbeatcounting: a cautionary tale about interoceptive awareness. *Biol. Psychol.* 104, 193–198
194. Desmedt, O. *et al.* (2018) The heartbeat counting task largely involves non-interoceptive processes: evidence from both the original and an adapted counting task. *Biol. Psychol.* 138, 185–188
195. Brener, J. and Ring, C. (2016) Towards a psychophysics of interoceptive processes: the measurement of heartbeat detection. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 371, 20160015
196. Whitehead, W.E. *et al.* (1977) Relation of heart rate control to heartbeat perception. *Biofeedback Self Regul.* 2, 317–392
197. Brener, J. and Kluvitse, C. (1988) Heartbeat detection: judgments of the simultaneity of external stimuli and heartbeats. *Psychophysiology* 25, 554–561
198. Ring, C. and Brener, J. (2018) Heartbeat counting is unrelated to heartbeat detection: a comparison of methods to quantify interoception. *Psychophysiology* 55, e13084
199. Garfinkel, S.N. *et al.* (2015) Knowing your own heart: distinguishing interoceptive accuracy from interoceptive awareness. *Biol. Psychol.* 104, 65–74
200. Khalsa, S.S. *et al.* (2008) Interoceptive awareness in experienced meditators. *Psychophysiology* 45, 671–677
201. Melloni, M. *et al.* (2013) Preliminary evidence about the effects of meditation on interoceptive sensitivity and social cognition. *Behav. Brain Funct.* 9, 47
202. Parkin, L. *et al.* (2014) Exploring the relationship between mindfulness and cardiac perception. *Mindfulness* 5, 298–313
203. Dirlich, G. *et al.* (1997) Cardiac field effects on the EEG. *Electroencephalogr. Clin. Neurophysiol.* 102, 307–315
204. Jousmaki, V. and Hari, R. (1996) Cardiac artifacts in magnetoencephalogram. *J. Clin. Neurophysiol.* 13, 172–176
205. Makeig, S. *et al.* (2002) Dynamic brain sources of visual evoked responses. *Science* 295, 690–694
206. Dirlich, G. *et al.* (1998) Topography and morphology of heart action-related EEG potentials. *Electroencephalogr. Clin. Neurophysiol.* 108, 299–305
207. Perez, J.J. *et al.* (2005) Suppression of the cardiac electric field artifact from the heart action evoked potential. *Med. Biol. Eng. Comput.* 43, 572–581
208. Park, H.D. *et al.* (2018) Neural sources and underlying mechanisms of neural responses to heartbeats, and their role in bodily self-consciousness: an intracranial EEG study. *Cereb. Cortex* 28, 2351–2364
209. Walker, B.B. and Sandman, C.A. (1982) Visual evoked potentials change as heart rate and carotid pressure change. *Psychophysiology* 19, 520–527
210. Edwards, L. *et al.* (2008) Pain-related evoked potentials are modulated across the cardiac cycle. *Pain* 137, 488–494

Appendix B: Does stroke volume influence heartbeat evoked responses?

Does stroke volume influence heartbeat evoked responses?

Anne Buot [1], Damiano Azzalini [1], Maximilien Chaumon [2], Catherine Tallon-Baudry [1]

1 Laboratoire de Neurosciences Cognitives et Computationelles, Département d'études cognitives, École normale supérieure, INSERM, PSL Research University, 75005 Paris, France.

2 Institut du Cerveau et de la Moelle épinière, ICM, Inserm U 1127, CNRS UMR 7225, Sorbonne Université, Paris, France

Corresponding author: Anne Buot
Email: anne.buot@gmail.com

Abstract

The heartbeat-evoked response (HER) is the most popular tool to assess how the brain transiently reacts to an incoming interoceptive signal, the heartbeat. HERs vary with cognitive parameters, but are also likely to be influenced by cardiac parameters. Here, we investigate to which extent the stroke volume, or volume of blood ejected at each heartbeat, measured with impedance cardiography, accounts for the amplitude of the HER measured non-invasively with magneto-encephalography, in a sample of 24 male and female participants at rest with eyes open. We find that HER's amplitude was correlated to stroke volume late in the cardiac cycle between 500 and 700ms post R-peak. We also find that once the cardiac field artefact is removed from MEG data using a cardiac-artefact tailored independent component analysis, HER no longer co-vary with stroke volume. Proper cardiac field artefact correction is thus mandatory to rule out stroke volume as a contributor to HERs. We also observed that cardiac parameters such as ECG amplitude and interbeat intervals, that are easy to measure, do correlate with stroke volume, and are more sensitive to stroke volume fluctuations than uncorrected MEG data. Cardiac parameters and ECG amplitude can thus be used as proxies when stroke volume measurements are not available.

Introduction

The heartbeat-evoked response (HER) is used to quantify how the brain transiently responds to an incoming interoceptive signal, the heartbeat. HER corresponds to electrophysiological (EEG or MEG) data averaged locked to one peak of the electrocardiogram, for instance the R peak that corresponds to ventricular contraction. Initially revealed by the group of Schandry in the late 80's (Schandry et al., 1986), it has gained a lot of interest in the recent years as it offers a window on heart to brain communication and its potential role in perception, emotion, cognition and consciousness (for review: Park and Tallon-Baudry, 2014; Azzalini et al., 2019; Park and Blanke, 2019).

The HER is a response to an internal, cyclical neuro-muscular event resulting in blood ejection. It has a number of specificities as compared to EEG or MEG responses evoked by external stimuli. The first one is that the neural response to heartbeats measured non invasively is contaminated by the cardiac field artefact, due to the fact that MEG or EEG sensors detect the electrical activity of the heart (Jousmäki and Hari, 1996; Dirlich et al., 1997). Independent component analysis (ICA) (Jung et al., 2001; Makeig et al., 1996) is efficient to attenuate, but probably not remove completely, the cardiac field artefact in MEG data (Vigario et al., 2000). Another characteristic is that HER amplitude is likely to be influenced by both stimulus parameters (here, cardiac parameters) and the nature of the task to be performed, similar to visual, auditory or somatosensory evoked responses. The experimenter cannot control cardiac parameters, but can measure them. Most recent studies in cognitive neuroscience compared HERs between two conditions, observed HERs differences and verified that neither the electrocardiogram (ECG) nor parameters derived from the ECG such as inter-beat interval (IBI) and inter-beat interval variability differed between conditions. The difference in HERs was thus attributed to a difference in the neural processing of heartbeat related to the task at hand, rather than to a difference in cardiac input (Babo-Rebelo et al., 2019, 2016; Canales-Johnson et al., 2015; Müller et al., 2015; Park and Tallon-Baudry, 2014; Pollatos and Schandry, 2004; Schulz et al., 2015; Sel et al., 2017). However, one important cardiac parameter, the stroke volume (SV), or volume of blood ejected by the left ventricle at each cycle, was not measured.

The stroke volume might contribute to HERs through physiological pathways, by inducing a change in the neural firing of mechanoreceptors located in the heart wall (Shepherd, 1985) or in blood vessels (Armour, 1973), which transduce the mechanical contraction of the heart and/or vessel distension into neural firing (Charkoudian et al., 2005; Vaschillo et al., 2012). Then, if the electrical activity of the heart varies with the stroke volume, it might also

influence the HER through the cardiac field artefact. Two experimental studies previously suggested that the HERs might be influenced by stroke volume. In a review paper, Schandry et al mentioned that in a 1minute EEG recording, HERs in participants with a larger stroke volume were of higher amplitude (Schandry and Montoya, 1996). In addition, Gray et al (Gray et al., 2007) show that cardiac output, which corresponds to the product between stroke volume and heart rate, is altered in patients with cardiac pathologies during cognitive stress and correlates with the amplitude of the HER over left temporal and frontal EEG electrodes.

The aim of this paper is to determine whether fluctuations in SV contribute to the HER measured with MEG, before and after ICA correction. The answer to this question has implications for the interpretation of HER fluctuations reported in the literature, as well as practical consequences for experimental set-up. We quantified the contribution of beat-to-beat fluctuations in stroke volume, measured non-invasively with impedance cardiography, to beat-to-beat fluctuations in HERs, measured with magnetoencephalography (MEG) in 21 healthy young adults.

Materials and Methods

Participants

24 right-handed volunteers were included in the study after giving written informed consent and were paid for their participation. The study was approved by the ethics committee CPP Ile de France III. Three participants were excluded from the study for extra systole (n=1), noisy impedance cardiography data (n=1) and outlying stroke volume values (n=1). The 21 remaining participants (10 males; age range [19 28]) were included in the analysis.

Procedure

Five minutes recordings were acquired during resting-state with eyes open. Participants were seated in front of a grey screen with a black fixation point. They were asked to remain silent, to relax, to avoid large eye or body movement, as well as repetitive physical or mental activity such as finger taping or counting.

Recordings

Continuous MEG data were acquired using a whole-head MEG system with 102 magnetometers and 204 planar gradiometers (Elekta Neuromag TRIUX, sampling rate of 1000 Hz, online low-pass filtered at 330 Hz). The ground electrode was located on the left costal margin.

ECG data were acquired on a separate amplifier (BIOPAC Systems, Inc., sampling rate of 1000 Hz, online band-pass filters 0.05 to 35 Hz) using 4 electrodes placed around the base of the neck (2 electrodes over the left and right clavicles, 2 electrodes above the left and right supraspinatus muscle) and referenced to a left abdominal location. Three different leads (I, II and III) were computed offline (Figure 1A).

Impedance cardiography data were acquired using the Biopac EBI100C module using four spot electrodes (Figure 1A). Two sources electrodes placed on the left side of the abdomen and neck, separated by a distance L (range: 27 to 34 cm), were used to inject a low intensity (400 μA rms) high frequency (12.5 kHz) current. Two monitoring electrodes placed 4 cm above and below the sources electrodes were used to measure the voltage across the tissue. The distance L, which is important when computing SV based on impedance cardiography data, was precisely measured in each subject.

Cardiac sounds were recorded by placing an a-magnetic homemade microphone (online band-pass filter 0.05-300 Hz) on the chest of the subject.

Gaze location and pupil diameter were monitored using an eye-tracker (EyeLink 1000, SR Research).

R peak detection

We detected the R-peaks by correlating the ECG with a template QRS complex defined on a subject-by-subject basis and identifying the local maximum within the episodes of correlation >0.6. R-peak detection was visually verified in all subjects. Q-peak was then identified as the minimum within 50ms before the R-peak, S-peak as the minimum within 100ms post R-peak, and T-wave as the maximum in the interval [S S+0.4s].

IBI consisted of the average time distance between two R-peaks and the heart rate variability corresponded to the standard deviation of the IBIs.

SV computation

For each participant, we calculated the beat-to-beat SV from the impedance cardiography data using the following Kubicek's formula (Kubicek et al., 1970):

$$SV = rho \times \frac{L^2}{Z0^2} \times LVET \times \left(\frac{dZ}{dt}\right)max$$

- Rho is blood resistivity, considered here to be 135 ohms.cm (Berntson et al., 2007). Although blood resistivity is known to vary from one participant to the other, and increases when current injection frequency decreases, we did not measure individual blood resistivity since we were interested in relative beat-to-beat fluctuations within each participant, rather than in stroke volume value *per se*.
- L is the distance in cm between the two recordings electrodes
- Z_0 is the average impedance, in ohm, during the left ventricular ejection
- LVET is the left ventricular ejection time in second.
- dZ/dt_max is the maximum value of the first derivative of the impedance measure in ohm/sec.

To estimate LVET (duration of the left ventricular ejection) at each beat, we identified for each heartbeat, the beginning (B-point) and end (X-point) of the ejection. The procedure followed to determine these key points is illustrated in figure 1B and 1C.

B point identification: Impedance cardiography data were low pass filtered at 15 Hz and the first derivative dZ/dt was computed. dz/dt was baseline corrected by removing the average signal in the [Q Q+0.05sec] interval (Sherwood et al., 1990), to limit respiratory influences on impedance measurements (Cybulski, 1988). The B point was then identified as the time point at which dZ/dt reaches 15% of its maximum in the [Q Q+0.4sec] window.

X point identification: The end of the ejection was defined as the time point at which dZ/dt reaches a minimum in the [B-point B-point+0.4sec] time window. If multiple local minima were observed in the time window of interest, we chose the one closest in time to the second heart sound that mainly reflects the closing of the aortic valve (Sherwood et al., 1990).

The identification of the B and X-point were performed using a homemade Matlab script (source code available at github.com/dnacombo/heart_functions), which included a visual verification step and detected outlying values for the ejection time, the delay between the R peak and the B point, and the delay between S2 (second heart sound) and the X point. An illustration of LVET determination is represented in Fig 1C.

MEG/ECG data preprocessing

Continuous MEG data were denoised using temporal signal space separation (as implemented in MaxFilter) and band pass filtered between 0.5 and 40 Hz using fourth-order Butterworth filter. MEG data contaminated by artefacts were detected through visual inspection and excluded from further analysis. MEG data contaminated by blink or saccade with amplitude >3degrees as indicated by the Eyelink were also excluded.

