# The Brain Valuation System and its role in decision-making

Alizée Lopez

## To cite this version:

Alizée Lopez. The Brain Valuation System and its role in decision-making. Neurons and Cognition [q-bio.NC]. UPMC - Université Paris 6 Pierre et Marie Curie, 2016. English. NNT : . tel-01522094v1

HAL Id: tel-01522094
https://hal.sorbonne-universite.fr/tel-01522094v1
Submitted on 12 May 2017 (v1), last revised 12 Jun 2017 (v2)

## THESE DE DOCTORAT DE
## L'UNIVERSITE PIERRE ET MARIE CURIE

Spécialité

Neurosciences cognitives

Ecole doctorale Cerveau Cognition Comportement

Présentée par

**Alizée Lopez**

Pour obtenir le grade de

DOCTEUR de l'UNIVERSITÉ PIERRE ET MARIE CURIE

Directeur de recherche : Mathias PESSIGLIONE
Co-directeur de recherche : Sacha BOURGEOIS-GIRONDE

# The Brain Valuation System and its role in decision-making

Présentée et soutenue publiquement le 9 Décembre 2016
Devant un jury formé de :

| | | |
|---|---|---|
| M. Mathias PESSIGLIONE | Université Pierre et Marie Curie | Directeur de Thèse |
| M. Emmanuel PROCYK | Université de Lyon | Rapporteur |
| M. Laurence HUNT | University College London | Rapporteur |
| M. Christian RUFF | Université de Zurich | Examinateur |
| Mme Julie GREZES | Ecole Normale Supérieure | Examinatrice |
| M. Olivier SIGAUD | Université Pierre et Marie Curie | Président du Jury |

# ACKNOWLEDGEMENTS

First, I would like to warmly thank the members of my jury, the referees Emmanuel Procyk and Laurence Hunt, the examiners Julie Grèzes and Christian Ruff and the president Olivier Sigaud for having accepted to evaluate this work. I hope it will meet their expectations.

I also would like to thank the institutions which made this work financially possible: la Direction Générale de l'Armement, du Ministère de la défense, which funded my three first years of PhD and le labex BioPsy, which founded my fourth year.

I am extremely grateful to my supervisor Mathias Pessiglione who gave me a priceless scientific expertise and rigor. Interacting with you was always intellectually stimulating, I truly admire the way you make things sound clear and simple despite their complexity. I consider myself very lucky to have worked with you along this road.

I would also like to thank Sacha Bourgeois-Gironde for his economist point of view on my work and the enjoyable exchanges we had during this PhD.

Obviously, I would like to thank the MBB team, which I saw changing a lot during these years. Thus, in the order of *MBB generations*, thanks to the 'old guys': Stephano, Maël and Florent who will always be the reference of the MBB spirit; my 'contemporaries': Lionel, Vasilisa, Marie, Chiara, Irma, Raphaëlle and of course Bastien who could be a member of all the *generations*, as Raphaël, with whom I really had good and fun time along this thesis and definitely Fabien: your presence has always been comforting, you are humanly and intellectually impressive. I also want to warmly thank the 'new not so new guys': Delphine, you made my space so *Feng Shui* that I was always happy to come to the lab knowing you would be here; Nicolas (the first version) thank you for your sensitivity, your laugh and your 'always after the gong' questions, it made me laugh a lot and I am laughing just thinking about it… Emmanuelle, thank you for sharing your good mood and your (sneaky) smile, and of course for sharing my *burrata and garlic passion* (not together), it would not have been the same without you. Caroline, my noisy but beloved neighbor, I miss you so much that I am going to join you in your new lab! Chen, thank you for your sensitive and good listening, you brought me warm comfort when I needed it during those last not so easy years. Huge thanks to the '*last generation*': Nicolas (2nd version) for standing by my side (literally) while writing this manuscript, thanks for your eternal good mood, it helped me a lot! Doug, we finally managed to understand each other (thanks to Jesse!) and it is finally for the best! Antonius, I truly enjoyed your smiley calm; and finally, Jules, my favorite 'happy *chouquette man*', your support was priceless, thank you! Thanks also to the 'young ones': Léa, Audrey, Jesse, David, Roeland, it was really cool having you in the team!

Special thanks to the two other PIs of the MBB: Jean, it was great and intellectually stimulating to interact with you, I truly enjoyed learning from you and I am sure everybody will see your influence in my work, I am grateful for

that. Sebastien, your smile, your good mood, your priceless advices and electrophysiological teaching made the five years spent in the MBB a great adventure which I will happily remember.

I also would like to thank our direct collaborators: Philippe, I learnt a lot from you and it was really enjoyable to work together, I will be very happy to continue working with you! Julien, thank you for having presented me the exciting world of human electrophysiology, it is priceless for me! Mathilde, thank you for your time and huge contribution to the amazing iEEG dataset I had the opportunity to work on. Katia, Karim, Anne, thank you for helping me with the patients and the data collection, and also for your time and good mood, it was really great interacting with you.

Merci à Sophie Pariente et Laure Rondi-Reig qui ont eu un rôle important dans mon parcours : j'éprouve envers vous une profonde gratitude.

Bien évidemment je voudrais remercier mes amis qui m'ont soutenue durant toutes ces années, des bancs de la fac jusqu'à aujourd'hui, avec une petite mais mémorable halte à San Francisco : Léa, Stéphanie, vous êtes là depuis le début et je vous remercie pour votre présence et tous les bons moments passés ensemble ! Les copains du Cogmaster, c'était fou de faire cette aventure avec vous : Antoine, tu nous auras fait rigoler tout du long ; Juliane, le beggae c'est toi en vrai ! Auréliane, de Samarcande à Colombo, tu nous auras vendu du rêve, on repart quand tu veux ; Mehdi, tu es la preuve vivante que tout est possible ! Les copains de SF : Pierre & Salomé, Badr & Pauline, Simon, Jérémy, Samy, Delphine, Gaétan, merci pour votre soutien et pour avoir ramené un bout de la Californie à Paris ! Les amis de l'ICM : Laurène, Sami, Morwena, Camille, Adrien, Philippe, ma vie à l'institut n'aurait pas été la même sans vous, merci pour tous ces moments de joie !

Un immense merci à ma Maman, qui connait la valeur des choses et la motivation ! Modèle de volonté, tu m'as fait comprendre que tout est possible et que l'effort paye. Merci de m'avoir toujours soutenue malgré les aléas de la vie et les difficultés qu'on y rencontre, merci d'avoir toujours cru en moi, je t'en serai éternellement reconnaissante.

Mille mercis à ma sœur Coralie, qui m'a soutenue et supportée pendant toutes ces années. Ton regard sur mon travail a toujours énormément compté, j'ai adoré débattre de science et 'pas science' avec toi !

Un MERCI tout particulier à Marwa, mon éternel binôme sans qui rien de tout ça n'aurai eu la même saveur. De la licence de biologie à aujourd'hui, on a partagé l'intégralité de nos vies, des aventures farfelues – de Gif-sur-Yvette à Beyrouth en passant par Londres – jusqu'à nos soutenances de thèse, tu as été le pilier de ma vie parisienne. De nouveaux horizons s'ouvrent maintenant à nous, je suis plus qu'heureuse de savoir qu'on les vivra encore en binôme !

Et pour finir, merci à Mehdi pour avoir été là pour moi, pour nos passionnantes discussions sur la vie, le monde et la science, pour la paix que tu m'as toujours apportée. Merci pour ces merveilleuses années de bonheur ! On a encore tout à découvrir ensemble, j'espère seulement que le temps ne passera pas trop vite à tes côtés…

# RÉSUMÉ

« Préférez-vous manger des chips ou un cookie ? » Alors que l'on peut très facilement répondre à cette question, les mécanismes cérébraux engagés dans la production d'une réponse à cette question sont loin d'être compris. Ils peuvent cependant être décomposés en plusieurs étapes : il s'agit premièrement d'assigner une « valeur » à chaque option, c'est-à-dire une quantification du désir d'obtenir chacune d'entre elles. Cette valeur est propre à chacun, elle est donc « subjective ». Ensuite, il faut comparer les valeurs subjectives de ces deux options pour pouvoir être ainsi capable de sélectionner celle qui a la plus grande valeur. L'assignation d'une valeur à un objet semble être effectuée par un réseau cérébral bien spécifique composé entre autres du cortex préfrontal ventro-médian et du striatum ventral. Ce réseau recoupe le réseau de la récompense identifié chez l'animal et il a été logiquement nommé le « système cérébral des valeurs ».

Le travail réalisé dans cette thèse s'intéresse à la notion de valeur et aux moyens d'y avoir accès, aux propriétés du réseau cérébral d'évaluation et à son implication dans le processus de décision. La première étude qui compare trois méthodes d'élicitation des valeurs a montré que les moyens utilisés pouvaient être considérés comme équivalents et que ces valeurs étaient relativement robustes aux méthodes. Ce résultat nous permet de valider les tâches utilisées dans les études suivantes. La deuxième étude, réalisée sur des données d'intra-électroencéphalographie humaine, a permis d'étudier la dynamique neurale du réseau cérébral d'évaluation, mais aussi de répliquer et développer ses propriétés établies en imagerie par résonance magnétique fonctionnelle (IRMf) : il encode notamment les valeurs de manière subjective (propre à chacun), générique (peu importe la nature de l'objet) et automatique (sans que l'on ait besoin de l'expliciter). La dernière expérience s'intéresse aux mécanismes de la décision et en particulier aux mécanismes de comparaison de deux options. Dans notre étude, nous testons l'hypothèse que nos préférences a priori sur les catégories des objets (préférer la nourriture salée à la nourriture sucrée) va définir une option « par défaut » (les chips). Au niveau comportemental, nous avons montré que les préférences a priori induisait un biais dans les choix en faveur de l'option par défaut qui était choisie plus souvent (et plus rapidement) que sa valeur ne le prédirait. Au niveau cérébral, établit en IRMf, nous avons trouvé que la valeur décisionnelle était encodée dans le cortex préfrontal ventro-médian (vmPFC) dans le référentiel de l'option par défaut versus l'option alternative. Nous pensons que cette étude propose une solution générale sur l'implémentation neurale du processus de décision et qu'elle révèle des

mécanismes sources de biais dans le comportement jusqu'ici inexplorés.

Les résultats de ces études considérés dans leur ensemble mettent en lumière certains mécanismes cognitifs de la prise de décision en explorant les propriétés neurales d'assignation de valeurs mais également en proposant un nouveau cadre d'implémentation de la décision elle-même.

# SUMMARY

‘Do you prefer a French fry or a cookie?’ While this question is simple for us to answer, what is the process engaged in the brain to allow us providing an answer is a central question in neuroscience of decision-making and many aspects of it remain unclear. A decomposition of this process might help us to understand the involved mechanisms. Indeed, first we need to assign what we will call a ‘subjective value’ to each option – i.e. the quantification of how much we like each of these options. Then, we need to compare those values to finally being able to select one of them. Assigning a value seems to be the function of an interesting brain network mainly composed of the ventro-medial prefrontal cortex (vmPFC) and the ventral striatum – a network which overlaps the reward circuitry identified in animal studies – and which is called the Brain Valuation System (BVS).

In the first study of this PhD thesis, we investigated and compared three behavioral ways to have an access to these ‘subjective values’. We found that subjective values were relatively robust to the way they were elicited. In the second study, we investigated the specific properties of the Brain Valuation System established through fMRI in humans in a large dataset of intra-EEG recordings in epileptic patients. We were able to replicate those properties and to provide insights on the underlying dynamics of this network. Finally, in the last study we investigated how this brain network was involved during a binary choice. We specifically invested whether prior preferences defined at the category level (savory versus sweet food for example) would define a default policy towards one or the other item (French fry versus cookie). At the behavioral level, we found that prior preferences induce a bias that leads participants to choose the default option more often (and faster) than its value would predict. At the neural level, we found that prior preference influenced the BVS baseline activity and that decision value was expressed by the vmPFC in a default versus alternative framing.

Altogether, these findings shed light on the distinct cognitive mechanisms underlying value-based decision-making i) by exploring the neural properties of value assignment and ii) by proposing a general solution to the neural implementation of the comparison between option values. We believe this demonstration points to hidden default policies as sources of bias in choices.

# TABLE OF CONTENTS

## *Why do we do what we do?*

The *'WHY'* aspect of this question has been assessed during centuries through the concept of motivation and decision-making. Indeed, motivation has been defined as the process which **drives** our behavior toward a **direction** with a specific **intensity**. Thus, to answer this question, we need to go further in the investigation of motivation. Naturally, one might rightfully ask what is this process, what is its inputs, and ***HOW*** does it work? While we are able to examine the outputs of this process through the observation of actions, knowing what is the substrate used by this process to generate actions is a much more complicated problem. Indeed, goals and needs are relatively abstract concepts which are difficult to formalize, however, they do have a critical feature which is used to orient and energize our behavior: a **subjective value**. Seen as a major key to understand behavior, subjective values are essential to decision-making since they allow us formulating preferences and orienting our behavior. Understanding the biological substrates of subjective valuation and decision-making has become one of the biggest challenges in the field of cognitive neuroscience and this is the topic of this manuscript. Indeed, I will focus on the neural bases of the subjective valuation process and then assess how those values are used by the brain to make decisions.

# Chapter 1: Theoretical background

In this chapter, I will first review the critical work done by philosophers, economists, neurophysiologists and psychologists to define the concept of subjective value, to formalize the properties associated to it and to decipher the cognitive mechanisms underlying decision-making based on these values.

Then, I will present what neuroscientists brought to the field by reviewing the literature about the 'Brain Valuation System', a critical brain network involved in the valuation process, which has also essential properties which helps understanding human behavior.

In the third part, I will review the neuroscientific work done around value-based decision-making to present the actual neuro-computational accounts proposed but also to target the debates and controversies in the field.

Finally, I expose the main questions I will assess in the following parts.

## 1. Subjective values & preferences

### A. Brief history of subjective value investigation

The concept of subjective value has been assessed for more than 2000 years and defined by three main disciplinary fields: philosophy with Aristotle, economics with Bernoulli and Kahneman among others and psychology of reinforcement learning with Thorndike, Pavlov, and Rescorla and Wagner.

### *Philosophy of pleasure*

Subjective values have been assessed through the prism of **pleasure** in the classic philosophical account. The central question was: “What is the nature of pleasure?” Pleasure needs to be distinguished from happiness, which is considered as a stable and long well-being. Pleasure can be seen as a feeling of an instant, but also as an experience, or as “a feeling we seek to bring into consciousness and retain there”, or even as the “motive power”(Sidgwick, 1907).

To make a long story very short and simple, I would propose than when philosophers discuss what we call **subjective value**, they mainly talk about an **amount of pleasure**, with the idea that pleasure is valuable while **pain** is disvaluable. Then, seeking to get **pleasure** or seeking to avoid pain becomes a **desire** (which we would call “**goal**”). John Stuart Mill proposed that a desire for an object is only caused by pleasure, or at least by the idea of experiencing pleasure (Mill, 1901). This last proposition overlaps hedonistic psychological theories, which refers to the idea that pleasure and pain are the *only* things of importance for motivating us to behave.

The unitary aspect of pleasure had also been central in the attempt to explain it. In other words, knowing whether pleasure is one single kind of feeling with different causes had been subject of debate, since its relations to awareness, attention and motivation are not simple and it can be divided into several kinds such as liking something or being aware of liking something.

Desire is also subject to controversy. Aristotle distinguishes **desire** into appetition and volition. The former being originating in the body and which can be seen as primary needs while the latter being more rational and involved in the commitment to an action toward a desire.

Thus, the nature of pleasure and desire are still hot topics in philosophy. However, the question of the “amount” or “strength” of pleasure and desire is rarely assessed in this field while it is precisely what we are interested in. We will thus define subjective value here as an internal representation of the amount of pleasure an object, an action or a state could provide and which is able to drive the behavior both in its **orientation** (make a choice) and in it **intensity** (put some effort in an action). The ‘orientation’ aspect has been largely assessed by the economic field, which I will briefly describe in the forthcoming section.

## *Value as subjective utility in economics*

The field of decision-making has been dominated by an economic perspective for a very long time and the first concept of the field is that humans should be rational and their decisions should serve to optimize utility. However, there are numerous violations of rational behavior (which we will see later in this chapter) and Daniel Bernoulli (1700-1782) was the first to formally describe why we need to define utility as a subjective value - as a value which is specific to each individual and/or contexts. Indeed, when formalizing the concept of **diminishing marginal utility** Bernoulli, (1954) distinguished the **objective value** of an item - its *price* - and the **subjective value** of it – its *utility*- with the idea that the price is equal for everyone and the utility depends on the person and circumstances. For him, ‘*there is no doubt that a gain of one thousand ducats is more significant to a pauper than to a rich man though both gain the same* *amount*'. From this statement, he proposed the first theory to explain the concept of diminishing marginal utility (even if he did not introduced the term itself) by formalizing that '*utility resulting from any small increase in wealth will be inversely proportionate to the quantity of goods previously possessed*'.

This formalism is introducing the concept of subjectivity in the value, or utility, which is used when making a choice. In 1879, William Stanley Jevons develops the **theory of utility**, and sees it as the quantitative measurement of pleasure: "*numerical expression of quantities of feeling*" (Jevons, 1879).

However, a critical difficulty is that utility cannot be measured directly. Indeed, it needs to be inferred from the behavior, either through the amount of effort or the willingness to pay to get access to a **cardinal** representation of utility or through choices between several options to get access to an **ordinal** representation of option utility. Measurements and representation of utility are two topics that I will detail later, but most of economics theories on utility have been developed by measuring utility through choices.

### *Values in reinforcement learning theory*

While philosophers and economists were focused on pleasure and its measurement through choices, some scientists and psychologists were investigating behavior to describe what drives it and what could explain its repetition. As hedonists, they rapidly converged on the concept of **reward**, an object associated to a **positive value**. The concept of reward is central to the physiological field of value investigation. While reward is an object or a situation, the associated pleasure can be seen as the subjective value defined by economists and philosophers. For them, reward is what allows us to learn.

Thorndike (1874-1949) was the first to introduce a formal learning theory with his famous 'connectionism theory'. He described it as the association between *'a certain act with a certain situation and resultant pleasure'* (Thorndike, 1898). Using puzzle boxes and animals (cats, dogs, and chicks), he showed that the time taken by animals to escape the box was decreasing with the number of trials until reaching a constant and a short response time. This form of learning was called trial-and-error learning and is instrumental, meaning that it is an association between an action and an outcome. From this experiment, Thorndike defined the **law of effect** (1905), which states that responses that produce a satisfying effect in a particular situation become more likely to occur again, while responses that produce a discomforting effect are less likely to be repeated.

In parallel, Pavlov (1849-1936) ran his well-known experiment on dogs and formalized the concept of **classical conditioning** (Pavlov, 1927). Pavlov presented a stimulus (a ring bell) and then gave food to a dog. After several repetitions, the dog started to salivate at the ring bell, thus in response to the stimulus while it was only in response to food before what he called the 'conditioning'. The conditional stimulus (the ring bell) had been associated to the food (unconditioned response), meaning that the conditioned stimulus took on the primary value of the food.

While philosophers and economists seemed to yield the question of *anticipated* and *experienced* pleasure as one unique concept. Physiologists investigated the link between them and showed how one could influence the other through learning.

Following Pavlov's work, Rescorla and Wagner developed the Rescorla-Wagner model in the early 1970s to formally describe the circumstances under which Pavlovian conditioning occurs (Rescorla et al., 1972). It introduced the idea that learning comes from the difference between what is **expected** to happen and what is actually **experienced**. The main principle is that the amount of surprise an organism experiences when facing an unconditioned stimulus (US) is assumed to be dependent on the summed associative value of all present cues during that event.

Thus, subjective values can drive the behavior and orient it toward cues, or objects in the environment. They are also critical for learning.

## B. Values in decision-making

How values are used to make decisions is a central question in my work. However, we need a framework to address this question; indeed, we need to define the different processes and steps involved in decision-making. We also need to define what kind of behavior and decision we are referring to. Finally, we need to distinguish the different types of subjective value that are involved in such processes. I will address these issues in the following section.

### *Decision-making steps*

A unified framework of value-based decision-making has been proposed by Rangel, Camerer and Montague (Rangel et al., 2008), it has the advantage to unify choices studied by economists (which we would label "goal-directed behavior") and processes assessed by reinforcement learning theory. Five major steps are described (see Figure 1):

- **Representation**, including the representation of the environment (external states), of the self (internal states) and of the available options;

- **Valuation**, which corresponds to the assignment of a utility to each action at stakes, given the internal and external states, which can also have a value. We will see later that there exists several kind of value here, such as the action value or the reward value.

- **Action selection** which corresponds to the decision itself, it is based on the previous critical step: computed values are compared and the best[1] option is selected, resulting in the selection and execution of an action. Several valuation systems (which I will describe later) can conflict at this stage.

- **Outcome evaluation** is the step where the output of the decision is evaluated and (sometimes) experienced.

- **Learning**, which can have an impact of the three first described steps since it allows updating the representation of the actions and states (internal and external), the value of the outcome through **prediction error**: the value of the prediction and the value of the outcome are compared; and finally updating the action-selection processes.

We can see in this simple framework that valuation is critical to decision-making since it takes place in several steps of the framework. Kahneman has been highly influential in the distinction of several kinds of values, which I will describe and define here for the clarity of what I am mostly going to present in this manuscript.

---

[1] ‘Subjectively supposed best option

### *What are the different types of value & how can we measure them?*

As I mentioned, the first type of value is the **predicted utility** (or anticipated value, expected reward value, goal value). It is also described as the stimulus value (Rangel and Hare, 2010), the subjective desirability or even affective forecasting. In formal terms, we can define it as the

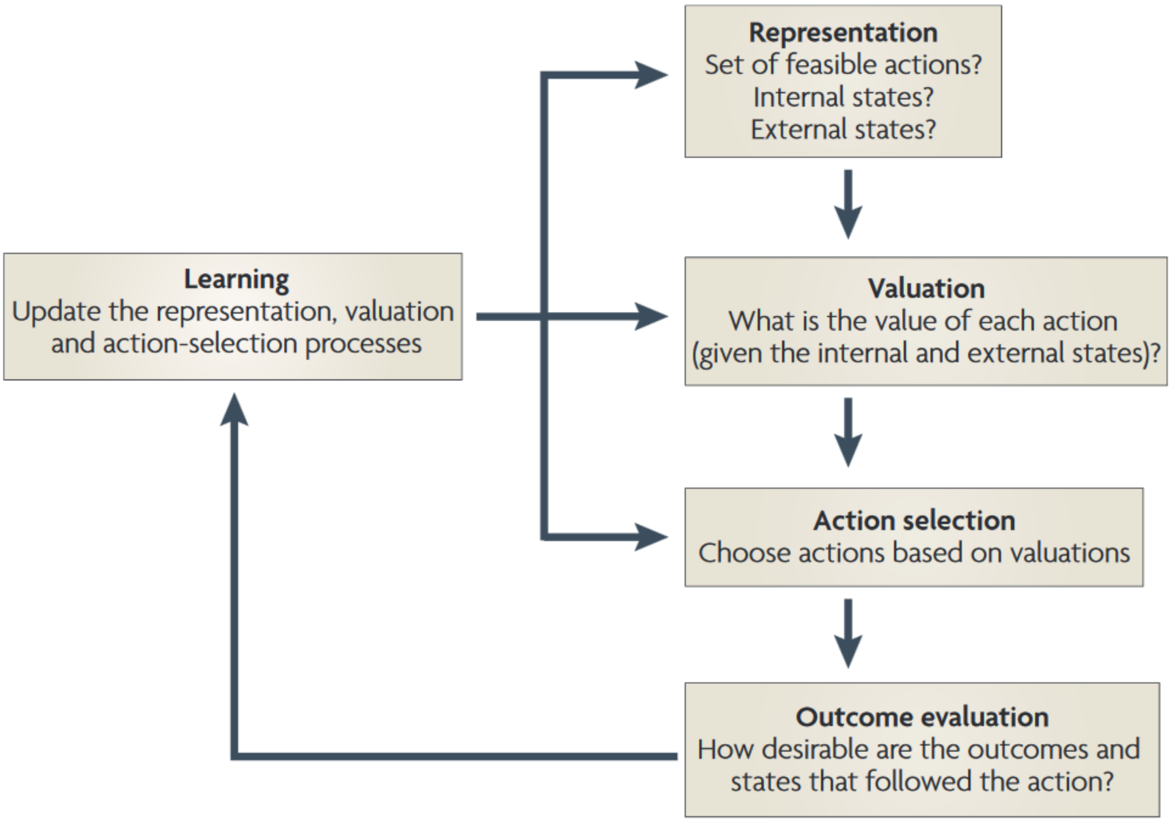


Figure 1 – Decision-making steps

Five basic computations can be described to provide a unified decision-making framework. First, the representation of elements which will be taken into account in the decision such as internal states like hunger. Second, the valuation of these elements (cost and benefits for example). Then, the choice itself, implying to select an action which will satisfy computations made on the valuation step. Fourth, the valuation of the outcome (experienced value) which assesses how desirable the outcome is. And finally, learning, which implies to compare the expected and the experienced value and to update the representation, the valuation and the choice processes. From Rangel et al., (2008).

expectation of how much an option will be liked when received[2].

**Costs** also need to be computed in order to make a decision. Several types of costs are distinguishable (Rangel and Hare, 2010): the price (giving up something valuable such as money to obtain the reward at stake), the effort, which are the cost of the action itself: moving one meter to get coffee is a small effort cost while having to climb one stair to get coffee is a higher effort cost; and the delay: having to wait to obtain a reward is indeed considered as costly.

Then, **decision utility**, or decision value which can be seen as the net value of the **action** since it includes the anticipated value and the cost (price / of the action) to obtain the outcome. It is the one which drives the choice and/or orient the action. Note that in the case of binary choices, decision value might refer to the difference between the expected subjective values of two options.

The instant value or what I called the **experienced value** refers to the pleasure resulting of the consumption of the (chosen) reward. Prediction error is computed with this value and it drives learning and future behavior.

In my manuscript, I will mainly focus on the expected subjective value of rewards since it is the one temporally preceding actions and choices on which the processes are going to be based on. However, as we saw the distinction between expected value and experienced value is critical since it allows learning. Moreover, in the context of value-based decision-making such as choosing between two kinds of juices, there are two expected values to be computed but only one to be experienced, thus the cognitive computation of this kind of values must be

---

[2] We can see here the multiplicity of terms used in the literature to describe it. I will often use the expression 'subjective value' to mention it.

distinguishable. Nevertheless, this distinction is not the topic of this manuscript.

### *How to measure subjective values?*

Now that we have defined the different types of values, I would like to describe how we can measure them. Four kinds of measure are dominant in the field. They are considered as equivalent to access those values but a systematic comparison has not been done. It will be the topic of my first study.

The first type of measure is what we call a **rating task** (Lebreton et al., 2009), it uses subjective reports. An item or situation is presented to a participant and she is asked to rate (i.e. evaluate) how much it would be pleasant to obtain the item using a scale. This kind of scale is also used to access cost value (by asking "how much it would be unpleasant").

The second and third types of measure use costs. The '**willingness to pay**' (Varian, 1992; Plassmann et al., 2007) is measured in auction bid task (Becker–DeGroot–Marschak - BDM, Becker et al., 1964) which asks participants to indicate on a scale how much they would be ready to pay to obtain the presented item. The '**effort task**' asks them to produce an effort to obtain a reward. The latter has different forms such as asking subjects to produce an effort to determine the probability of receiving a reward (Pessiglione et al., 2007; Cléry-Melin et al., 2011), or asking subjects if they are willing to produce an amount X of effort to obtain a reward Y (Bonnelle et al., 2015). This measure can be considered as indirect since it is supposed that participants will minimize their costs to obtain a reward and that subjective value would be therefore revealed.

The last one is the well-known **binary choice task**, which imposes participants to make choices between two options. Contrary to the others, this measure is not ordinal: the value of each option can be inferred from its rank, meaning that if A is chosen against B and B against C, we can say that A has a greater value than B and B than C but we do not know the subjective values of those three options. However, measuring reaction time of this kind of choice can indicate whether the two values are close or not (a choice between items with similar values will be more difficult and will take more time). Furthermore, the notion of distance between them can also be assessed by investigating the choice rate (frequency of choosing A over B when the choice is repeated). It is also different from the other tasks since it focuses on the decision value (difference between the anticipated values) while the other ones focus on the anticipated value or the costs.

Cardinal measures of subjective values are closely linked to the range of the scales proposed to subjects. Moreover, the use of the scales can differ according to subjects, thus it is important mentioning here that even if we refer to the word 'cardinal' for these measures, this is still subjective values reported *relatively* to a given scale.

The tasks used in the field differ in terms of cost, scales and frames and to my knowledge it has not been assessed whether those measures were equivalent. I will test the validity of these measurements in the first study of my thesis.

### *Several valuation systems*

In order to simplify the framework of my studies, I also need to mention that valuation can take place in distinct types of behavior. Three valuation systems can be described: the Pavlovian system, the habitual system and the goal-directed system, respectively (Rangel et al., 2008).

### The Pavlovian system

As we saw in the beginning of this chapter (page 14), Pavlov had a strong impact on the field by describing what we now call 'Pavlovian' behaviors. It relates to behaviors which are often be seen as 'innate', such as starting to salivate when seeing food. The Pavlovian system assigns values to environmental cues, which can be pre-determined or induced by repetitive exposition (Pavlov, 1927). However, it concerns a small set of behaviors which are usually triggered by appetitive unconditioned stimuli (Clark et al., 2012) and it is not instrumental, as opposed to the two other systems, i.e. the learning is not about actions but about external stimuli naturally predictive of a reward.

### The habitual system

In comparison, the habitual system can be applied to any type of behaviors as soon as it is learned through repeated training. Thus, this system is going to assign a value to an option through stimulus-response association and by trial-and-error experience. This system is fundamentally based on the link between a **stimulus** and an **action** leading to a reward and is involved in learning (Balleine and Dickinson, 1998).

### The goal-directed system

In opposition to the habitual system, the goal-directed system assigns values to actions and creates associations between **actions** and **outcomes**. Thus, any behavior driven by this system would normally be directed toward a specific reward, or 'goal'. Such a system would also quickly react to a change in the environment or in the internal state such as satiety which would decrease the value of food and consequently decrease the approaching behavior toward it.

## Interactions between the valuation systems

Those systems are supposed to work in parallel and to assign different values to the same options, sometimes leading to a conflict (or a competition) between them. For example, Wayne Hershberger (Hershberger, 1986) showed that chicks were not able to learn to run away from food to get access to it. This example is an illustration of the failure of the instrumental system against the Pavlovian one (Dayan et al., 2006). Thus, the co-existence of these systems might (partially) explain why we often observe inconsistencies in behavior and choices.

A more recent view opposes the habitual system to the goal-directed system in a distinction between **model-free** and **model-based** valuations (Daw et al., 2005) to explain discrepancies in choices. The main idea underlying this framework is that the low cost model-free system computes value of options only on a reinforcement learning mode (only based on the previous output of the choice) while the model-based system would be more computationally demanding since it supposes that values are assigned in accordance to a learned model of the world. The knowledge of the environment and the internal states are comprised in this system which would correspond to the goal-directed system. This view allows explaining unexpected behaviors and supposes that the two systems can compete to assign values to options. Moreover, it has been shown that a mixture of those two systems is used by subjects to solve tasks designed to test those systems (see Daw et al., 2005).

In conclusion, values can be assigned with different systems but when we focus on binary choice tasks which require no learning, we can suppose that the goal-directed system is mainly involved, especially in the context of the laboratory (in which there is no habit). However, before entering the process of decision in this kind of task, we need to describe some properties of subjective values.

## C. Subjective value properties

Many times, it has been observed than human is not rational, i.e. his choices cannot be predicted at 100% even if the subjective values of considered options have been reported. Numerous of violations of rationality have been described and it has been shown that values assigned to options can be modulated by factors such as uncertainty, risk, time or context. Interactions between values and these variables have been well described and formalized by economists in order to explain inconsistencies and biases in choices and why human is not purely rational. The next part addresses the impact of various variables on subjective valuation.

### *Diminishing marginal utility*

As we saw in the historical section of value investigation, there is a nonlinear relationship between the amount of reward and the subjective value (Bernoulli, 1954). This phenomenon has been described as the 'diminishing marginal utility' and imposes a concave utility function to the amount of reward (Figure 2). Thus, we can define the utility as:

$$U = f(R)$$

With U being the Utility, or subjective value associated to the reward R and f a concave function. This is classically explaining non-linearity observed between amounts of reward and subjective values.

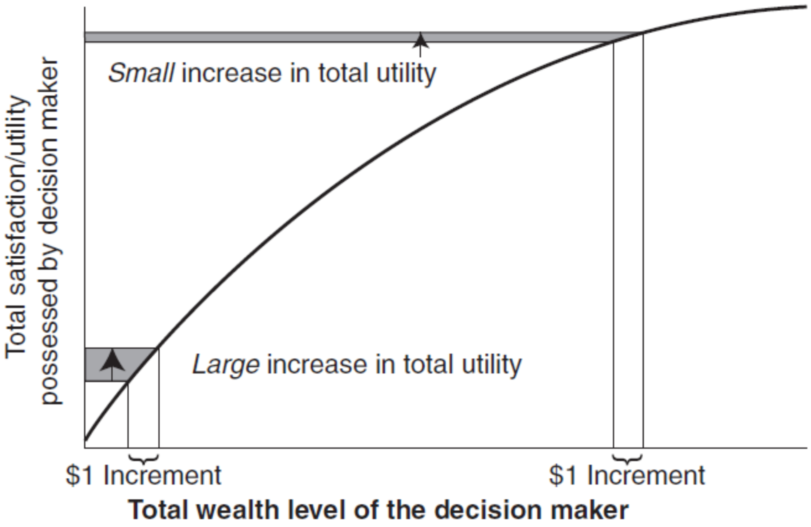


Figure 2 – Diminishing marginal utility

Relationship between the value of an object and the amount of this object the decision-maker possesses. The utility of possessing the object increases less when the owner already has a lot of it. From Glimcher and Fehr, (2013).

## *Value & Uncertainty*

However, even if we take into account the fact that utility is concave, it does not explain why some choices are not rational. One important point is that events and environment are not always deterministic of reward. Instead, most of the time, getting a reward is probabilistic. Then, utility of a reward is going to be dependent on the probability associated to it, such as:

$$U(R) = f(R) * p(R)$$

With p being the probability of getting the reward R and U as the utility. Von Neumann and Morgenstern, (1947) developed the classic **Expected Utility Theory** (EUT) from this observation. The main idea is that if an agent satisfies the four following axioms, then there exists a utility function assigning values to lotteries (reward associated to a probability):

- *Completeness*: For any lotteries A and B, one of these proposition is true: $A > B, A < B \text{ or } A \sim B$, with $<$ and $>$ meaning '*is preferred to*' and ~ meaning '*is equivalent to*'.
- *Transitivity*: if $A > B, B > C, \text{then } A > C$
- *Continuity*: if $A > B > C$, then, there exists a probability $p$ ($\in [0, 1]$) such that $pA + (1 - p)C \sim B$
- *Independence*: If $A > B$, then for any $C$ and $p$ ($\in [0, 1]$): $pA + (1 - p)C < pB + (1 - p)C$. It means that $A > B$ would be true even in the presence of other options.

However, even with this theory, there were still unexplained anomalies observed in human behavior, especially concerning the perception of probabilities, which is deformed: people tend to see low probabilities higher than they really are and high probabilities lower than they truly are. Savage (1954) included the subjective probability (P) of receiving an outcome instead of the real probability of getting it (p).

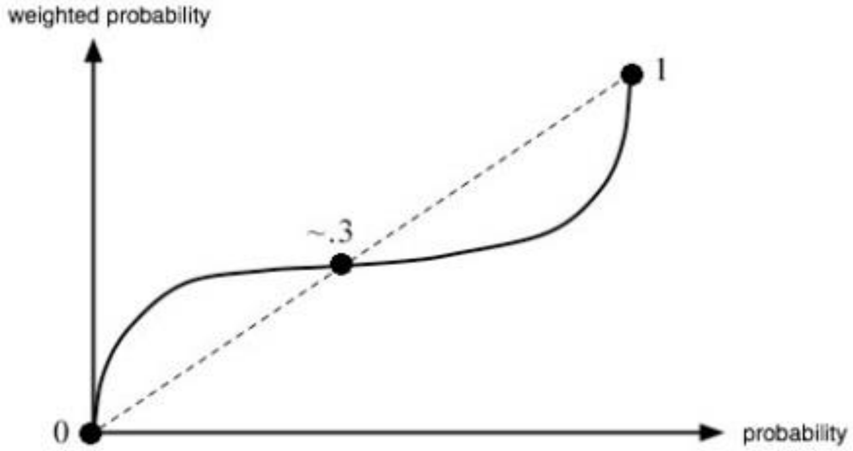


Figure 3 – Subjective probabilities

We tend to perceive low probabilities as higher than they are (winning to Loto), and high probabilities lower than they are (get a disease).

Then, utility of a reward becomes:

$$U(R) = f(R) * P(R)$$

Finally, another component can impact the utility of a reward: the risk aversion: we would usually prefer having 10€ at 100% probability to having 20€ at 50% even if the utility of the two options is the same according to the previous statements. In other words, it seems that we tend to prefer sure outcomes compare to uncertain ones. This phenomenon can be captured with a utility function varying from convexity (risk seeking) to concavity (risk aversion, the more common case) through linearity (risk-neutral) (Figure 4). The EUT proposes that this parameter is comprised in the subjective utility function assigning values to lotteries (Pratt, 1964).

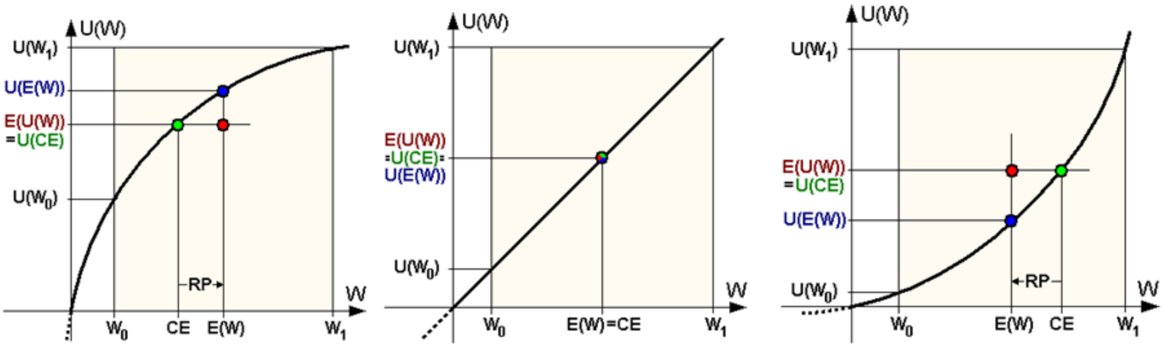


Figure 4 – Utility functions associated to three risk attitudes.

Utility (U) of an uncertain outcome W according to W. E(U(W)) is the expected utility of the uncertain outcome. U(CE) is the utility of a certain outcome, U(E(W)) is the utility of the expected value of the uncertain outcome. **Left**: Risk-aversion. A concave utility function is the most common risk attitude. The expected utility of an uncertain outcome is lower than an equivalent certain outcome. **Middle**: Risk-neutral. Uncertain outcome is equivalent to a certain lowest outcome. **Right**: Risk-loving. Convex utility function for people who tend to prefer an uncertain outcome compare to a certain lowest one.

### *Value of losses*

Even when using utility functions to take into account nonlinearity of the perception of reward, there were still anomalies in behavior regarding the perception of gain and losses. The domain of losses is asymmetrical compare to the domain of gains: the subjective utility for losses is convex (losing 5 euros when you already lose 1000 is higher than losing 5 euros when you already lost 5 euros), but it is also steeper, for example, losing 50 euros is more painful than winning 50 euros is enjoyable, i.e. losses hurt more than gains feel good (Figure 5). To compensate for this effect, Kahneman & Tversky developed the **Prospect Theory** (Kahneman and Tversky, 1979). Contrary to the expected utility theory which is normative ('how should be the behavior'), the prospect theory is descriptive. The key parameter in the prospect theory is the existence of a reference point, separating gains from losses and pleasure from pain. The utility function applied to gains and losses is asymmetrical around this point and creates a separation between the domain of gains and the domain of losses.

We should also note that the cumulative prospect theory was developed later, adding the deformation of perception of probabilities to the prospect theory (Tversky and Kahneman, 1992).

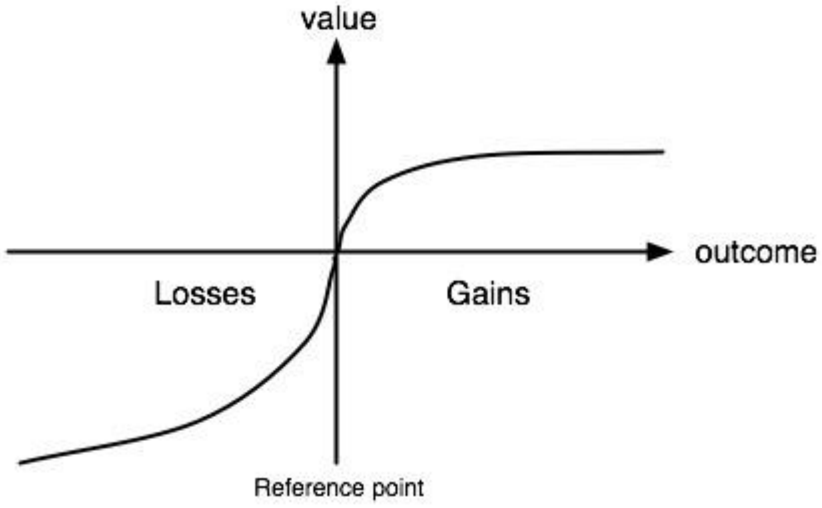


Figure 5 – Utility value function in the prospect theory.

The utility function of gains and losses is asymmetrical around a reference point. The slope of losses is steeper for losses.

## *Value & Time*

Utility functions can also be affected by time. Indeed, receiving a reward tomorrow or in one year are options with distinct subjective values. Usually, the immediate reward is preferred. This effect is called **delay discounting** and is classically measured with intertemporal choices which consist of choosing between a delayed large reward and an immediate smaller reward (Frederick et al., 2002). Economists (Jevons, 1871) and psychologists know this effect quite well and observed it in the context of instrumental learning in pigeons (McDiarmid and Rilling, 1965) but also goal-directed behavior (Prevost et al., 2010; Paglieri, 2013). The decrease of value with time is not linear; instead, it has an interesting concavity showing that utility falls off rapidly for short delays while more gradually for longer delays. One might think that this effect could be due to the fact that uncertainty increases with temporal distance and that time only affects the probability of getting the reward. However, there is supporting evidences in behavioral studies (see Kalenscher and Pennartz, 2008 for an influential review) which show that the two phenomena seem distinct. Samuelson (Samuelson, 1937), developed the Discounted Utility Theory (DUT), based on the EUT with axioms specific to time such as stationarity (indifference between two options should depend only on the difference on the delays, and this should be true at any time point) or constant discount rate across time. However, numerous violations of this theory and especially of the stationary axiom persist and several authors showed that data were better approximated with a hyperbolic discounting compare to an exponential one, as proposed by the DUT (Figure 6).

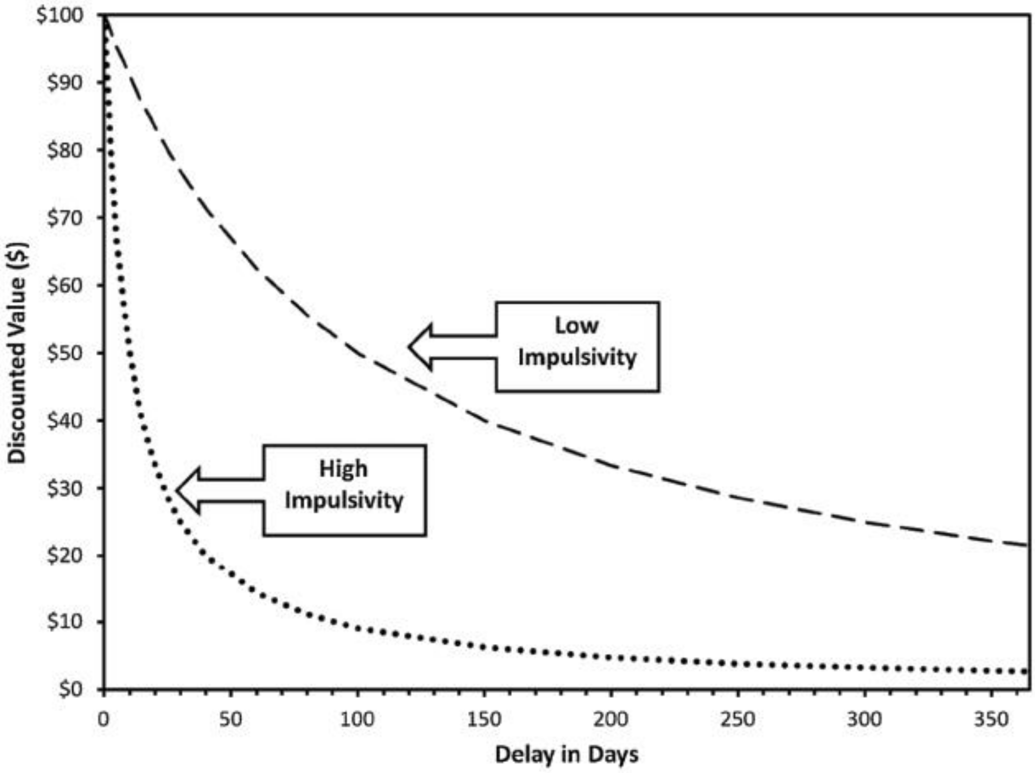


Figure 6 – Delay discounting

Subjective value of a reward decreases with time. With a high (low) discounting factor, value decreases faster (slower) and behavior is qualified with high (low) impulsivity.

To sum up, we have seen that economists tried to develop normative and descriptive theories of utility to explain why choices seem incongruent with objective reward values. Thus, utility is affected by probability, time, risk attitude, quantity of ownings, but also by the frame of it (loss and gain domain). Moreover, there are a lot of evidences showing that utility is affected by the context, such as internal or states, but also social environment.

### *Value & context*

Context is a non-specific concept which is usually used to point everything which is surrounding a process. Regarding valuation, context can be internal states, but also external environment and social situations.

## Internal states

An internal state refers to any state related to internal events such as satiety, fatigue, beliefs or mood. I will focus here on mood and beliefs as illustrative examples.

### *Mood*

Mood is known to influence cognitive processes and behavior. It is described as an affective state lasting for a relatively long period of time, contrary to emotions which are in reaction to external events. Mood has a value itself and people will tend to increase or maintain a positive level of mood while they will try to avoid a decrease in the level of mood. The Mood-Behavior Model (MBM) (Gendolla, 2000) posits that mood can have an *informational impact* by affecting judgments and a *directional impact* by influencing preferences. The informational impact is usually seen as congruent with the mood: if someone is in a good mood, valuation tends to be more optimistic with the reverse effect for a negative mood (Clore et al., 1994). It impacts valuation in the intensity of the judgment. On the other hand, Gendolla also describes what he calls the directive impact of the mood on behavior: he supposes that people are going to make qualitatively different choices according to their mood in order to maintain it if it is good ("mood-maintaining behaviors") or to increase it if it is bad ("mood-repairing behaviors").

A recent model (Eldar et al., 2016) posits mood as an overall '**momentum of reward in the environment**' useful to learning. The idea is that if there is a change in the environment (let's use their example of an increased sunshine) which will affect several opportunities to get rewards (each tree would produce more fruits), then it makes no sense to update each opportunity separately instead of updating the value of the whole environment. Consequently, they suppose that this is precisely the function of mood, with the result of having an increase of positive surprise when more positive events are faced (which is explaining the informational effect described by Gendolla).

Thus, there are supporting evidences for an impact of the mood on valuation. While mood can be qualified as an affective modification of internal states, there are also examples showing that beliefs can change the way we valuate things.

### *Beliefs*

Beliefs (or expectations[3]) can affect valuation. One striking example is how much we like wine is influenced by its price. A study (Plassmann et al., 2008) interestingly showed that telling subjects that one wine was more expensive than it truly was decreased the subjective pleasantness subjects attributed to the wine. Thus, the impact of prior knowledge and its use in judging the quality of an experience or an object needs to be considered when one is assessing how the valuation process works.

Moreover, as we saw earlier, the prospect theory supposes the existence of a **reference point**. This reference point distinguishes losses from gains but it is not necessary set to an absolute *zero*. Meaning that losses and gains are relative to a reference point which can be for example what I already earn (this phenomenon is also called *status quo*, or *endowment effect*). This is classically seen in the asymmetry between the willingness to accept a payment for selling (WTA) an owned item and the willingness to pay (WTP) for this same item: WTA exceeds WTP (explained by the asymmetry between perception of losses and perception of gains). Thus, it is also supposed that the valuation of options depends on a reference point, which is thought to be set up according to beliefs, as proposed Koszegi & Rabin (Koszegi and Rabin,

---

[3] I do not make any distinction between belief and expectation, even if this could be discussed.

2006): *'a person's reference point is her recent beliefs about future outcomes'*. In other words, they suppose that the reference point is dependent of expectations. An interesting property predicted by their model is that willingness to pay for a good does not reflect subjective value of the good but instead reflects beliefs about whether the subject is going to buy it and if so, how much he is going to pay for it.

To put it in a nutshell, internal states like mood, beliefs, knowledge of what is already earned and expectations about future can affect valuation. We will see in the next part what is known about the external environment impact on valuation and decision-making.

### External environment

The impact of external environment is mainly seen in choices and violations of rationality seem to increase when the **number of options** available in the environment increases (DeShazo and Fermo, 2002). This can be explained by a cost to evaluate a lot of options: Tversky and Shafir (1992) found that individuals tend to defer making decisions when the set is too large or when the decision is difficult, which suggest the existence of a cost. Moreover, it can also be explained by the decrease in discriminability between options, which decreases the precision of the utility given to each option (DeShazo and Fermo, 2002).

Another external factor which can affect valuation is the impact of options' attributes and the choice set used. Three main effects have been well described in economy: the attraction, the compromise and the similarity effects. All of those effects occur when two equally preferred options (choice is 50% for each) varying on two dimensions are presented with a third option. The effects are dependent on the quality of the third option. The **attraction effect**, discovered by Huber, Payne and Puto (Huber et al., 1982) corresponds to the situation were a third option added is similar to B but inferior to it on one of the two attributes, this addition results in a probability of choosing B higher than 50% (Figure 7, left). The **compromise effect** (Simonson, 1989) arises when the third option has the same global utility as A and B but more extreme than B. In this case, the percentage of choice is also biased toward B (Figure 7, middle). Finally, the **similarity effect** occurs when the third option is very similar to B: the preference is shifted toward A (Figure 7, right). The mechanisms underlying those effects are still debated but they are critical examples of how the context can influence the valuation process.

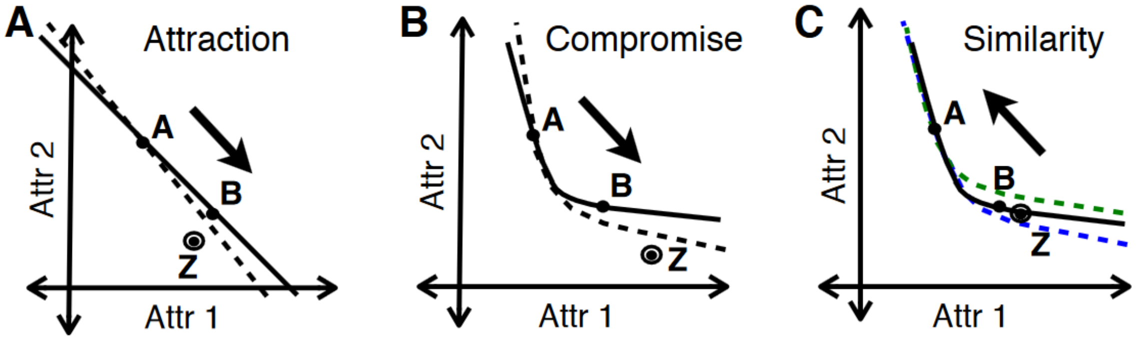


Figure 7 – Dependence of value alternatives in multi-alternative choice.

Options A and B are equally preferred when presented alone but A is higher on attribute 2 and B is higher in attribute 1. When a third option Z is added to the choice set, preferences are shifted toward one option (arrows), depending on the value of Z on attributes 1 and 2 relative to A and B. (A)Attraction effect, (B) Compromise effect, (C) Similarity effect. From (Shenoy and Yu, 2013).

Finally, **framing effects** concern all the situations in which the way of presenting options influence the decision. Tversky and Kahneman (1981, 1985) give a lot of examples in their review and I will present only one here to illustrate the main idea of what is a framing effect. Suppose you are a hero and you have to choose between alternative strategies in two different frames:

- Frame 1:
    - Option A: 200 people will be saved
    - Option B: 1/3 probability 600 people saved and 2/3 probability nobody saved
- Frame 2
    - Option C: 400 people will die
    - Option D: 1/3 probability nobody will die and 2/3 probability 600 will die

In their study, 72% of people choose the option A in the frame 1 and 78% the option D in the frame 2. This result is the illustration of the asymmetry between gains and losses: usually, people's behavior is risk averse in the gain domain and risk seeking in the loss domain. Thus, according to the frame of presentation, preferences can reverse.

### Social influence

Another strong impact on behavior and on decision-making is social influence. Asch was the first to demonstrate those effects (Asch, 1951) with its famous experiments on social conformity. One of the main results is occurring when subjects are presented a card with a line on it which is followed by another card with three lines A, B and C with one being the same size as the line on the first card. Subjects were changing their response if the (fake) group of participants around them were giving a common wrong response. This example is based on perceptual decision-making and it is the expression of the choice which is affected more than the estimation process. However, strong effects of social influence have been found in the domain of value-based decision-making and mainly in economic game theory.

For example, the phenomenon of **goal contagion** has been addressed in a neuro-imagery study in the team (Lebreton et al., 2012). They showed at the behavioral level that desirability of objects was increased when the same object was shown as the goal of another agent (by showing a hand taking the object in a short video) (Figure 8). Thus, here it is the subjective value (independently of any choice) which is affected by the social context.

In economic game theory, several classic experiments are used to measure how much players would sacrifice their own gains in favor of equity or moral judgment, such as the **dictator game** (Kahneman et al., (1986) participants have to freely split an endowment with another player without consequences), the **ultimatum game** (Güth et al., (1982) in which participants have to propose a split which is accepted or rejected

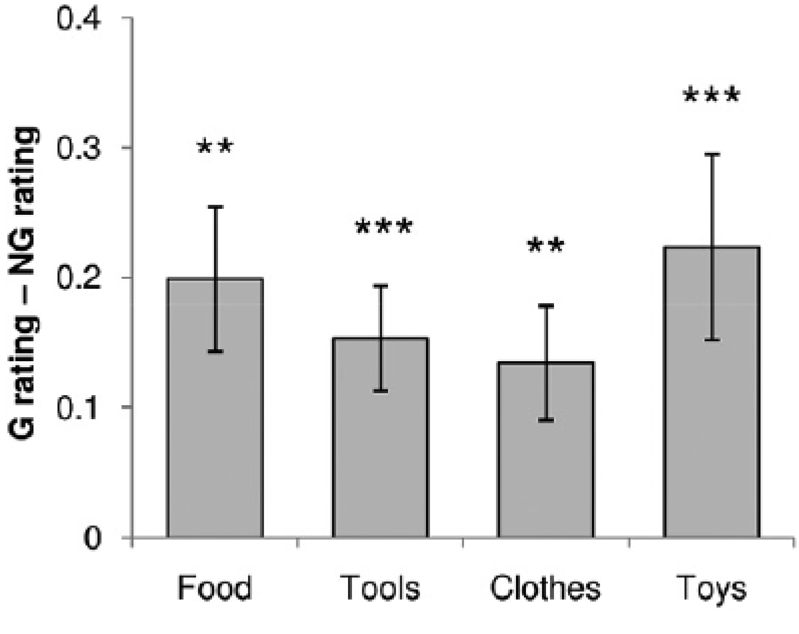


Figure 8 – Goal contagion illustration

In Lebreton et al., (2012), subjects were presented items from different categories (Food, Tools, Clothes and Toys) and had to rate them in two conditions: G, for Goal, in which the item was presented as the goal of someone else; and NG, for No Goal. The difference in ratings G rating – NG rating was shown to be significant for all kinds of goods.

by another player, if rejected, neither player receives any money) and **third party punishment games** (Fehr and Fischbacher (2004), they include a third player to the previous games who can punish players who violates social norms). Those games show violations of neoclassic economic theories about decision-making in which people are thought to be self-interested and to try to maximize their own utility.

We have seen that external factors such as available options, frames or social norms are able to modify the way we evaluate options, but also internal factors such as mood and beliefs. However, another factor needs to be taken into account when one is investigating the valuation process: confidence. Indeed, a reported subjective value has a weak meaning if the 'reporter' is not confident in its own judgment.

### *Value & confidence*

The link between valuation and confidence has gained interest in the last years. Confidence can be considered as a degree of certainty on a judgment but also as a probability of being correct. It can be expressed in most of the steps of the decision process: representation ('how sure I am this item is a fruit?'), valuation ('How sure I am to like this item?'), decision ('How sure I am to have chosen the item I really preferred?'), but also outcome evaluation ('How sure am I to enjoy the output of my decision?'). I will focus here on the confidence related to the valuation process. We will see later how confidence may relate to the decision process.

First, it is worth trying to define what confidence is. Coming up with a clear definition is struggling and debates are still existing. (Pouget et al., 2016) recently proposed to distinguish confidence from certainty in the sense that confidence is a subjective probability that a decision or a proposition is correct given some evidence. Certainty, on the other hand, would concern the encoding of all other kinds of probability, such as sensory evidence of other cognitive variables. Thus, confidence is seen as a **belief**[4] of correctness, as a special form of certainty. They propose an important computational distinction between certainty and confidence. While certainty would be an intrinsic part of the value in a Bayesian framework (as we see later) and would be integrated in the valuation process, confidence would express on the decision process, or even at the end of the process.

However, when subjects are evaluating options in a classic rating task, they are making a decision on the position of the cursor and there is consequently a computation of the probability of being correct (giving the real internal subjective value). In this case, certainty and confidence on value are the same concepts (quantitatively[5] speaking).

In the Bayesian framework confidence is considered as a Bayesian probability (Meyniel et al., 2015) and relates to a specific object. It has two forms: a distributional form which is a probabilistic representation and a summary form, corresponding to the second moment of the distribution. In other words, we can consider that subjective value is an internal representation (distribution) which has a mean (close to the **reported** subjective value) and a variance (reported uncertainty/inverse of confidence). The variance is consequently the inverse precision of the value representation in this framework (Figure 9, middle). Thus, when a subjective value is reported, it is supposed to be drawn from a distribution of mean $\mu$ and variance $\sigma$ .

---

[4] And as we seen, beliefs can influence subjective valuation.

[5] I want to insist on the quantitative similarity between uncertainty and confidence since the former would express before the latter if any dynamics is involved and then they would be qualitatively two different things.

Lebreton et al., (2015) proposed a formal model to account for the U-shape relationship existing between value[6] and confidence (Figure 9, left). First, the model supposes that participants try to minimize the mismatch between their subjective report and their internal judgment. Second, the internal judgment has to be mapped on the scale proposed by the experimenter. Finally, to express their response, subjects have to integrate the noisy information of value on the mapped scale. The analytical solution of this model shows that the given rating corresponds to the first moment of the probability distribution of the mapped judgment while the confidence corresponds to the inverse of the variance. Critically, the model predicts that confidence is a **quadratic function** of the rating (the middle of the scale is used for uncertain judgments). Moreover, it can apply to any kind of judgment such as likeability but also age evaluation of probability estimation. Thus, it can

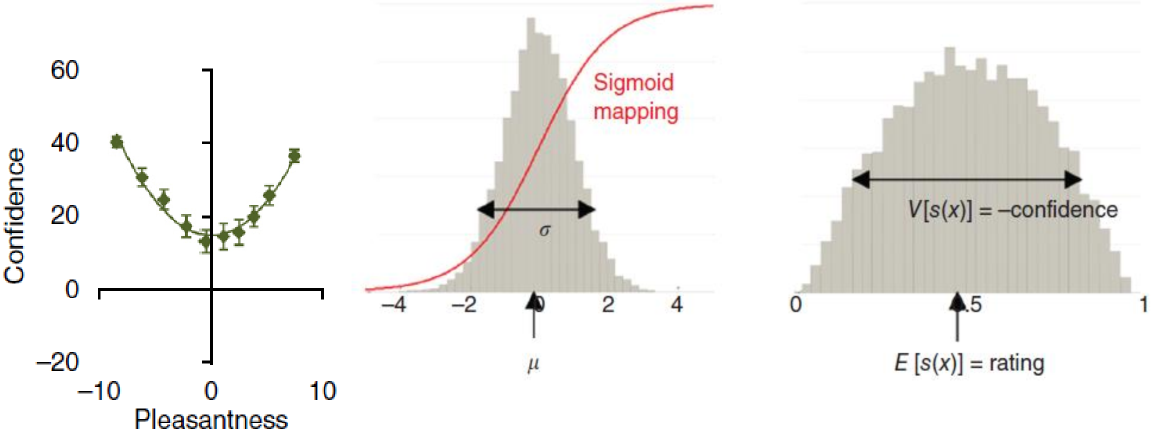


Figure 9 – Confidence as a quadratic function of the rating

**Left**: Classic quadratic relationship between confidence and pleasantness in rating tasks. Middle: Model simulations representing agent's probability distribution over its internal judgment x with a mean $\mu$ and a variance $\sigma$. **Right**: Projection (or mapping) of the subjective probability distribution on the given scale (sigmoidal projection) with the first-order moment E equivalent to the reported subjective value and the second-order moment V corresponding to the reported confidence in judgment.

---

[6] They extend the concept to judgments and confidence about general judgments.

be defined as a second-order judgment or in other words, as a **meta-value**.

On this line, confidence can also be considered itself as a hedonic value. Indeed, we tend to be happy when confident about something and we often try to avoid situations in which we are not confident. Thus, confidence might also be simply seen as a quantification of how much we are happy of our judgement/action.

To sum up, we have seen that values are subjective predictors of rewards which drive the behavior. However, they are not a linear representation of the objective value of the reward (because of many factors such as decreasing marginal utility), they integrate the subjective probability of wining the reward, they are asymmetric according if we are going to win or to lose the reward (gain and loss asymmetry), and they are decreased by the delay to obtain the reward (delay discounting). On top of those economics properties, we have also seen that they are largely affected by internal context (satiety, mood, beliefs), as well as external context (number of options, available alternative, social influence). Finally, studies on confidence allow us bringing precision to the definition of the subjective value by considering it as a probability distribution in which confidence would inform us on the precision of the estimate of the reported subjective value.

Now that I have presented the concept of subjective values, I will review the models developed to explain how those values are used to make decisions.

# D. Models of decision-making

Values are needed to make a decision. Indeed, if I have to make a choice between an item A and an item B, I need to compute the value of A and the value of B, with each value being affected by factors seen in the last part; then, I will need to compare the two values in order to select the item which has the highest value. However, I often mentioned that economists developed their theories to explain inconsistencies in choices. Indeed, observed choices often violate the principle of choosing the highest subjective value ('argmax' rule: choosing the maximum). Thus, how we compare values and how a decision is made is a topic which has yielded a growing interest in the last decades and several kinds of models have been developed to propose psychological mechanisms underlying decision-making. I will focus on two types of models, widely used in the field. The first kind is what I will call '**static models**' which proposes mechanisms to explain choices and the second kind concerns models which explain both decision and time of decision, I will refer to it as '**dynamical models**'. Those two types of models have their origins in psychophysics and they were originally focusing on perceptual decision-making but they have been recently extended to value-based decision-making.

## *Static models of decision-making*

### Basic fitting

As we saw, choices cannot be explained and modelled by a simple deterministic '*argmax*' rule since all factors such as context cannot be taken into account in the computation of the subjective value. Then, one solution is to use probabilistic methods to explain them by fitting a sigmoidal function such as the **softmax function**, already used in learning theories (Luce, 1959). These methods allow drawing psychometric functions which capture the fact that the number of inconstancies in choices increases when values are close (Snodgrass et al., 1985). I will use this kind of logistic fit in most of the experiment I will present to analyze choices, since it is one of the first manipulation check one should do when investigating choices. From this kind of fit, we can have access to two parameters: an intercept indicating if there is any bias in choices and a temperature usually called β, which is capturing the noise in choices[7]. However, this function is not bringing any mechanistic view on how choices are made, and that is why more mechanistic models have been developed.

### Signal Detection Theory (SDT)

The initial purpose of the Signal Detection Theory was to explain **sensory decisions** with uncertainty (Peterson et al., 1954; Tanner Jr and Swets, 1954). It assumes that there are four kinds of possible answers when we have to detect the presence of a stimulus: **hit** (respond yes when the stimulus is present), **correct rejection** (respond no when the stimulus is absent), **false alarm** (respond yes when the stimulus is absent), and **miss** (respond yes when the stimulus is present). From the proportion of those kinds of responses, one can compute the sensitivity *d′* (or discriminability) to the stimulus, which can be seen as a pure perceptual parameter, and a criterion $\beta$, sometimes called 'response bias', which can inform whether the subject is 'conservative' (tendency to say 'no': she wants to avoid false alarms, with the risk of missing some detection) or 'liberal' (tendency to say 'yes': she wants to avoid misses, with the risk of saying yes while the stimulus is absent) (see Figure 10).

---

[7] This same parameter in the context of learning captures a ratio of exploration/exploitation.