ICA procedure to correct for the cardiac field artifact

Independent component analysis (ICA), as implemented in the Fieldtrip toolbox (Oostenveld et al., 2011), was used to attenuate the cardiac field artifact, for both magnetometer and gradiometer signals. MEG and ECG data were first high-pass filtered at 0.5 Hz using fourth-order Butterworth filter and then epoched in [-200 to +200 ms] segments centered on the R-peaks. MEG data segments free from artifact, blink and large saccade were then decomposed in independent component using the ft_componentanalysis fieldtrip function (which is using

the runica function from EEGlab, (Delorme and Makeig, 2004)). Because temporal signal space separation induces rank deficiency, we defined the number of ICA components by first computing a principal component analysis (PCA). To detect the components corresponding to cardiac activity, we computed the mean pairwise phase consistency (ppc; Vinck et al., 2010) between each independent component and the lead II ECG signal, in the 0–25 Hz range. Components with a ppc superior to the mean+3std across ppcs were selected as cardiac components. The component selection process was repeated iteratively until no ppc was exceeding the mean+3std threshold, or until the maximum number of 3 components was reached. The cardiac components (between 1 and 3 per subject) were then removed from MEG data using the ft_rejectcomponent fieldtrip function. Before removing the cardiac components, their topographies and timecourses were visually checked.

Cardiac components extraction

To isolate the timecourses of the cardiac components, we used the same ICA decomposition as described above, selected all components that were not correlated to cardiac activity and removed them from the ICA-uncorrected data using the ft_rejectcomponent fieldtrip function. This results in one continuous signal at each sensor corresponding to all cardiac components initially rejected.

HERs

To perform the analysis on HERs, MEG data were epoched in segment of [-0.1 to 0.7sec] around the R peak. Epochs containing an artifact, a blink or a large saccade, epochs containing a cardiac cycle shorter than 0.7 s, or with corresponding SV values exceeding a [mean+/-3std] threshold, were discarded from further analysis. On average, 37% of HER was discarded (min: 5%, max: 74%), resulting in a mean number of valid HER per subject of 205 (min: 108; max: 314). On average, 3% (min:0% max: 19%) of the trials were discarded due to artefact contamination, 26% (min: 1% max: 59%) due to blink, 2% (min: 0% max: 11%) due to saccade, 5% (min: 0% max: 47%) due to cardiac cycle duration and 1% (min: 0% max: 2%) due to outlying SV values. Gradiometers data were then recombined using the fieldtrip function ft_combineplanar. The exact same selection procedure was applied to ECG data.

Correlation analysis

For each subject, each channel and each time point, the Pearson correlation coefficient between HER and SV values was calculated and Fisher-transformed. To estimate chance level correlation in each subject, channel and time point, we created shuffled data where we permuted the association between SV and HER (e.g., SV in trial i re-allocated to HER in trial j) and recalculated the Fisher-transformed correlation coefficient. This operation was performed 100 of times and the median value of the random correlation coefficient distribution was extracted. This procedure resulted in one empirical correlation and one chance level correlation, at each channel and time point in each participant. Empirical and chance level correlation could then be statistically compared at the group level with a cluster-based permutation procedure (Maris and Oostenveld, 2007) detailed below.

Median-split analysis

For each subject, the median SV value was extracted. For each channel, high (above median) and low-SV (below median) HERs were then computed by averaging MEG data across segments according to their categories. Here again, the significance of the difference was assessed at the group level using a cluster-based permutation t test.

Group-level statistics: Cluster-based permutation procedure

Group level statistics relied on a cluster-based permutation procedure (Maris and Oostenveld, 2007). This method does not require the definition of any a priori spatial or temporal regions and intrinsically corrects for multiple comparisons in time and space. Briefly, for each test (correlation/median-split), a statistical value t that quantifies the effect is computed between the two distributions being compared (empirical versus random r for the correlation, high versus low-SV HERs for the median-split). Individual samples with a statistical value corresponding to a P value below a selected threshold ($P<0.05$, two tailed) are clustered together based on temporal and spatial adjacency. The cluster is characterized by the sum of the t values of the individual samples. To establish the likelihood that a cluster was obtained by chance, we shuffled the labels (empirical versus random or high versus low-SV) 10,000 times and repeated the clustering procedure selecting the maximum positive and the minimum negative cluster-level statistic across the tests. The Monte Carlo P value corresponds to the proportion of elements in the distribution of maximal (or minimal) cluster-level statistics that exceeds (or is inferior to) the originally observed cluster-level test statistics.

Effect size

Effect size was estimated using Cohen's d definition for paired samples $d=t/\sqrt{N}$, where N is the number of participants and t the paired t-test statistics obtained when comparing high versus low-SV IBIs, ECG and HERs.

Results

Validating the stroke volume measure in the MEG at 12.5 kHz

Impedance cardiography necessitates the injection of a small high-frequency current to quantify impedance changes due to blood ejection at each cardiac cycle. We found that the most commonly employed current frequency (100 kHz) is not compatible with good quality MEG recordings, because current injection at 100 kHz generated large spectral peaks at subharmonics of stimulation frequency. we thus estimated beat-to-beat stroke volume with a 12.5 kHz current, where spectral perturbations were not visible to the naked eye. Using the Kubiceck formula (see Materials and Methods), we obtained a mean SV of 29.98 ± 1.74mL (SE) and a mean SV-variability of 3.25 ± 0.24mL. An example distribution of the stroke volume in one subject (S12) is represented on figure 2A.

In our sample, SV magnitudes were low compared to those reported in textbooks (e.g., 60-100 mL in Turner, 2000). We suspected an effect of the stimulation frequency. We recorded, in five subjects, five minutes at rest with 100 kHz stimulation and with 12.5 kHz stimulation, and compared the values of stroke volume thus obtained. In all five subjects, the measured stroke volume was smaller at 12.5 kHz (30.6±2.8 ml) than 100 kHz (48.2 ± 4.0 ml). We then verified the incidence of the stimulation frequency on each parameter used to compute the stroke volume. The left ventricular ejection time was similar at 12.5 and 100 kHz (mean LVET ± SE at 12.5 kHz = 0.288±0.007 s; at 100 kHz = 0.283±0.008 s), whereas the mean body impedance (mean Z0±SE at 12.5 = 59.45±4.05 Ohms; at 100 kHz = 39.74±2.35 Ohms) and the maximum value of the first derivative of the impedance (max(dZ/dT) ± SE at 12.5 = 3.40±0.34; at 100 kHz = 2.47±0.26) were systematically larger at 12.5 than 100 kHz. The distance between the electrodes remaining constant in each subject across stimulation frequencies, this confirmed that the stimulation frequency influenced the basal level of the impedance, leading to smaller stroke volume magnitude without affecting ejection time. Since all further analysis are based on beat-to-beat fluctuations in stroke volume, rather than on the absolute value of stroke volume, this systematic underestimation of stroke volume has little incidence.

SV and HERs from MEG data

We then evaluated whether variations of stroke volume were associated with modulations of HER amplitude computed on the ICA-uncorrected magnetometers data, on the cardiac components isolated using ICA, and on the ICA-corrected data. We performed two types of

analysis: a correlation between SV and MEG data at each time point, and a direct comparison of HERs computed in high- versus low-SV cardiac cycles. In both analyses, cluster-based permutation tests (Maris and Oostenveld, 2007) were applied to identify clusters of channels and time points on which the correlation or the difference was significant.

ICA-uncorrected magnetometer data revealed a significant negative correlation between HER and SV from 534 to 617ms after the R-peak in a cluster of 41 bilateral occipital and occipito-temporal channels (figure 3A; sum(t)=-2886, Monte Carlo p=0.037). The median-split analysis showed very similar results with a significant difference between the high- and low-SV HERs ranging from 569 to 694ms post R-peak in a cluster of 39 channels, with a similar topography (figure 3B; sum(t)=-4113, Monte Carlo p=0.002). A second negative subthreshold cluster was observed with a high tvalue, but in which the difference between the high- and low-SV HERs did not reach significance (sum(t)=-1991, Monte Carlo p=0.07). This candidate cluster ranged from 472 to 536ms post R-peak, and included 27 bilateral occipital channels.

The cardiac components identified with ICA and locked to heartbeats negatively correlated with SV from 489 to 700ms post R-peak in a cluster of 38 left-lateralized channels (figure 3C; sum(t)=-9889, Monte Carlo p=0.002). A significant difference between the high- and low HERs was also observed from 595 to 700ms post R-peak, in a cluster including 39 channels (figure 3D; sum(t)=-4419, Monte Carlo p=0.030).

Finally, HERs computed on ICA-corrected magnetometer data did not significantly correlate with the stroke volume (figure 3E), and the difference between high- and low-SV HERs was not longer significant (figure 3F). This suggests that ICA successfully removed the cardiac artefact related to stroke volume. Still, a negative candidate cluster, very similar to the subthreshold cluster observed on ICA-uncorrected data, was observed but again did not reach significance (sum(t)=-1820, MonteCarlo p=0.069). This cluster ranged from 464 to 532ms post R-peak and included 28 bilateral occipito-temporal channels.

The exact same analysis was performed on HERs computed from the combined gradiometers. In opposition to the magnetometers, we did not observe any significant correlation or any significant difference between the low- and high-SV HERs on any types of data (ICA-uncorrected data, cardiac component, ICA-corrected data).

In summary, the HERs computed on combined gradiometers are not sensitive to stroke volume, but HERs computed on magnetometer data are sensitive to fluctuations in stroke

volume. ICA correction strongly attenuates the contribution of stroke volume to HERs computed on magnetometers but might not entirely remove it.

Electrocardiogram and heart rate co-vary with stroke volume

Given previous results, it seems that HERs amplitude might still be weakly influenced by the magnitude of the stroke volume, even after ICA correction. This means that recording stroke volume could be mandatory to conclude on the origin of HERs amplitude differences, except if stroke volume fluctuations could be captured through fluctuations of the classical – and more readily accessible - measures of cardiac activity. To verify this assumption, we analysed whether the ECG itself and the instantaneous heart rate, or interbeart interval (IBI), were sensitive to stroke volume fluctuations.

At the group level, we observed a significant positive correlation between IBI and stroke volume (figure 2C, mean correlation coefficient ±sd = 0.28±0.16, one-sample t test comparing the Fischer-transformed correlation coefficient to 0, two-sided $t(20)=7.61$, $p<10^{-6}$). The correlation was confirmed at the individual level with a significant positive correlation observed in 86% (18 out of 21) of the subjects. One example of co-variation between SV and IBI in one subject (S12) is shown on figure 2B.

The comparison between high and low-SV trials analysis revealed a similar pattern with significantly longer IBIs in cardiac cycles with a larger stroke volume (figure 2D, mean difference between the high- and low-SV cardiac cycles ±sd = 29.70±18.45ms; one-sample t test, $t(20)=7.37$, $p<10^{-6}$). Individual data revealed a significant difference of IBI between high- and low-SV cycles in 81% (17 out of 21) of the subjects.

In conclusion, fluctuations in stroke volume were significantly associated with fluctuations of the IBI, with larger stroke volume being associated with longer cardiac cycle.

We then assessed the relationship between stroke volume and ECG. As ECG traces vary depending on the location of the recording electrodes, the analysis was performed on three different ECG derivations, named lead I, lead II and lead III in reference to the Einthoven triangle (figure 1A). As shown in detail in table 1 and illustrated on figure 4, all ECG leads significantly correlated with stroke volume at the group level, at different latencies contained in the 125-540ms time window post R peak. In other words, spontaneous fluctuations in stroke volume are associated with modulations of ECG amplitude, during the ejection of the blood from the left ventricle, but also later in the cardiac cycle during diastole.

Table 1. Summary of the correlation and median-split analysis on the ECG

	Lead I			Lead II			Lead III		
	Latency (ms)	Sign of the correlation	Cluster statistics (sum(t), MonteCarlo p)	Latency (ms)	Sign of the correlation	Cluster statistics (sum(t), MonteCarlo p)	Latency (ms)	Sign of the correlation	Cluster statistics (sum(t), MonteCarlo p)
Correlation with SV	139-264	-	sum(t)=-399 p=0.01	276-543	+	sum(t)=-1126 $p<10^{-3}$	156-306	+	sum(t)=367 p=0.036
	441-534	+	sum(t)=294 p=0.026				316-527	+	sum(t)=821 p=0.002
SV High vs. Low	141-207	-	sum(t)=-160 p=0.044	361-522	+	sum(t)=599 p=0.001	126-297	+	sum(t)=563 p=0.005
	444-527	+	sum(t)=285 p=0.007				347-516	+	sumt(t)=561 p=0.005

ECG is more sensitive to SV than MEG data

We have shown that HERs computed from magnetometers, as well as instantaneous heart rate and ECG itself are sensitive to the stroke volume. Because measuring stroke volume by impedance cardiography requires a specific set-up and induces additional noise on electrophysiological data, it would be interesting to be able to use the ECG or the IBI as a proxy to detect stroke volume differences between conditions. To determine whether this is a possibility, we computed the Cohen's d effect size for the difference between high- and low-SV variables (table 2).

Table 2. Summary of the effect sizes

	Raw	ICA cardiac components	ECG lead I (horizontal)	ECG lead II (vertical)	ECG lead III (vertical)	IBI
Mean across all channels	0.61	0.62	0.54	0.83	0.74	1.61
Mean on max channel	0.64	0.64	0.70		0.73	

The highest Cohen's d was observed for IBI with an effect size of 1.61. The second highest Cohen's D were observed for ECG-lead II with an effect size of 0.83. Effect sizes on HERs computed from magnetometers were systematically smaller, reaching 0.72 when selecting the channel showing the largest effect. In conclusion, IBIs and ECG lead II are more sensitive to stroke volume fluctuations than MEG data.

Discussion

In this study, we found that in healthy participants at rest, stroke volume contributed to HER amplitude in occipital and bilateral occipito-temporal areas in ICA-uncorrected MEG data. The latency of the influence of SV on MEG data was late in the cardiac cycle, between 500 and 700ms post R-peak. However, using an ICA procedure to remove signals covarying with the ECG, we found that the influence of SV disappeared in ICA-corrected MEG data, but was present in the cardiac-related ICA components of MEG signals. Because stroke volume requires a specific equipment to be recorded, and can induce noise in electromagnetic recordings, we also explored the link between the stroke volume and two easily accessible and commonly studied measures of cardiac activity: the IBI and the ECG. We found that the stroke volume was positively correlated with the IBI, and co-varied with the ECG at earlier latencies than ICA-uncorrected MEG data, ranging from 150 to 550ms post R-peak depending on ECG lead. Finally, we found that the influence of SV on ICA-uncorrected MEG data had a smaller effect size than the influence of SV on ECG lead II and on IBI, which was by far the most sensitive measure. To our knowledge, this is the first study exploring the link between the stroke volume and HER's amplitude on a beat-to-beat basis and it is also the first study evaluating the incidence of cardiac ICA correction on it. Studying whether or not SV has an influence on ECG data, as well as the incidence of ICA correction on SV effect may help us in understanding the origin of the SV effect.

Origin of the SV effect

Before ICA correction, MEG data do co-vary with stroke volume at late latencies. Beat-to-beat fluctuations in stroke volume are susceptible of contaminating HER computed from MEG data through multiple pathways. First, we have shown that the ECG, which contaminates MEG data, varies with the SV. HERs variations related to the SV might thus come from the cardiac field artifact. This is coherent with the observation that ICA procedure is effective at removing the SV effect from MEG data, since the cardiac ICA components are identified based on their correlation with the ECG. However, if SV affected MEG data through the cardiac field artefact, we would expect SV to co-vary with MEG at the same latencies than with ECG, which was not the case in our study. Still, the fact that the SV effect is removed from MEG data by the cardiac ICA procedure strongly indicates that the source of SV influence is statistically dependent on ECG signal.

A puzzling observation about the ICA correction is that, despite the fact that ICA components were computed from segments of data ranging from -200 to +200ms around the R peak, it modulated the SV effect that occurred at later latencies, from +500 to +700ms around the R peak. This suggests a statistical dependency of MEG signal locked to the R-peak between these two time-windows

Another candidate to account for the SV effect on ICA-uncorrected MEG data is the pulse artifact. The origin of this artifact is still unclear but has been related to different factors such as the small movement of body parts linked to the pulsation of blood vessels and the movement of blood particles (Debener et al., 2010) that have different magnetic properties depending on their oxygenation level (Tank et al., 1992). From EEG studies, the pulse artifact shares similar characteristics with the SV effect since it varies within subjects across cardiac cycles, and also with the duration of the cardiac cycle (Debener et al., 2010). Regarding MEG data, the influence of blood flow is not clear. Whereas one study suggested that MEG data were not contaminated by blood flow (Jousmäki and Hari, 1996), another study suggested that a late artifact might be related to the flow of blood hitting the aortic arch (Rodin et al., 2005). However, in both MEG and EEG data, the pulse artifact seems to occur around 200ms after the R-peak (Allen et al., 1998; Rodin et al., 2005), at the time of the ejection of blood from the left ventricle, which does not correspond to the latencies of the SV effect.

Finally, the late latency of the effect might points toward a more integrated pathway with the observed SV effect being a brain response to some cardiac events. Whereas some studies postulated that baroreceptor's information should reach the central nervous system between 400 and 800ms after the R-peak (Gray et al., 2007), other studies estimated that HER's latency should range from 280 to 370ms after the R-peak given that the pressure peak in the heart and aorta occurs at 180 and 220ms respectively (Schandry and Weitkunat, 1990). It has to be mentioned that a great variety of arterial baroreceptors exists, with both myelinated and non-myelinated pathways (Shepherd, 1985) and various number of synapses (Malliani et al., 1986). Additionally, baroreceptors are not only present in the arteries but also in the heart atria and ventricles (Armour, 1973). Activation of these baroreceptors could occur not only at the end of the systole when the blood is ejected from the left ventricle but at any time during the heart cycle, while the heart is filling. Additional pathways could also be recruited such as the skin somatosensory pathway (Khalsa et al., 2009) or the recently suggested vasculo-neuronal coupling (Kim et al., 2016). Under the hypothesis of the SV effect being a neural response, ICA

procedure would remove more than the cardiac field artifact from the data. In fact any signals with a strong statistical dependency with the ECG might be removed from the data irrespective of its origin.

What to do in the absence of SV monitoring?

The modulation of HERs amplitude according to SV means that a difference in HERs between two conditions, even if it occurs late in the cardiac cycle when the heart is electrically silent, can still be related to a difference in cardiac activity and not solely to a difference in the way cardiac information are integrated by the central nervous system. To eliminate or at least lower the influence of cardiac activity on HER amplitude, correction method such as ICA are often used to remove the so-called cardiac field artifact from the brain data. Here, we showed that the stroke volume effect was present on the cardiac ICA components and not anymore on the ICA-corrected data. This highlights the importance of applying a correction procedure on brain data when studying HERs and shows the efficiency of the ICA in removing SV effects. However it must be kept in mind that the ICA is a statistical procedure, which results might slightly vary from one computation to the other. Additionally, the ICA procedure used in our study is not the standard textbook procedure since we segmented the data in 400ms time-window centered on the R-peak. Applying the same filters on ECG and MEG data might also be important to select the cardiac components, even more in case of noisy datasets. Thus we cannot rule out that in some instances, depending on ICA parameters and efficiency, an effect of stroke volume might remain in ICA-corrected data.

Luckily, stroke volume-related modulations were also observed on the ECG and on the IBI, with significant correlations as well as differences in high- versus low-SV trials. The highest effect size was observed for the IBI and then for the lead II of the ECG. This means that in the absence of stroke volume monitoring, comparing ECG waveforms (with a preference for the lead II of the Einthoven triangle) and IBI lengths between two experimental conditions can be used as a proxy to evaluate stroke volume differences between the two conditions. In the absence of differences between IBI or between ECG, it seems very unlikely that differences in brain response would be linked to a difference in cardiac activity and especially in stroke volume. It must be noted that the stroke volume effect on the ECG is observed both early and late in the cardiac cycle, and that SV-related difference in ECG amplitude is fairly small and undetectable by eye. Thus, to rule out an effect of cardiac parameters between two conditions using the ECG, we recommend performing statistical test on the entire ECG timeseries. Even

if we cannot conclude here on a potential beat-to-beat effect of the SV on EEG recordings, the recommendations on ECG and IBI processing still hold for EEG studies to rule out an effect of experimental manipulation on heart activity.

Conclusion

We have shown that the stroke volume, a measure of cardiac activity that might play an important role in the generation of HER, modulates the amplitude of the HER. This finding suggests that not only the occurrence of a heartbeat but also the parameters of the heartbeats might be important in shaping the brain response to the heartbeat. Our results also highlight the fact that careful assessment of cardiac activity and controls are necessary when studying HER.

References

Allen, P.J., Polizzi, G., Krakow, K., Fish, D.R., Lemieux, L., 1998. Identification of EEG events in the MR scanner: the problem of pulse artifact and a method for its subtraction. Neuroimage 8, 229–239. https://doi.org/10.1006/nimg.1998.0361

Armour, J., 1973. Physiological behavior of thoracic cardiovascular receptors. American Journal of Physiology-Legacy Content 225, 177–185. https://doi.org/10.1152/ajplegacy.1973.225.1.177

Azzalini, D., Rebollo, I., Tallon-Baudry, C., 2019. Visceral Signals Shape Brain Dynamics and Cognition. Trends Cogn. Sci. (Regul. Ed.) 23, 488–509. https://doi.org/10.1016/j.tics.2019.03.007

Babo-Rebelo, M., Buot, A., Tallon-Baudry, C., 2019. Neural responses to heartbeats distinguish self from other during imagination. NeuroImage 191, 10–20. https://doi.org/10.1016/j.neuroimage.2019.02.012

Babo-Rebelo, M., Richter, C.G., Tallon-Baudry, C., 2016. Neural Responses to Heartbeats in the Default Network Encode the Self in Spontaneous Thoughts. The Journal of Neuroscience 36, 7829–7840. https://doi.org/10.1523/JNEUROSCI.0262-16.2016

Berntson, G.G., Quigley, K.S., Lozano, D., 2007. Cardiovascular Psychophysiology, in: Handbook of Psychophysiology- Third Edition. Cambridge University Press, pp. 182–210.

Canales-Johnson, A., Silva, C., Huepe, D., Rivera-Rei, Á., Noreika, V., Garcia, M. del C., Silva, W., Ciraolo, C., Vaucheret, E., Sedeño, L., Couto, B., Kargieman, L., Baglivo, F., Sigman, M., Chennu, S., Ibáñez, A., Rodríguez, E., Bekinschtein, T.A., 2015. Auditory Feedback Differentially Modulates Behavioral and Neural Markers of Objective and Subjective Performance When Tapping to Your Heartbeat. Cereb. Cortex 25, 4490–4503. https://doi.org/10.1093/cercor/bhv076

Charkoudian, N., Joyner, M.J., Johnson, C.P., Eisenach, J.H., Dietz, N.M., Wallin, B.G., 2005. Balance between cardiac output and sympathetic nerve activity in resting humans: role in arterial pressure regulation. J. Physiol. (Lond.) 568, 315–321. https://doi.org/10.1113/jphysiol.2005.090076

Cybulski, G., 1988. Computer method for automatic determination of stroke volume using impedance cardiography signals. Acta Physiol Pol 39, 494–503.

Debener, S., Kranczioch, C., Gutberlet, I., 2010. EEG Quality: Origin and Reduction of the EEG Cardiac-Related Artefact, in: Mulert, C., Lemieux, L. (Eds.), EEG - FMRI:

Physiological Basis, Technique, and Applications. Springer, Berlin, Heidelberg, pp. 135–151. https://doi.org/10.1007/978-3-540-87919-0_8

Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of Neuroscience Methods 134, 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

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

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

Jousmäki, V., Hari, R., 1996. Cardiac artifacts in magnetoencephalogram. J Clin Neurophysiol 13, 172–176.

Jung, T.-P., Makeig, S., McKeown, M.J., Bell, A.J., Lee, T.-W., Sejnowski, T.J., 2001. Imaging Brain Dynamics Using Independent Component Analysis. Proc IEEE Inst Electr Electron Eng 89, 1107–1122. https://doi.org/10.1109/5.939827

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

Khalsa, S.S., Rudrauf, D., Feinstein, J.S., Tranel, D., 2009. The pathways of interoceptive awareness. Nature Neuroscience 12, 1494–1496. https://doi.org/10.1038/nn.2411

Kim, K.J., Ramiro Diaz, J., Iddings, J.A., Filosa, J.A., 2016. Vasculo-Neuronal Coupling: Retrograde Vascular Communication to Brain Neurons. The Journal of Neuroscience 36, 12624–12639. https://doi.org/10.1523/JNEUROSCI.1300-16.2016

Kubicek, W.G., From, A.H., Patterson, R.P., Witsoe, D.A., Castaneda, A., Lillehei, R.C., Ersek, R., 1970. Impedance cardiography as a noninvasive means to monitor cardiac function. J Assoc Adv Med Instrum 4, 79–84.

Makeig, S., Bell, A.J., Jung, T.-P., Sejnowski, T.J., 1996. Independent component analysis of electroencephalographic data. Advances in Neural Information Processing Systems 8, 145–151.