This model has been largely used in the field of psychophysics. However, if we keep the idea that subjective value is an internal distribution of probabilities, then when we have to compare two values, the closer they are and the less discriminable they will be (*d′* is small, then the percentage of choosing the wrong option increases), in these cases, misses and false alarms are just one type of error: not choosing the option with the highest value. The decision bias would express (or not) on one or the other option, depending on the context or on priors. This can be transposed to risky decisions: if someone is risk averse, he will

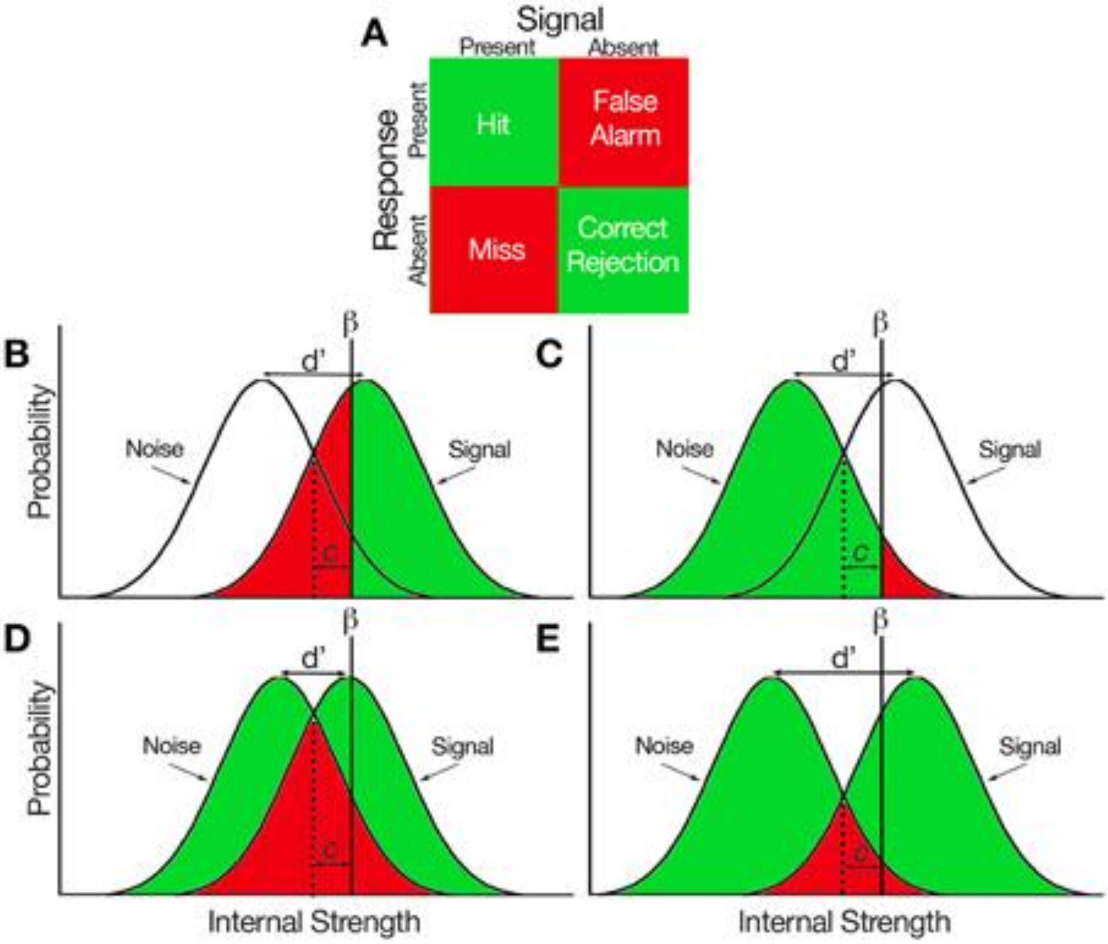


Figure 10 – Signal Detection Theory

(A) Response matrix of possible answers according to the signal. Green means correct responses and red represents errors. (B) & (C): illustration of the response according to internal representation of Noise and Signal in the SDT.(B) Representation of hits and misses, (C) representation of correct rejections and false alarms. In B and C, the criterion $\beta$ is favoring acceptance, thus there are more false alarms than misses. (D) & (E): Effect of d′ on error rate. (D) A small d′ induces a large proportion of errors while (E) a large d′ reduces errors. From (Anderson, 2015) tend to have a bias toward the safe option while if he is risk seeker, the criterion would tend on the riskier option.

We can consider signal detection theory as the mathematical basis for perceptual decision-making. However, since it has been developed to explain decisions related to stimulus detection among noise, it has its limits. Moreover, it does not take into account the dynamics of the decision process. Indeed, it is now well-known that the more difficult a decision is (perceptual or not), the more time it will take. In the next part, I will present how the signal detection theory has been extended to dynamical forms which are compatible with value-based decision-making.

## *Dynamic models of decision-making*

SDT can extent to dynamical models through the idea that at each micro-time point of the general decision process, if the stimulus is present, a piece of information will be added to the previous one, and so on. **Sequential sampling models** have been developed on this idea of **accumulation of evidence**.

### Sequential sampling models

Sequential sampling models (Figure 11) suppose that information about an option accumulates over time until a particular threshold (or boundary) is reached. The time to which this threshold is reached corresponds to the reaction time (RT) and the decision criterion which is reached defines the choice which is made (if there are several boundaries, as we will see later). The accumulation starts at a starting point $S$ and increases with a drift rate $\mu$, corresponding to the available amount of evidence. Crucially, this process is stochastic, which makes the accumulation noisy, thus the accumulation process will vary across repetitions. Then, it is usually choice probabilities and RT distributions which are inferred from this kind of dynamic models. A non-decision time (*Tnd*) can be introduced before the accumulation starts, which can capture the time needed by decision-unrelated processes such as stimulus identification or memory access. Finally, another but more intuitive view of the role of these parameters can be seen as follow: an increase in difficulty would decrease the drift rate $\mu$, an increase in time pressure (or speed-accuracy tradeoff) would decrease the decision criterion A (and B), an *a priori* on evidence would modify the value of the starting point S (such as beliefs), and finally, a stimulus difficult to categorize would increase the non-decision time.

Several versions of sequential sampling models have been developed in the field, I will detail the two main ones: the **Drift Diffusion Model** and the **Race Model** and quickly present some variants.

### *The Drift Diffusion Model*

The Drift Diffusion Model (DDM), also called related random walk model, is a classic type of sequential sampling model, often used in cognitive neuroscience. The main principle of this model is that choices proceed from an accumulation process driven by the difference between option values with a rate $\mu=\Delta V$ $(\Delta V=V_A-V_B)$, (Ratcliff, 1978; Ratcliff and McKoon, 2008). This model can be implemented only with two alternative (A and B) options and has been widely applied to fit behavior during perceptual and memory decision-making (Kinchla and Smyzer, 1967; Link and Heath, 1975; Pearson et al., 2009) .

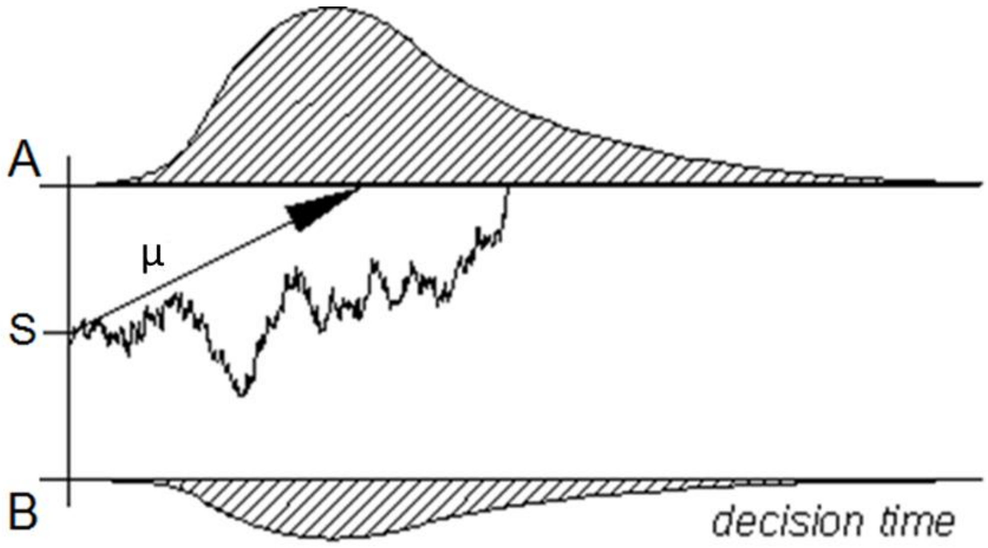


Figure 11 – Sequential sampling model

Schematic representation of a sequential sampling model with two boundaries A and B. Accumulation starts at a point S (here it is not biased toward A or B) and goes on stochastically with a drift rate μ until a boundary (here A) is reached. Reaction Time distributions are drown for each response A and B.

### *Race models*

Race models assume that evidence in favor of each option is accumulated in separate accumulators, or modules (LaBerge, 1962; Vickers, 1970). A decision is made whenever any of the accumulators reaches a boundary. The response time corresponds to the time at which this boundary is reached.

Race models have the advantage to be suitable for situations where there is more than two alternatives, even if the drift diffusion model with more than two alternatives can have a drift rate reduced to $\mu=V_{best}-V_{second\ best}$ (Krajbich and Rangel, 2011) or to $\mu=V_{best}-V_{average}$ (McMillen and Holmes, 2006).

Both Drift Diffusion Models and Race models have already been successfully applied to perceptual but also economic decision-making tasks (Basten et al., 2010; Gold and Shadlen, 2007; Krajbich et al., 2010, 2012; Milosavljevic et al., 2010). Many variants have been developed to account for some discrepancies between model predictions and observations.

### *Variants: collapsing boundaries, decay, competition & attentional effects*

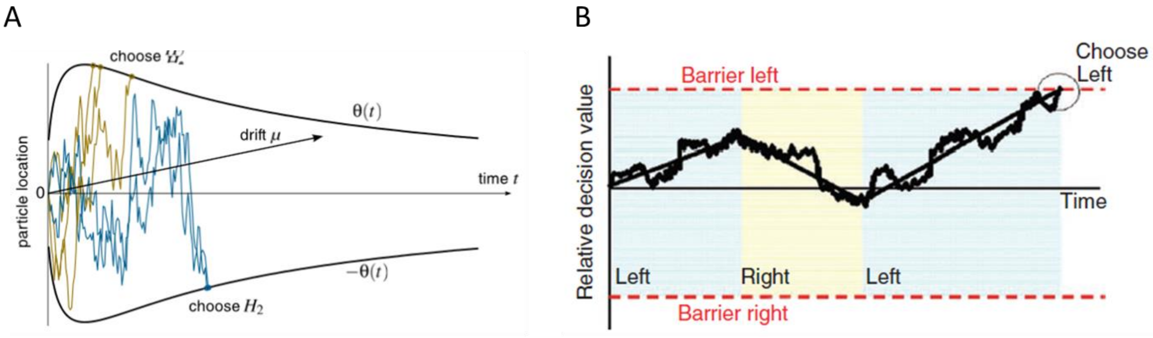
One of the variants of diffusion models is the ‘**collapsing boundaries’**, or ‘urgency signals’ diffusion model which supposes that less evidence is required to reach a decision as time goes. This variant allows accounting for the fact that it is costly to spend time acquiring information to reach a decision. It supposes that boundaries are collapsing as time passes during the decision process and allows the accumulator reaching a boundary even if there is a really small noise and if the drift is equal to zero (Drugowitsch et al., 2012) (Figure 12, left).

The second variant I want to mention here is the ‘**Leaky Competing Accumulator**’ model (LCA) (Tsetsos et al., 2012; Usher and McClelland, 2001). It is a kind of race model (with an accumulator for each option) which supposes that accumulated information is subject to leakage and is consequently degraded over time. It also supposes that each accumulator competes with the others through a process of lateral inhibition. In other words, the higher an accumulator is, the stronger it will inhibit other accumulators. As the collapsing boundaries model, it accounts for imperfections in the information accumulation process and for a choice mechanism which can integrate multiple alternative options.

Finally, several attempts to integrate attention in the diffusion process have been done. The oldest is the **Decision Field Theory** (DFT) (Busemeyer and Townsend, 1993), it is based on the fluctuation of attention and supposes that attention will first focus on the more *important* attribute of the attended option. It also comprises lateral inhibition which is dependent on the distance between options attributes (while in the LCA, inhibition is independent from psychological distances). It has been developed to account for the similarity, the compromise and the attraction effects described earlier (page 33).

More recently, the **attentional Drift Diffusion Model** (aDDM) (Krajbich et al., 2010, 2012) has been proposed and posits attention as the driver of the accumulation process, or more precisely, fixations as a guide for the comparison. It supposes that the drift rate of the accumulation process will be equal to the value of the fixated option minus the value of the unfixated option down-weighted by a parameter $\theta$ (Figure 12, right). This model makes several predictions regarding the link between fixation pattern and choices: for example, the most fixated option has a higher probability to be chosen, but also that the last fixated one.

I have focused here on the dynamic versions of sequential sampling models. Most of these models have an analytical solution allowing us to fit them to behavioral data. Among others, the Linear



Collapsing boundaries

aDDM

Figure 12 – Drift Diffusion Model variants

(A) Collapsing boundaries. Boundaries are decreasing with time. Here the decrease is nonliear, from Drugowitsch et al., (2012). (B) Attentional Drift Diffusion Model: The drift rate is driven by the relative decision value framed according to which option is fixated, left (blue) or right (yellow). The down-weighting of the unfixated option is not represented here (it is in the drift), from Krajbich et al., 2010.

Ballistic Accumulator (LBA) is one analytic version of the race model (Brown and Heathcote, 2008) and the EZ(2)-diffusion (Wagenmakers et al., 2007; Grasman et al., 2007, 2009) is an analytic version of the drift diffusion model. I will essentially use the latter to fit the behavioral data of the third study because of its simplicity.

## Integrating confidence in dynamic models of decision-making

Several attempts to integrate confidence in choice in models of decision-making have been accomplished and we can try to sum up them in two main views. The first view sees confidence as a by-product of the decision process itself and simultaneously to the choice itself. The second view tends to conceptualize confidence as a second order judgment occurring after the decision. I will briefly present these two views.

### *Confidence as a part of the decision process*

Confidence has been conceptualized as a part of the decision process within the framework of the signal detection theory. Indeed, SDT is able to predict proportion of 'yes' responses according to the distance between noise and stimulus distribution, but also confidence in the response by the distance between the mean of the distribution and the criterion (Erev et al., 1994; Budescu et al., 1997). Moreover, confidence has also been formalized in the sequential sampling framework as a product of the drift-rate and the bound (Heath, 1984; Link, 2003; Kiani and Shadlen, 2009), or as the difference in the evidence accumulated by different 'racers' in the context of a race model (De Martino et al., 2013). In the Bayesian framework, a Bayesian version of the Drift Diffusion Model developed by Bitzer et al., (2014) showed that the drift and diffusion parameters can be seen as 'internal uncertainties', which might correspond to a form of confidence.

### *Confidence as a post-decisional process: the 'two-Stage Dynamic signal detection theory'*

In 2010, Pleskac and Busemeyer (Pleskac and Busemeyer, 2010) proposed the **Two-Stage Dynamic signal detection theory**. Starting from the assumption that choices, decision times and confidence ratings are observations from the same unique decisional process, they point the fact that none of the existing models of decision-making are able to predict (or at least to explain) more than two of these primary variables. Indeed, SDT is able to predict proportion of 'yes' responses and the confidence in the response. However, it does not claim anything on decision time. On the other hand, sequential sampling models can predict responses and decision time but does not account for confidence ratings. They propose a model which combines both sequential sampling concept and Signal Detection Theory: the Dynamic Signal Detection Theory. The model first behaves like a classic Drift Diffusion Model, but once a boundary is reached at time point $t_1$, the accumulation goes on and confidence can be read out from the distribution of evidence accumulated (Figure 13) until a second time point $t_2$. It makes the strong assumption that confidence is post-decisional.

### *The problem with confidence in choices*

Those models are considering confidence in the choice itself. However, as we have seen earlier, confidence can express on values, then we have two type of confidence that we can measure when we observe a choice: the confidence in values which are compared and the confidence in the choice. It is still unknown how those two types of confidence relate. Usually, confidence in choice is supposed to rely on a balance between the evidence on the considered options but more recently, it has been suggested that confidence in choice was more weighted by the evidence favoring the chosen option compared to the unchosen one (Zylberberg et al., 2012). However, this asymmetry has not been assessed in the context of value-based decision-making.

All these models, originating mainly from psychophysics mechanisms, have been extended to value-based decision-making. However, they yield some criticisms and they are still not completely satisfying to explain such decisions. Indeed, one would rightfully argue that we do not need to accumulate evidence with a drift equal to value if we already know the values. Another example is that some models are based on strong assumptions: fixations are assumed to be random for the attentional drift diffusion model but their randomness is unproven. Together, the multiplicity of models and their limits show that the mechanisms underlying decision-making are still misunderstood.

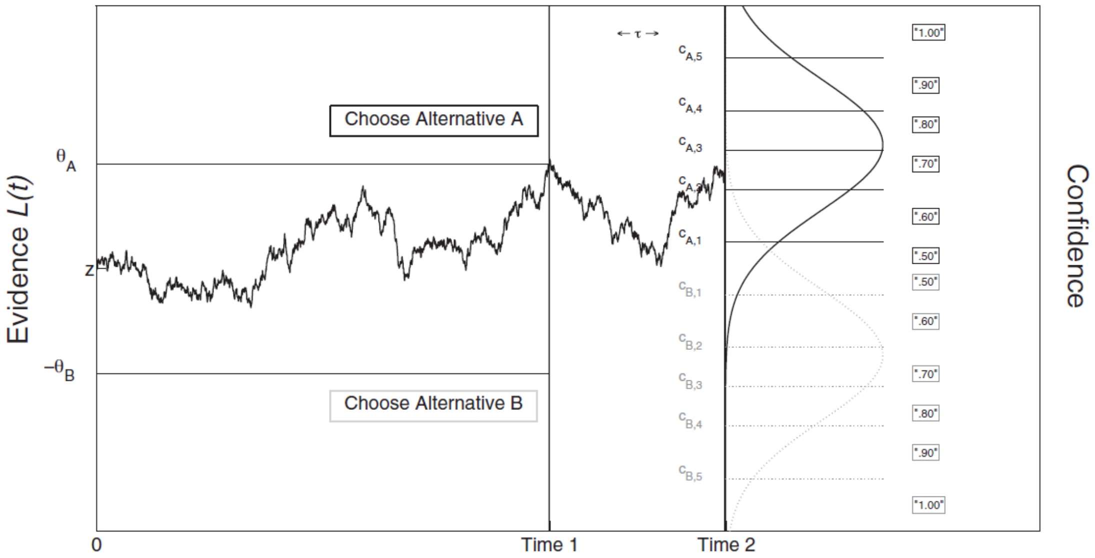


Figure 13 - Two-Stage Dynamic signal detection theory

Until time 1, the model is equivalent to a drift diffusion model with evidence being accumulated until reaching the boudary A. However, accumulation continues after a choice is made and evidence is read out at time 2, following the signal detection theory.

## E. Summary & open questions

In this first section, we have seen that subjective value could be defined as the quantification of how much we desire/like an object. We have also seen to which factors value could be sensitive to and what were the cognitive models proposed to explain how they could be integrated in decision-making processes.

Subjective value sensitivity to external and internal factors might be an issue when investigating them through different measures. Thus, before investigating the neural processes underlying valuation and decision-making, we wanted to verify whether the measures we are using (presented page 20) are inducing deformations in subjective values. Then, rightful questions regarding this issue can be listed as follow:

- Are **value functions** obtained from measurements such as ratings, efforts and choices equivalent?
- If so, is there any deformation induced by these measures in **values**?
- Which one should we use to invest valuation processes?

We addressed these questions in the first experimental study of my PhD using three tasks (rating, effort and choices) and a model-based approach. I will present the design and results in the second chapter of this manuscript (page 118).

In the following section, I will review the main results on the neural processing of values and present the main neural network involved: the Brain Valuation System, as well as its known properties.

# 2. Neural substrate of values

In this section, I will first introduce the fundamental basis of neuroanatomy related to the reward circuitry, then I will present the measurements one can do when investigating neural activity since I will use several neuroimaging techniques in my experimental work. In the two next parts, I will review the literature linked to value coding in the brain and to the established properties of the involved areas. This section ends with a presentation of the questions related to the Brain Valuation System I will address in the second study of this thesis.

## A. Neuroanatomy

Several main brain structures will be studied in this thesis through fMRI techniques and intra-electroencephalography recordings. In order to get a broad picture of the structural and functional organization of the fronto-striatal circuit, which has been identified as the reward network, I will shortly present the basal ganglia, its interaction with cortex and I will finally focus on our main cortical brain area of interest: the ventro-medial prefrontal cortex.

### *Basal ganglia*

The basal ganglia circuit is involved in a large variety of motor and cognitive behaviors, making these brain structures a central point of interest in decision-making. Initially, the function associated to this circuit was to provide a pathway whereby information coming from the associative cortex was transmitted to the motor cortex (Evarts and Thach, 1969; Kemp and Powell, 1971). However, additional circuits organized in parallel of the motor pathway have been discovered (Alexander et al., 1986; Alexander and Crutcher, 1990) and cognitive functions have been associated to each of those circuits. More specifically, the limbic circuit (or 'emotive circuit') was found to be involved in a motivational system related to reward seeking (Mogenson et al., 1980).

Basal ganglia circuits are anatomically organized around several nuclei. The **striatum** (composed of the **caudate**, **putamen** and **nucleus accumbens**/ventral striatum) receives inputs from cortical areas and projects outputs to the **globus pallidus** (and **substantia nigra**), which in turn projects to the **thalamus**. The thalamus then projects to the cortex, which is completing the corticostriatal loop (Figure 14). Two pathways are classically described in this loop to explain effects of striatal activity on thalamus: the **direct** and the **indirect** pathways. The first one can be seen as the 'Go' pathway: the striatum inhibits the internal segment of the globus pallidus (GPi), which tonically inhibits the thalamus. This double inhibition results in an activation (more precisely, in a

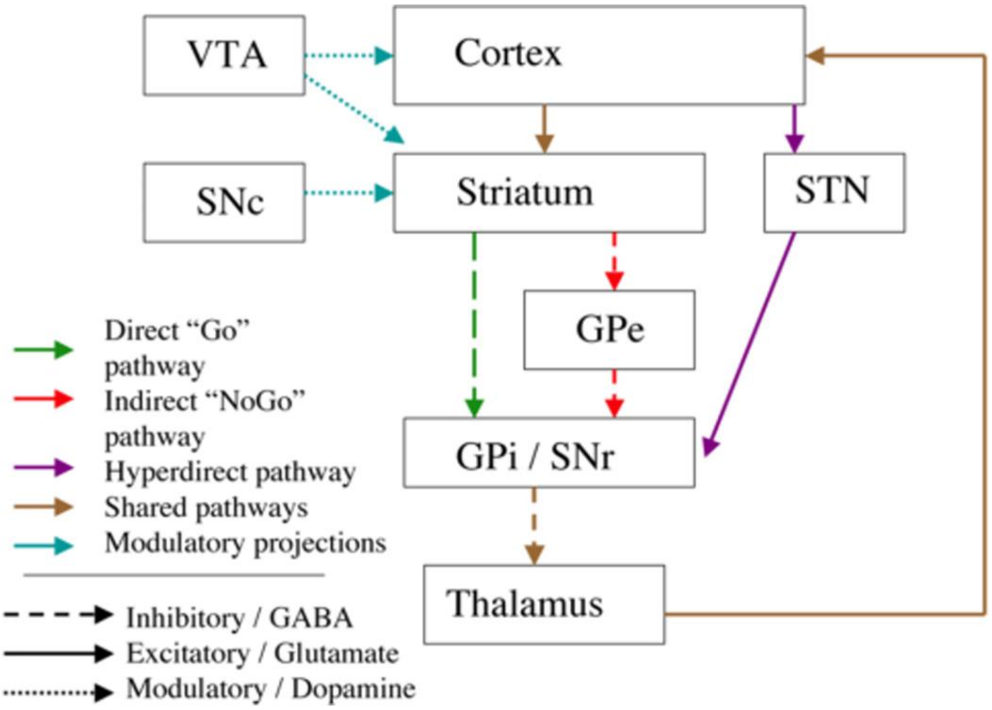


Figure 14 – Cortico - Basal ganglia circuit

Representation of the direct (green), indirect (red) and hypoerdirect (purple) pathway in the cortio-basal ganglia circuit. From (Seger and Peterson, 2013).

disinhibition) of the thalamus. In the indirect pathway, the striatum inhibits the external segment of the globus pallidus (GPe) which is itself inhibiting the GPi (Sato et al., 2000). This creates an increase in the tonic inhibition of the Thalamus by the GPi. Since the thalamus is inhibited by this pathway, the indirect pathway is usually called the 'NoGo' pathway.

Crucially, it is important to precise that the nature of the dopaminergic receptor is different in the two pathways: dopamine receptor of type D1, excited by dopamine, are most prevalent in the direct pathway while D2 receptors, inhibited by dopamine, are most prevalent in the indirect pathway (Aizman et al., 2000). Thus, dopamine has a positive modulatory power on striatum and action selection ('GO' selection).

On top of those two pathways, a third '**hyperdirect**' pathway is worth to mention. First assigned to the indirect pathway (Alexander et al., 1986), the **subthalamic nucleus (STN)** has been reclassified as being a part of the hyperdirect pathway (Nambu et al., 2002), which is also involved in action selection. The STN receives direct cortical projection from the frontal lobe and has an excitatory output on the GPi. Then, the activation of the STN produces an inhibition of the thalamus and slows down the action selection process operated by the direct and indirect pathways. Several studies support the idea that the STN is involved in conflict detection (Bogacz and Gurney, 2007; Frank, 2005) and through its activation, it provides more time to take a decision.

Now that we have a clearer understanding of basal ganglia structure and functional organization, I will present the different cortico-striatal loops and the organization of the cortico-subthalamics inputs.

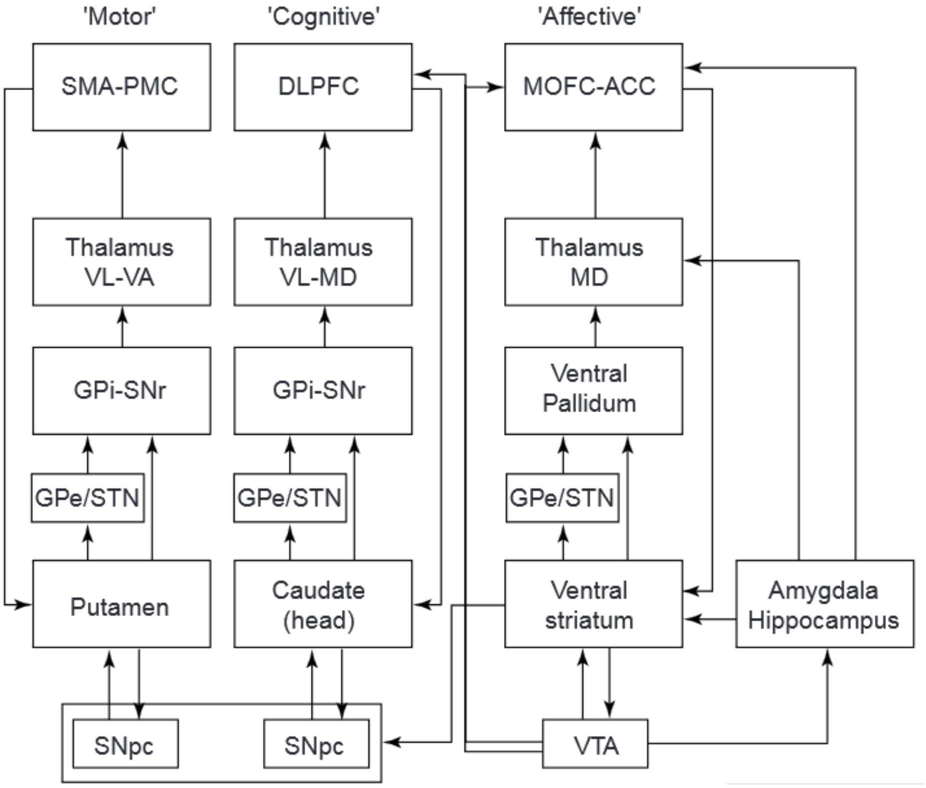


Figure 15 – Schematic representation of Motor, Cognitive & Affective circuits between basal ganglia & cortical areas.

(MD, mediodorsal; VA, ventroanterior; VL, ventrolateral; VM, ventromedial; from (Brown and Pluck, 2000).

## *Connecting the cortex to basal ganglia*

Studies on anatomical and functional connections of cortical areas to basal ganglia have revealed a tripartite organization of the cortico-basal ganglia-thalamo-cortical circuits (Alexander and Crutcher, 1990; Parent, 1990; Parent and Hazrati, 1995). Indeed, three main functional circuits have been defined. The **motor circuit**, involving the dorso-lateral part of the striatum, also called the putamen and the ventro-lateral-anterior (VL-VA) thalamus, receives cortical projections from the Supplementary Motor Area (SMA) and the premotor cortex. A somatotopic organization had been found in the different stages of the motor circuit and it is clearly involved in movement preparation and execution. The **cognitive circuit**, involving the dorso-medial part of the striatum (caudate) and the ventro-lateral/medio-dorsal (VL-MD) area of the thalamus, mainly receives projection from the dlPFC and is associated to executive functions such as working memory and cognitive flexibility. Finally, the **limbic (or affective) circuit**, comprising the ventral striatum and the medio-dorsal thalamus, receives projections from the medial OFC and the anterior cingulate cortex. It is linked to emotional and reward responses. Those different cortical areas project on the basal ganglia but also receive outputs from those circuits to form a cortico-basal ganglia-cortical loop.

Additionally, basal ganglia also receive projections from the **substantia nigra compacta** (SNc) and the **ventral tegmental area** (VTA) Figure 16). Those two nuclei located in the midbrain are formed by dopaminergic neurons and make the dopamine a central neurotransmitter in decision-making.

### *The dopaminergic system*

Dopamine, described as a critical neurotransmitter in the central nervous system (Carlsson, 1959), is highly involved in motivation and movement control. The substantia nigra *pars compacta* (SNpc) and the ventral tegmental area (VTA) contain the cellular bodies of dopaminergic neurons. The SNpc, through a nigro-striatal and meso-striatal pathways mainly innervates the basal ganglia while the VTA mostly projects on prefrontal areas through the meso-cortical pathway and the subcortical limbic system (ventral striatum and amygdala) through the mesolimbic pathway (Figure 16).

Dopamine is known to influence striatal activity through two distinct modes: a phasic activity, which correspond to a transient release caused by dopamine neuron firing; and a tonic mode, corresponding to a background and sustained ongoing activity of dopaminergic neurons

(Grace, 1991).

### *The prefrontal cortex*

We saw that the prefrontal cortex (PFC) is highly implicated in the cortico-striatal circuits and receives projections from the dopaminergic neurons. This makes this cortical brain region particularly interesting for us. Moreover, there are strong evidences supporting the implication of the prefrontal cortex in motivation and decision-making. Indeed, clinical manifestations of prefrontal lesions are highly informative and allow us to distinguish several functional parts in the PFC (Catani and Thiebaut de Schotten, 2012; Stuss, 2011). First, the dorso-lateral parts of the PFC,

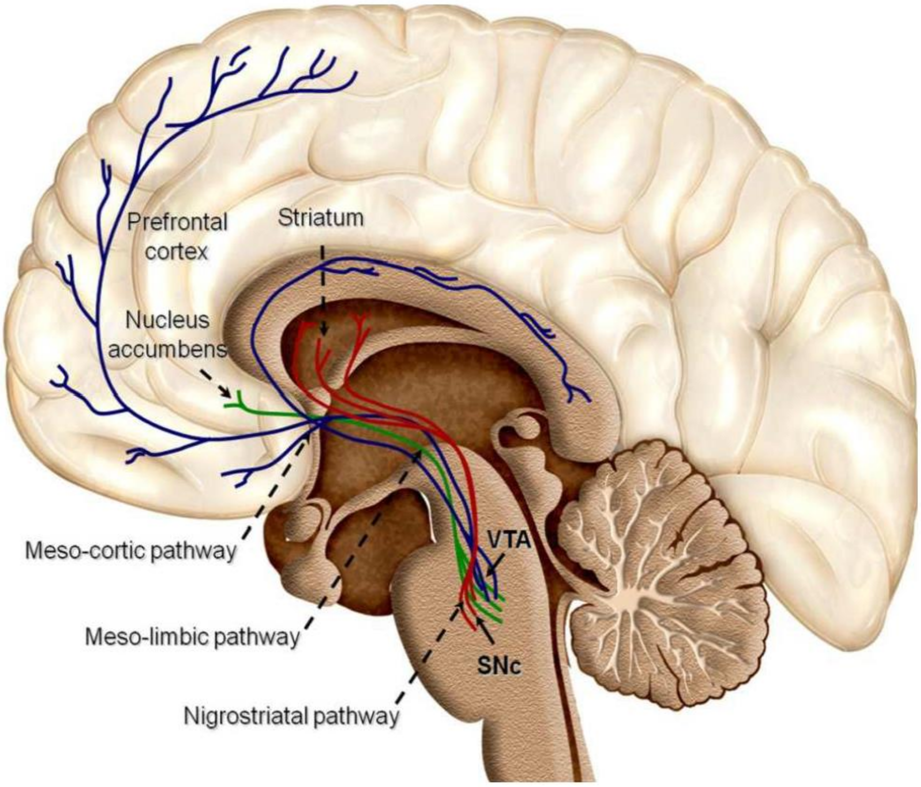


Figure 16 – Schematic representation of the dopaminergic projections

Dopaminergic neurons are located in VTA and SNc and project to the striatum (red), to the frontal cortex (blue) and to the nuclus accumbens (green). From Arias-Carrión et al., (2010) when impaired, induce deficits in executive functions for goal-directed behavior such as planning, rule learning, focusing, hierarchical organization, switching or monitoring. Lesions to the medial prefrontal cortex induce apathy. The affective and cognitive components of apathy are more affected when lesions are anterior while the motor component is impaired when lesions are more posterior. It has been suggested that this medial PFC is associated to an energization process of both cognitive/affective and executive processes (Stuss, 2011). Finally, orbitofrontal cortex (OFC) lesions induce personality changes such as disinhibitions, social inappropriateness, sexual preoccupations but also impair reward and risk processing (Bechara et al., 1999) and affect mood (Catani and Thiebaut de Schotten, 2012). Thus, the anatomical projections and the deficits observed with damages to the PFC make this brain area of particular interest when investigating decision-making processes.

Value-based decision-making studies particularly focus on the orbital and ventromedial part of the frontal lobe. Divisions of the ventral parts of the prefrontal cortex need to be clarified here since I will refer to them in the rest of this manuscript. In humans, it is mostly the medial wall of the ventral part of the prefrontal cortex which is mentioned as the **vmPFC** (ventromedial prefrontal cortex) in reward-related studies. However, in monkey, this region has an equivalent and it corresponds to areas 14m, 32 and 24/25 (Figure 17). In monkey studies, the OFC is most commonly mentioned, but the same areas exist in humans. OFC refers to the more ventral part of the PFC. It is classically divided into the **medial OFC** (mOFC, areas 14c, 14r and 11m, below the vmPFC, see Figure 17), the **central OFC** (cOFC, areas 11 and 13) and the **lateral OFC** (lOFC, areas 12m, 12o) (Mackey and Petrides, 2010).