Malliani, A., Lombardi, F., Pagani, M., 1986. Chapter 4 Sensory innervation of the heart, in: Progress in Brain Research. Elsevier, pp. 39–48. https://doi.org/10.1016/S0079-6123(08)62755-7

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

Müller, L.E., Schulz, A., Andermann, M., Gäbel, A., Gescher, D.M., Spohn, A., Herpertz, S.C., Bertsch, K., 2015. Cortical Representation of Afferent Bodily Signals in Borderline Personality Disorder: Neural Correlates and Relationship to Emotional Dysregulation. JAMA Psychiatry 72, 1077–1086. https://doi.org/10.1001/jamapsychiatry.2015.1252

Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.-M., 2011. FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. Computational Intelligence and Neuroscience 2011, 1–9. https://doi.org/10.1155/2011/156869

Park, H.-D., Blanke, O., 2019. Heartbeat-evoked cortical responses: Underlying mechanisms, functional roles, and methodological considerations. Neuroimage 197, 502–511. https://doi.org/10.1016/j.neuroimage.2019.04.081

Park, H.-D., Tallon-Baudry, C., 2014. The neural subjective frame: from bodily signals to perceptual consciousness. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 369, 20130208. https://doi.org/10.1098/rstb.2013.0208

Pollatos, O., Schandry, R., 2004. Accuracy of heartbeat perception is reflected in the amplitude of the heartbeat-evoked brain potential. Psychophysiology 41, 476–482. https://doi.org/10.111/1469-8986.2004.00170.x

Rodin, E., Funke, M., Haueisen, J., 2005. Cardio-respiratory contributions to the magnetoencephalogram. Brain Topogr 18, 37–46. https://doi.org/10.1007/s10548-005-7899-7

Schandry, R., Montoya, P., 1996. Event-related brain potentials and the processing of cardiac activity. Biological Psychology 42, 75–85. https://doi.org/10.1016/0301-0511(95)05147-3

Schandry, R., Sparrer, B., Weitkunat, R., 1986. From the heart to the brain: a study of heartbeat contingent scalp potentials. Int. J. Neurosci. 30, 261–275.

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

Schulz, A., Köster, S., Beutel, M.E., Schächinger, H., Vögele, C., Rost, S., Rauh, M., Michal, M., 2015. Altered patterns of heartbeat-evoked potentials in depersonalization/derealization disorder: neurophysiological evidence for impaired

cortical representation of bodily signals. Psychosom Med 77, 506–516. https://doi.org/10.1097/PSY.0000000000000195

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

Shepherd, J.T., 1985. The heart as a sensory organ. Journal of the American College of Cardiology 5, 83B-87B. https://doi.org/10.1016/S0735-1097(85)80533-7

Sherwood(Chair), A., Allen, M.T., Fahrenberg, J., Kelsey, R.M., Lovallo, W.R., Doornen, L.J.P., 1990. Methodological Guidelines for Impedance Cardiography. Psychophysiology 27, 1–23. https://doi.org/10.1111/j.1469-8986.1990.tb02171.x

Tank, D.W., Ogawa, S., Ugurbil, K., 1992. Mapping the brain with MRI. Current Biology 2, 525–528. https://doi.org/10.1016/0960-9822(92)90011-X

Turner, M.A., 2000. Impedance cardiography: a noninvasive way to monitor hemodynamics. Dimens Crit Care Nurs 19, 2–9; quiz 10–12.

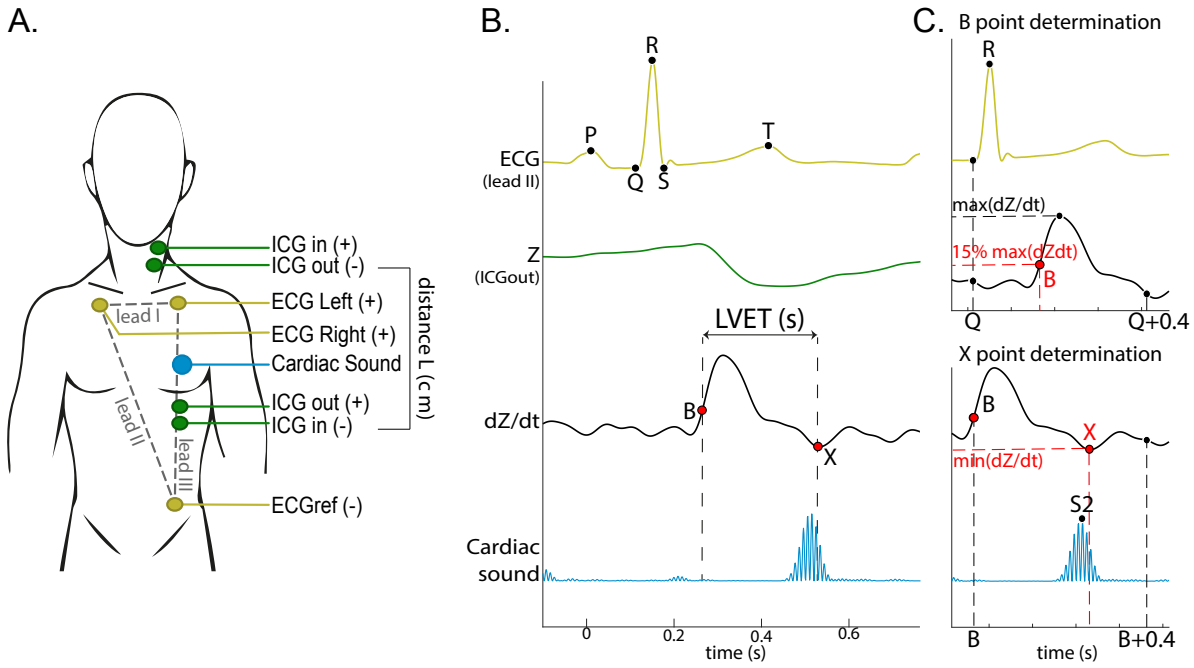
Vaschillo, E.G., Vaschillo, B., Buckman, J.F., Pandina, R.J., Bates, M.E., 2012. Measurement of vascular tone and stroke volume baroreflex gain. Psychophysiology 49, 193–197. https://doi.org/10.1111/j.1469-8986.2011.01305.x

Vigario, R., Sarela, J., Jousmiki, V., Hamalainen, M., Oja, E., 2000. Independent component approach to the analysis of EEG and MEG recordings. IEEE Transactions on Biomedical Engineering 47, 589–593. https://doi.org/10.1109/10.841330

Vinck, M., van Wingerden, M., Womelsdorf, T., Fries, P., Pennartz, C.M.A., 2010. The pairwise phase consistency: A bias-free measure of rhythmic neuronal synchronization. NeuroImage 51, 112–122. https://doi.org/10.1016/j.neuroimage.2010.01.073

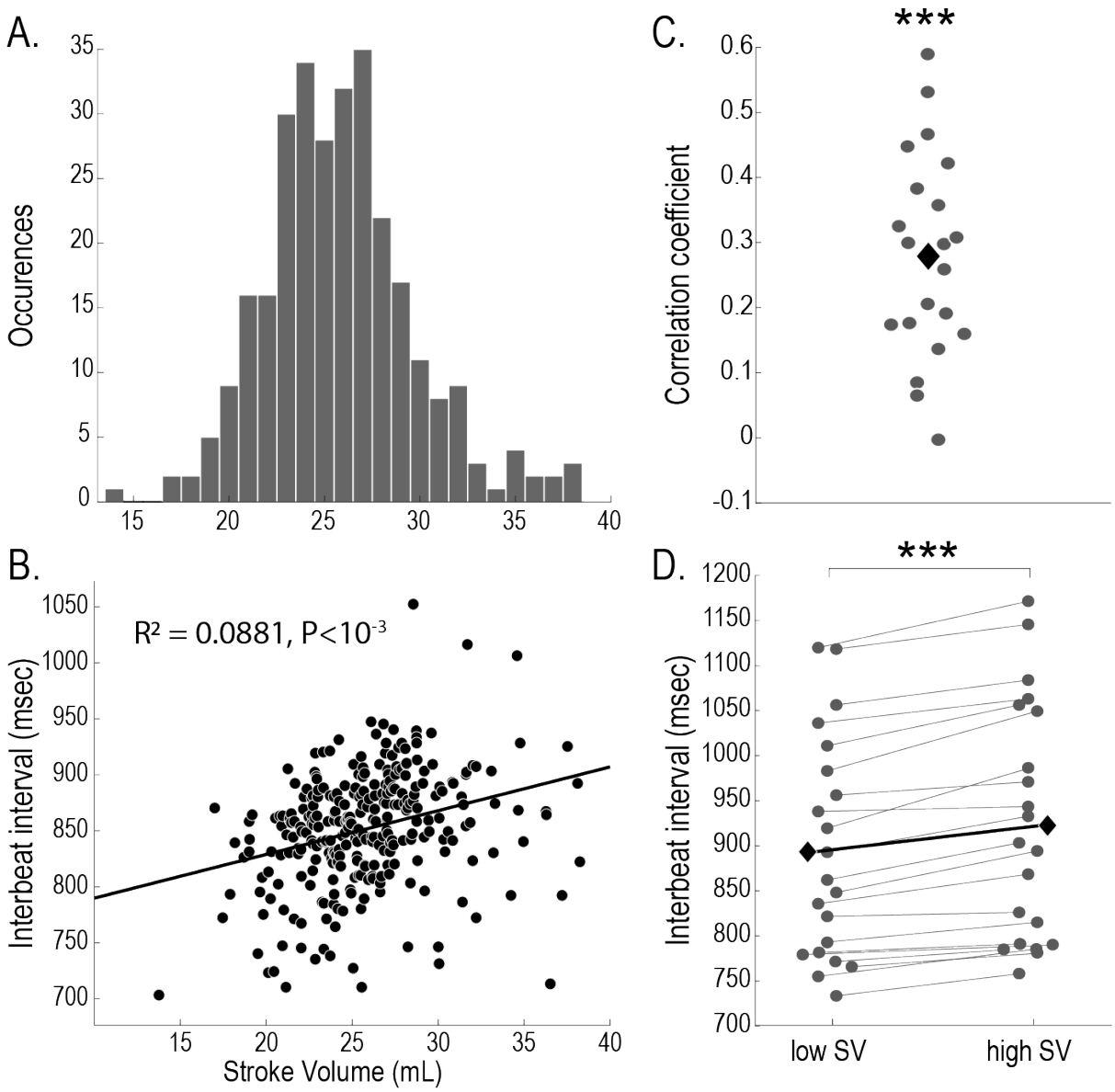
Figures

Figure 1. Experimental set up and stroke volume computation



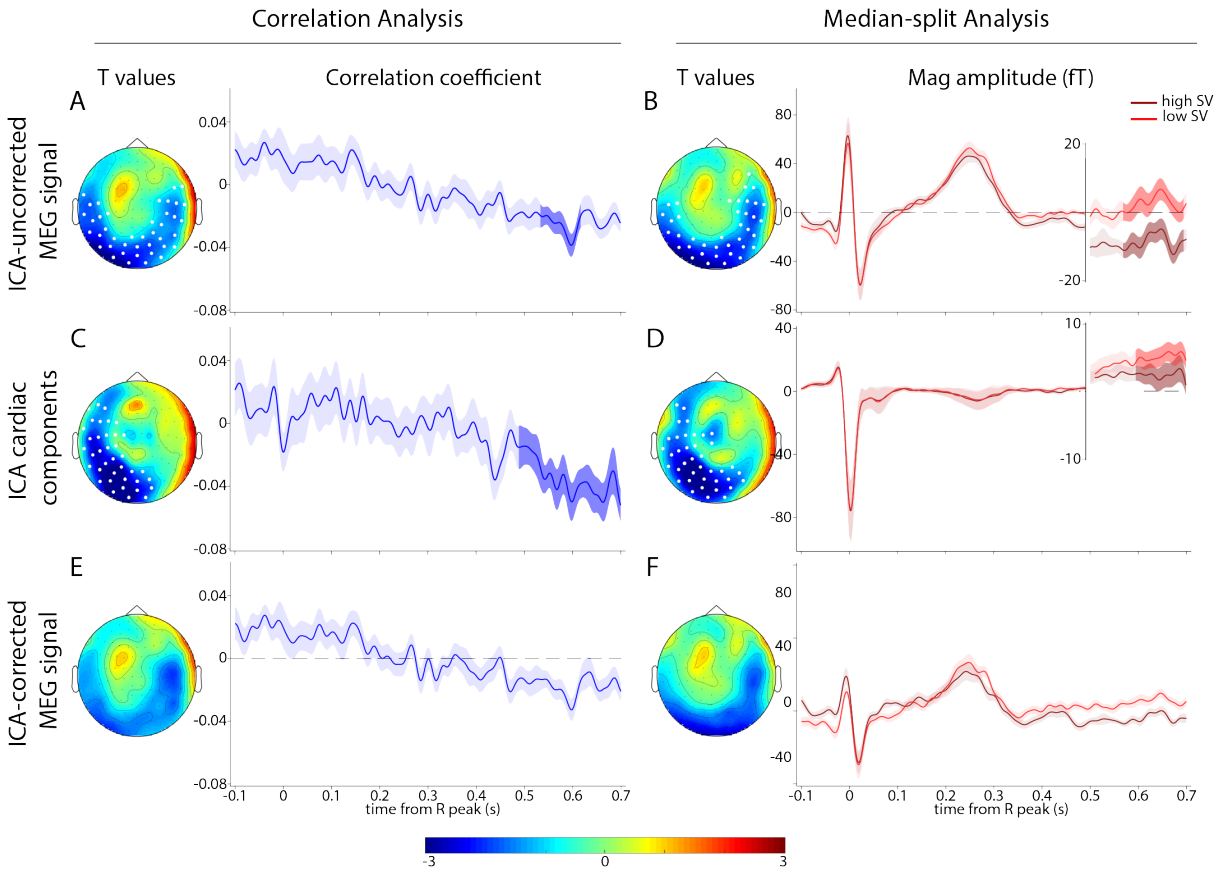
A. Experimental set up used to record the ECG (yellow), the ICG (impedance cardiography, green; ICG in, current injection; ICG out, measurement), and the cardiac sound (blue) **B.** The different signals used in impedance cardiography, with the ECG, time-varying impedance Z and its first derivative dZ/dt, and cardiac sound. The Kubicek's formula (see Material and Methods) used to compute the stroke volume at each cardiac cycle requires to identify the left ventricular ejection time, marked in gray. **C.** Methodology to compute LVET using the ECG, the ICG and the cardiac sound. **Right upper panel**: B point (LVET onset) determination. The B point marks ejection onset and is defined as the time point when the 1st derivative of the impedance reaches 15% of its maximum in the time interval between Q and Q+400ms. **Right lower panel**: X point (LVET offset) determination. The X point marks the end of the ejection and is defined as the time point when the 1st derivative of the impedance reaches a minimum in the time interval [B to B+400ms].

Figure 2. Cardiac parameters: Stroke volume and IBI



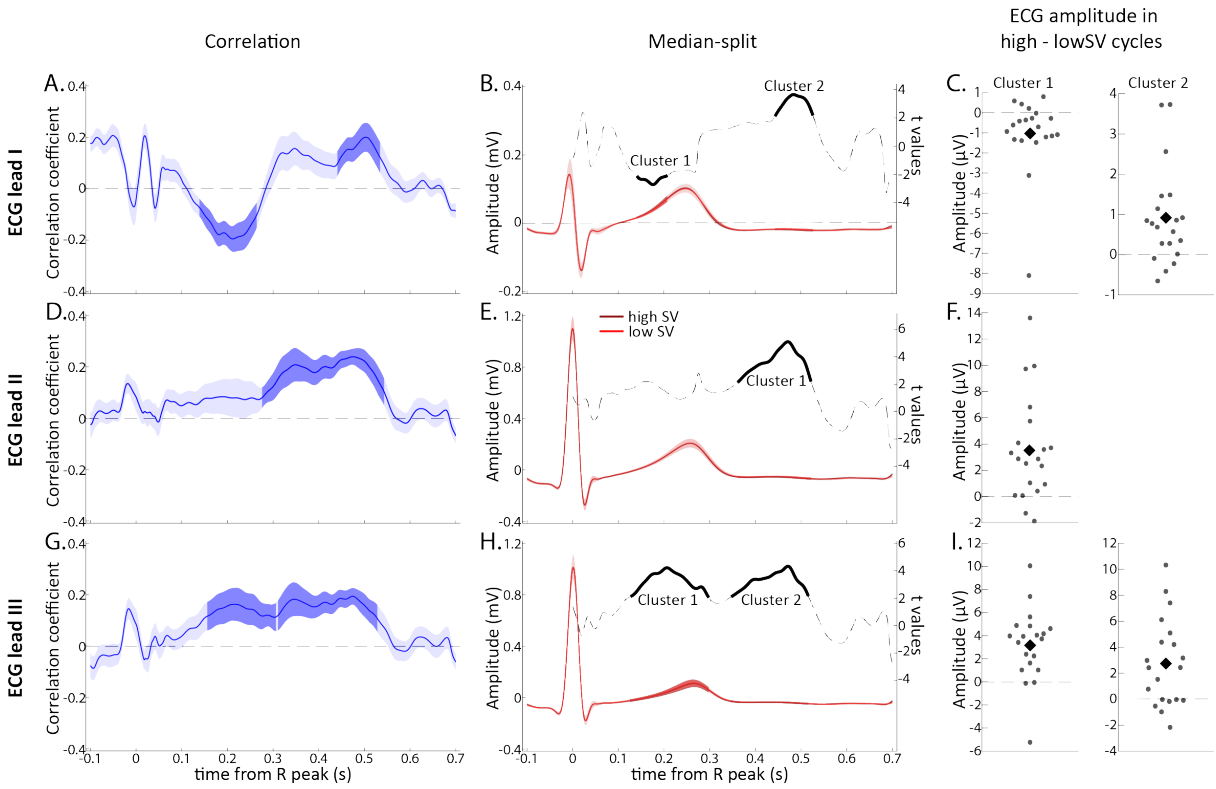
A. Stroke volume distribution in one subject. **B.** Correlation between IBI and stroke volume in the same subject (R^2=0.0881, $P<10^{-3}$). **C**. Correlation coefficients between IBI and stroke volume in each subject (gray dots) and mean across participants (black diamond) (mean±sd = 0.28±0.16, One sample t test on Fischer-transformed correlation coefficients against 0: t(20)=7.61, $P<10^{-3}$). **D.** Mean IBI in low and high-SV cardiac cycles in each subject (gray dots and thin lines) and average across participants (black diamonds and thick line) (median-split based on stroke volume; mean IBI in low-SV cardiac cycles ±se = 894±27ms, mean IBI in high-SV cardiac cycles ±se = 923±29ms, Paired t test: t(20)=7.37, $P<10^{-3}$). *** indicates $P<10^{-3}$.

Figure 3. Correlation and median-split analysis of the MEG data



A. Correlation analysis between HERs computed from ICA-uncorrected MEG signal and SV, grand average across participants. The topography of the cluster with a significant correlation is shown on the left, with significant channels highlighted in white. The time course of the correlation is shown on the right, with significant points in bold. **B.** Comparison between low- and high-SV HERs computed from ICA-uncorrected MEG signal. The topography of the cluster with a significant difference is shown on the left, with significant channels highlighted in white. The timecourses of the HERs are shown on the right, with significant points in bold. **C and D** represents the same analysis performed on the cardiac ICA components. **E and F** represents the same analysis performed on the ICA-corrected MEG-signal.

Figure 4. Correlation and median-split analysis of ECGs



A, D, C. Time-course of the correlation coefficient between the stroke volume and ECG data (A, ECG lead I; D, ECG lead II; G, ECG lead III), grand average across participants. Significant clusters are indicated in bold . **B, E, H.** ECG data locked on the R peak, grand-averaged across participants, in low (light red) and high (dark red) SV trials, and t-values (black) of the comparison between the two. Significant clusters are indicated in bold . **C, F, I.** Difference in ECG amplitude in high – low-SV trials for each subject (gray dots), in each cluster, and grand average across participants (black diamond).

References

Abitbol R, Lebreton M, Hollard G, Richmond BJ, Bouret S, Pessiglione M (2015) Neural Mechanisms Underlying Contextual Dependency of Subjective Values: Converging Evidence from Monkeys and Humans. J Neurosci 35:2308–2320.

Abraham A, Schubotz RI, von Cramon DY (2008) Thinking about the future versus the past in personal and non-personal contexts. Brain Res 1233:106–119.

Adler D, Herbelin B, Similowski T, Blanke O (2014) Breathing and sense of self: Visuo-respiratory conflicts alter body self-consciousness. Respir Physiol Neurobiol 204:131–137.

Ai S, Yin Y, Chen Y, Wang C, Sun Y, Tang X, Lu L, Zhu L, Shi J (2018) Promoting subjective preferences in simple economic choices during nap. Elife:1–21.

Al E, Iliopoulos F, Forschack N, Nierhaus T, Grund M, Motyka P, Gaebler M, Nikulin V V., Villringer A (2020) Heart-brain interactions shape somatosensory perception and evoked potentials. Proc Natl Acad Sci U S A 117:10575–10584.

Allard E, Canzoneri E, Adler D, Morélot-Panzini C, Bello-Ruiz J, Herbelin B, Blanke O, Similowski T (2017) Interferences between breathing, experimental dyspnoea and bodily self-consciousness. Sci Rep 7:1–11.

Amanzio M, Torta DME, Sacco K, Cauda F, D'Agata F, Duca S, Leotta D, Palermo S, Geminiani GC (2011) Unawareness of deficits in Alzheimer's disease: Role of the cingulate cortex. Brain 134:1061–1076.

Amassian VE (1951) CORTICAL REPRESENTATION OF VISCERAL AFFERENTS. J Neurophysiol 14:433–444.

Aspell JE, 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. Psychol Sci 24:2445–2453.

Azzalini D, Buot A, Palminteri S, Tallon-Baudry C (2020) Responses to heartbeats in ventromedial prefrontal cortex contribute to subjective preference-based decisions. bioRxiv:776047.

Azzalini D, Rebollo I, Tallon-Baudry C (2019) Visceral Signals Shape Brain Dynamics and

Cognition. Trends Cogn Sci 23:488–509.

Babo-Rebelo M, Buot A, Tallon-Baudry C (2019) Neural responses to heartbeats distinguish self from other during imagination. Neuroimage 191:10–20.

Babo-Rebelo M, Richter CG, Tallon-Baudry C (2016a) Neural Responses to Heartbeats in the Default Network Encode the Self in Spontaneous Thoughts. J Neurosci 36:7829–7840.

Babo-Rebelo M, Wolpert N, Adam C, Hasboun D, Tallon-Baudry C (2016b) Is the cardiac monitoring function related to the self in both the default network and right anterior insula? Philos Trans R Soc B Biol Sci 371:20160004.

Bakkour A, Palombo DJ, Zylberberg A, Kang YH, Reid A, Verfaellie M, Shadlen MN, Shohamy D (2019) The hippocampus supports deliberation during value-based decisions. Elife 8:1–28.

Bang D, Fleming SM (2018) Distinct encoding of decision confidence in human medial prefrontal cortex. Proc Natl Acad Sci 115:6082–6087.

Bartra O, McGuire JT, Kable JW (2013) The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. Neuroimage 76:412–427.

Bechara A, Damasio AR (2005) The somatic marker hypothesis: A neural theory of economic decision. Games Econ Behav 52:336–372.

Becker GM, DeGroot MH, Marschak J (1964) Measuring utility by. Behav Sci:226–232.

Beissner F, Meissner K, Bar K-J, Napadow V (2013) The Autonomic Brain: An Activation Likelihood Estimation Meta-Analysis for Central Processing of Autonomic Function. J Neurosci 33:10503–10511.

Benoit RG, Gilbert SJ, Volle E, Burgess PW (2010) When I think about me and simulate you: Medial rostral prefrontal cortex and self-referential processes. Neuroimage 50:1340–1349.

Benoit RG, Szpunar KK, Schacter DL (2014) Ventromedial prefrontal cortex supports affective future simulation by integrating distributed knowledge. Proc Natl Acad Sci U S A 111:16550–16555.

Bernoulli D (1954) Exposition of a New Theory on the Measurement of Risk. Econometrica 22:23.

Berridge KC (2004) Motivation concepts in behavioral neuroscience. Physiol Behav 81:179–209.

Bicanski A, Burgess N (2018) A neural-level model of spatial memory and imagery. Elife 7.

Birch J, Schnell AK, Clayton NS (2020) Dimensions of Animal Consciousness. Trends Cogn Sci 24:789–801.

Birznieks I, Boonstra TW, Macefield VG (2012) Modulation of human muscle spindle discharge by arterial pulsations - functional effects and consequences. PLoS One 7.

Bishop VS, Malliani A, Thorén P (2011) Cardiac Mechanoreceptors. In: Comprehensive Physiology. Hoboken, NJ, USA: John Wiley & Sons, Inc.

Blanke O, Metzinger T (2009) Full-body illusions and minimal phenomenal selfhood. Trends Cogn Sci 13:7–13.

Blanke O, Slater M, Serino A (2015) Behavioral, Neural, and Computational Principles of Bodily Self-Consciousness. Neuron 88:145–166.

Blood AJ, Zatorre RJ, Bermudez P, Evans AC (1999) Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. Nat Neurosci 2:382–387.

Bornemann B, Singer T (2017) Taking time to feel our body: Steady increases in heartbeat perception accuracy and decreases in alexithymia over 9 months of contemplative mental training. Psychophysiology 54:469–482.

Bouret S, Richmond BJ (2010) Ventromedial and Orbital Prefrontal Neurons Differentially Encode Internally and Externally Driven Motivational Values in Monkeys. J Neurosci 30:8591–8601.

Bowren MD, Croft KE, Reber J, Tranel D (2018) Choosing spouses and houses: Impaired congruence between preference and choice following damage to the ventromedial prefrontal cortex. Neuropsychology 32:280–303.

Brehm JW (1956) Postdecision changes in the desirability of alternatives. J Abnorm Soc Psychol 52:384–389.