Before presenting the results regarding value coding in the brain, I will briefly introduce three types of signal one can measure in the brain and the link between them: the firing rate, local field potentials (LFP) and the BOLD signal recorded with fMRI.

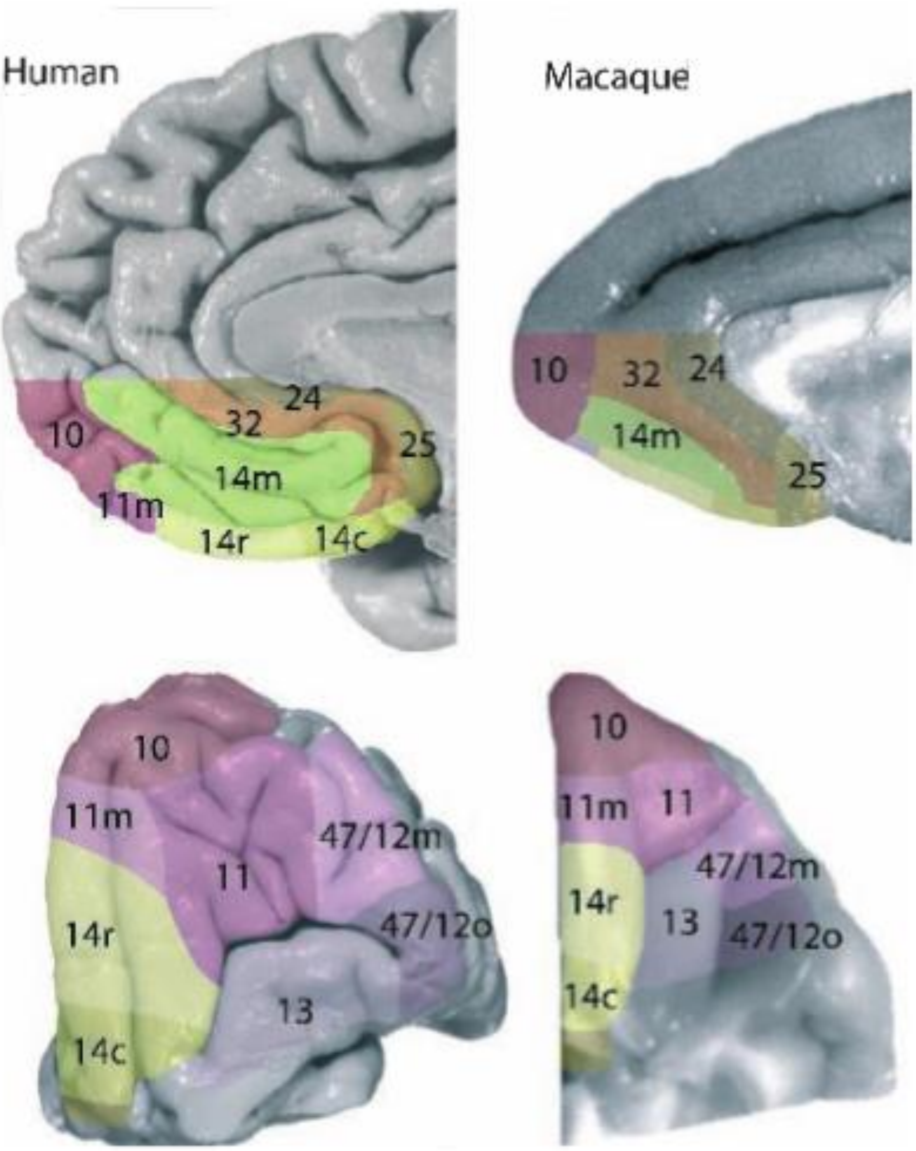


Figure 17 – Architectonic parcellation of the orbital & ventromedial surface of human & monkey brain.

From (Mackey and Petrides, 2010)

## B. Brain activity measurements

*Brain activity* usually refers to **electrical activity** but it can be assess through different levels: the level of neurons as single units, the level of the population of neurons, in which we observe the electrical activity of an ensemble of neurons or the level of a brain area, usually indirectly assessed with fMRI.

### *Single-unit recordings: the firing rate*

Neurons are cells which are communicating through electrical pulses called **action potentials**, or more commonly in the lab '**spikes'**. I won't go into the details of how they are generated since it is not the topic of this part but what we should keep in mind is that one of the smallest unit of 'brain activity' we can record. To record it, we use implanted electrodes in a brain region of interest and record the '**firing rate'**, i.e. the number of spikes per second (through single unit or multiple units (MUA)). According to the type of structures and neurons, the firing rate and the waveform of spikes vary. For example, an excitatory (glutamatergic) neuron will tend to have a different pattern of discharge than an inhibitory (GABAergic) neuron (see the introduction of Mitchell et al., 2007 for more details). As we saw, the pattern of firing can also vary, some being more tonic and other more phasic. Usually, neurons changes their firing rate in response to external stimuli (Schultz, 1997) and it is highly admitted that information about stimulus, but also any kind of behavior is encoded in the pattern of action potentials generated by the brain. However, even if this measure is really precise, it does not allow generating the 'broad picture' of the brain activity and the general dynamics of an area.

At the level of population of neurons, neuronal activity can be more or less synchronized following low or high frequencies of synchronization. It is what we call neural oscillations.

### *Local Field Potentials*

While neuronal spiking is about single or multiple (but still a few) neuron activities at the same time, they do not inform on the activity of neural ensembles. Using bigger electrodes (‘**macroelectrodes’**) allows having access to the ‘Local Field Potentials’ (LFP). LFP are characterized by an oscillatory activity which can be decomposed in frequency ranges from 0.5 to 200Hz. This decomposition includes ***delta*** wave (0.5-4Hz), ***theta*** (6-10Hz) classically found in the hippocampus, ***alpha*** (7.5-12.5Hz) commonly found in the occipital lobe when eyes are closed, ***mu***, which is in the same frequency but classically found in the motor cortex and suppressed when an action is performed. Higher frequencies such as ***beta*** (12.5-30Hz), ***gamma*** (30-50Hz) and ***high gamma*** (50-150Hz)[8] are usually associated to attention, consciousness but their roles are still under debate. Even the fact that they *do* have a function per se is debated. However, neuronal oscillations have been supposed to play a role in neuronal communication (Fries, 2005). Neuronal oscillations are characterized by a phase and amplitude. Fries proposed that brain structures need to flexibly communicate to achieve complex functions and that they might do so through phase-locking (‘Communication-Through-Coherence’ theory)[9]. Moreover, even if he proposed that distant neural ensembles are communicating through lower frequencies, he supposed more recently (Fries, 2015) that lower frequencies such as alpha and beta are involved in top-down mechanisms while gamma and high-gamma would be more involved in bottom-up processes (Engel et al., 2001; Buzsaki, 2004).

---

[8] Ranges are approximate and might depend on the investigated processes and species

[9] It has also been proposed that this phase-locking was used as a ‘tag’ to bind features of objects together and to construct an appropriate representation (Singer and Gray, 1995). This hypothesis considers oscillations as a part of the representational code while the CTC hypothesis does not.

Then, if high frequencies are supposed to be closer to the activity of a local neuronal population, one might rightfully ask what is the link between the spiking activity and the frequencies observed in LFP recordings. This link has been investigated and it has been found that high-gamma oscillations are closely linked to the occurrence of spikes (Ray et al., 2008a, 2008b) *(*Figure 18*)*. From these studies, it has been concluded that high-gamma oscillations could be a useful indicator of the firing dynamics of the recorded population.

High frequencies such as high gamma band are not accessible through classical techniques to record human brain activity like EEG or MEG (Electro-encephalography and Magneto-encephalography) because of artifacts such as electromyographic signals which classically require applying a low-pass filter around 70Hz. Then it is classically investigated in animal studies in monkey or in rodents. However, drug-resistant epileptic patients are sometimes implanted with invasive techniques such as ECoG or iEEG in order to identify the locus of the seizure. ECoG stands for Electrocorticography and uses electrodes which are placed on the surface of the brain, under the skull. It has the advantage to record higher frequencies but the localization of the recorded signal faces the same problem as in EEG or MEG since it needs source reconstruction techniques. iEEG stands for intracranial electro-encephalography and is highly invasive since electrodes are implanted inside the brain, usually in the cortex and in the hippocampus. This kind of data has a no temporal resolution limitation, it can record high frequencies (more than 200Hz if needed[10]) and has a spatial resolution inferior to 1 centimeter. This kind of data is quite rare and the access to the patients to run experiments is obviously very limited.

---

[10] But irrelevant for us: frequencies higher than 200Hz (ripples) are supposed to be a characteristic of seizures (Timofeev and Steriade, 2004).

fMRI has been extensively used in neuroscience to get access to the brain activity with a high spatial resolution and to localize brain processes.

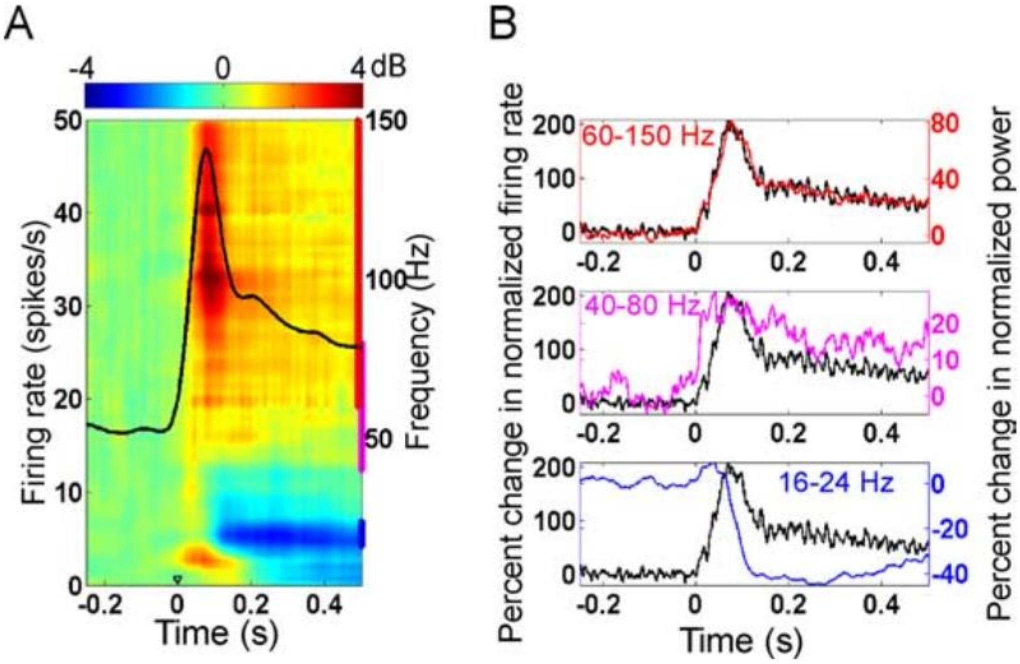


Figure 18 - Firing rate and its relation to frequency bands

A) Time-frequency plot of the change of LFP power relative to baseline. The firing rate of the neuronal population is plotted in black. B) The percent change in normalized power in the high gamma (red), low gamma (magenta) and beta (blue) bands is plotted in function of time. The overlap between the firing rate (black line) and the high gamma power is an argument to consider high gamma power as a proxy for local activity of a neuronal population. From (Ray et al., 2008a).

### *fMRI*

There is a lot to say about neurovascular coupling. fMRI (functional Magnetic Resonance Imaging) is a technique which allows visualizing **indirectly** the brain activity. Indeed, it records variation in the level of oxygen in the blood through the BOLD signal (Blood-Oxygen-Level-Dependent). It is supposed to measure the neural activity since it is based on the assumption that if neurons become active they will consume oxygen. Then, the local blood flow will increase and the hemoglobin will become deoxygenated, which can be measured with the BOLD signal (Huettel et al., 2004). One of the biggest limits of this technique is the temporal resolution. Indeed, the hemodynamics is quite slow and it takes around 2 seconds to hemoglobin to become deoxygenated and the BOLD signal will peak around 6-8 seconds.

Studies combining fMRI and electrophysiological recordings showed that BOLD activity was more related to LFP than single-unit or multiple-unit recordings (Logothetis et al., 2001). However, it has been shown that the high gamma activity was reflected in the BOLD signal, filling a gap between spiking activity and BOLD signal (Mukamel, 2005; Nir et al., 2007).

To conclude, we have seen that various techniques exist to investigate brain activity. Some have a good temporal resolution but a weak spatial resolution, for other it is the opposite. Techniques combining a good spatial and temporal resolution are invasive and rare in humans. I will try to present in the following parts the results concerning the valuation process in the brain through the prism of the three potential existing signals (Single-unit or MUA, LFP, BOLD).

## C. Value coding in the brain

The first studies mentioning a link between reward and brain were self-stimulation studies (Sidman et al., 1955; Valenstein and Beer, 1962) and lesion studies (Butter et al., 1969; Mora et al., 1979; Rolls et al., 1980) in rodents and monkeys. Self-stimulation studies revealed that the

OFC seemed to be involved in reward evaluation. As an example, the rate of self-stimulation of OFC in monkeys increased more when the monkey was hungry compared to when sated (Mora et al., 1979). This illustrates the idea of a ***subjective*** value coding compared to an ***objective*** value coding: food has a lower subjective value when we are hungry compared to when we are sated. On the other hand, OFC lesion studies revealed perseverative behaviors (Iversen and Mishkin, 1970) and an increase in approaching non-food items (Butter et al., 1969). These results suggest a lack of inhibition induced by OFC lesions. However, following studies will show that it is the association between stimulus and reward which is impaired and which leads to response to non-rewarding stimulus and items (Jones and Mishkin, 1972). Later, the first single unit recordings in OFC showed that neurons were responding to the removal of food (Mora et al., 1979) and were assumed to code availability of the reward (Rosenkilde et al., 1981) since their spiking rate was increasing with the presence of reward obtained with a correct response in a short-term memory task.

Since the beginning of the brain reward system investigation, a lot of studies have been published on the topic and I will try to focus on two types of value coding: the expected reward value and the experienced reward value (since both considered as subjective values), in both animal[11] electrophysiological studies and in neuroimaging human studies[12]. As we saw in the neuroanatomical part of this chapter, dopamine is central to the brain reward circuitry, consequently, I will start by presenting results from single unit recordings of dopaminergic neurons. Then, I will focus on electrophysiological data from the

---

[11] Mainly monkey studies for OFC since I do not discuss homologies between human and rodent OFC.

[12] I will focus on studies investigating value coding in the framework of value-based decision making in the last part of this chapter since we are focusing here on the valuation process and not on the comparison (and decision) process

striatum and the OFC. Finally, I will present some influential studies investigating subjective value coding in the whole brain and especially in the vmPFC through fMRI.

### *Dopaminergic value coding*

Midbrain dopaminergic neurons recorded in the VTA are well-known to show phasic response to stimuli predicting a reward in a huge variety of tasks (Schultz, 1997). They specifically code a **reward prediction error** (expected reward – experienced reward (Schultz, 1998)).

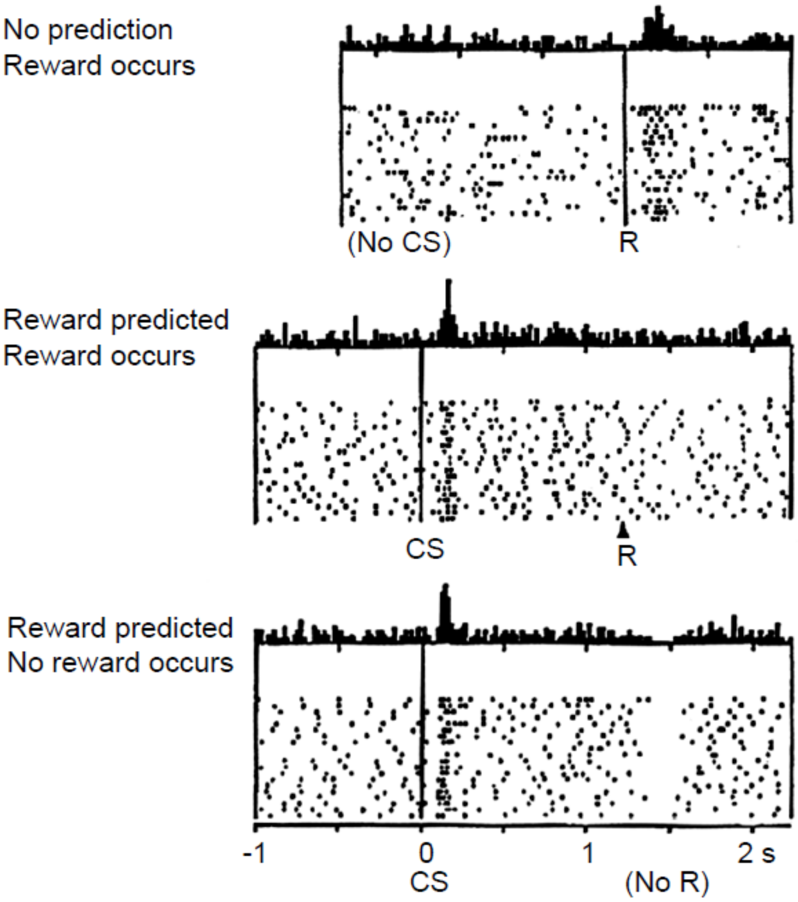


Figure 19 – Reward Prediction Error in dopaminergic neurons.

**Top**: Positive prediction error. When unexpected, reward delivery triggers an increase in spiking activity. **Middle**: Predictive activity & no prediction error. When predicted, the reward delivery does not change spiking activity. **Bottom**: Predictive activity and negative prediction error. When the reward is predicted but not delivered, the firing rate decreases at the time when the reward was expected.

The most famous experiment regarding this property is the study done by Schultz, Dayan and Montague (Schultz et al., 1997). They recorded dopaminergic neurons while monkeys were receiving a reward which was predicted or not by a cue (conditioning task). First, they showed that reward delivery, when unpredicted, was triggering an increase in dopaminergic neurons firing rate. Second, when the reward is predicted, neurons increased their firing rate at the cue predicting the reward but not anymore at the reward delivery. Finally, when the predicted cue occurred but the reward was not delivered (omission), a decrease in the firing rate was observed at the time when the reward was expected (Figure 19). Taken together, these three results provided a strong evidence for a **reward prediction error** being coded by dopaminergic

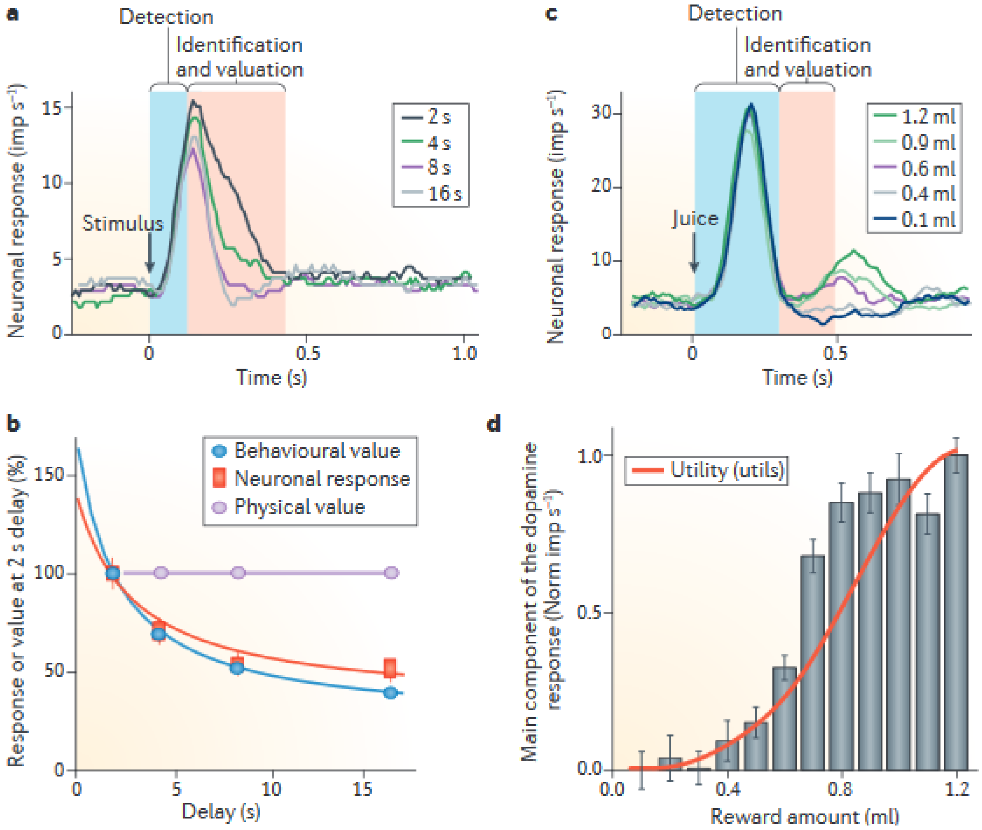


Figure 20 – Properties of dopaminergic neurons responding to reward.

(a & c) Illustration of the two component in the dopaminergic response: detection (blue) and identificiation & valuation (red) for delay (a) and rewards (c). (b) shows the behavioral (blue) and neuronal (red) responses to delay and (d) shows the non linearity relationship between the objective value of the reward (ml) and the subjective value coded by dopaminergic firing rate. From (Schultz, 2016a).

neurons.

The dopaminergic reward signal is characterized by two components: the first one is sensitive to the intensity of the stimulus while the second is sensitive to the value of the reward (respective to its prediction) (Schultz, 2016a; Stauffer et al., 2016) (Figure 20). Moreover, it is important to mention that dopaminergic neurons do not encode the amount of reward but the **subjective value** of the reward. Indeed, it has been shown that the value signal integrated risk (Stauffer et al., 2014), delay (Kobayashi and Schultz, 2008), weighting probabilities (Fiorillo, 2003) and concavity of utility (diminishing marginal utility, in Stauffer et al., 2014, Figure 20).

### *Value coding in striatal neurons*

As we have seen, basal ganglia are central in the Brain Valuation System and are receiving inputs from dopaminergic neurons. One critical nucleus in the basal ganglia showing responses to reward is the striatum, and more precisely, the **ventral striatum** (or Nucleus Accumbens). This nucleus is now well-known to be involved in Pavlovian learning (Day and Carelli, 2007, and page 22) and in creating associations between events and their predictors, in particular between stimuli and reward value. I will first review results from single-unit recordings, then results from LFP recordings and finally I will focus on results about value coding in the human ventral striatum.

#### Single-unit recordings

The recorded striatal signals are less homogeneous than those recorded in the VTA. Indeed, some neurons respond selectively to reward-predicting stimuli (Khamassi et al., 2008) while some respond selectively to reward type, independently of the predicting stimulus

(Tremblay et al., 1998). The dynamic of responses also varies among striatal neurons: some increase their activity until the reward is delivered while others would respond transiently to the predictive stimulus. The large variety of reward-related neuronal response of the ventral striatum suggest that its role is beyond a simple hedonic response (Day and Carelli, 2007).

Value coding in the ventral striatum is however closely linked to the dopaminergic signal. Indeed, ventro-striatal neurons showed response to predictive cues in several studies (Day et al., 2006; Khamassi et al., 2008; van der Meer, 2009), suggesting the ability of these neurons to provide an anticipatory reward signal.

Other studies have shown that striatal neurons were coding reward together with specific **actions**, suggesting that they are able to inform about the value of a chosen action (Dickinson and Balleine, 1994; Schultz, 2016b).

### LFP results

Many studies reported various results about local field potentials recorded in the ventral striatum in rodents (Van Der Meer and Redish, 2009; Kalenscher et al., 2010) and in humans (Cohen et al., 2009a; Lega et al., 2011). One converging result is the involvement of high frequency bands in relation to reward. For example Berke (Berke, 2009) showed that high gamma frequency band (70-90Hz) was significantly increased when a **reward was received** by rats (Figure 21, bottom). More precisely, the spiking activity of some striatal neurons seems to be locked to the phase of the high gamma frequency band (Figure 21, top).

Another study (Kalenscher et al., 2010) investigated gamma oscillations and spiking activity in ventral striatum in rats and replicated those results. Moreover, they also showed that low gamma oscillations

(50Hz) were discriminating reward type (higher power for chocolate mousse compared to sucrose and vanilla) when approaching the reward sites, while high gamma (80Hz) was mainly decreasing after the arrival at the reward sites. However, the results are heterogeneous, and other studies (Van Der Meer and Redish, 2009) showed various different involvement of high and low gamma frequency bands in approaching and receiving reward. Thus, even if it seems clear that spiking activity is closely related to high gamma frequencies, the role of rhythms and their involvement in subjective value coding is not clear.

In humans, in the work of Cohen (Cohen et al., 2009b, 2009c, 2009a), intra-electroencephalographic signals have been recorded in the ventral striatum of patients with major depression while they were performing learning and risky decision-making tasks. These studies provide evidences suggesting that the ventral striatum is reflecting the value of expected reward at the cue onset and the value of the experienced reward at the feedback (Cohen et al., 2009b) in ERPs. Moreover, when investigating the time-frequency spectrum of those ERPs during a similar task (Cohen et al., 2009c), they found an involvement of gamma band (40-80Hz), as in rodents studies. They also showed that bursts of gamma band were coupled to alpha oscillations (8-12Hz), differentially for gains (following alpha bursts) and losses (preceding alpha bursts).

Even if these studies bring huge insights about the ventral striatal response to reward, it is important to precise that the interpretation of 'value coding' in the ventral striatum is inferred from contrasts such as "winning versus losing" and not from true correlations (or regressions) between the observed signals (iERP or frequency band power). Moreover, those results are observed during learning tasks involving actions and decisions about cues and not items with a 'natural' subjective value.

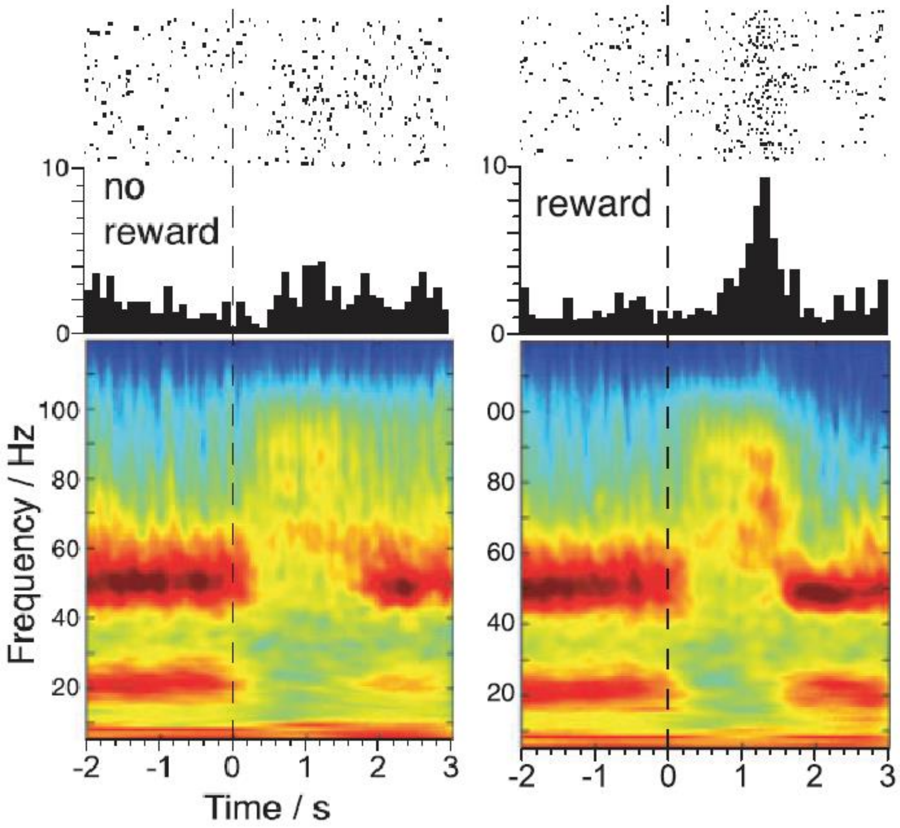


Figure 21 – Firing rate and LFP signal related to reward in the monkey ventral striatum.

When a reward is received (right), there is an increase in high gamma activity to which spiking activity seems to be loked on. From (Berke, 2009)

## *Orbitofrontal Cortex & Subjective value coding*

### Single-unit recordings

The first studies with monkey OFC neuron recordings showed that their activity was related to the **consumption of a reward** (Thorpe et al., 1983) and that the spiking activity was varying according to satiety, suggesting that they were coding for the subjective value of a reward (Yaxley et al., 1988). Moreover, the decrease in spiking activity with satiety for one type of reward had been shown to be specific to this type of reward (Rolls et al., 1989). Since these first results, a lot of studies investigated how OFC neurons were coding subjective values.

As in the ventral striatum, OFC neurons respond to the cue and seems to code for the **expected value** of the reward (Tremblay and Schultz, 2000a, 2000b), but also to the **experienced value** of the reward (Kennerley and Wallis, 2009). However, distinctions between medial and lateral part of the OFC have been investigated and it seems that while both medial (mOFC) and lateral (lOFC) OFC neurons are responding to visual cues predicting rewards, mOFC neurons appeared to be more sensitive to internal factors such as satiety and lOFC neurons were more sensitive to external factors such as the cue (Bouret and Richmond, 2010).

Moreover, if lOFC and striatal neurons have similar responses to reward and predictive cues, one would rightfully ask how their responses differ. One interesting study (Simmons et al., 2007) investigated this question in monkeys performing a reward schedule task (discriminating green and red colors and being rewarded for correct trials in a complete schedule). They replicated the result about neurons responding in both areas to expected and experienced reward. Moreover, they also found that reward information was coded earlier in lOFC neurons compared to ventral striatal neurons but also that the OFC response peak was occurring in the 'wait' period of the trial while in the ventral striatum, it was during the 'action' period. These results provide insights in the information flow of reward value coding and processing but also support the idea that OFC would be linked to Pavlovian conditioning (stimulus-outcome association) while ventral striatum might be more linked to instrumental conditioning (action-outcome association) (Ostlund and Balleine, 2007).

### Local Field Potential results

Similarly to the ventral striatum, LFP recorded in the OFC in rats revealed a response of gamma frequency to reward anticipation with a phase coupling with lower frequency oscillations such as delta (Donnelly et al., 2014) or theta frequency band (van Wingerden et al., 2010). A coupling between theta band and spiking activity of rats OFC neurons when a reward was received by rats has also been found by Horst and Laubach, (2013), suggesting that experienced reward also involved OFC neurons. In monkey, reward anticipation was found to be coded in the high gamma band (50-100Hz), but also in the beta band (15-29Hz) but with a reverse coding: power in beta band was higher for low rewards compare to high rewards (Zhang et al., 2016). LFP recordings in human medial and lateral OFC showed similar results with both an increase in LFP amplitude with reward probability and experienced reward value, especially in the lateral OFC for the latter (Li et al., 2016).

Thus, one converging result in OFC is an involvement of high frequencies such as high gamma in reward and outcome evaluation. This high gamma coding has often be associated to a decrease in lower frequencies such as in the beta band, but it has been characterized as a less specific signature of value coding compare to high gamma (Jung et al., 2011).

### *Investigating the Brain Valuation System with fMRI*

While electrophysiological studies on subjective value and its neural processes have the advantage to bring precise results on the dynamics and on the neuronal activity at a local level, fMRI studies have the advantage to group those areas in a brain networks since it allows investigating BOLD response in the whole brain.

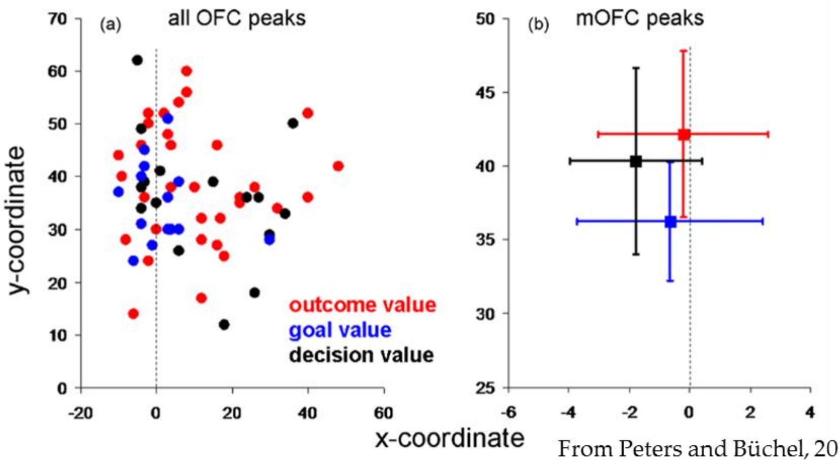
A large number of studies and reviews have shown that the BOLD signal in the vmPFC and in the ventral striatum is correlated with subjective values of choice alternatives (see the last part of this chapter, page 102), of items to evaluate or to bid on (Lebreton et al., 2009;

Plassmann et al., 2007), but also with the experienced subjective value of the reward when received (O'Doherty et al., 2001). Bartra et al., (2013) conducted a qualitative meta-analysis confirming that the vmPFC and the ventral striatum, but also the Posterior Cingulate Cortex (PCC) BOLD signal positively correlates with expected and experienced subjective values in a linear fashion. However, the vmPFC seems to respond stronger in the valuation stage compare to the receipt time. The ventral striatal activity presents the same positive correlation for item values with positive valence but it is the opposite for penalties, suggesting a U-shape relationship between item value and BOLD response. Moreover, there was no evidence of stronger response at one stage or another of the task trials (evaluation vs receipt).