Brener J, Ring C (2016) Towards a psychophysics of interoceptive processes: The measurement of heartbeat detection. Philos Trans R Soc B Biol Sci 371.

Brown SD, Heathcote A (2008) The simplest complete model of choice response time: Linear

ballistic accumulation. Cogn Psychol 57:153–178.

Buckner RL, Andrews-Hanna JR, Schacter DL (2008) The brain's default network: Anatomy, function, and relevance to disease. Ann N Y Acad Sci 1124:1–38.

Cai X, Padoa-Schioppa C (2012) Neuronal encoding of subjective value in dorsal and ventral anterior cingulate cortex. J Neurosci 32:3791–3808.

Camerer C, Babcock L, Loewenstein G, Thaler R (1997) Labor Supply of New York City Cabdrivers: One Day at a Time. Q J Econ 112:407–441.

Camille N, Griffiths CA, Vo K, Fellows LK, Kable JW (2011) Ventromedial frontal lobe damage disrupts value maximization in humans. J Neurosci 31:7527–7532.

Canales-Johnson A, Silva C, Huepe D, Rivera-Rei Á, Noreika V, Del Carmen Garcia M, Silva W, Ciraolo C, Vaucheret E, Sedeño L, Couto B, Kargieman L, Baglivo F, Sigman M, Chennu S, Ibáñez A, Rodríguez E, Bekinschtein TA (2015) Auditory feedback differentially modulates behavioral and neural markers of objective and subjective performance when tapping to your heartbeat. Cereb Cortex 25:4490–4503.

Cannon WB (1927) The James-Lange theory of emotions: A critical examination and an alternative theory. Am J Psychol 39:106–124.

Carmichael ST, Price JL (1995) Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. J Comp Neurol 363:642–664.

Cechetto DF, Saper CB (1987) Evidence for a viscerotopic sensory representation in the cortex and thalamus in the rat. J Comp Neurol 262:27–45.

Chammat M, El Karoui I, Allali S, Hagège J, Lehongre K, Hasboun D, Baulac M, Epelbaum S, Michon A, Dubois B, Navarro V, Salti M, Naccache L (2017) Cognitive dissonance resolution depends on episodic memory. Sci Rep 7:1–10.

Chen MK, Risen JL (2010) How Choice Affects and Reflects Preferences: Revisiting the Free-Choice Paradigm. J Pers Soc Psychol 99:573–594.

Chib VS, Rangel A, Shimojo S, O'Doherty JP (2009) Evidence for a Common Representation of Decision Values for Dissimilar Goods in Human Ventromedial Prefrontal Cortex. J Neurosci 29:12315–12320.

Christoff K, Cosmelli D, Legrand D, Thompson E (2011) Specifying the self for cognitive neuroscience. Trends Cogn Sci 15:104–112.

Clithero JA, Rangel A (2014) Informatic parcellation of the network involved in the computation of subjective value. Soc Cogn Affect Neurosci 9:1289–1302.

Couto B, Adolfi F, Velasquez M, Mesow M, Feinstein J, Canales-Johnson A, Mikulan E, Martínez-Pernía D, Bekinschtein T, Sigman M, Manes F, Ibanez A (2015) Heart evoked potential triggers brain responses to natural affective scenes: A preliminary study. Auton Neurosci 193:132–137.

Craig AD (2009) How do you feel — now? The anterior insula and human awareness. Nat Rev Neurosci 10:59–70.

Critchley HD, Harrison NA (2013) Visceral Influences on Brain and Behavior. Neuron 77:624–638.

Critchley HD, Wiens S, Rotshtein P, Öhman A, Dolan RJ (2004) Neural systems supporting interoceptive awareness. Nat Neurosci 7:189–195.

D'Argembeau A (2013) On the role of the ventromedial prefrontal cortex in self-processing: The valuation hypothesis. Front Hum Neurosci 7:372.

Damasio A (2003a) Feelings of Emotion and the Self. Ann N Y Acad Sci 1001:253–261.

Damasio A (2003b) Mental self: The person within. Nature 423:227–227.

De Martino B, Camerer CF, Adolphs R (2010) Amygdala damage eliminates monetary loss aversion. Proc Natl Acad Sci U S A 107:3788–3792.

De Martino B, Fleming SM, Garrett N, Dolan RJ (2013) Confidence in value-based choice. Nat Neurosci 16:105–110.

De Martino B, Kumaran D, Seymour B, Dolan RJ (2006) Frames, biases and rational decision-making in the human brain. Science (80-) 313:684–687.

Denny BT, Kober H, Wager TD, Ochsner KN (2012) A Meta-analysis of Functional Neuroimaging Studies of Self- and Other Judgments Reveals a Spatial Gradient for Mentalizing in Medial Prefrontal Cortex. J Cogn Neurosci 24:1742–1752.

Desender K, Boldt A, Yeung N (2018) Subjective Confidence Predicts Information Seeking in Decision Making. Psychol Sci 29:761–778.

Dirlich G, Dietl T, Vogl L, Strian F (1998) Topography and morphology of heart action-related EEG potentials. Electroencephalogr Clin Neurophysiol - Evoked Potentials 108:299–305.

Dirlich G, Vogl L, Plaschke M, Strian F (1997) Cardiac field effects on the EEG. Electroencephalogr Clin Neurophysiol 102:307–315.

Dixon ML, Thiruchselvam R, Todd R, Christoff K (2017) Emotion and the prefrontal cortex: An integrative review. Psychol Bull 143:1033–1081.

Drugowitsch J, Wyart V, Devauchelle AD, Koechlin E (2016) Computational Precision of Mental Inference as Critical Source of Human Choice Suboptimality. Neuron 92:1398–1411.

Dum RP, Levinthal DJ, Strick PL (2009) The spinothalamic system targets motor and sensory areas in the cerebral cortex of monkeys. J Neurosci 29:14223–14235.

Dworkin BR, Elbert T, Rau H, Birbaumer N, Pauli P, Droste C, Brunia CH (1994) Central effects of baroreceptor activation in humans: attenuation of skeletal reflexes and pain perception. Proc Natl Acad Sci 91:6329–6333.

Edwards L, Ring C, McIntyre D, Winer JB, Martin U (2009) Sensory detection thresholds are modulated across the cardiac cycle: Evidence that cutaneous sensibility is greatest for systolic stimulation. Psychophysiology 46:252–256.

Egan LC, Bloom P, Santos LR, Egan LC (2010) Choice-induced preferences in the absence of choice: Evidence from a blind two choice paradigm with young children and capuchin monkeys. J Exp Soc Psychol 46:204–207.

Elliott R, Graf V (1972) Visual sensitivity as a function of phase of cardiac cycle. Psychophysiology 9:357–361.

Fellows LK, Farah MJ (2007) The role of ventromedial prefrontal cortex in decision making: Judgment under uncertainty or judgment per se? Cereb Cortex 17:2669–2674.

Findling C, Skvortsova V, Dromnelle R, Palminteri S, Wyart V (2019) Computational noise in reward-guided learning drives behavioral variability in volatile environments. Nat Neurosci 22:2066–2077.

Fleming SM (2020) Awareness as inference in a higher-order state space. Neurosci Conscious 2020:1–9.

Fukushima H, Terasawa Y, Umeda S (2011) Association between interoception and empathy: Evidence from heartbeat-evoked brain potential. Int J Psychophysiol 79:259–265.

Garfinkel SN, Critchley HD (2016) Threat and the Body: How the Heart Supports Fear

Processing. Trends Cogn Sci 20:34–46.

Garfinkel SN, Minati L, Gray MA, Seth AK, Dolan RJ, Critchley HD (2014) Fear from the Heart: Sensitivity to Fear Stimuli Depends on Individual Heartbeats. J Neurosci 34:6573–6582.

Gentsch A, Sel A, Marshall AC, Schütz-Bosbach S (2018) Affective interoceptive inference: Evidence from heart-beat evoked brain potentials. Hum Brain Mapp.

Goldstein RZ, Alia-Klein N, Tomasi D, Carrillo JH, Maloney T, Woicik PA, Wang R, Telang F, Volkow ND (2009) Anterior cingulate cortex hypoactivations to an emotionally salient task in cocaine addiction. Proc Natl Acad Sci U S A 106:9453–9458.

Gray MA, Beacher FD, Minati L, Nagai Y, Kemp AH, Harrison NA, Critchley HD (2012) Emotional appraisal is influenced by cardiac afferent information. Emotion 12:180–191.

Gray MA, Taggart P, Sutton PM, Groves D, Holdright DR, Bradbury D, Brull D, Critchley HD (2007) A cortical potential reflecting cardiac function. PNAS 104:6818–6823.

Grueschow M, Polania R, Hare TA, Ruff CC (2015) Automatic versus Choice-Dependent Value Representations in the Human Brain. Neuron 85:874–885.

Hansson SO, Grüne-Yanoff T (2018) Preferences. In: The Stanford Encyclopedia of Philosophy, Summer 201. (Zalta EN, ed). Metaphysics Research Lab, Stanford University.

Hare TA, Schultz W, Camerer CF, O'Doherty JP, Rangel A (2011) Transformation of stimulus value signals into motor commands during simple choice. Proc Natl Acad Sci U S A 108:18120–18125.

Heeger DJ (1992) Normalization of cell responses in cat striate cortex. Vis Neurosci 9:181–197.

Heilbronner SR, Hayden BY (2016) Dorsal Anterior Cingulate Cortex: A Bottom-Up View. Annu Rev Neurosci 39:149–170.

Hey JD, Orme C (1994) Investigating generalizations of expected utility theory using experimental data. Econometrica 62:1291–1326.

Hornsby AN, Love BC (2020) How decisions and the desire for coherency shape subjective preferences over time. Cognition 200:104244.

Huber J, Payne JW, Puto C (1982) Adding asymmetrically dominated alternatives: Violations

of regularity and the similarity hypothesis. J Consum Res 9:90–98.

Hull CL (1943) Principles of behavior. Appleton-century-crofts New York.

Hunt LT, Behrens TEJ, Hosokawa T, Wallis JD, Kennerley SW (2015) Capturing the temporal evolution of choice across prefrontal cortex. Elife 4:1–25.

Hunt LT, Hayden BY (2017) A distributed, hierarchical and recurrent framework for reward-based choice. Nat Rev Neurosci 18:172–182.

Hunt LT, Kolling N, Soltani A, Woolrich MW, Rushworth MFS, Behrens TEJ (2012) Mechanisms underlying cortical activity during value-guided choice. Nat Neurosci 15:470–476.

Izuma K, Matsumoto M, Murayama K, Samejima K, Sadato N, Matsumoto K (2010) Neural correlates of cognitive dissonance and choice-induced preference change. Proc Natl Acad Sci U S A 107:22014–22019.

James W (1890) The Principles of Psychology. In. Henry Holt.

Jenkins AC, Mitchell JP (2011) Medial prefrontal cortex subserves diverse forms of self-reflection. Soc Neurosci 6:211–218.

Joshi S, Gold JI (2020) Pupil Size as a Window on Neural Substrates of Cognition. Trends Cogn Sci 24:466–480.

Juechems K, Balaguer J, Herce Castañón S, Ruz M, O'Reilly JX, Summerfield C (2019) A Network for Computing Value Equilibrium in the Human Medial Prefrontal Cortex. Neuron 101:977-987.e3.

Juechems K, Balaguer J, Ruz M, Summerfield C (2017) Ventromedial Prefrontal Cortex Encodes a Latent Estimate of Cumulative Reward. Neuron 93:705-714.e4.

Juechems K, Summerfield C (2019) Where does value come from? Trends Cogn Sci:836–850.

Kable JW, Glimcher PW (2007) The neural correlates of subjective value during intertemporal choice. Nat Neurosci 10:1625–1633.

Kandel ER, Schwartz JH, Jessell TM, of Biochemistry D, Jessell MBT, Siegelbaum S, Hudspeth AJ (2000) Principles of neural science. McGraw-hill New York.

Keramati M, Gutkin B (2014) Homeostatic reinforcement learning for integrating reward

collection and physiological stability. Elife 3:1–26.

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.

Khalsa SS et al. (2018) Interoception and Mental Health: A Roadmap. Biol Psychiatry Cogn Neurosci Neuroimaging 3:501–513.

Kim J, Park HD, Kim KW, Shin DW, Lim S, Kwon H, Kim MY, Kim K, Jeong B (2019) Sad faces increase the heartbeat-associated interoceptive information flow within the salience network: a MEG study. Sci Rep 9:1–15.

Kim KJ, Diaz JR, Iddings JA, Filosa JA (2016) Vasculo-neuronal coupling: Retrograde vascular communication to brain neurons. J Neurosci 36:12624–12639.

Kolling N, Behrens TEJ, Mars RB, Rushworth MFS (2012) Neural mechanisms of foraging. Science (80-) 335:95–98.

Krajbich I, Armel C, Rangel A (2010) Visual fixations and the computation and comparison of value in simple choice. Nat Neurosci 13:1292–1298.

Kühn S, Gallinat J (2011) Common biology of craving across legal and illegal drugs - a quantitative meta-analysis of cue-reactivity brain response. Eur J Neurosci 33:1318–1326.

Kurtz-David V, Persitz D, Webb R, Levy DJ (2019) The neural computation of inconsistent choice behavior. Nat Commun 10:1583.

Lacey JI (1967) Somatic response patterning and stress: Some revisions of activaction theory. In: Psychological stress: Issues in research, Appelton-C. (Appley MH, Trumbull R, eds), pp 14–42. New York.

Lang PJ, Davis M (2006) Chapter 1 Emotion, motivation, and the brain: Reflex foundations in animal and human research. Prog Brain Res 156:3–29.

Latty T, Beekman M (2011) Irrational decision-making in an amoeboid organism: transitivity and context-dependent preferences. Proc R Soc B Biol Sci 278:307–312.

Lebreton M, Abitbol R, Daunizeau J, Pessiglione M (2015) Automatic integration of confidence in the brain valuation signal. Nat Neurosci 18.

Lebreton M, Jorge S, Michel V, Thirion B, Pessiglione M (2009) An Automatic Valuation

System in the Human Brain: Evidence from Functional Neuroimaging. Neuron 64:431–439.

LeDoux JE, Hofmann SG (2018) The subjective experience of emotion: a fearful view. Curr Opin Behav Sci 19:67–72.

Levy DJ, Glimcher PW (2011) Comparing Apples and Oranges: Using Reward-Specific and Reward-General Subjective Value Representation in the Brain. J Neurosci 31:14693–14707.

Levy DJ, Glimcher PW (2012) The root of all value: A neural common currency for choice. Curr Opin Neurobiol 22:1027–1038.

Lichtenstein S, Slovic P (2006) The Construction of Preference (Lichtenstein S, Slovic P, eds). Cambridge University Press.

Litt A, Plassmann H, Shiv B, Rangel A (2011) Dissociating valuation and saliency signals during decision-making. Cereb Cortex 21:95–102.

Lockwood PL, Wittmann MK, Apps MAJ, Klein-Flügge MC, Crockett MJ, Humphreys GW, Rushworth MFS (2018) Neural mechanisms for learning self and other ownership. Nat Commun 9.

Loewenstein G (1996) Out of control: Visceral influences on behavior. Organ Behav Hum Decis Process 65:272–292.

Loomes G (2005) Modelling the stochastic component of behaviour in experiments: Some issues for the interpretation of data. Exp Econ 8:301–323.

Lopez-Persem A, Bastin J, Petton M, Abitbol R, Lehongre K, Adam C, Navarro V, Rheims S, Kahane P, Domenech P, Pessiglione M (2020) Four core properties of the human brain valuation system demonstrated in intracranial signals. Nat Neurosci 23:664–675.

Lopez-Persem A, Rigoux L, Bourgeois-Gironde S, Daunizeau J, Pessiglione M (2017) Choose, rate or squeeze: Comparison of economic value functions elicited by different behavioral tasks. PLoS Comput Biol 13:1–18.

Luce RD (1959a) On the possible psychophysical laws. Psychol Rev 66:81–95.

Luce RD (1959b) Individual Choice Behaviour. Hoboken, NJ: John Wiley.

Luft CDB, Bhattacharya J (2015) Aroused with heart: Modulation of heartbeat evoked potential by arousal induction and its oscillatory correlates. Sci Rep 5:15717.

Macefield VG (2003) Cardiovascular and respiratory modulation of tactile afferents in the human finger pad Translation and Integration Experimental Physiology : The human finger pad is highly vascularized so it might be expected that the on-going cardiac pulsations in the. Exp Physiol 88.

Mackey S, Petrides M (2014) Architecture and morphology of the human ventromedial prefrontal cortex. Eur J Neurosci 40:2777–2796.

Marshall AC, Gentsch A, Jelinčić V, Schütz-Bosbach S (2017) Exteroceptive expectations modulate interoceptive processing: repetition-suppression effects for visual and heartbeat evoked potentials. Sci Rep 7:16525.

Marshall AC, Gentsch A, Schröder L, Schütz-Bosbach S (2018) Cardiac interoceptive learning is modulated by emotional valence perceived from facial expressions. Soc Cogn Affect Neurosci.

Martinelli P, Sperduti M, Piolino P (2013) Neural substrates of the self-memory system: New insights from a meta-analysis. Hum Brain Mapp 34:1515–1529.

Mazor M, Friston K, Fleming S (2020) Distinct neural contributions to metacognition for detecting, but not discriminating visual stimuli. Elife 9:1–34.

McFadden D (2001) Economic choices. Am Econ Rev 91:351–378.

McNamee D, Rangel A, O'Doherty JP (2013) Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex. Nat Neurosci 16:479–485.

Meer L Van Der, Vos AE De, Stiekema APM, Pijnenborg GHM, Tol M Van, Nolen WA, David AS, Aleman A (2012) Insight in Schizophrenia : Involvement of Self-Reflection Networks ? :1–11.

Moeller SJ, Goldstein RZ (2014) Impaired self-awareness in human addiction: Deficient attribution of personal relevance. Trends Cogn Sci 18:635–641.

Moeller SJ, Konova AB, Parvaz MA, Tomasi D, Lane RD, Fort C, Goldstein RZ (2014) Functional, structural, and emotional correlates of impaired insight in cocaine addiction. JAMA Psychiatry 71:61–70.

Morales J, Lau H, Fleming SM (2018) Domain-General and Domain-Specific Patterns of Activity Supporting Metacognition in Human Prefrontal Cortex. J Neurosci 38:3534–3546.

Morgenstern O, Von Neumann J (1953) Theory of games and economic behavior. Princeton university press.

Neubert F-X, Mars RB, Sallet J, Rushworth MFS (2015) Connectivity reveals relationship of brain areas for reward-guided learning and decision making in human and monkey frontal cortex. Proc Natl Acad Sci U S A 112:E2695-704.

Noonan MP, Walton ME, Behrens TEJ, Sallet J, Buckley MJ, Rushworth MFS (2010) Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. Proc Natl Acad Sci U S A 107:20547–20552.

O'Doherty JP (2004) Reward representations and reward-related learning in the human brain: Insights from neuroimaging. Curr Opin Neurobiol 14:769–776.

O'Doherty JP (2014) The problem with value. Neurosci Biobehav Rev 43:259–268.

Ojemann JG, Akbudak E, Snyder AZ, McKinstry RC, Raichle ME, Conturo TE (1997) Anatomic localization and quantitative analysis of gradient refocused echo-planar fMRI susceptibility artifacts. Neuroimage 6:156–167.

Padoa-Schioppa C (2013) Neuronal origins of choice variability in economic decisions. Neuron 80:1322–1336.

Padoa-Schioppa C, Assad JA (2006) Neurons in the orbitofrontal cortex encode economic value. Nature 441:223–226.

Palminteri S, Khamassi M, Joffily M, Coricelli G (2015) Contextual modulation of value signals in reward and punishment learning. Nat Commun 6:8096.

Palminteri S, Lefebvre G, Kilford EJ, Blakemore SJ (2017) Confirmation bias in human reinforcement learning: Evidence from counterfactual feedback processing. PLoS Comput Biol 13:e1005684.

Park H-D, Bernasconi F, Bello-Ruiz J, Pfeiffer C, Salomon R, Blanke O (2016) Transient Modulations of Neural Responses to Heartbeats Covary with Bodily Self-Consciousness. J Neurosci 36:8453–8460.

Park H-D, Bernasconi F, Salomon R, Tallon-Baudry C, Spinelli L, Seeck M, Schaller K, Blanke O (2017) Neural Sources and Underlying Mechanisms of Neural Responses to Heartbeats, and their Role in Bodily Self-consciousness: An Intracranial EEG Study. Cereb Cortex 28:1–14.

Park H-D, Correia S, Ducorps A, Tallon-Baudry C (2014) Spontaneous fluctuations in neural responses to heartbeats predict visual detection. Nat Neurosci 17:612–618.

Park H-D, Tallon-Baudry C (2014) The neural subjective frame: from bodily signals to perceptual consciousness. Philos Trans R Soc Lond B Biol Sci 369:20130208.

Park HD, Blanke O (2019a) Coupling Inner and Outer Body for Self-Consciousness. Trends Cogn Sci 23:377–388.

Park HD, Blanke O (2019b) Heartbeat-evoked cortical responses: Underlying mechanisms, functional roles, and methodological considerations. Neuroimage 197:502–511.

Paulus MP (2007) Decision-Making Dysfunctions in Psychiatry—Altered Homeostatic Processing? Science (80-) 318:602–606.

Pelletier G, Fellows LK (2019) A critical role for human ventromedial frontal lobe in value comparison of complex objects based on attribute configuration. J Neurosci 39:2969–18.

Pessiglione M, Schmidt L, Draganski B, Kalisch R, Lau H, Dolan RJ, Frith CD (2007) How the brain translates money into force: A neuroimaging study of subliminal motivation. Science (80-) 316:904–906.

Petzschner FH, Weber LA, Wellstein K V., Paolini G, Do CT, Stephan KE (2019) Focus of attention modulates the heartbeat evoked potential. Neuroimage 186:595–606.

Petzschner FH, Weber LAE, Gard T, Stephan KE (2017) Computational Psychosomatics and Computational Psychiatry: Toward a Joint Framework for Differential Diagnosis. Biol Psychiatry 82:421–430.

Plassmann H, O'Doherty J, Rangel A (2007) Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. J Neurosci 27:9984–9988.

Plassmann H, O'Doherty J, Shiv B, Rangel A (2008) Marketing actions can modulate neural representations of experienced pleasantness. Proc Natl Acad Sci U S A 105:1050–1054.

Polanía R, Krajbich I, Grueschow M, Ruff CC (2014) Neural Oscillations and Synchronization Differentially Support Evidence Accumulation in Perceptual and Value-Based Decision Making. Neuron 82:709–720.

Polanía R, Moisa M, Opitz A, Grueschow M, Ruff CC (2015) The precision of value-based choices depends causally on fronto-parietal phase coupling. Nat Commun 6.

Polanía R, Woodford M, Ruff C (2019) Efficient coding of subjective value. Nat Neurosci

22:134–142.

Pollatos O, Kirsch W, Schandry R (2005) Brain structures involved in interoceptive awareness and cardioafferent signal processing: A dipole source localization study. Hum Brain Mapp 26:54–64.

Pollatos O, Schandry R (2004) Accuracy of heartbeat perception is reflected in the amplitude of the heartbeat-evoked brain potential. Psychophysiology 41:476–482.

Pramme L, Larra MF, Schächinger H, Frings C (2014) Cardiac cycle time effects on mask inhibition. Biol Psychol 100:115–121.

Pramme L, Larra MF, Schächinger H, Frings C (2016) Cardiac cycle time effects on selection efficiency in vision. Psychophysiology 53:1702–1711.

Qin P, Northoff G (2011) How is our self related to midline regions and the default-mode network? Neuroimage 57:1221–1233.

Rangel A (2013) Regulation of dietary choice by the decision-making circuitry. Nat Neurosci 16:1717–1724.

Rangel A, Camerer C, Montague PR (2008) A framework for studying the neurobiology of value-based decision making. Nat Rev Neurosci 9:545–556.

Rangel A, Clithero JA (2012) Value normalization in decision making: Theory and evidence. Curr Opin Neurobiol 22:970–981.

Rangel A, Clithero JA (2013) The Computation of Stimulus Values in Simple Choice. In: Neuroeconomics: Decision Making and the Brain: Second Edition, pp 125–148. Elsevier Inc.

Ratcliff R (1978) A theory of memory retrieval. Psychol Rev 85:59.

Ratcliff R, Smith PL, Brown SD, McKoon G (2016) Diffusion Decision Model: Current Issues and History. Trends Cogn Sci 20:260–281.

Resulaj A, Kiani R, Wolpert DM, Shadlen MN (2009) Changes of mind in decision-making. Nature 461:263–266.

Rich EL, Wallis JD (2016) Decoding subjective decisions from orbitofrontal cortex. Nat Neurosci 19.

Ring C, Brener J (2018) Heartbeat counting is unrelated to heartbeat detection: A comparison

of methods to quantify interoception. Psychophysiology 55:1–10.

Ring C, Brener J, Knapp K, Mailloux J (2015) Effects of heartbeat feedback on beliefs about heart rate and heartbeat counting: A cautionary tale about interoceptive awareness. Biol Psychol 104:193–198.

Rolls ET, Grabenhorst F, Parris BA (2008) Warm pleasant feelings in the brain. Neuroimage 41:1504–1513.

Rouault M, Dayan P, Fleming SM (2019) Forming global estimates of self-performance from local confidence. Nat Commun 10.

Roy M, Shohamy D, Wager TD (2012) Ventromedial prefrontal-subcortical systems and the generation of affective meaning. Trends Cogn Sci 16:147–156.