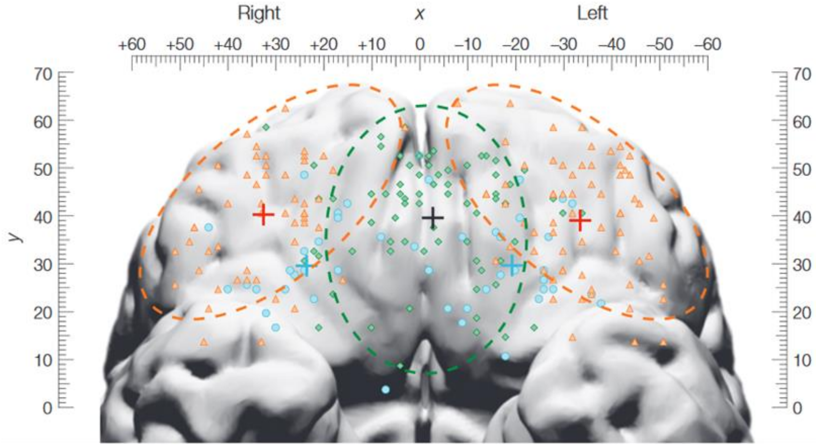
In the following parts, I will first focus on the OFC and point the difference between the medial and lateral part of it, then the differences between the OFC and the ventral striatum to get a clearer view on the Brain Valuation System main components. Then, I will present the most common 'partners' of the vmPFC and ventral striatum to define the BVS in its entirety.

### Medial & Lateral OFC: the debate

In Bartra's meta-analysis, the lateral OFC did not show up in the results of subjective value coding. However, we have seen that numerous electrophysiological studies found neurons coding reward value. Another meta-analysis (J. Peters and Büchel, 2010) investigated the percentage of contrasts revealing subjective value coding in mOFC, lOFC and ventral striatum in humans among 45 studies. All type of subjective values together, they found that vmPFC BOLD signal was correlated with values in 76% of contrasts investigated, while 43.5% for lOFC and 28.3% for ventral striatum. However, no clear distinction could be found between lOFC and mOFC according to the type of value (expected reward, experienced reward or even subjective utility including costs) or to the sensory aspect of items (gustatory, olfactory, primary, abstract…) (Figure 22, top). This is contradicting a previous meta-analysis (Kringelbach, 2005)(Figure 22, bottom) suggesting a distinction of value coding according to the valence, with the vmPFC involved for positive values while the lOFC involved for negative values associated to punishment. This distinction has been interpreted as the vmPFC being



From Peters and Büchel, 2010



From Kringelbach, 2005

Figure 22 – Similarities & Differences between vmPFC & OFC value coding.

**Top**: Peak coordinates (left) and average peak coordinates (right) for acitvations related to outcome value, goal value and decision value in the human OFC. **Bottom**: Medio-lateral distinction between reward value coding (green) and punishment (orange) and antero-posterior distinction for abstract (green) and primary (blue) reinforcers.

monitoring and holding in mind reward values while lateral OFC as being involved in a more transient valuation process through response suppression (Elliott, 2000; O'Doherty et al., 2001; Kringelbach, 2004). However, there are a lot of studies showing that this distinction does not hold and other theories have been developed. One interesting study suggested that lateral OFC would code the value of **external** stimuli while more medial parts would be involved in evaluating the value of information derived from the structure of the task (Rich and Wallis, 2014). Another group used multivariate pattern analysis to investigate value coding in lOFC and vmPFC (Howard et al., 2015) and found that value could be decoded from both parts of the OFC but the pattern of lOFC activation was heterogeneous (not linearly varying with value) and was signaling both the identity and the value of items[13].

Finally, discrepancies observed between lOFC and vmPFC between species and techniques might have several origins. Wallis (Wallis, 2011) proposed that it could be due to function, as we have seen (such as response suppression, external information, pattern coding…), but also to the nature of the tasks (learning tasks for animals, with animals being trained for a long time) or to the technique itself (BOLD could increase because of an increase in inhibitory neurons activity). Indeed, the methodology used to investigate value coding differs dramatically between humans and monkeys. One could suppose that fMRI is more sensible to value signals in vmPFC than in OFC because of artifacts due to air-tissue boundary close to OFC. fMRI studies also investigate OFC signal at the group level, leading to a decreased sensitivity for regions showing high inter-individual variability, thus, if OFC activity is more variable across subject, one would not see any activation of this region at the group level. Another difference lay on the

---

[13] Making it invisible to classical parametric analysis aimed at looking for correlation with value.

measure itself: electrophysiologists usually investigate single-unit recordings, which correspond to action potentials of individual neurons, while in human studies the BOLD (blood oxygen level-dependent) response, close to local field potential (LFP), is investigated and reflects activity of an area instead of single neurons. Because of the excitatory or inhibitory nature of neurons composing these brain regions, the interpretation of activation through LFP or BOLD signal can be compromised and the conclusions coming from the different measures can be opposed. Finally, it has been shown that OFC neurons can have different selectivity, meaning that some of them are going to respond positively to reward stimulus and others are going to respond negatively. Such a pattern of firing could also lead to a cancellation of signal in fMRI.

### OFC & ventral striatum

On the other hand, Peters and Büchel, (2010) found that the ventral striatum had a different pattern of response according to the type of value: while 0% of the contrasts for expected reward value revealed the ventral striatum 72.7% revealed it for the decision value (expected reward with costs).

fMRI studies which investigated **reward prediction error** correlations with BOLD signal often found an engagement of the ventral striatum, without a co-involvement of the vmPFC (McClure et al., 2003; Abler et al., 2006; Hare et al., 2008; Pessiglione et al., 2006; Lin et al., 2011).

Electrophysiological studies showed that midbrain dopamine neurons are expressing reward prediction error and few showed that striatal neurons were coding it. Thus the observed response in BOLD signal to reward prediction error in the ventral striatum is likely to be due to the dopaminergic projections the ventral striatum is receiving.

Motivation paradigms, in which the expected reward is coupled with behavior activation such as effort tasks also showed ventral striatum activations without a coupling with the OFC (Knutson et al., 2001; Schmidt et al., 2012).

Thus, it seems that even if both OFC and ventral striatum are representing subjective values, ventral striatal activity seems closer to the action value and decision time than OFC value coding related activity.

## The Brain Valuation System & its partners

Depending on contexts and design of all the studies which investigated the Brain Valuation System, several partners of the vmPFC and ventral striatum have been identified: the posterior cingulate cortex (PCC), the hippocampus, the dorso-lateral prefrontal cortex (dlPFC) and the dorso-anterior cingulate cortex (dACC) for the most common areas found involved in the valuation process.

### *Posterior Cingulate Cortex*

The most common brain area responding with vmPFC and ventral striatum to subjective value is the PCC (Bartra et al., 2013). Its activity has been found to correlate positively with the subjective value of items in many rating and choice tasks.

The role of PCC is not clear in the neuroscientific literature. It is a central part of the Default Mode Network (DMN), meaning that it is systematically activated when subjects have to rest in the scanner. It has been proposed that it is linked to the retrieval of autobiographical memories, but also to the focus of attention (Leech and Sharp, 2013). It is also known to have a strong connection to the vmPFC through the cingulum bundle (Greicius et al., 2009). Together, those results allow us to emit several hypotheses on the involvement of this area with the Brain Valuation System: it could help to retrieve personal information on items to compute a subjective value, or help focusing attention of items which are preferred[14]. Another hypothesis is that PCC is processing sensory signals to recognize items and connectivity between PCC and vmPFC would allow computing the subjective value of those items. This hypothesis is supported by a study using tACS[15] on frontal and parietal cortex (Polanía et al., 2015) which wanted to assess whether fronto-parietal communication was causally involved in decision-making. They showed that when frontal and parietal cortices were desynchronized, value-based choices were less precise than when they were synchronized. The localization of the stimulations is not precise with this technique but it allows proposing functional hypothesis on the couple vmPFC/PCC. Further investigation is needed to help understanding its precise role in the valuation process.

### *Hippocampus*

Even if the hippocampus is not revealed by the previously mentioned meta-analysis, this brain structure seems to be an important partner of the Brain Valuation System. Indeed, it has been shown that hippocampus has a critical role in the valuation of imagined outcomes (Barron et al., 2013; Lebreton et al., 2013; Peters and Büchel, 2010).

Usually, the hippocampus is known for its role in episodic memory coding and retrieval (Huijgen and Samson, 2015) but it has been suggested that the hippocampus can bias decisions by creating associations between items and events occurring together and then reactivating neural representation of a rewarding event A when an event

---

[14] Even if those processes do not trivially imply a correlation between the level of activity and the subjective value.

[15] Transcranial alternating current stimulation : used to induce specific neural oscillations

B is presented because they were previously associated (Wimmer and Shohamy, 2012). It has also been supposed that hippocampus biases decisions toward more remembered items through an increase in connectivity between hippocampus and ventral striatum (Gluth et al., 2015).

In a more general manner, it is supposed that the vmPFC and the hippocampus are interacting when elaborating values which require episodic retrieval or simulation.

### *Dorso-lateral Prefrontal Cortex*

The dorso-lateral prefrontal cortex is usually found to be involved in cognitive control (Kouneiher et al., 2009), such as working memory (Barch et al., 1997) and emotion regulation (Goldin et al., 2008). In value-based decision-making, sometimes we have to exert control to choose the option which is best for us (for example healthy food) and not the option that we really prefer (chocolate bar). How the values of options are modified by the control we need to exert has been investigated by Hare, Camerer and Rangel (Hare et al., 2009). They showed that people who integrated healthy features in their food valuation had their vmPFC activity which was correlated to the global utility (taste and healthy) conversely to other people. They also observed that dlPFC activity increases when people exert self-control in a decision between a tasty and not healthy item and a healthy but less tasty item. Moreover, they found that the vmPFC was functionally connected to dlPFC through a third area close to the dlPFC (Inferior frontal Gyrus / BA46) (Figure 23). More recently, another but related interaction between dlPFC and vmPFC had been shown (Rudorf and Hare, 2014) in a task where the valuation of options was dependent of the context.

However, the results regarding the dlPFC activity involvement in valuation are mainly observed during choice tasks (Figner et al., 2010) and more rarely during simple rating tasks. Thus, even if the dlPFC activity could affect the vmPFC activity in the valuation process, it is not clear if this is the valuation or the decision process which is affected by the dlPFC.

### *Dorso-anterior cingulate cortex / Dorso-medial prefrontal cortex*

In Bartra's meta-analysis, the positive effect of subjective value revealed an involvement of the vmPFC. However, this activity extended to the anterior cingulate cortex (ACC) or more precisely, the dorsal ACC (dACC) (Figure 24, A). Interestingly, the negative effect of subjective value revealed the dorso-medial prefrontal cortex area (dmPFC, Figure 24, B) and the conjunction with the first contrast showed an area in between: the dorso-anterior prefrontal cortex (Figure 24, C).

Already mentioned in 2000 by Schultz (Schultz, 2000) as an area with neurons responding to reward delivery, dACC and its link with value-based decision-making and valuation have been extensively investigated. Numerous evidences in non-human primate studies have suggested that dACC neurons were responding to expected and experienced value of rewards (Amiez et al., 2005, 2006; Wallis and Kennerley, 2011), but also to action values (Rudebeck et al., 2008; Rushworth et al., 2011) and costs (Sallet et al., 2007).

In human studies, a large number of experiments have suggested that the dACC was involved in the valuation of alternative options (when engaged in a default state for example) and in the exploration/ exploitation tradeoff by coding exploration value (Boorman et al., 2013; B. Y. Hayden et al., 2011; Kolling et al., 2012; Quilodran et al., 2008; Tsetsos et al., 2014). The dACC also seems to respond to conflict and difficulty and it has been proposed that it is involved in the computation of an '**expected value of control**' (Shenhav et al., 2013, 2014; Shenhav and Botvinick, 2015), with the idea that it provides the amount of control which is necessary to exert in order to complete a more or less difficult task.

As the dlPFC, dACC's role in the valuation process is not clear since often associated to a choice. Situations where the dACC is involved without having to make a choice or an effort are rare and that is why I will come back on the role of this area in the last part of the chapter,

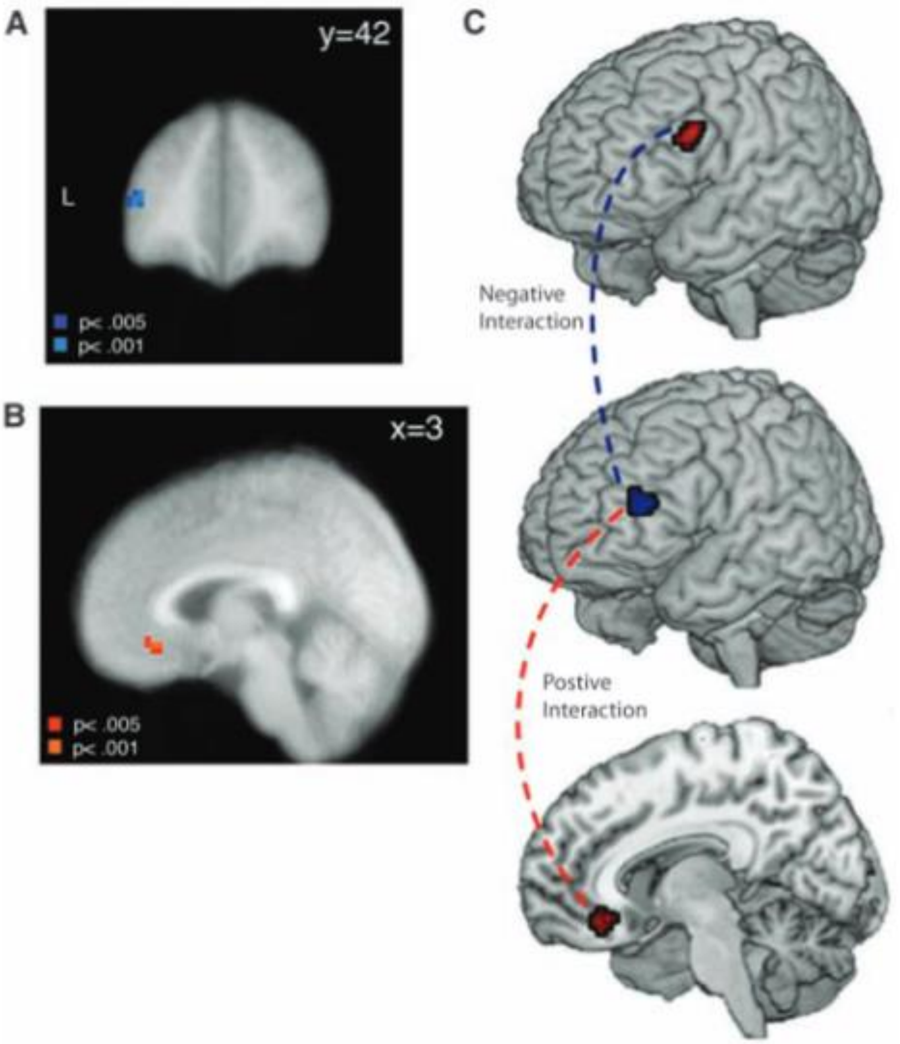


Figure 23 – Interaction between dlPFC and vmPFC during a choice task.

A) Left IFG showed a negative connectivity with the left dlPFC during decision about unhealthy items. B) vmPFC showing subjective value coding and positive connectivity with the left IFG. C) Schema representating the modulation of the vmPFC activity by the dlPFC through the IFG.

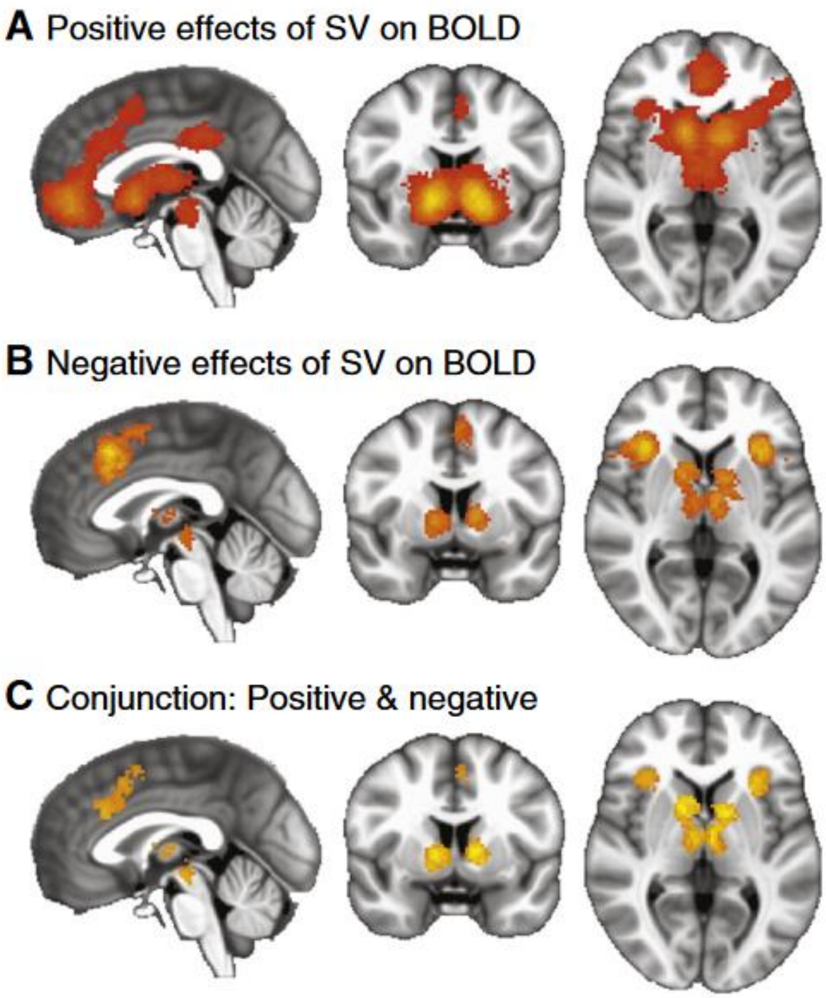


Figure 24 – Networks involved in subjective valuation.

From Bartra's meta-analysis (Bartra et al., 2013). A) Positive effects of subjective value; B) Negative effects; C) Conjunction between positive and negative effects.

which presents the problematic of the neural bases of the comparison process.

### *The Mirror Neuron System & its interaction with the BVS*

As mentioned page 35, social context can have an influence on subjective valuation. In Lebreton et al., 2012, they showed that a desired object by another agent would be more valuable than the same object not desired. The effect is called 'Mimetic desire' and they showed that it was involving an interaction between the BVS and another network: the Mirror Neuron System (MNS), composed of parietal lobules and premotor cortex among other such as the temporo-parietal junction, an area classically involved in social tasks (Grèzes et al., 2003). For more details on the neural basis of subjective values in social decision-making, see Ruff and Fehr, (2014).

Overall, these findings suggest that there exists a brain network composed critical for valuation at different stages of the decision process. This network can be legitimately called the 'Brain Valuation System'. It is composed of the ventral striatum, which seems to code values closer to actions than to stimulus; of the OFC mostly investigated in animal studies, which seems to be involved in associating a subjective value to an external stimulus and in which the parametric coding of subjective value is still debated; and of the vmPFC mostly studied in human experiments (with fMRI) and which might be the 'hub' of the subjective value processing. In other words, it seems that this area is integrating both external information about reward and internal information such as satiety. Close links exist between those areas and other brain regions such as the PCC, the hippocampus, the dlPFC or the MNS and we can suppose that external information the vmPFC needs to integrate is provided by those partners. However, it is still not clear how this network is involved during decision-making. Nevertheless, interesting properties have been found regarding it, and I will present them in the following part.

### D. Specific Properties of the Brain Valuation System

As we have seen, there is a large number of studies supporting the idea that the vmPFC, the ventral striatum and the PCC, sometimes together with the hippocampus are forming the Brain Valuation System. This network, and particularly the vmPFC, has several interesting properties which I will discuss in this section. Indeed, it is **generic**, meaning that it codes the value of items independently of the category they belong to, it also seems to be **automatic**: it responds to subjective value even if participants are not engaged in a valuation process or value-based decision. Finally, it is also highly dependent of the external and internal **context**.

### *'Genericity': Common neural currency*

One striking emergent observation from the previous mentioned studies is that the Brain Valuation System seems to encode subjective values of any kind of category of objects. This hypothesis has been tested specifically in a study where subjects had to make choices between trinkets, snacks or money goods and a reference amount of money. They found that the value of items was correlated to the vmPFC BOLD signal independently of their category (Chib et al., 2009) (Figure 25, A). This result has been replicated in other studies (Levy and Glimcher, 2011), but also with faces, paintings and houses (Lebreton et al., 2009)(Figure 25, B), and even to self-generated stimuli (which could be objects, people or situations in Cunningham et al., 2011)(Figure 25, C). Social stimuli value have also been shown to be correlated with the vmPFC activity in several studies (Lin et al., 2011) and also with the ventral striatal activity (Izuma et al., 2008). Levy and Glimcher (Levy and Glimcher, 2012) proposed that altogether, these results were strongly supporting the idea that a '**common neural currency**'[16] is used by the brain to allow comparison between goods in a same 'unit', it can also be considered as the neural signature of the economic utility described in the first part of this

---

[16] I often use the neologism 'genericity' to refer to this property.

chapter.

This property has been extended to the level of modalities, showing that subjective value associated to **odors** (Howard et al., 2015) and **music** (Abitbol et al., 2015) was also correlated with the vmPFC activity.

However, there are some studies revealing a category-dependent coding of value in the human vmPFC. One of them showed that it is possible to decode value independently from its category in the vmPFC BOLD signal, as previously shown, but it was also possible to decode category dependent values in a more ventral part of the vmPFC (mOFC) (McNamee et al., 2013). Moreover, they found clusters in the lateral OFC

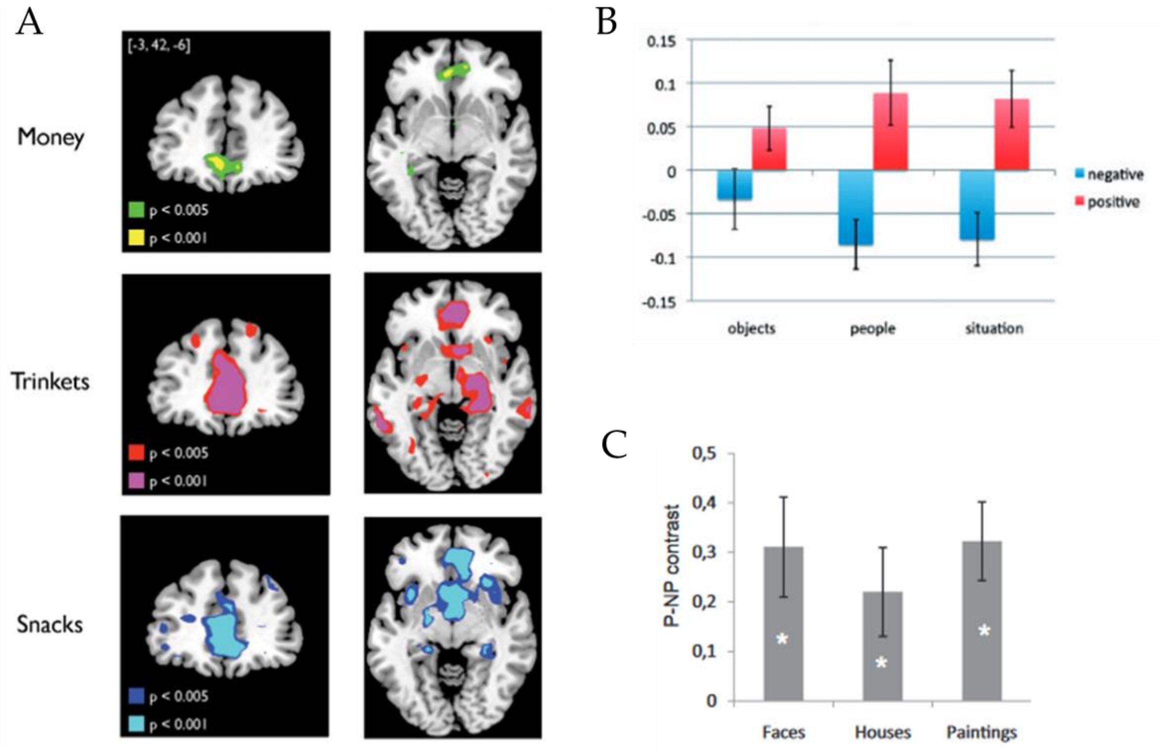


Figure 25 – Subjective value coding is generic in the vmPFC

(A) Areas with activity correlating with subjective values in three conditions: money, trinkets and snacks. vmPFC is the only area involved in the valuation of the three categories of items, (from Chib et al., 2009). (B) vmPFC BOLD response to three categories of items: objects, people and situation for two conditions: positive and negative items (from Cunningham et al., 2011). (C) vmPFC BOLD response for preferred minus less preferred items from three different categories: faces, houses and paintings (from Lebreton et al., 2009).

in which value coding was category-dependent, which has been replicated in another study (Howard et al., 2015) and confirmed in a meta-analysis investigating the processing of primary and secondary reward in the human brain (Sescousse et al., 2013). However, in this last meta-analysis, they showed an anteroposterior gradient coding depending on the nature of the reward, with primary reward such as erotic images being coded in more posterior parts while more abstract rewards such as money are coded in more anterior parts of the OFC. This result is supported by another study from the same group (Sescousse et al., 2010) and seems to be more specific to the lOFC compared to the mOFC (Kringelbach, 2004).

At the neuronal level, this property seems less univocal than at the fMRI level. Indeed, one of the most important study in the field (Padoa-Schioppa and Assad, 2006) showed that OFC neurons (lOFC) could be divided in three categories: **offer values cells**, responding to values of specific juice; **taste cells**, responding in a binary manner to the identity of the chosen juice and **chosen value cells**, coding the value of the chosen juice, independently of its category (negatively or positively) (Figure 26). Thus we see here that some cells in the lOFC have a generic coding of value while others have a category-dependent value coding. Interestingly, the proportion of cells presenting the latter profile is higher than the former in this study; consequently, it seems congruent with what had been observed in fMRI studies.

The generic property of the subjective value coding in the OFC is supported by another interesting result from a study of Tremblay and Schultz, (1999) and go further: they showed that an OFC neuron responding to a preferred good A compared to B was responding less to B, but more to B when opposed to a less preferred good C. This is suggesting that OFC neurons can adapt their firing to the context (and I will come back later on that) but most importantly, it shows that the same neurons are responding to different type of goods. However, there are neurons coding the value of specific juice which have been found to respond independently of which other juice was presented with the first

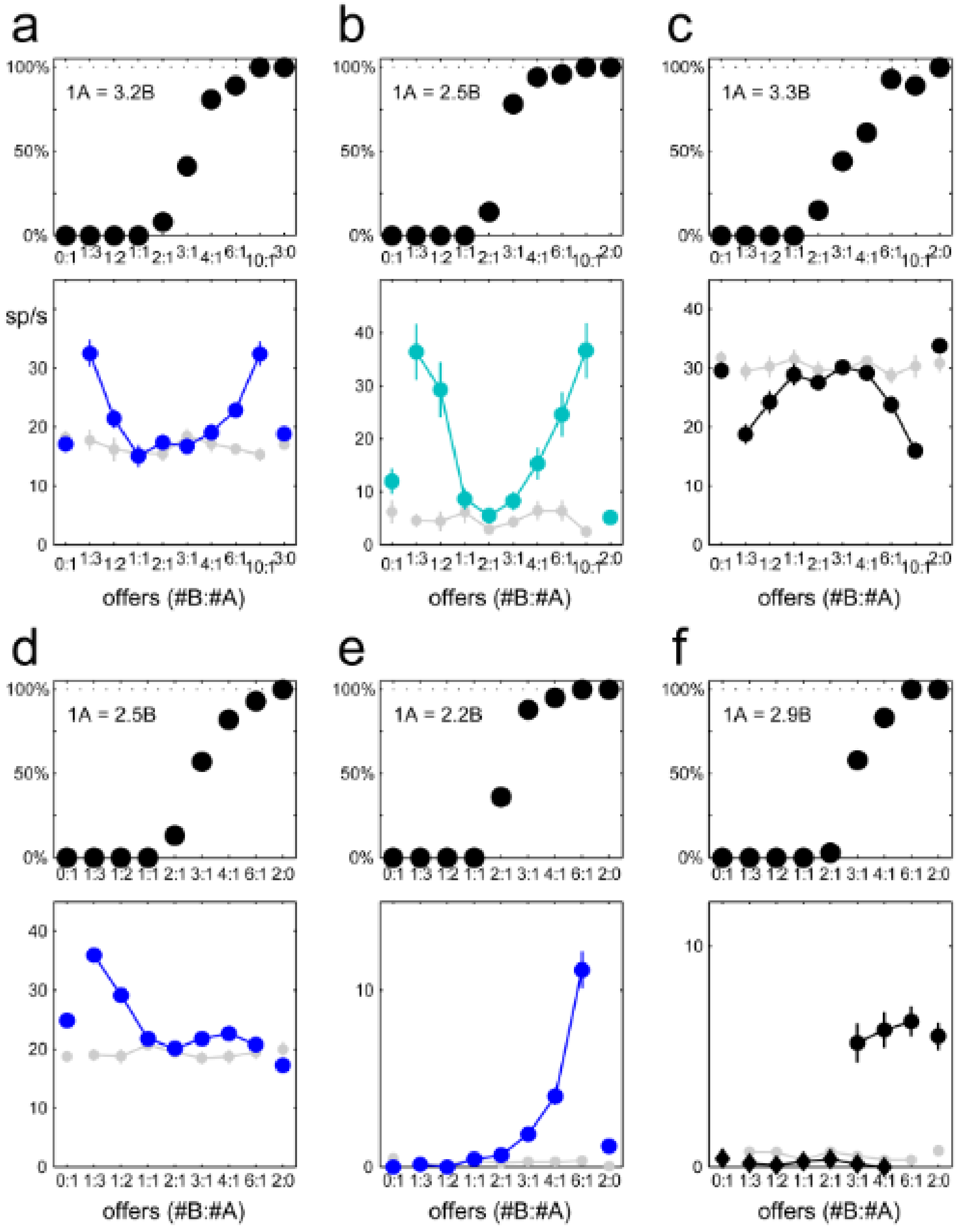


Figure 26 – Three types neuronal responses in the monkey OFC

Six neuron responses (bottom) associated to their choice pattern (top) according to the relative value. (a-c) Cells with a 'chosen value' profile, with (a) and (b) responding positively and (c) negatively. (d-e) Cells showing a 'offer value' profile. (f) Cell showing a 'taste' profile with a binary response (from Padoa-Schioppa and Assad, 2006).

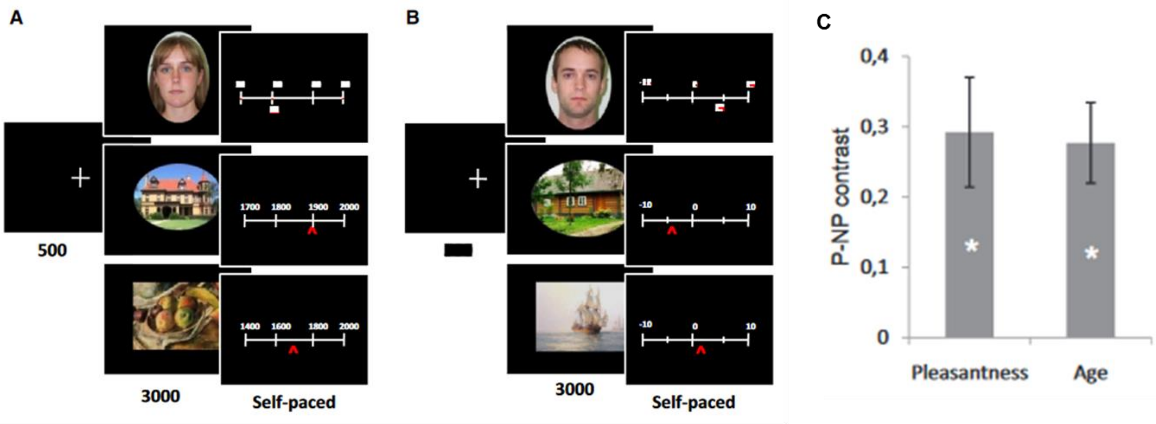
one (Padoa-Schioppa and Assad, 2008).

Thus, it seems that the lOFC have multiple responses in both humans and monkeys and that the code used by the lOFC seems specific to each category of object. Consequently, even if the value of all kinds of items is represented in this region, it is not done in the same way. However, results in the vmPFC indicate that value of all stimuli is linearly represented in this brain region. This allows reaching a more consensual conclusion about the vmPFC being representing a **common neural currency** during valuation process.

### *Automaticity*

We have seen that the Brain Valuation System is computing the subjective values assigned to items. However, most of the studies investigating this topic are using tasks related to subjective values such as rating tasks, choice tasks or effort tasks. Thus, it is unclear whether the BVS is coding value because it is behaviorally relevant for the task or because it is its function, independently of the task, as suggested by Padoa-Schioppa and Cai (Padoa-Schioppa and Cai, 2011) for the OFC.

In order to specifically test the hypothesis that value coding is automatically done by the BVS, a study has been conducted in the team (Lebreton et al., 2009) in which subjects had first to evaluate the age of faces, houses and paintings and then (in another block) to indicate how much they liked the items using rating scales. When investigating the BOLD signal responding to preferred versus non-preferred items during the age evaluation task, they found that the vmPFC activity was significantly stronger for preferred items, even if subject where asked to evaluate the age of the items (Figure 27) .



Age task Pleasantness task vmPFC SV coding

Figure 27 - Subjective value coding is automatic in the vmPFC

A) Age evaluation task: subjects had to indicate the age of faces, houses or paintings using a scale. B) Pleasantness evaluation task: subject had to indicate how much they liked presented items. C) Contrast value for preferred versus not preferred items in both task.