Rushworth MFS, Kolling N, Sallet J, Mars RB (2012) Valuation and decision-making in frontal cortex: one or many serial or parallel systems? Curr Opin Neurobiol 22:946–955.

Salomon R, Ronchi R, Dönz J, Bello-Ruiz J, Herbelin B, Martet R, Faivre N, Schaller K, Blanke O (2016) The Insula Mediates Access to Awareness of Visual Stimuli Presented Synchronously to the Heartbeat. J Neurosci 36:5115–5127.

Saper CB (2002) The Central Autonomic Nervous System: Conscious Visceral Perception and Autonomic Pattern Generation. Annu Rev Neurosci 25:433–469.

Sass LA, Parnas J (2003) Schizophrenia, Consciousness and the Self. Schizophr Bull 29:427–444.

Schandry R, Montoya P (1996) Event-related brain potentials and the processing of cardiac activity. Biol Psychol 42:75–85.

Schandry R, Sparrer B, Weitkunat R (1986) From the heart to the brain: A study of heartbeat contingent scalp potentials. Int J Neurosci 30:261–275.

Schandry R, Weitkunat R (1990) Enhancement of Heartbeat-related Brain Potentials Through Cardiac Awarness Training. Int J Neurosci 53:243–253.

Schoenbaum G, Takahashi Y, Liu TL, Mcdannald MA (2011) Does the orbitofrontal cortex signal value? Ann N Y Acad Sci 1239:87–99.

Schuck NW, Cai MB, Wilson RC, Niv Y (2016) Human Orbitofrontal Cortex Represents a Cognitive Map of State Space. Neuron 91:1402–1412.

Schultz W, Apicella P, Scarnati E, Ljungberg T (1992) Neuronal activity in monkey ventral striatum related to the expectation of reward. J Neurosci 12:4595–4610.

Sel A, Azevedo RT, Tsakiris M (2017) Heartfelt Self: Cardio-Visual Integration Affects Self-Face Recognition and Interoceptive Cortical Processing. Cereb Cortex 27:5144–5155.

Shafir S, Waite TA, Smith BH (2002) Context-dependent violations of rational choice in honeybees (Apis mellifera) and gray jays (Perisoreus canadensis). Behav Ecol Sociobiol 51:180–187.

Sharot T, Martino B De, Dolan RJ (2009) How choice reveals and shapes expected hedonic outcome. J Neurosci 29:3760–3765.

Sharot T, Velasquez CM, Dolan RJ (2010) Do decisions shape preference? Evidence from blind choice. Psychol Sci 21:1231–1235.

Shenhav A, Botvinick MM, Cohen JD (2013) The expected value of control: An integrative theory of anterior cingulate cortex function. Neuron 79:217–240.

Shenhav A, Straccia MA, Cohen JD, Botvinick MM (2014) Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value. Nat Neurosci 17:1249–1254.

Shoemaker JK, Wong SW, Cechetto DF (2012) Cortical Circuitry Associated With Reflex Cardiovascular Control in Humans: Does the Cortical Autonomic Network “Speak” or “Listen” During Cardiovascular Arousal. Anat Rec 295:1375–1384.

Sierra M, David AS (2011) Depersonalization: A selective impairment of self-awareness. Conscious Cogn 20:99–108.

Simonson I, Tversky A (1992) Choice in context: Tradeoff contrast and extremeness aversion. J Mark Res 29:281–295.

Soltani A, de Martino B, Camerer C (2012) A range-normalization model of context-dependent choice: A new model and evidence. PLoS Comput Biol 8.

Stalnaker TA, Cooch NK, Schoenbaum G (2015) What the orbitofrontal cortex does not do. Nat Neurosci 18:620–627.

Stein BE, Stanford TR (2008) Multisensory integration: Current issues from the perspective of the single neuron. Nat Rev Neurosci 9:255–266.

Stephan KE, Manjaly ZM, Mathys CD, Weber LA, Paliwal S, Gard T, Tittgemeyer M,

Fleming SM, Haker H, Seth AK, Petzschner FH (2016) Allostatic Self-efficacy: A Metacognitive Theory of Dyshomeostasis-Induced Fatigue and Depression. Front Hum Neurosci 10.

Strait CE, Sleezer BJ, Hayden BY (2015) Signatures of value comparison in ventral striatum neurons. PLoS Biol 13:1–22.

Strohminger N, Nichols S (2015) Neurodegeneration and identity. Psychol Sci 26:1469–1479.

Sui J, Enock F, Ralph J, Humphreys GW (2015) Dissociating hyper and hypoself biases to a core self-representation. Cortex 70:202–212.

Sui J, He X, Humphreys GW (2012) Perceptual effects of social salience: Evidence from self-prioritization effects on perceptual matching. J Exp Psychol Hum Percept Perform 38:1105–1117.

Sui J, Rotshtein P, Humphreys GW (2013) Coupling social attention to the self forms a network for personal significance. Proc Natl Acad Sci U S A 110:7607–7612.

Suzuki K, Garfinkel SN, Critchley HD, Seth AK (2013) Multisensory integration across exteroceptive and interoceptive domains modulates self-experience in the rubber-hand illusion. Neuropsychologia 51:2909–2917.

Suzuki S, Cross L, O'Doherty JP (2017) Elucidating the underlying components of food valuation in the human orbitofrontal cortex. Nat Neurosci 20:1780–1786.

Tallon-Baudry C, Campana F, Park H-D, Babo-Rebelo M (2018) The neural monitoring of visceral inputs, rather than attention, accounts for first-person perspective in conscious vision. Cortex 102:139–149.

Tanner Jr WP, Swets JA (1954) A decision-making theory of visual detection. Psychol Rev 61:401.

Thayer JF, Åhs F, Fredrikson M, Sollers JJ, Wager TD (2012) A meta-analysis of heart rate variability and neuroimaging studies: Implications for heart rate variability as a marker of stress and health. Neurosci Biobehav Rev 36:747–756.

Tom SM, Fox CR, Trepel C, Poldrack RA (2007) The neural basis of loss aversion in decision-making under risk. Science (80-) 315:515–518.

Tversky A, Kahneman D (1981) The Framing of Decisions and the Psychology of Choice. Science (80-) 211:453–458.

Usher M, McClelland JL (2001) The time course of perceptual choice: the leaky, competing accumulator model. Psychol Rev 108:550.

Vaccaro AG, Fleming SM (2018) Thinking about thinking: A coordinate-based meta-analysis of neuroimaging studies of metacognitive judgements. Brain Neurosci Adv 2:1–14.

Vaidya AR, Fellows LK (2020) Under construction: ventral and lateral frontal lobe contributions to value-based decision-making and learning. F1000Research 9:1–8.

van der Meer L, Costafreda S, Aleman A, David AS (2010) Self-reflection and the brain: A theoretical review and meta-analysis of neuroimaging studies with implications for schizophrenia. Neurosci Biobehav Rev 34:935–946.

van der Meer L, de Vos AE, Stiekema APM, Pijnenborg GHM, van Tol M-J, Nolen WA, David AS, Aleman A (2013) Insight in Schizophrenia: Involvement of Self-Reflection Networks? Schizophr Bull 39:1288–1295.

Vanhaudenhuyse A, Demertzi A, Schabus M, Noirhomme Q, Bredart S, Boly M, Phillips C, Soddu A, Luxen A, Moonen G, Laureys S (2010) Vanhaudenhuyse et al. - 2010 - Two Distinct Neuronal Networks Mediate the Awareness of Environment and of Self.pdf. :570–578.

Vickers D (1970) Evidence for an accumulator model of psychophysical discrimination. Ergonomics 13:37–58.

Vigário R, Särelä J, Jousmäki V, Hämäläinen M, Oja E (2000) Independent component approach to the analysis of EEG and MEG recordings. IEEE Trans Biomed Eng 47:589–593.

Vinckier F, Rigoux L, Kurniawan IT, Hu C, Bourgeois-Gironde S, Daunizeau J, Pessiglione M (2019) Sour grapes and sweet victories: How actions shape preferences. PLoS Comput Biol 15:1–24.

Vinckier F, Rigoux L, Oudiette D, Pessiglione M (2018) Neuro-computational account of how mood fluctuations arise and affect decision making. Nat Commun 9.

Vogt BA, Derbyshire SWG (2009) Visceral circuits and cingulate-mediated autonomic functions. Cingulate Neurobiol Dis:219–236.

Vöröslakos M, Takeuchi Y, Brinyiczki K, Zombori T, Oliva A, Fernández-Ruiz A, Kozák G, Kincses ZT, Iványi B, Buzsáki G, Berényi A (2018) Direct effects of transcranial electric

stimulation on brain circuits in rats and humans. Nat Commun 9.

Wagenmakers E-J, Van Der Maas HLJ, Grasman RPPP (2007) An EZ-diffusion model for response time and accuracy. Psychon Bull Rev 14:3–22.

Wallis JD (2012) Cross-species studies of orbitofrontal cortex and value-based decision-making. Nat Neurosci 15:13–19.

Wang X-J (2002) Probabilistic Decision Making by Slow Reverberation in Cortical Circuits. Neuron 36:955–968.

Wang XJ (2008) Decision Making in Recurrent Neuronal Circuits. Neuron 60:215–234.

Webb R, Glimcher PW, Louie K (2020) The Normalization of Consumer Valuations: Context-Dependent Preferences From Neurobiological Constraints. Manage Sci 79:1–33.

Wilkinson M, McIntyre D, Edwards L (2013) Electrocutaneous pain thresholds are higher during systole than diastole. Biol Psychol 94:71–73.

Wimmer GE, Shohamy D (2012) Preference by Association: How Memory Mechanisms in the Hippocampus Bias Decisions. Science (80-) 338:270–273.

Yoo SBM, Hayden BY (2018) Economic Choice as an Untangling of Options into Actions. Neuron.

Yoo SBM, Hayden BY (2020) The Transition from Evaluation to Selection Involves Neural Subspace Reorganization in Core Reward Regions. Neuron 105:712-724.e4.

Zamariola G, Maurage P, Luminet O, Corneille O (2018) Interoceptive accuracy scores from the heartbeat counting task are problematic: Evidence from simple bivariate correlations. Biol Psychol 137:12–17.

Zhang Z, Fanning J, Ehrlich DB, Chen W, Lee D, Levy I (2017) Distributed neural representation of saliency controlled value and category during anticipation of rewards and punishments. Nat Commun 8.

RÉSUMÉ

Décider selon nos préférences implique d'évaluer subjectivement des options, c'est à dire de déterminer ce que ces options valent pour soi-même. En d'autres termes, il faut mettre en lien chaque option avec une certaine représentation de soi-même. Dans ma thèse, j'ai montré comment le suivi cérébral des signaux physiologiques, la forme plus simple de représentation de soi, contribue à l'évaluation subjective même pour des choix qui ne comportent aucune conséquence physiologique. Dans ma première expérience, j'ai montré comment les réponses neuronales aux battements cardiaques, un index du suivi cortical de l'activité cardiaque, sont liées au processus de référence à soi impliqué pendant les choix basés sur des préférences et comment elles influencent l'évaluation subjective au niveau neural et comportemental. Dans ma deuxième expérience, j'ai montré que les réponses neuronales aux battements cardiaques permettent aussi la mise en place d'un référentiel commun qui rend possible la comparaison d'options caractérisées par des attributs très différents en déterminant la précision des choix. Les changements d'excitabilité corticale, d'état d'éveil et des paramètres cardiaques ne peuvent pas expliquer ces résultats. En conclusion, dans ma thèse, j'ai montré comment deux caractéristiques centrales des décisions selon les préférences, leur nature subjective et l'habileté de comparer des options très différentes, reposent sur le même phénomène neural : la référence à soi à travers les réponses neurales aux battements cardiaques.

MOTS CLÉS

Réponse évoquée aux battements cardiaques, électrophysiologie, subjectivité, prise de décision, cœur.

ABSTRACT

Preference-based decisions entail subjective evaluation, that is, a self-referential process through which we estimate the meaning that an item has to us. In my PhD, I investigated whether the monitoring of physiological signals, the evolutionary most ancient and simplest form of self-representation, supports the self-referential process at play in preference-based decisions - even when the decisions do not bear any direct physiological consequence. In a first experiment, I show that heartbeat-evoked responses (HERs), a cortical signature of the neural monitoring of cardiac inputs, index the self-reflection at play during preference-based decisions. In addition, HER fluctuations predict subjective evaluation both neurally and behaviourally. The results from my second experiment reveal that HERs provide a common signal allowing the coordination of different feature spaces, needed to compare attributes of goods of different natures, a mechanism that further predicts choice precision. All these results could not be trivially accounted for by changes in arousal, cortical states or changes in cardiac parameters. Altogether, in my PhD, I showed how two core features of preference-based decisions, namely their inherent subjective nature and the remarkable ability of choosing between goods of different kinds, are well accounted for by the same underlying phenomenon: the implementation of self-referential processing through neural responses to heartbeats.

KEYWORDS

Heartbeat-evoked responses (HER), electrophysiology, subjectivity, decision-making, heart.