This result has been supported by a previous study where the vmPFC responded to face values even in a task not related to value. The ventral striatum had the same pattern of activation except that it was true only if it was the first time that faces were seen, contrary to the vmPFC (Kim et al., 2007). Moreover, it has been replicated in other studies using music CD cover and DVD cover (Levy et al., 2011) or even images of cars (Tusche et al., 2010) replicating in the same time the genericity property of the Brain Valuation System.

However, this property has to be cautiously considered since there are studies supporting the opposite claim. Indeed, Grueschow et al., 2015 showed a dissociation between value coding in the PCC and in the vmPFC with the former representing values in a similar manner when they are relevant or not for the task and the latter representing values only when a choice has to be made. Another study showed that the vmPFC response to value was reduced when attention was focused on another feature of the item than pleasantness (Grabenhorst and Rolls, 2008). Then, the role of attention in the valuation process might be of importance, and I will come back on this aspect in the last part of this chapter with the neural implementation of the attentional Drift Diffusion Model (presented page 47). Finally, it has also been shown that the vmPFC was not coding value during forced bid trial comparing to free bid trials (Plassmann et al., 2007, 2010) which is against the 'automaticity' property.

To conclude, there are evidences toward and against an automatic value coding in the vmPFC and the origin of these discrepancies between results has not been explained yet. While attention could be one explanation, it is hard to find clear differences in the attentional processes involved in the mentioned studies. On the other hand, the term 'automaticity' might be somehow misleading and we do not claim that the vmPFC response to value cannot be suppressed. However, one potential explanation is that studies which do not observe this automatic property are using additional stimuli (as forced bids) which might also trigger valuation (since automatic (?)) and blur the correlation with the first stimulus value. Another group of results supporting the claim for an automatic value coding in the vmPFC regards the effect of context on the vmPFC activity, which I present in the next part.

### *Contextual influence*

As in the behavioral effect of context on valuation and decision-making (page 30-35), several types of contexts can be discussed. I will first focus on the **internal context** and then on the **external context** effects on neural value computation.

#### Baseline effects

Ongoing activity of a brain region recorded by electrophysiology has been shown to explain variability in cortical evoked response following a stimulus presentation (Arieli et al., 1996). Moreover, spontaneous fluctuations in firing rate and in gamma LFP are likely to be at the origin of observed fMRI fluctuations (Nir et al., 2008) and the BOLD response to an event seems to be dependent on the **spontaneous fluctuations** (Fox et al., 2006). In relation to this spontaneous variability, one might suppose that the variability observed in the behavior for different processes (from perception to decision-making) might be related to the cortical spontaneous variability. This hypothesis has been tested by Hesselmann et al., (2008) in a perceptual decision task. They used the famous Rubin's vase-face picture and asked participants to indicate whether they were perceiving a face or a vase while they were in a scanner. They found that the pre-stimulus activity in the FFA (Fusiform Face Area, commonly known to respond to faces) was higher when participants perceived a face compared to when they were perceiving a vase (Figure 28). This result is supported by other studies in somatosensory perception (Boly et al., 2007) and auditory detection

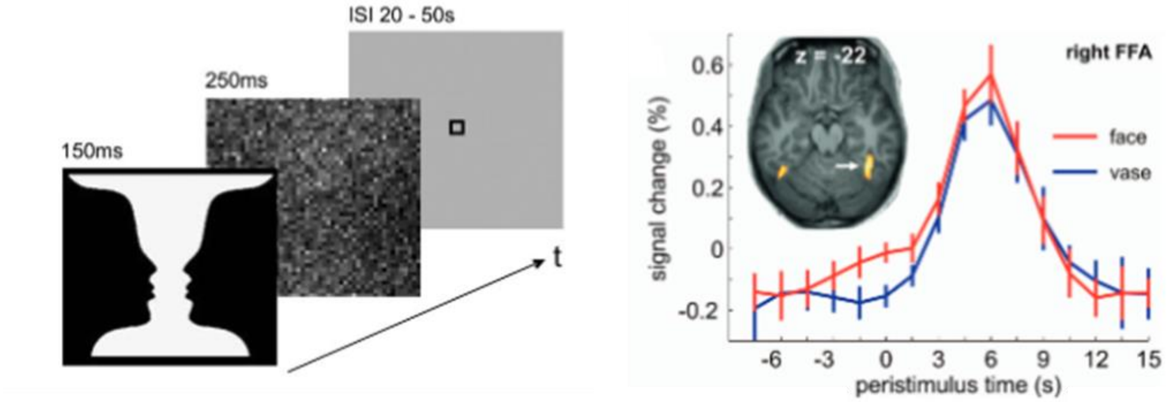


Figure 28 – Pre-stimulus activity impact on perception.

Left: Subjects saw the Rubin's vase-face picture for 150ms and were asked to indicate whether they perceived a vase or a face. Right: Pre-stimulus activity (-1s and 0s) in the FFA was higher when a face (red) was perceived compared to when it was a vase (blue). From Hesselmann et al., (2008) (Sadaghiani et al., 2009). It has also been shown that baseline activity could predict choices independently of subjective values in ‘offer value’ cells (Padoa-Schioppa, 2013) and choices in more executive tasks (Soon et al., 2013).

Consequently, if the Brain Valuation System has spontaneous fluctuations, as the other brain regions, and if it is computing subjective values, we would expect to observe a baseline-dependency of subjective value coding in this network. This hypothesis has been tested in the team in both human with fMRI and monkeys with electrophysiology (Abitbol et al., 2015). In humans, the vmPFC baseline activity was manipulated by music (which is replicating the automaticity property) and it created a systematic bias in subjective valuation (the higher the baseline was, the higher the rating) (Figure 29, top). In monkeys, the baseline activity was affected by the number of trial (satiety or fatigue effect) and induced the same kind of bias. Thus, this study is showing that across species, baseline activity has an influence on the expressed subjective value and that baseline is affected by context, both internal and external such as satiety and music respectively.

As mentioned earlier, mood is known to influence judgment. Then, we should observe a link between the Brain Valuation System and the mood. An interesting study showed that it was the case, with positive mood increasing the activity specifically in this network during reward anticipation (Young and Nusslock, 2016)[17].

Taken together, those results regarding spontaneous fluctuations, integration of fatigue or satiety and mood support the idea that the Brain Valuation System is integrating information about internal context and that it might affect the computation of subjective value (See Figure 29,

---

[17] A paper by Vincquier et al is in preparation in the team to show that mood relies on vmPFC (and anterior Insula) baseline activitiy.

bottom for a proposed mechanistic account by Abitbol et al., 2015).

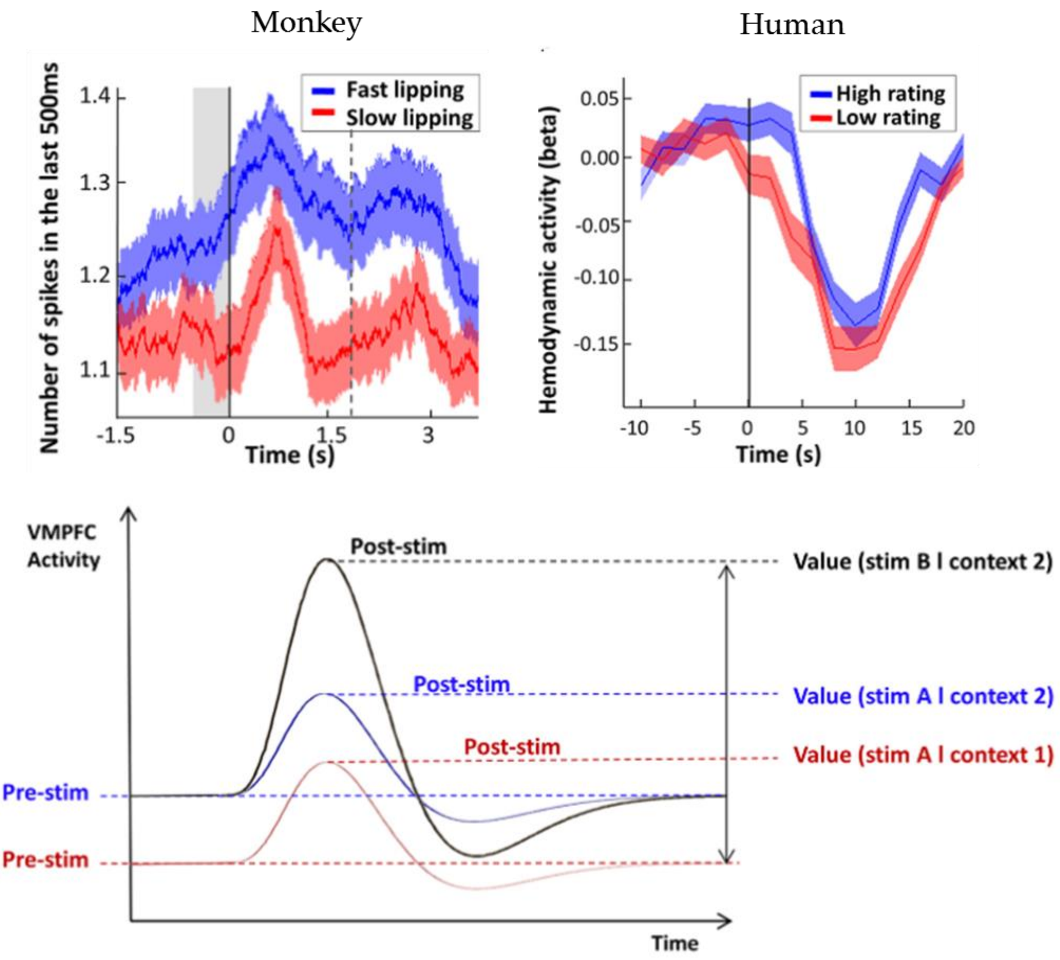


Figure 29 – Dependency between ongoing & evoked activity in the context of subjective valuation.

Top: Baseline activity is higher for high (blue) subjective value in monkey spiking activity (left) and human hemodynamic activity (right). Bottom: Schematic summary of the theoretical vmPFC baseline effect on subjective valuation. If pre-stimulus activity is increased by a context (2, blue), the effect persists in the evoked response and induces a shift in the value of the stimulus A compared to when it is presented in another context (1, red). (From Abitbol et al., 2015).

### Form range dependence value coding to divisive normalization

One simple but not trivial question one might raise is how the Brain Valuation System represents value for different ranges of values. In other words, how the same system can represent values of cars and values of candies while the ranges are very different. A **cardinal** representation of utility as the price of objects seems to increase the difficulty in representing value since there is an infinite number of potential outcomes to evaluate. Conversely, the brain has a limited capacity and if a neuron is able to code for the whole range of possible values, its discriminability would be very weak. Thus, one hypothesis is that value is coded in an **adaptive** way, using a representation of value according to the context, or to the considered range.

Tobler et al. (2005) tested this hypothesis by recording midbrain dopamine neurons in monkeys while looking at visual stimuli associated to reward and probability. They found that the activation of dopaminergic neurons increased with the expected liquid volume and responded to reward prediction error, replicating past results. Moreover,

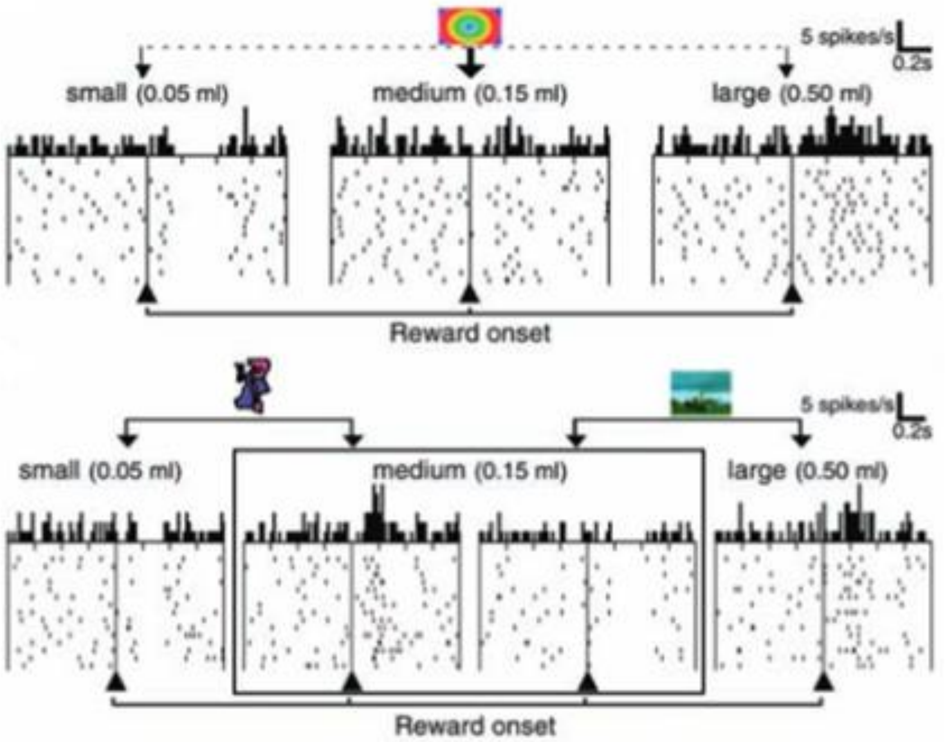


Figure 30 – Adaptive coding of reward value by dopamine neurons.

**Top:** Reward prediction error coding: dopamine neuron response at the reward onset, following a cue indicating a medium reward. Its response decreases if the reward is smaller than expected (left) and increases if the reward is higher (right). It does not change if the reward is medium (middle). **Bottom:** Adaptive coding: According to the reward associated to the cues, a same neuron will respond positively or negatively for a medium amoun of reward. From Tobler et al. (2005)

when they investigated spiking activity in response to a stimulus predicting either a small or a medium reward compared to a stimulus predicting either a medium or a large reward, they found that in both case, the delivery of a larger reward (than expected) increased the spiking activity and the delivery of a smaller one decreased it. Thus, in one case, the receipt of the medium reward increased the spiking activity while in the other case, it decreased it (Figure 30). This result suggests an adaptive coding of reward in midbrain dopaminergic neurons.

The range adaptation property has been extended to the expected reward value coding in monkey OFC neurons (Kobayashi et al., 2010; Padoa-Schioppa, 2009), and it has been recently proposed that it was an ubiquitous property of value coding in brain, with the **divisive normalization theory** (Louie et al., 2011, 2013; Glimcher, 2014). The main idea of the divisive normalization coding is that the neural representation of the value of an item is dependent on the value of other available options in the context, creating a close dependence between the considered item and its environment.

As the influence of spontaneous activity, range adaptation through normalization has been observed in a various set of processes and brain areas such as response in primary visual cortex (Heeger, 1992) and auditory cortex (Rabinowitz et al., 2011) but also in higher processes such as multisensory integration (Ohshiro et al., 2011). It has been supposed that subjective value could be represented at the neural level in the same way that sensory information: through a relative coding. Louie et al., (2011) showed that it was the case in lateral intraparietal (LIP) neurons in monkey performing a perceptual decision-making task; indeed, in this study, LIP neurons represented saccadic values through divisive normalization.

## Effect of priors and external context on valuation

Other effects of context have been observed at the neural level, such as effect of prior information. For example, when participants are presented a cheese-like odor in a fMRI scanner and told that it was either cheese or body odor, they were preferring the odor with the label cheese, and vmPFC activity was responding higher to this label compared to the 'body odor' one (de Araujo et al., 2005). The same kind of results have been observed in another study (Plassmann et al., 2008) in which subjects were had to evaluate the taste of three different wines, associated to low or high prices. They found that reported subjective pleasantness was modulated both by taste and prices, and more importantly, that the vmPFC activity was higher for the higher price compared to the lower price for the same wine (Figure 31).

Thus, beliefs about value, or at least information given prior to the delivery of the reward, can induce a change in the subjective value representation in the vmPFC and in the behavioral report.

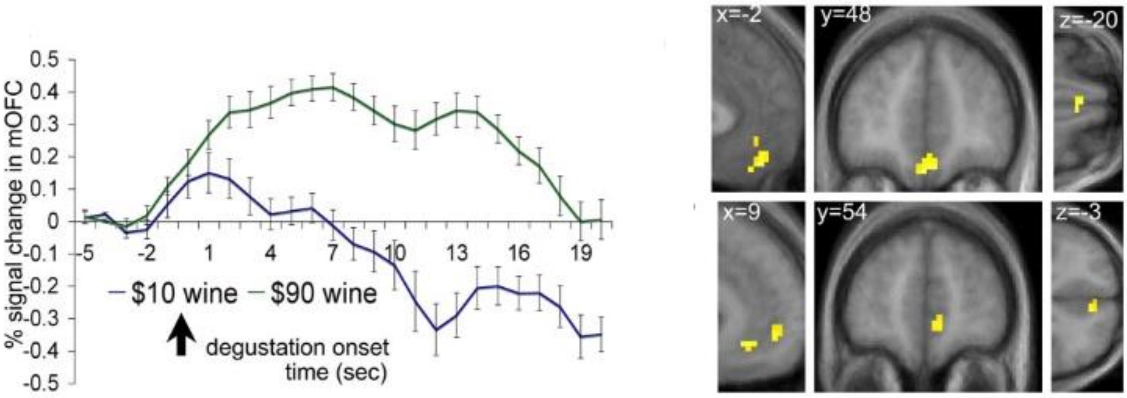


Figure 31 – Effect of prior on vmPFC activity

The vmPFC activity is increased by the consumption of a wine A. The increase is higher when subjects believe it is a 90$ wine compared to when they believe it costs 10$. From (Plassmann et al., 2008)

As proposed in the first part of the chapter, confidence in closely linked to the valuation and decision processes (see page 37). Thus it can be supposed that confidence expresses on top of the value in the Brain Valuation System.

### *Confidence in value: a meta-value expressed in the Brain Valuation System?*

Studies on confidence essentially focused on decision confidence, i.e. confidence in choices. Confidence can be measured through subjective reports or inferred from reaction time, performance or decision difficulty (Festinger, 1943; Kepecs and Mainen, 2012). Indeed, it is classically observed that confidence decreases with difficulty and that reaction time are inversely proportional to confidence. Then, when looking for confidence in the brain, one can look for correlation with easiness of the choice or reaction time but this is an indirect measurement. Several studies suggested that confidence is represented in the vmPFC (De Martino et al., 2013; Schnyer et al., 2005) but other found confidence coding in lateral intra-parietal (LIP) neurons (Kiani and Shadlen, 2009) which makes this result possibly dependent on the nature of the task, of the stimuli, or even on the definition of confidence itself (certainty or second order judgment for example).

A recent study in the team suggested that vmPFC was involved in confidence coding as a meta-value (Lebreton et al., 2015). They showed that the signal observed in the vmPFC during likeability rating has a profile which seems to combine a linear value coding and a quadratic value coding (which would correspond to a confidence signal) resulting in an asymmetric U-shape (Figure 32). This particular profile was observed only with subjective value since the vmPFC activity had a quadratic coding with probability evaluation or age estimation without a linear coding.

Thus, even if this needs to be replicated and further investigated, we can add the 'quadratic coding' property to the genericity, the automaticity and the anticipation properties of the vmPFC.

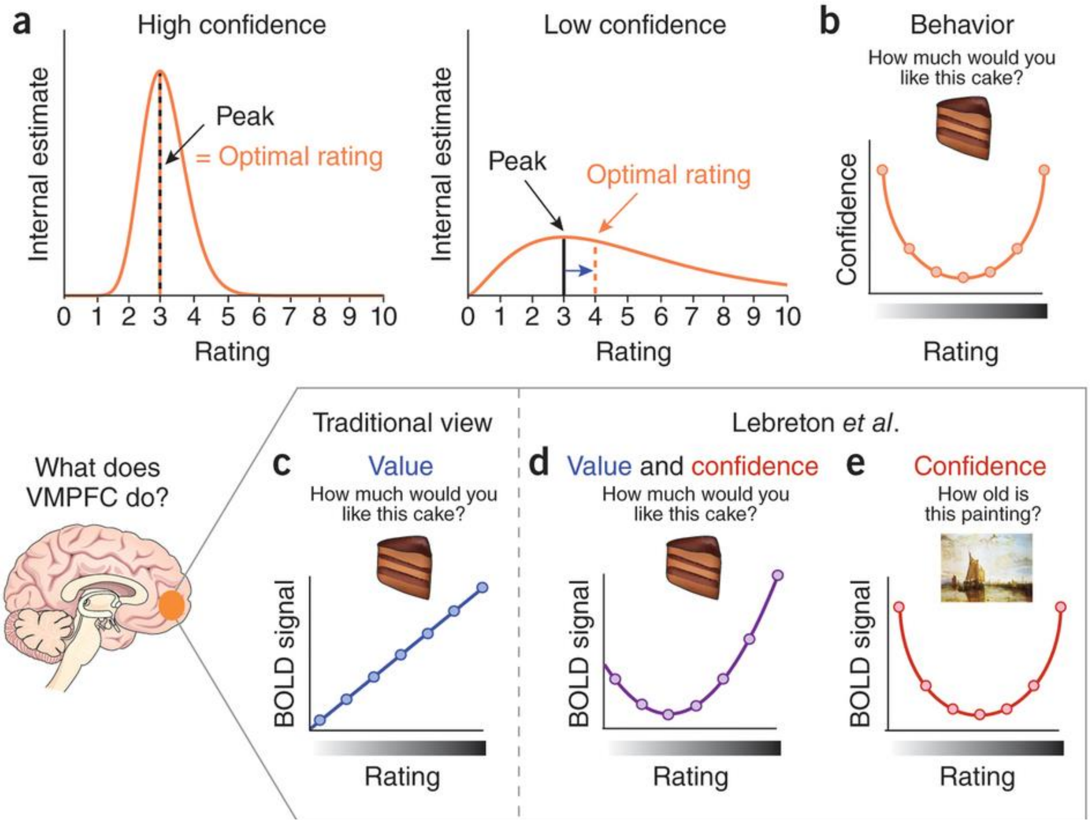


Figure 32 – Quadratic value coding in the vmPFC

When we have to rate how much we like an item, we need to report our internal representation on a scale. (a) If we are confident, the peak of the internal estimate will be close to the optimal rating we would provide. If our confidence is low, the potentially provided rating will be less close to the peak of our internal estimate and then the error between the peak and the provided rating will be higher. This error can be minimized by moving the rating toward the center of the scale (toward the optimal rating). This leads to a quadratic relationship between the confidence report and the rating (b). We know that the vmPFC activity is correlated with subjective value (c). Lebreton et al., 2015 showed that the signal in the vmPFC could be modeled as a mixture between a quadratic coding (e) and a linear coding (c), resulting in an assymetric U-shape (d). (Image from Barron et al., 2015)

## E. Summary & open questions

We have seen that the two main components of the human Brain Valuation System are the vmPFC and the ventral striatum. It is also composed of the PCC and the hippocampus and might interact with other brain network such as the Mirror Neuron System or the dlPFC and dACC which we could label the Control Network.

It has interesting properties which are still debated due to discrepancies in the literature or to the novelty of the result (quadratic coding for example). Moreover, we have seen that the OFC was highly involved in value coding in animal studies while rare in human studies, leading to an (unfair?) omission of this area in the BVS. Finally, the results found in humans were for the vast majority established through fMRI, which is blind to the dynamics of the signal and to eventual different contributions of frequencies in the valuation process. Thus, several open questions arise from those observations:

- What is the dynamics of the valuation signal in the vmPFC?
- Are the established properties true at the level of neuronal population?
- What is the dynamics of the OFC signal during subjective valuation in human?
- What are the frequencies involved in value coding in the vmPFC and in its partners?

As I mentioned (page 61 in the signal measurement part), getting an access to Local Field Potentials in human is really rare because highly invasive and restricted to a reduce number of patients. However, during my PhD, I had the opportunity to investigate these questions in 36 epileptic patients who performed rating tasks and choice tasks. I will present this study and the results of it in the second chapter of this manuscript (page 147).

The following part of this chapter moves on to assess a critical question in the field: ***How subjective values are used by the brain to make decisions?*** I will present the main results and actual theories in the field regarding this question in what follows.

## 3. Using values in choices

In the previous part of this chapter we have seen that subjective values were represented in several brain regions such as the ventral striatum, the vmPFC and the OFC. However, I focused on the valuation aspect of the decision-process, either preceding or following a decision. One might think that once subjective values are computed, the brain 'simply' needs to compare them and to select the highest one. However, it is not a trivial question and there is no consensus in the literature on how the brain compares and selects an option. I will first focus on studies on perceptual decision-making which brought important insights to the understanding of value-based decision-making and I will then present the issues under debate in the field of value-based decision-making.

## A. The case of simple choices

As mentioned in the first part of this chapter (page 42), Signal Detection Theory (SDT) has been developed to understand mechanisms underlying perceptual decision-making. It reflects sensitivity to a stimulus, but also the way the decision maker is considering the available information. In the context of perceptual decision-making, neural signals can be interpreted through the prism of SDT. Indeed, To my knowledge, one of the first to have found neural signals relevant in the context of SDT is Newsome and colleagues (Newsome et al., 1989). They used the classic random dot motion task, where numerous small white dots are moving on a black background. The key manipulation is that there is a degree of coherence between the movements of each dot, from no coherence at all to all dots moving in the same direction. The task monkeys have to perform is to decide the direction of the coherence in motion by making a saccade to the location indicated by the direction of the motion. While monkeys were performing this task, neurons were recorded in MT (for Middle Temporal, also called V5), a brain area pertaining to the visual dorsal stream known to have neurons coding the direction of motion in various stimuli. Each recorded neuron was selective for one direction and the extent to which one could infers what decision was going to be made given the response of the recorded neuron was computed in a '**neurometric**' function which was compared against the '**psychometric**' function computed from the behavior. They found that the two curves were matching quite well, suggesting that perceptual performance could rely on the neuronal activity of MT and that both a sensitivity and a criterion parameters were implemented in the spiking response.

## B. Sequential sampling models & neural implementation

As we saw, SDT does not make any prediction about decision time. However, this variable is critical since it brings information on how the decision process is implemented in the brain. For example, a classic observation is that the performance in perceptual decision tasks increases with the duration of the stimulus presentation. This is not explained by SDT or by the neural signals corresponding to this supposed mechanism. Sequential sampling models, which I also presented in the first part of this chapter (page 44), posit a theoretical dynamical mechanism underlying decision-making. Neural signals congruent with this 'accumulation to bound' theory. Indeed, neuronal activity increasing until the response similar to a '**ramping**' signal have been found in the lateral intraparietal area (LIP) while monkeys were performing a random dot motion task (Roitman and Shadlen, 2002) (Figure 33). This result has been replicated several times (Gold and Shadlen, 2007). However, a recent study challenged this idea by showing that the average of the spiking activity across trials was 'looking like' a ramping activity but that this average could be the result of both neurons spiking in a ramping mode and neurons spiking in a stepping mode (Latimer et al., 2015). Indeed, averaging the activity of stepping neurons with different stepping time would result in a 'ramping signal'. To disentangle this issue, the authors compared models of neurons 'stepping' versus 'ramping' and found that more than two thirds of the neuronal activity was better explained by a stepping model. Nevertheless, even if neurons do not present a 'ramping' activity, the average populational signal does, thus, it does not prevent us to use the drift diffusion model and the proxy of ramping activity as an 'as if' mechanism to investigate the dynamics of some neural signals.

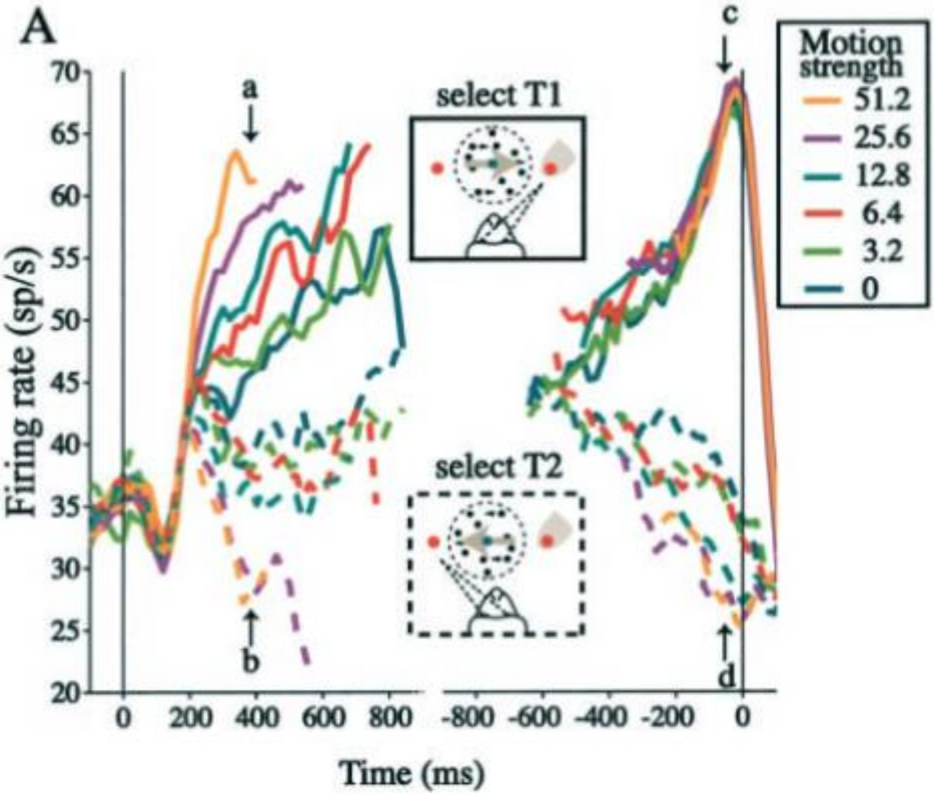


Figure 33 – Time course of average LIP neuron activity in a random dot motion discrimination task.

Typical ‘ramping’ activty. On the left, firing rate of neurons is averaged and locked on the stimulus onset and show an increase in activity proportional to the motion strength and specific to the given response. On the right, signal is locked on the response. It returns to baseline as soon as a response is given. From (Roitman and Shadlen, 2002).

In the next part, we will see an interesting attempt to extend this model to perceptual decision-making proposed by Wang which can be extended to value-based decisions, with an accumulator for each option with the drift rate corresponding to the value of each option. Then, we will see the assumed neural signature(s) of a comparison process and the related issues: notably i) the role of the vmPFC in the comparison process and ii) the frame of computation for the decisional value.

### *Wang model: a biophysically realistic model*

Sequential sampling models are intuitively appealing and have been used to fit psychophysical data. However, their biophysically implementation is not clear. Indeed, how is preset the threshold to cross in order to trigger a decision is not specified, and more importantly, how the difference A-B to compute the value of the drift rate is implemented is not specified and remains an open question.

Wang's model (Wang, 2002) is a biophysically realistic cortical

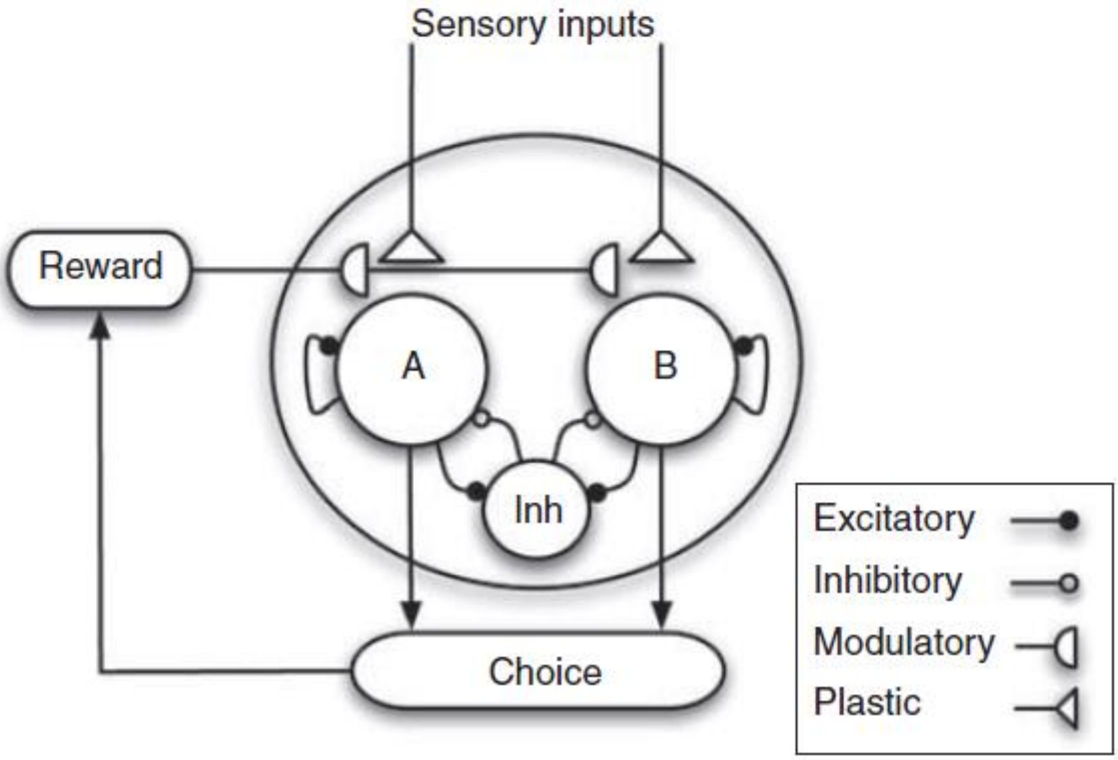


Figure 34 – Wang's Neural Circuit Model adapted for decision making with two alternatives.

Two pools of excitatory neurons A and B have strong recurrent excitatory connections and are selective to each option (sensory inputs). They compete through feedback inhibition from a pool of interneuron (Inh). The outcome signal from the reward is supposed to modulate hebbian plasticity of input synapses in A and B. Once the competition is over (in a 'winner take all' fashion) a choice is made. Image from (Glimcher and Fehr, 2013).

network model which has been used to explain perceptual decision-making. The main idea is that each stimulus with a value (or strength, in the case of perceptual decision-making) selectively activates a pool of pyramid cells, called **attractors**. In each pool, there is recurrent excitation between cells of similar selectivity, involving that the highest the value of stimulus input is, the fastest the activity of the pool will increase. This creates a phenomenon of 'ramping activity' for each pool of pyramidal cells (Figure 34). The recruitment of feedback inhibition from interneurons will create an effective inhibition between the two attractors. At the end of the process, one pool will end up with a high firing state (**chosen option**) while the other one will end up with a low firing state (**unchosen option**). This model is biophysically realistic since the biophysical properties of neurons and synapses are included in the parameters.

### *Neural signature of the comparison process*

Recently, a study investigated the neural predictions of Wang's model on a magnetoencephalographical (MEG) dataset recorded from human participants while choosing between two goods (Hunt et al., 2012). They found that the theoretical observed signal resulting from the simulation would correspond first to a response to the sum of the two considered options which would be followed by the signed difference between the chosen and the unchosen option values. They found a signal close to these predictions in a low frequency range (3-9Hz) in the frontal pole region, identified with source reconstruction as the vmPFC.

This model and results suggest that neurons in the vmPFC are coding the subjective value of options but also implementing the comparison process. This result is supported by other studies such as

(Strait et al., 2014) in which they recorded vmPFC neurons in monkey and showed that neuronal activity was consistent with the prediction of this model.

However there are some inconsistencies in the literature since another group of studies supports the idea that the signal observed in an area which would implement the comparison process in the frame of sequential sampling models for value-based decision-making would look like a signed difference between the unchosen and the chosen option values. Indeed, Wunderlich et al., (2009) developed the Competition Difference Model (CDM) to show that the theoretical signal we should observe in an area implementing a comparison would be '$V_{Unchosen}$-$V_{Chosen}$'. As Wang's model, it is biologically inspired and they modelled it as a simple neural network which carries out value comparisons by stochastic mutual inhibition between two populations of neurons: one encoding the value of each option. In the supplemental information of this last paper, the authors discuss the neural validity of the Drift Diffusion Model and conclude that it does not provide a good account for the value comparison signal they assign to a network comprising the dACC, the dlPFC and bilateral anterior insula (Figure 35). However, another study (Basten et al., 2010) had arguments suggesting that a brain region which would implement a drift diffusion model mechanism would have a higher activity for hard decisions and would consequently negatively correlate with the unsigned difference between option values (which is highly correlated to $V_{unchosen}$-$V_{chosen}$).

Finally, we could note that this network could simply signal values with a negative correlation even if it would be not evolutively adapted to have to networks computing the same signal but in an opposite manner.

Despite this discussion about model predictions on the neural signature a comparator region should express, we could suppose that the vmPFC would implement the value computation process and provide the 'drift rate value' to the comparator network.

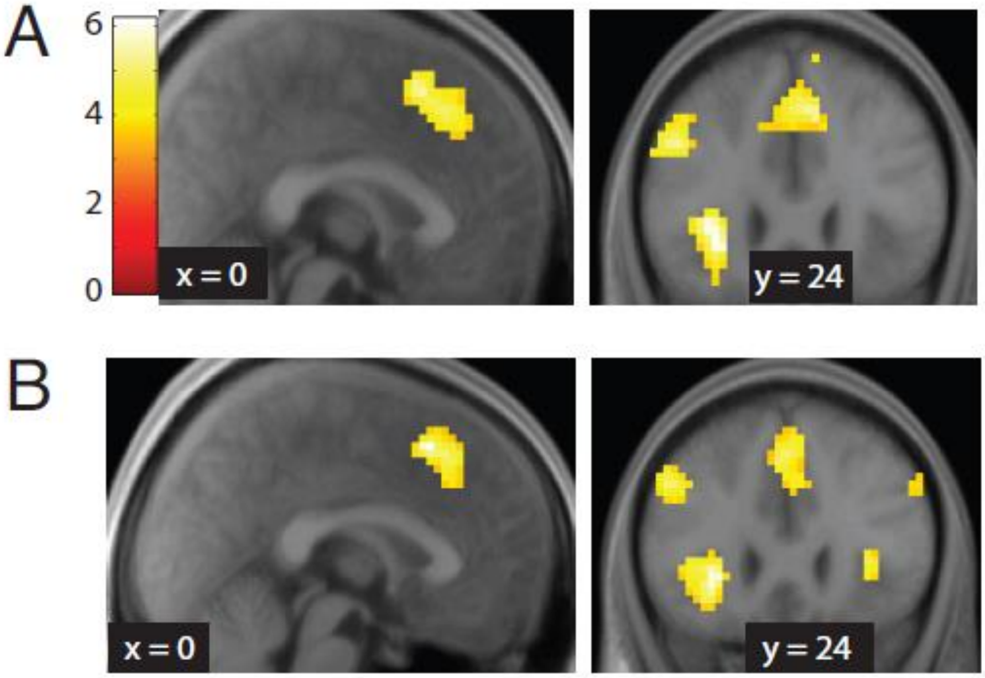


Figure 35 – The comparator network

Classical areas revealed by the contrast '$V_{Unchosen}$-$V_{Chosen}$', but also by the unsigned difference between these values and by the difficulty of the decision. Here, the map is from Wunderlich et al., (2009) and genereated by the contrast '$V_{Unchosen}$-$V_{Chosen}$'.

## C. Debates around the value-based decision-making network(s)

### *vmPFC: the frame problem*

We have seen that the vmPFC activity is correlated with the subjective value of any kind of object during pleasantness rating task, but also during age rating task. The idea that the vmPFC and the OFC encode value during decision-making is highly intuitive but how the values of the considered options are represented in order to make a choice is subject of debate. Indeed, there is a lot of evidence suggesting that OFC encodes the highest relative option value (Daw et al., 2006; O'Doherty, 2007; Roesch, 2004), but also evidences suggesting that the **chosen value** is represented in this brain region (Palminteri et al., 2009; Wunderlich et al., 2010; Kolling et al., 2012), or the difference **chosen-unchosen** (Boorman et al., 2009; Philiastides et al., 2010; Boorman et al., 2013), and sometimes, the **unsigned difference** (FitzGerald et al., 2009), or even the **sum** (Palminteri et al., 2009; Wunderlich et al., 2010; Hare et al., 2011).

One possible explanation proposed by Rangel and colleagues is that the vmPFC is encoding the value of the **attended option** minus the value of the **unattended option** value (Lim et al., 2011). This assumption is based on a model developed by Krajbich and Rangel (Krajbich et al., 2010; Krajbich and Rangel, 2011; Krajbich et al., 2012), the attentional Drift Diffusion Model which I presented in the first part of this chapter (in Variants: collapsing boundaries, decay, competition & attentional effects of DDM, page 47). Using a binary choice task where subjects had to fixate options according to a displayed instruction, they showed that the vmPFC signal was following the imposed fixation pattern: when option A was fixated, vmPFC activity was correlated with the value of A minus the value of B and when it was B, the sign of the difference was reversed (Figure 36).

Through the prism of this result, they re-interpret previous results regarding the coding of the chosen value: since the chosen item is fixated longer than the unchosen one (Krajbich et al., 2010; Krajbich and Rangel, 2011; Krajbich et al., 2012), the value of the chosen option during the

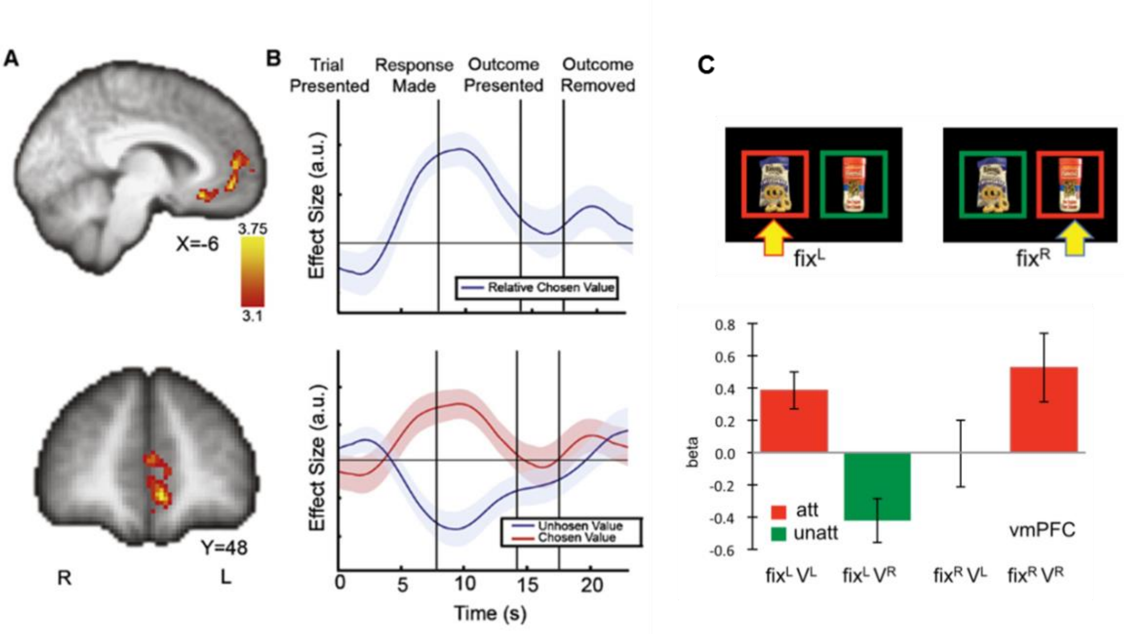


Figure 36 – Frames of the vmPFC signal during binary choices.

A & B) Regression estimates of values in the vmPFC. The chosen value (red) is positively correlated with the vmPFC activity while the unchosen value is negatively correlated (blue). This is resulting in a positive correlation with the signed difference and the BOLD signal (top B), from Boorman et al., (2009). C) An alternative pre-choice frame is proposed by Lim et al., (2011) who showed that the vmPFC activity was positively correlated with the value of the fixated option (red) and negatively with the unfixated (green).

choice process is more correlated to the attended one, thus there is no contradiction. However, this frame (attended-unattended) has the advantage to precede the choice itself, meaning that it provides a decision value on which to rely on to make a decision.

The problem of the 'frame' of the decision value will be assess in the third experiment in which I test the hypothesis of the 'default' frame coding, suggesting that the vmPFC uses prior information to define a default option and compute a decision value '**default-alternative**'.

The default frame hypothesis is following a group a studies investigating foraging in primates. One of the main results of this field is that the dACC is involved in exploratory behavior, or sometimes more precisely in 'switching from the default' kind of behavior. Interestingly, this is the same area found to express a neural signature congruent with a comparator, but also with a conflict solver.

### *dACC: Comparator, Explorator or Conflict Solver ?*

As already mentioned, the dACC has been previously associated with the comparator network with the dlPFC and anterior insula. However, another interpretation for observing a signal correlated to $V_{unchosen}$-$V_{chosen}$ during a binary choice task is that dACC is engaged in foraging by representing the value of **alternatives** (environment) and the cost of foraging (Blanchard and Hayden, 2014; Kolling et al., 2012), in other words, the value of switching to an alternative option. This view has been supported by other studies suggesting that dACC is coding the value of 'fictive reward': reward observed but not experienced (as alternative or unchosen) (Hayden et al., 2009), or the value of the best counterfactual outcome (Boorman et al., 2011).

However, this theory has been challenged by another group proposing that the dACC does not reflect the value of a non-default option but instead choice **difficulty** (Shenhav et al., 2014). They claim that foraging value is confounded with choice difficulty in the design created by Kolling et al., (2012) and that this result is supporting the theory of **Expected Value of Control** (for more details, see Botvinick, 2007; Shenhav et al., 2013). This theory proposes that the dACC is integrating the expected value of the outcome, the amount of control to get this outcome and the cost in terms of cognitive effort in order to determine how much control needs to be implemented in the considered action. It is proposed to explain why its activity is correlated to difficulty (Kouneiher et al., 2009), conflict (Botvinick et al., 2004), error (Amiez et al., 2005), or even to surprise (B. Y. Hayden et al., 2011; O'Reilly et al., 2013).

Thus, this theory supposes that the dACC is not specific to decision-making process but logically involved in this process. However, it does not explain why a foraging signal could be observed in the dACC if not confounded with difficulty. To assess this question, Kolling et al., (2016) responded back to Shenhav et al., (2014) and proposed that dACC was carrying multiple signals such as internal model of the world, difficulty but also search/foraging value and that a distinction could be made with a difficulty signal being in a more dorsal part of the dACC (close to the dmPFC) compared to the search value signal.

Nevertheless, it should be mentioned that these results are obtained and debated in the context of fMRI studies. dACC neurons recorded in monkey have more various responses: they tend to respond to the expected reward value relatively to alternatives options (Sallet et al., 2007). However, it seems that there is several type of cells and signals since the proportion of cell responding positively to reward expected value was the same that the ones responding negatively. Moreover, another studies showed that dACC neurons were more likely to code multiple decision signals such as positive or negative expected reward value compared to OFC and LPFC (Kennerley et al., 2011).

### D. Summary & open questions

To conclude, we saw that there are two main theories on how choices are processed by the brain but the 'full picture' is still not complete. The first one supposes that both valuation and selection processes are implemented by the vmPFC while the second theory supposes that the vmPFC only implements the valuation process and computes the decision value but does not implement per se the selection process which could be taking place in more dorsal areas such as the dACC. Given the vmPFC properties reviewed earlier, our studies rely more on the second framework, with the idea that vmPFC would automatically compute all kind of values such as option values or decision values during a binary choice and that it would provide information on value to a comparator, or 'executor' (see Rushworth et al., 2012 for an interesting review on this topic). This last theory is in line with a study investigating neuronal activity with simultaneous recordings in areas involved in decision-making. They showed that during a simple binary choice task between juice rewards, OFC neurons were coding reward value before dlPFC and dlPFC neurons seemed to code for the response on top of the reward value compared to OFC neurons (Wallis and Miller, 2003). More recently, another study elegantly decomposed LFP (from monkey electrophysiology and MEG recordings in humans) signals with Principal Component Analysis (PCA) (Hunt et al., 2015) and showed that the dlPFC would be a good candidate for implementing a competitive dynamics similar to the one the network model (presented page 106) would predict. Moreover, the dlPFC activity seemed to be preferentially influenced by OFC activity during delay-based decisions and by ACC during effort-based decisions. Nevertheless, the signal which could correspond to a competitive process appeared to be distributed across multiple areas and authors proposed that it is highly probable that parallel competitions might take place in various reference frames.

We briefly pointed that a parallel (but close) literature on foraging is interested by what is considered as 'natural' choices: staying or switching from a default option (Stephens and Krebs, 1986). This idea has been explored to investigate the neural basis of sequential foraging decisions (Hayden et al., 2011) and the role of the dACC in such decisions (Kolling et al., 2012). Those studies allow providing a clear definition of what is called a **default option** in a choice set: it should be chosen faster and more often than other options. We will use this concept in the third experiment of my PhD thesis. As we saw, the actual hypothesis on the decision value signal implemented in the vmPFC is either '**chosen-unchosen**' but it is by definition post-decisional and does not provide any decision value prior to the choice; either '**attended-unattended**', but it makes the decision-value signal entirely dependent on attention and does not take into account natural situations in which we are already engaged in or prior information we can use to make our choices. Thus, in the third study, we decided to test the hypothesis of a decision value which would follow the frame '**default-alternative**', with the idea that the default is defined by priors. The aim of it was to answer the following questions:

- Are priors able to set up a 'default option'?
    - If yes, we should be able to observe evidence in the behavior such as a faster response for a default option (which would also be more often chosen)
- If so, how priors are implemented in the brain to bias decisions toward the default?
- And finally, what is the frame of the decision value in the vmPFC when a default is set up? i.e. Does the vmPFC decision value signal corresponds to the frame 'default-alternative' ?

So far this chapter has first focused on the characterization of the subjective value, second on its implementation in the Brain Valuation System and finally on its use in the neural processes of decision-making. The following chapter will present the work I conducted in my PhD with a first behavioral study investigating measures of values, then an intra-EEG study focused on the dynamics of the Brain Valuation System and its properties and finally a fMRI study investigating the neural basis of value-based decision-making.

# Chapter 2: Experimental studies

In this chapter, I will present the experimental studies I performed during my PhD thesis. It is composed of three main studies:

- A behavioral investigation of subjective value measurements, titled '***Choose, rate or squeeze: comparison of economic value functions elicited by different behavioral tasks***'. The aim of this study is to compare and discuss the existing measurements of subjective value in the field.
- An intracranial EEG investigation of the Brain Valuation System and its properties, called: '***Testing the core properties of the Brain Valuation System: an Intracranial EEG investigation***'. In this study, we used a large dataset of human electrophysiological recordings from epileptic patients in which we invested how subjective value is coded by the vmPFC and its partners in Local Field Potentials.
- An fMRI study titled '***How prior preferences determine decision-making frames and biases in the human brain***' in which we investigate the vmPFC activity during binary choices.

In this section, I will introduce and discuss each of these studies with respect to their own specific content.

# 1. Behavioral investigation of subjective value measurements

## A. Introduction

As mentioned in the first chapter, several tasks can be used to measure subjective values. How to choose the task highly depends on the process you are targeting. However, if one wants to get access to a subject's value, knowing whether to use a 'willingness to pay' task or a rating task might be tricky and we often rely on what have been done (and worked) in the lab before. For example, our group classically uses two kind of measures: rating tasks in which participants simply used a scale to report their subjective value regarding objects (Abitbol et al., 2015; Lebreton et al., 2009, 2012, 2013, 2015) and effort tasks, classically more used to investigate cost-benefit trade-off but with the idea that the amount of provided effort will be proportional to the benefit estimated by subjects (Cléry-Melin et al., 2011; Meyniel et al., 2013; Meyniel and Pessiglione, 2013; Schmidt et al., 2007, 2009, 2012). For a second example, another group like O'Doherty's one will classically use auction bids to measure willingness to pay (Chib et al., 2009; McNamee et al., 2013; Plassmann et al., 2007, 2010). This method is also called Becker–Degroot–Marshak (BDM) auction (Becker et al., 1964) and consists in asking subjects to state their willingness to pay (WTP) for objects. The trick is to inform them that a random trial will be selected at the end of the experiment and the experimenter will generate a random selling price. If the price is less than or equal to the subject's WTP, the subject has to purchase it; if the price is greater, the subject keeps its money and does not purchase the item. It is supposed to be a measure in which the subject unique best strategy is to state its true WTP. Nevertheless, an unpublished comparison has been conducted in the lab (to set up an experiment) and no difference has been observed between BDM auctions and likeability ratings. A last measure consists of ranking the set of proposed items. The ranking can be done by the subject himself (Milosavljevic et al., 2012) but it implies to deal with a small set of items (ranking more than 20 items might be difficult and long); or by making subjects choose between items. Binary choice tasks are used to infer subjective values and are classically used in animal studies (Padoa-Schioppa et al., 2006; Padoa-Schioppa, 2009, 2013; Padoa-Schioppa and Assad, 2008; Tremblay and Schultz, 1999).

In human studies, ratings or auction bids are often used to predict choices and the congruence between those measures is rarely above 80%. Moreover, we do not know whether it is due to the measure itself or to the noise in utility. To my knowledge, a systematic comparison between subjective value measures has not been conducted yet. In the following study, we propose to compare three of those measures (rating, effort, choice) through a model-based approach and to answer three questions:

- Is it the same utility function which underlies the subjective values revealed by the three tasks?
- If so, are there differences between the subjective values induced by the tasks?
- What are the methodological advantages and disadvantages of each task?

## B. Article

# Choose, rate or squeeze: comparison of economic value functions elicited by different behavioral tasks.

A. Lopez-Persem(1,2), L. Rigoux(3,4), S. Bourgeois-Gironde(5), J. Daunizeau(1,2), M. Pessiglione(1,2)

(1) Motivation, Brain and Behavior team, Institut du Cerveau et de la Moelle épinière (ICM), Paris, France
(2) ICM, INSERM UMRS 1127, CNRS UMR 7225, Université Pierre et Marie Curie UPMC-Paris 6, Paris, France
(3) Max Planck Institute for Metabolism Research Cologne, Gleueler Str. 50, 50931 Cologne, Germany
(4) Institute for Biomedical Engineering, University of Zürich and ETH Zürich, Translational Neuromodeling Unit, Wilfriedstrasse 6, 8032 Zürich, Switzerland
(5) Institut Jean Nicod, Ecole Normale Supérieure, Paris, France

## *Abstract*

A standard view in neuroeconomics is that to make a choice, an agent first assigns subjective values to available options, and then compare them to select the best. In choice tasks, cardinal values are typically inferred from the preference expressed by subjects between options presented in pairs. Alternatively, cardinal values can be directly elicited by asking subjects to place a cursor on an analog scale (rating task) or to exert a force on power grip (effort task). These tasks can vary in many respects: they can be more or less costly, implicit and consequential. Here, we compared the value functions elicited by choice, rating and effort tasks on bidimensional options composed of a monetary amount for the subject (gain) and one for a charity (donation). Bayesian model selection showed that despite key differences, the three tasks elicited a same value function with similar weighting of gain and donation. Our finding therefore suggests that previous studies using the different tasks to elicit values are comparable. Moreover, we report slight differences in the convexity of value functions and in the computational efficiency of parameter estimation that may guide the design of future studies.

## *Keywords*

Effort, rating, choice, utility function, computational modeling, adaptive design

## *Introduction*

Value (or utility) functions have been defined to account for preferences revealed in choice tasks (Samuelson, 1938). One basic principle is that if an agent prefers A over B, then for this agent the value of A is higher than the value of B. Assuming basic axioms of expected utility theory, cardinal functions have been described, such that option values can be positioned on a numeric scale (Von Neumann and Morgenstern, 1947). Cardinal values rely on the notion that choice probability depends on the distance between option values, as well as on their distance from a reference point (Kahneman and Tversky, 1979). Value functions can be parameterized when choice options are combinations of objective quantities, typically the probability and magnitude of monetary payoff. The parameters can then be estimated through fitting procedures that maximize the likelihood of observed choices under the valuation model. Fitting choices involves specifying a function relating choice probability to option values, generally a softmax rule (Luce, 1959).

Alternatively, a more direct approach has been used in the neuroeconomics literature, using behavioral tasks in which subjects assign cardinal values to available options, instead of inferring value functions from their choices. One possibility is to ask subjects to rate on analog scale the desirability (or likeability) of the outcomes associated to every option. Another possibility is to ask subjects to express the maximal cost (e.g. price, effort or delay) they are willing to endure in order to get these outcomes. The aim of the present study was to compare the value functions obtained with these direct cardinal measures, and the value functions obtained from fitting choice data. We selected, in addition to a standard binary choice task where subjects state their preference between two options, a subjective rating task where subjects express the desirability of the option being actualized (Lebreton et al., 2009) and an effort production task where the probability of the option being actualized depends on the force produced with a handgrip (Pessiglione et al., 2007).

There are a priori reasons why these value functions should differ. Regarding rating tasks, a first difference is that they are thought to measure expected values (i.e., anticipated pleasure), while choices rely on decision values. These two types of values may be calculated in different brain systems (Rangel et al., 2008). A second difference is that choices are consequential, whereas rating tasks have no consequence for the subject. Thus subjects could pretend to have values different from their true values, particularly when options include moral concerns (Polnaszek and Stephens, 2013). A third difference is that the rating scale is somewhat arbitrary, and may yield distortions of value functions due to anchoring phenomena (Lynch, Jr. et al., 1991).

Regarding effort tasks, a first difference is that the behavior is driven by the net value, in which the cost must be subtracted from the option value. This may affect the value function if effort cost has particular properties, such as non-linearity (Rigoux and Guigon, 2012). A second difference is that effort cost is more implicit than the cost of making a choice, which is forgoing one of the two options. A last difference is that fatigue may increase with the number of performed trials and influence effort cost, and hence the expressed values (Le Bouc et al., 2016).

The different elicitation procedures were compared using a same set of 121 options, which involved moral concerns that may be traded against costs. Each option combined some money for the subject and some money for helping a charity NGO. We explored the value functions elicited by the choice, rating and effort tasks. The same valuation model provided the best fit of behavior in the three tasks, with slight differences in model parameters and in the time needed for parameter estimation.

## *Results*

### Basic results

Subjects (n=19) participated to three tasks allowing measuring subjective values about bi-dimensional options composed of one gain for themselves and one donation for a charity organization they choose prior to the experiment (Figure 1, top). In the rating task, participants rated how much they would like to obtain the composite option using a scale graduated from 0 to 10. The feedback was probabilistic and they won the option 70% of the time. In the force task, subjects had to squeeze a handgrip and the feedback (probability of winning the option) was determined by the percentage of their maximal force they produced during the trial. Previous experiments in the lab using the grip task showed that subjects produce on average about 70% of their maximal force (Cléry-Melin et al., 2011; Pessiglione et al., 2007; Schmidt et al., 2007, 2009, 2012). In the choice task, participants had to choose between two composite options and they won the selected 70% of trials. The choice task followed an adaptive design in which options were proposed so as to optimize the parameterization of an 'a priori' utility function (linear integration of gain and donation with their interaction).

As expected, subjects produced force and rated the options proportionally to the incentives (Figure 1, bottom). Subjective values inferred from choices were also proportional to the composite options (Figure 1, bottom). A linear regression indicated that both gain (G) and donation (D) had a significant impact on ratings ($\beta_R$(G)=0.07±6.10-3, t(18)=11.5, p=1.10-9; $\beta_R$(D)=0.06±7.10-3, t(18)=8.2, p=1.10-7), on exerted effort ($\beta_F$(G)=0.05±6.10-3, t(18)=8.5, p=1.10-7; $\beta_F$(G)=0.05±6.10-3, t(18)=7.2,

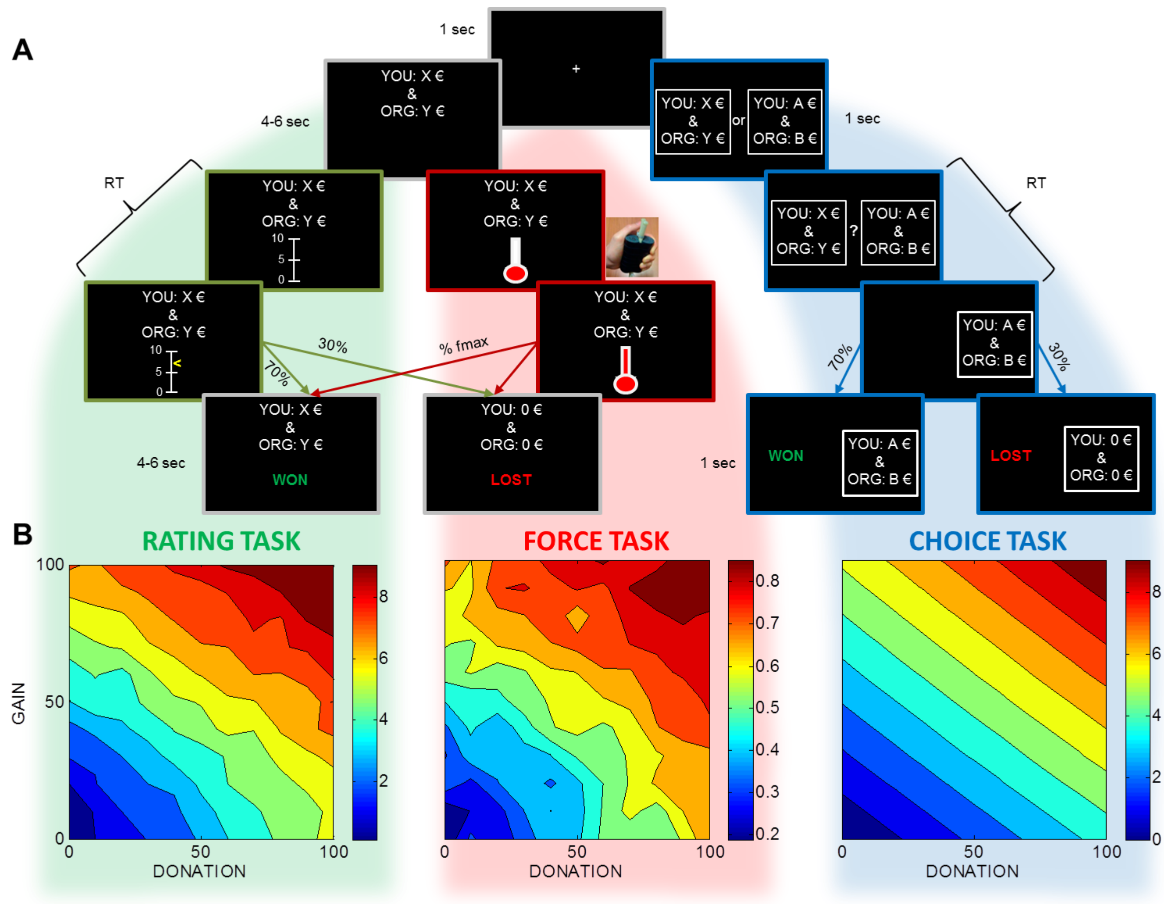


**Figure 1. Task design and behavioral results**

**A**. From top to bottom, successive screenshots of example trials are shown for the three tasks (left: rating task, middle: force task, right: choice task). Every trial started with a fixation cross. In the force and in the rating tasks a single composite proposition with an amount X of gain for the subject (YOU) and Y for the charity organization (ORG) was displayed for 4 to 6 seconds on the screen. In the rating task (green) and in the force task (red) a scale and a thermometer respectively appeared on the screen indicating subjects that they could provide an answer (rating or effort). A feedback was then displayed for 4 to 6 seconds with the probability of winning fixated to 70% in the rating task and determined by the amount of effort produced in the force task. In the choice task, two composite options were displayed for 2 seconds. Subjects knew they could answer when the term 'or' switched to '?'. They won the option that they chose on 70% of the trials. **B**. Average ratings (left), forces (right) and values inferred from choices (right) as functions of the amount of gain and donation. Cold to hot colors indicate low to high values.

p=9.10-7) and on choices ($\beta_C(G)=0.16\pm0.03$, t(18)=5.6, p=2.10-5; $\beta_C(G)=0.12\pm0.02$, t(18)=5.4, p=4.10-5). Interaction terms were significant for the rating and the force task but not for the choice task ($\beta_R(G*D)$=-2.10-4±9.10-5, t(18)=-2.7, p=0.01; $\beta_F(G*D)$=-3.10-5±1.10-5, t(18)=-2.6, p=0.02;

$\beta_C$(G*D)=1.10-5±2.10-4, t(18)=0.1, p=0.95). In none of the task we found a significant difference between the weights for gain and for donation despite trends in the choice and in the rating tasks (R: t(18)=1.79, p=0.0897; F: t(18)=1.10, p=0.286; C: t(18)=2.05, p=0.0549) which is suggesting that subjects were in average using both attributes to express their judgment with no bias toward one or the other dimension. However, those interpretations are based on absence of significance, thus we cannot confirm it. To get more details on those results, we used a model-based approach.

## Modelling

In order to further investigate how those measures could affect the subjective value of each option, we defined a set of twelve utility functions which could explain the observed behavior in each task (see methods). Data for each subject and each task were fitted using Variational Bayesian Analysis (VBA) for all utility functions. Averaged determination coefficients across subjects were comprised between 0.43 and 0.70 in the force task, between 0.57 and 0.85 in the rating task and between 0.45 and 0.85 in the choice task, suggesting that all models could correctly account for the behavior.

### *Comparison of utility functions*

First, using Bayesian model comparison, we examined whether subjective values from both tasks could be explained by the same utility function. This comparison first shows that the family of models with the same utility function for the three tasks is winning against the family of models with different utility function (Ef=0.95, Xp=1, Figure 2A).

Second, we found that the utility function called ‘Constant Elasticity of Substitution’ (CES) is the one which better explains the data in the three tasks (Model comparison inside the ‘same’ family: Ef=0.61, Xp=1, Figure 2C).

### *Comparison of free parameters*

The CES utility function is characterized by two main parameters: a selfishness parameter $\alpha$ comprised between 0 and 1 (the closer to 1 the more selfish) and a convexity parameter $\delta$ (above 1 indicates a higher sensitivity to high values in a composite proposition). We examined which free parameters of this model best explained the differences between the subjective values of the three tasks. We found that convexity parameters were significantly different across tasks ($F(2,54)=3.72$, $p=0.03$), and more precisely, the convexity parameter in the choice task was significantly lower than in the two other tasks ($\delta_F=1.98\pm0.31$; $\delta_R=1.60\pm0.20$; $\delta_C=1.10\pm0.14$; $\delta_F$ vs $\delta_R$: $t(19)=1.55$, $p=0.14$; $\delta_F$ vs $\delta_C$: $t(19)= 2.9$, $p=9.10^{-3}$; $\delta_R$ vs $\delta_C$: $t(19)= 2.56$, $p=0.02$). Moreover, there is no significant difference between the selfishness parameters of the three tasks ($F(2,54)=0.09$, $p=0.91$) (Figure 2E, Left). Nevertheless, since the absence of significance does not allow us to conclude that there is no difference and because the choice selfishness parameter is significantly different from 0.5 ($\alpha_C=0.58\pm0.04$, $t(19)=2.16$ $p=0.049$) while the other are not ($\alpha_F=0.56\pm0.05$, $t(19)=1.29$, $p=0.21$; $\alpha_R=0.58\pm0.05$, $t(19)=1.84$ $p=0.082$) (Figure 2D, Left), we assessed this question with another model comparison.

In order to test whether there was no difference between the selfishness parameter of each task, we fitted the CES models simultaneously to the three tasks. We then compared a model including one single selfishness parameter for all the tasks to a model including three different selfishness parameters. We also included the intermediate variants in the comparison ($\alpha_F=\alpha_R\neq\alpha_C$; $\alpha_F=\alpha_C\neq\alpha_R$; $\alpha_R=\alpha_C\neq\alpha_F$; see Methods). Among those five models, the model with a unique selfishness parameter for all the tasks won the comparison (Ef=0.52, Xp=0.97, Figure 2D, Right). We found that this common selfishness parameter is

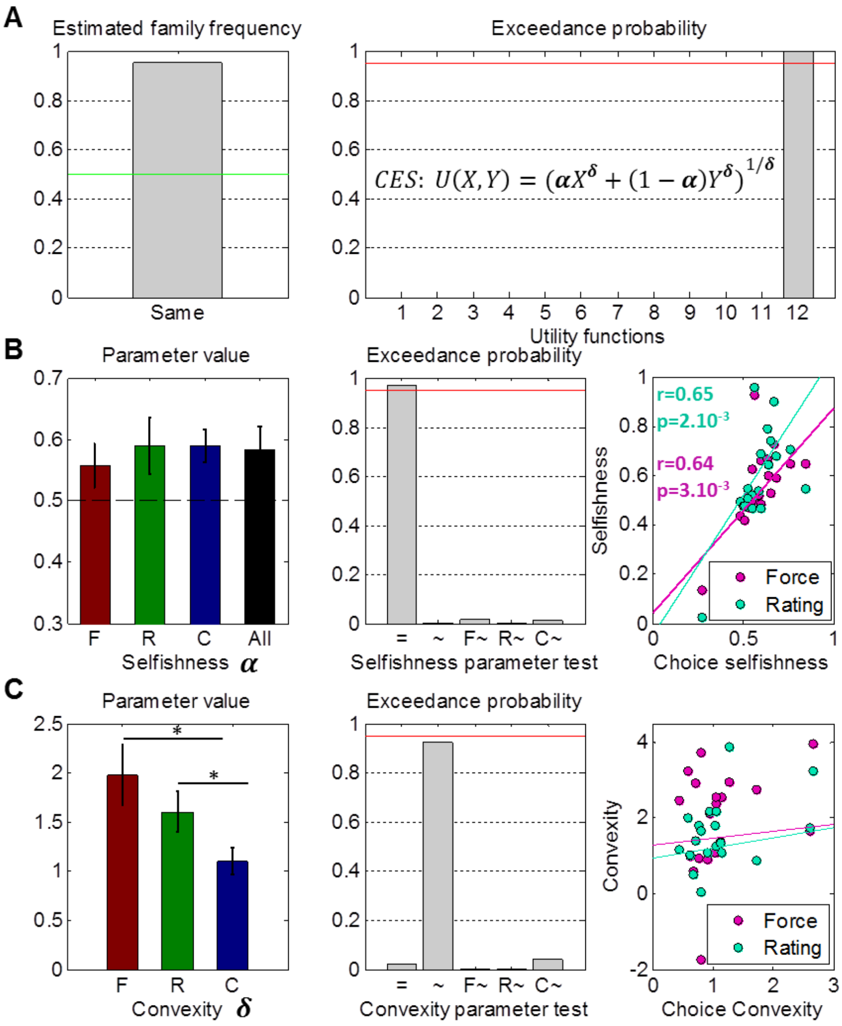


**Figure 2. Comparison of utility functions and their parameters**

A. Investigation of the utility underlying behavior in the three tasks. Left: Estimated family frequency for the family of models in which the three tasks are explained by the 'same' utility function (compared to the family of models in which they are explained by 'different' utility functions). Green line indicates the threshold for the null hypothesis (50%). Right: Comparison of the twelve models (utility functions) belonging to the 'same' family. The CES utility is winning the comparison (12) and its function is reported on the panel, with U(X,Y) the utility of the gain X and the donation Y, $\alpha$ the selfishness parameter and $\delta$ the convexity parameter. **B.** Left: Selfishness parameter values in the three tasks (F, R and C) and with the three tasks together (All). The dark dashed line indicates no bias (0.5) toward one or the other dimension (gain or donation). Error bars indicates S.E.M. Middle: Exceedance probabilities of models including one single selfishness parameter for the three tasks (=), three different selfishness parameters (~), or only one different from the other (F~, R~, C~, with 'X~' standing for 'task parameter different from the others'). Right: Correlation of the choice selfishness with the force selfishness (turquoise) and with the rating selfishness (Magenta) across subjects. **C.** Same analysis as B but for the convexity parameter. Left: The choice convexity parameter is significantly different from the force and the rating convexity parameters. Stars indicate significant differences between tasks. Middle: The model with one different convexity parameter for each task is winning the model comparison. Right: There is no significant correlation between the convexity parameters of each task across subjects.

significantly favoring the individualistic gain in the proposition ($\alpha=0.58\pm0.04$, t(19)=2.16, p=0.044) (Figure 2D, Left).

We ran the same analysis on the convexity parameter to invest whether the rating and the force task could be explained by a unique parameter since the difference between them was not significant. The winning model (Ef=0.46, Xp=0.92, Figure 2E, Right) was the model with the convexity parameter being different in each task, suggesting that despite the absence of significant difference between $\delta_F$ and $\delta_R$ in average, the data are better explained with different parameter values.

This second analysis allowed us to conclude that the task used to access subjective values had an impact on the convexity of the utility function but not on the weight given to the attributes.

### *Comparison of estimation efficiency*

Then, we wanted to assess to what extent choices could be predicted from the other measures. Thus, we compared the proportion of choices predicted by the utility computed from the rating and the force tasks and found no significance difference between them (t(19)=0.82, p=0.42), with a good prediction score for each of them (Force: 77±3%; Rating: 78±2%) compared to the balanced accuracy from the function inferred from choices (84±2%). Moreover, when fitting a logistic regression on choices with the rating and force utilities, we could not find any significance difference in the temperature ($\beta_F$=1.13±0.46; $\beta_R$=0.86±0.40; t(19)=1.36, p=0.19), suggesting that both measures were equally good to predict choices (Figure 3, A & B).

Then, in order to further investigate the efficiency of parameter estimation, we decided to examine noise, task duration and number of trials.

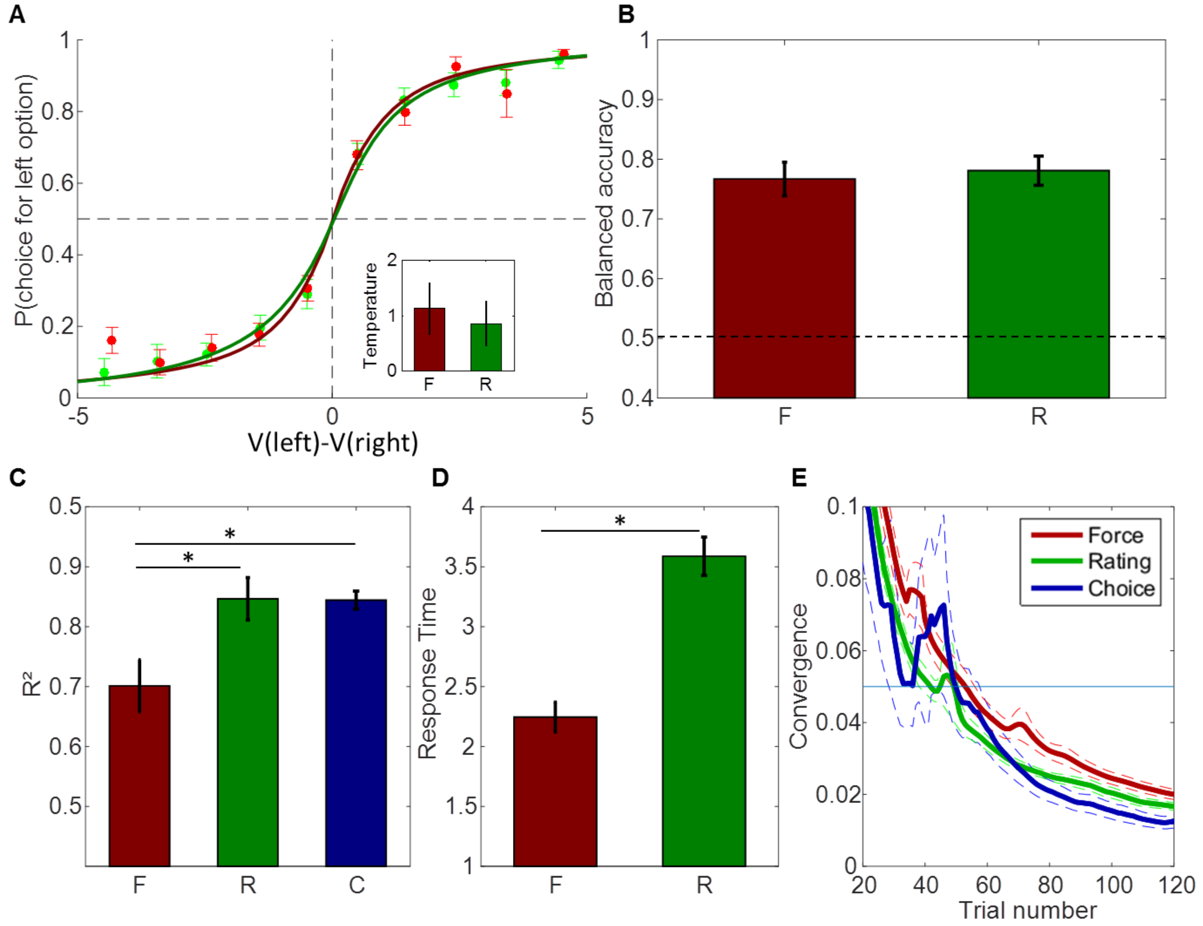


**Figure 3. Comparison of estimation efficiency**

A. Proportion of choices for the left option in the choice task according to the difference between the left and the right option values computed with the CES utility inferred from the force task (red) and the rating task (green). Inset represents the temperature estimate of the logistic fits and show that there is no difference between the estimations of the force and the rating task. B. Balanced accuracy according to the CES utility inferred from forces (red) and ratings (green). C. Coefficient of determination $R^2$ for the fit of each task (Force, Rating and Choice). The $R^2$ in the force task is significantly lower than in the other tasks. D. Response time in the force task and in the rating task. Star indicates a significant difference between the two tasks. E. Convergence measure according to the trial number with optimized trial order in the force (red), rating (green) and choice (blue) task. Error bars indicate S.E.M.

First, we compared the quality of fit of the three tasks with the CES utility model. We found that the CES utility function provided a better fit for the rating and the choice task compared to the force task ($R^2_F$=0.70±0.04; $R^2_R$=0.85±0.04; $R^2_C$=0.84±0.02; $R^2_F$ vs $R^2_R$: t(19)=-4.78, $p=1.10^{-4}$; $R^2_F$ vs $R^2_C$: t(19)=-3.87, $p=1.10^{-3;}$ $R^2_c$ vs $R^2_R$: t(19)=-0.09, p=0.93). There was no significant difference of fit quality between the choice and the rating task (Figure 3C). This suggests that the force data were noisier than the two other tasks.

Then we compared the time needed to provide an answer in the rating and in the force task. We found that the response time in the force task was shorter than in the rating task ($RT_F$=2.24 ± 0.12 sec; $RT_R$=3.59±0.16; t(19)=6.01; p=$1.10^{-5}$). Thus, for a same number of trials, the force task was overall shorter to run than the rating task. We did not include the choice task in this analysis because of the difference in time display of options. Moreover, the number of option to consider is not the same and it makes the comparison less relevant (Figure 3D).

In our design, 121 options were presented to the subjects in the force and in the rating tasks. The number of trials presented in the choice task was fixated to compute reliable parameter estimation. Nevertheless, we wanted to assess the minimum number of trial needed in each task to compute correct parameter estimates. We computed a convergence measure (see methods) to determine the trial number in which the information gain was below our convergence criterion of 5%. The rating and the force task did not need to be run with design optimization conversely to the choice task but in order to get a proper comparison and to know whether design optimization should be applied to the force and the rating tasks, we computed our convergence measure in those two task in two different manner: the first by using the random order of options as during the experiment and the second by using options bringing more information first (as in the choice task, see methods for more details). We found that the rating task needed in average 46±4 (range: 12 to 96 trials) trials in the random order and 42±2 (range: 28 to 67 trials) in the optimized order, with no significant difference between the two estimations (t(18)=0.81, p=0.43). In the force task, 57±2 (range: 38

to 81 trials) were needed in the random order and 57±3 (range: 44 to 89 trials) in the optimized order, with no significant difference between them (t(18)=0.08, p=0.94). Finally, in the choice task we found that parameter estimates were converging in 54±8 (range: 7 to 154 trials) with the optimized design for the linear and interaction utility function and in 60±7 (range: 15 to 145 trials) with the optimized design for the CES utility function, with no significant difference between them (t(18)=0.51, p=0.62). (Figure 3, bottom right).With the optimized order, we found a significant difference of the convergence trial number only between the rating and the force task ($\gamma_R$ vs$\gamma_F$: t(18)=5.57, p=$4.10^{-5}$; $\gamma_R$ vs$\gamma_C$: t(18)=0.12, p=0.9; $\gamma_C$ vs$\gamma_F$: t(18)=1.75, p=0.10). Without order optimization, we could not find any difference (all p>0.05) (Figure 3E). The trend was nonetheless that the force task required more trials than the rating task for converging on parameter estimation.

## *Discussion*

In this study, we showed that three tasks varying on several features elicited the same utility function driving participants' behavior. Moreover, the most critical parameter weighing the two dimensions of the options (selfishness) was similar in the three tasks. However, we found some differences in the concavity in the utility functions. In addition, the different tasks presented advantages and disadvantages that should be taken into account when selecting a particular elicitation procedure.

We showed with a Bayesian model comparison that the same utility function could account for the three types of behavior. This utility function, called 'Constant Elasticity of Substitution' (Armington, 1969), has been previously used in a study investigating the rationality of altruism (Andreoni and Miller, 2003). They showed that this utility function could account for choices made by participants in various experiments involving sharing money with others, which is consistent with what we found here. It has the advantage of being simple and relying on two parameters only: a selfishness parameter (more generally, a relative weighting of dimensions) and a convexity parameter (specifying the interaction between dimensions).

The three tasks not only shared the same utility function, but also elicited similar selfishness parameters. Thus, the differences in consequentiality, in implicitness and in the nature of cost did not impact the weights assigned to the gain and donation dimensions. It is remarkable given that exhibiting altruism comes for free (with no cost) in the rating task but not in the other tasks. This result suggests some stability across elicitation procedures in how dimensions are integrated. It also means that the results reported in the literature using the different tasks can be compared. However, the relative stability was observed here within subjects; results might have been more variable if we had tested separate groups of subjects on the different tasks. Another limitation is that the result could be specific to the dimensions that needed to be integrated in our paradigms, i.e. to selfish versus altruistic behavior. Further experiments would be needed for a generalization to other multi-attribute options as seen in risky or inter-temporal choices, or to more natural multidimensional options such as food items.

Even if the same utility function and the same selfishness parameter could explain the behavior in the three tasks, we found a significant difference between tasks in the convexity parameter. Indeed, the choice task did not reveal any convexity, indicating no interaction between dimensions, while the force task, and to a lesser extent the rating task, revealed a concavity, suggesting sensitivity to high amounts of money irrespective of the receiver. It remains difficult to conclude whether the concavity seen in rating and force tasks denotes an artifactual distortion of the actual value function or a better sensitivity to actual values, perhaps because choice tasks are more complex with four numbers to be integrated. Note that even if no concavity was observed on average in the choice task, the model with a concavity parameter (CES) nevertheless won the Bayesian selection. This means that some subjects exhibited concave and others convex utility shapes. This inter-subject variability possibly reflects differences in the sensitivity to equity (same amount for you and for charities).

Independently of the elicited utility function, we assessed how the tasks differed in terms of precision and speed of parameter estimation. The choice and rating tasks were better fitted, with higher coefficients of determination than the force task. However, the utility functions inferred from the rating and force tasks were equally capable of predicting choices. It was therefore not that the utility function elicited with the force task was distorted or variable, but simply that the force data were noisier. Thus, if the objective is to predict choices and not forces, there is no reason to prefer the rating task over the force task on the basis of the fit criterion.

On the other hand, response times recorded in the force task were shorter than in the rating task. Moreover, without design optimization, there was no significant reduction in the number of trials needed for stabilizing parameter estimation with the rating task compared to the force task. Thus, the speed criterion seems to be in favor of the force task. Finally, running a choice task implies to posit priors on utility functions and on parameters in order to adapt the design (i.e., the selection of choice options), whereas the other tasks can be run in a model-free manner. Thus, the experimentally simplest way to measure subjective value functions might not be the binary choice task.

## *Conclusion*

To our knowledge, this is the first study comparing direct elicitation of cardinal values (rating and force tasks) to ordinal ranking (choice task) for a same set of options. Those tasks are widely used in neuroeconomics and it is somewhat comforting that they reveal the same utility function driving the behavior despite differences many differences. They nonetheless present different advantages and drawbacks that may guide the design of future studies.

## *Methods*

### Participants

The study was approved by the Pitié-Salpétrière Hospital ethics committee. All subjects were recruited via e-mail within an academic database and gave informed consent before participation in the study. They were right-handed, between 20 and 30 years old, and had normal vision, no history of neurological of psychiatric disease. They were not informed in the recruitment that they could give money to a charity, in order to avoid a bias in the sample. Nineteen subjects (10 females; age, 22.2 ± 1.4) were included in the study. They believed that the money won while performing the task would be their remuneration for participating, but eventually their payoff was rounded up to a fixed amount (100€).

### Behavioral Tasks

Subjects performed the three tasks, with order of performing balanced across subjects for the force and the rating task. The choice task was always performed after the two others which were performed in a

MRI scanner for another unpublished study. The force task was preceded by a measurement of the maximal force for the right hand (Cléry-Melin et al., 2011; Pessiglione et al., 2007; Schmidt et al., 2007, 2009, 2012). Participants were verbally encouraged to squeeze continuously as hard as they could until a line growing in proportion to their force reached a target displayed on a computer screen. Maximal force was defined as the maximal level reached on three recordings. Then subjects were provided a real-time feedback about the force produced on the handgrip, which appeared as a red fluid level moving up and down within a thermometer, the maximal force being indicated as a horizontal bar at the top. Subjects were asked to try outreaching the bar and state whether it truly corresponded to their maximal force. If not, the calibration procedure was repeated.

In the force and the rating tasks, 121 trials were presented in a random order across three sessions of 40 or 41 trials. Each trial corresponds to one of the 121 combinations of the experiment design (eleven incentives for themselves by eleven incentives for charity donation from 0€ to 100€ with steps of 10€). Subjects performed the three sessions with the right hand, with short breaks between sessions to avoid muscle exhaustion.

In the force task, after baseline recording of the pressure at rest, each trial started by revealing the two monetary incentives with the inscriptions “YOU” followed by the amount for the subject, and “ORG” followed by the amount for the charity donation. The stimulus was displayed between 4 and 6 seconds jittered. Subjects knew that the probability to win each trial was proportional to the effort they would produce after the display of the thermometer on the screen. The probability to win the trial was equal to the percentage of their maximal force they produced. They were provided with online feedback on the exerted force (with a fluid level moving up and down within a thermometer), and they were informed that they had to produce a minimal effort at each trial (10% of their maximal force) and that the trial would be over when they stop exerting a force on the handgrip. Each trial ended with a 4 to 6 seconds display of the result of their effort with a mention “WON” or “LOST” (Figure 1, top)

During the rating task, subjects were presented the same 121 trials as the force task in three sessions. The successive events of each trial were similar as in the force task except for moment of answer. Instead of a thermometer display, a vertical rating scale from 0 to 10 units appeared after the stimulus presentation. Subjects were asked to rate the desirability of the composite proposition on the screen by pressing buttons with the right hand. They were asked to use the whole scale across the task and they were informed that the rating they gave would not have an impact on the feedback following the answer. They were rewarded with the “WIN” mention 70% of the time, according to the mean proportion of “won” trials in the force task (Figure 1, top).

In the binary choice task, 200 trials of two composite options were presented to the participants in each side of the screen. After watching the options for 2 seconds, subject could indicate the option they would prefer to win using their right hand. In 70% of the trials a positive feedback was displayed with the selected earnings while in the other trials, a negative feedback with 0€ for both receivers was displayed on the screen. Given the number of options in our design, there are $121^2$ (14641) possible binary choices. Constraints can be applied to reduce this number: choices are informative only if they are crossed (attributes never dominate on both dimensions), if there is no common dimension for both options, and they should not be repeated. However, those constraints only reduce the number of choices to 3025. Thus, knowing that some options are more informative than others to estimate a utility function, we used an online optimization design. At each trial, the design is optimized over a single dimension (gain or donation), the chosen combination is the one that minimize the posterior covariance matrix over the parameters of an a priori utility function defined as followed: f(G,D) = βc(G)*G+ βc(D)*D + βc(G*D)*G*D. Contrary to the force and the rating tasks, since options were optimized for each subjects, the amount for charity and subjects could vary from 0€ to 100€ but with steps of 1€.

## Data analysis

To investigate how both attributes were integrated in a subjective value in the three tasks, we compared 12 models with different utility functions on behavioral data that predict the same pattern of answers.

For the force and the rating tasks, answers were fitted as follow:

$$R(X,Y) = aU(X,Y) + b$$

With R being the exerted force or the rating given for an option composed of an amount X for the subject and Y for the donation and *a* and *b* scaling parameters. U(A) is the utility of option A computed from the utility functions which will be described later in this section. For the choice task, answers were fitted as follow:

$$P(X_1, Y_1) = \frac{1}{1 + e^{-\frac{U(X_1,Y_1) - U(X_2,Y_2)}{\beta}}}$$

With P$(X_1, Y_1)$ the probability of choosing option 1 when opposed to option 2, U$(X_1, Y_1)$ and U$(X_2, Y_2)$ the utility of each option and β the temperature of the softmax.

Utility functions U corresponded to twelve different ways to integrate the two amounts of money X (selfish gain) and Y (donation).

The sixt first models are based on the study of Park et al, 2011. They were initially used to study integration of negative and positive value into an overall subjective value. We adapted it by considering the integration of allocating money and receiving money as two separated values which need to be integrated into one subjective value. These models integrated donation and gain either independently or they additionally assumed an interaction between both dimensions. Behavioral studies have suggested nonlinear value functions that allow concavity for positive values and convexity for negative values (Kahneman and Tversky, 1979). For completeness, utility functions integrated dimensions in a linear or nonlinear manner with similar or dissimilar non-linearity between dimensions. We refer to these models as (1) linear independent, (2) nonlinear - similar independent, (3) nonlinear independent, (4) linear interactive, (5) nonlinear - similar interactive and (6) nonlinear interactive.

(1) Linear independent:

$$U(X, Y) = \alpha X + \beta Y$$

(2) Non-linear-similar independent:

$$U(X, Y) = \alpha X^{\delta} + \beta Y^{\delta}$$

(3) Non-linear independent:

$$U(X, Y) = \alpha X^{\delta} + \beta Y^{\varepsilon}$$

(4) Linear interactive:

$$U(X, Y) = \alpha X + \beta Y + \gamma XY$$

(5) Non-linear-similar interactive:

$$U(X,Y) = \alpha X^{\delta} + \beta Y^{\delta} + \gamma X^{\delta} Y^{\delta}$$

(6) Non-linear interactive:

$$U(X,Y) = \alpha X^{\delta} + \beta Y^{\varepsilon} + \gamma X^{\delta} Y^{\varepsilon}$$

In order to complete those six models, we added a model integrating the absolute difference of winnings and giving. This model is used to reflect the equity effect on utility and it is suggested by equity theory (Adams, 1965). We called this model Linear-equity (7).

(7) Linear-equity:

$$U(X,Y) = \alpha X + \beta Y + \gamma |X - Y|$$

We also included simple models based on only one dimension, either by being based on the minimum value (8) 'Mini', also called 'Leontief utility' (Andreoni and Miller, 2003) and on the maximum value (9) titled 'Maxi'.

(8) Mini / Leontief:

$$U(X,Y) = \min(\alpha X, \beta Y)$$

(9) Maxi:

$$U(X,Y) = \max(\alpha X, \beta Y)$$

In the literature, some utility functions have been developed to explain inequity aversion. Among them, one was proposed by (Fehr and Schmidt, 1999)(10):

(10) Fehr & Schmidt:

$$U(X,Y) = x - \alpha \max(Y - X, 0) - \beta \max(X - Y, 0)$$

Other utility functions such as the Cobb-Douglas production function (11) are functions which integrate value of factors into one single utility in a multiplicative and non -linear manner.

(11) Cobb-Douglas:

$$U(X,Y) = X^{\delta} * Y^{1-\delta}$$

Finally, the CES utility is a common type of function in consumer theory (Andreoni and Miller, 2003) which combines two or more types of goods into a common quantity (Leontief, Linear and Cobb-Douglas utility functions are special cases of the CES utility function). The share parameter $\alpha$ indicates selfishness; $\delta$ captures the convexity of preferences.

(12) Constant Elasticity of Substitution (CES):

$$U(X,Y) = (\alpha X^{\delta} + (1-\alpha)Y^{\delta})^{1/\delta}$$

### *Model fitting and comparison*

Each model was fitted separately for each individual to ratings, forces and choices using Matlab VBA-toolbox (available at http://mbb-team.github.io/VBA-toolbox/), which implements Variational Bayesian analysis under the Laplace approximation (Daunizeau et al., 2014). This iterative algorithm provides a free-energy approximation for the model evidence, which represents a natural trade-off between model accuracy (goodness of fit) and complexity (degrees of freedom) (Friston et al., 2007; Penny, 2012). Additionally the algorithm provides an estimate of the posterior density over the model free parameters, starting with Gaussian priors. Individual log-model evidences were then taken to group-level random-effect Bayesian model selection (BMS) procedure

(Penny et al., 2010). BMS provides an exceedance probability (Xp) that measures how likely it is that a given model (or family of models) is more frequently implemented, relative to all the others considered in the model space, in the population from which participants were drawn (L Rigoux et al., 2014; Stephan et al., 2009).

### *Between-tasks comparison*

We first wanted to determine whether the same utility function was used across the two tasks. To do this, $12^3$=1728 models with each possible combination of models for each task were needed to be compared. However, in order to assess this large model space, we can consider the three tasks as three conditions and first estimate the parameters of each model for each subject and condition. Then, as proposed by (L. Rigoux et al., 2014) one can think of the conditions as inducing an augmented model space composed of model 'tuples' that encodes all combinations of candidate models and conditions. We then have 1728 tuples, where the log-evidence of each tuple can be derived by appropriately summing up the log evidences over conditions. Then we can split the tuple space in two subsets: one in which the same model is underlying all conditions ('same' family) and one with the remaining tuples ('different' family).

To assess whether the same parameter value could be used for the three tasks for the selfishness and convexity parameter of the CES utility, we need to compare models that differ in terms of whether or not a given additional parameter for each task are set to zero. Thus, we fitted five models to the three tasks simultaneously with the parameter of interest $\varrho$ (either selfishness of convexity) varying as follow:

- $\varrho_F \neq \varrho_R \neq \varrho_C$

- $\rho_F = \rho_R \neq \rho_C$
- $\rho_F \neq \rho_R = \rho_C$
- $\rho_F = \rho_C \neq \rho_R$
- $\rho_F = \rho_R = \rho_C$

### *Convergence assessment*

To compute a convergence measure, we estimated the parameter estimates of the CES utility separately for each task by including one more trial in each iteration. We monitored at each step the increase in estimation precision $\gamma$such as:

$$\gamma_t = \frac{\sigma_{t-1} - \sigma_t}{\sigma_{t-1}}$$

With $\sigma_t$ the mean posterior variance of all parameters at trial t. This measure allows tracking the information gain afforded by each trial. The convergence criterion was defined as 5% and the minimum number of trial was defined as the last trial in which the convergence measure was above 5%. In this measure, the trial order was random as during the experiment. We ran the same analysis but with the trial order optimized with the VBA toolbox. The first eleven trials were set such as the limits of the option space were covered: combinations of 0, 50 and 100€ for both dimensions and two intermediate options composed of 30 and 70€. The order of these first options was random. Then, an option was selected at each trial such as the trace of the expected posterior matrix would be minimized. For the choice task, we computed the convergence measure according to the effective order of options as in the optimized design for the linear-interactive utility function and we also computed the convergence measure with an order optimized for the CES utility function.

# *References*

Andreoni, J., Miller, J., 2003. Giving according to GARP: An experimental test of the consistency of preferences for altruism. Econometrica 70, 737–753.

Armington, P.S., 1969. A Theory of Demand for Products Distinguished by Place of Production. Staff Pap. 16, 159–178. doi:10.2307/3866403

Cléry-Melin, M.-L., Schmidt, L., Lafargue, G., Baup, N., Fossati, P., Pessiglione, M., 2011. Why Don't You Try Harder? An Investigation of Effort Production in Major Depression. PLoS ONE 6, e23178. doi:10.1371/journal.pone.0023178

Daunizeau, J., Adam, V., Rigoux, L., 2014. VBA: A Probabilistic Treatment of Nonlinear Models for Neurobiological and Behavioural Data. PLoS Comput. Biol. 10, e1003441. doi:10.1371/journal.pcbi.1003441

Fehr, E., Schmidt, K.M., 1999. A theory of fairness, competition, and cooperation. Q. J. Econ. 114, 817–868.

Friston, K., Mattout, J., Trujillo-Barreto, N., Ashburner, J., Penny, W., 2007. Variational free energy and the Laplace approximation. NeuroImage 34, 220–34. doi:10.1016/j.neuroimage.2006.08.035

Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decision under risk. Econom. J. Econom. Soc. 263–291.

Le Bouc, R., Rigoux, L., Schmidt, L., Degos, B., Welter, M.-L., Vidailhet, M., Daunizeau, J., Pessiglione, M., 2016. Computational Dissection of Dopamine Motor and Motivational Functions in Humans. J. Neurosci. 36, 6623–6633. doi:10.1523/JNEUROSCI.3078-15.2016

Luce, R.D., 1959. On the possible psychophysical laws. Psychol. Rev. 66, 81.

Lynch, Jr., J.G., Chakravarti, D., Mitra, A., 1991. Contrast Effects in Consumer Judgments: Changes in Mental Representations or in the Anchoring of Rating Scales? J. Consum. Res. 18, 284. doi:10.1086/209260

Penny, W.D., 2012. Comparing dynamic causal models using AIC, BIC and free energy. NeuroImage 59, 319–30. doi:10.1016/j.neuroimage.2011.07.039

Penny, W.D., Stephan, K.E., Daunizeau, J., Rosa, M.J., Friston, K.J., Schofield, T.M., Leff, A.P., 2010. Comparing families of dynamic causal models. PLoS Comput. Biol. 6, e1000709. doi:10.1371/journal.pcbi.1000709

Pessiglione, M., Schmidt, L., Draganski, B., Kalisch, R., Lau, H., Dolan, R.J., Frith, C.D., 2007. How the Brain Translates Money into Force: A Neuroimaging Study of Subliminal Motivation. Science 316, 904–906. doi:10.1126/science.1140459

Polnaszek, T.J., Stephens, D.W., 2013. Why not lie? Costs enforce honesty in an experimental signalling game. Proc. R. Soc. B Biol. Sci. 281, 20132457–20132457. doi:10.1098/rspb.2013.2457

Rangel, A., Camerer, C., Montague, P.R., 2008. A framework for studying the neurobiology of value-based decision making. Nat. Rev. Neurosci. 9, 545–556. doi:10.1038/nrn2357

Rigoux, L., Guigon, E., 2012. A Model of Reward- and Effort-Based Optimal Decision Making and Motor Control. PLoS Comput. Biol. 8, e1002716. doi:10.1371/journal.pcbi.1002716

Rigoux, L., Stephan, K.E., Friston, K.J., Daunizeau, J., 2014. Bayesian model selection for group studies - Revisited. NeuroImage 84, 971–85. doi:10.1016/j.neuroimage.2013.08.065

Rigoux, L., Stephan, K.E., Friston, K.J., Daunizeau, J., 2014. Bayesian model selection for group studies — Revisited. NeuroImage 84, 971–985. doi:10.1016/j.neuroimage.2013.08.065

Samuelson, P.A., 1938. The numerical representation of ordered classifications and the concept of utility. Rev. Econ. Stud. 6, 65–70.

Schmidt, L., Clery-Melin, M.-L., Lafargue, G., Valabregue, R., Fossati, P., Dubois, B., Pessiglione, M., 2009. Get Aroused and Be Stronger: Emotional Facilitation of Physical Effort in the Human Brain. J. Neurosci. 29, 9450–9457. doi:10.1523/JNEUROSCI.1951-09.2009

Schmidt, L., d'Arc, B.F., Lafargue, G., Galanaud, D., Czernecki, V., Grabli, D., Schupbach, M., Hartmann, A., Levy, R., Dubois, B., Pessiglione, M., 2007. Disconnecting force from money: effects of basal ganglia damage on incentive motivation. Brain 131, 1303–1310. doi:10.1093/brain/awn045

Schmidt, L., Lebreton, M., Cléry-Melin, M.-L., Daunizeau, J., Pessiglione, M., 2012. Neural Mechanisms Underlying Motivation of Mental Versus Physical Effort. PLoS Biol. 10, e1001266. doi:10.1371/journal.pbio.1001266

Stephan, K.E., Penny, W.D., Daunizeau, J., Moran, R.J., Friston, K.J., 2009. Bayesian model selection for group studies. NeuroImage 46, 1004–17. doi:10.1016/j.neuroimage.2009.03.025

Von Neumann, J., Morgenstern, O., 1947. Theory of games and economic behavior.

## C. Discussion

In this study, we showed that a rating task, a force task and a choice task were equally able to reveal the utility function underlying the behavior related to subjective valuation. Nevertheless, we showed that the convexity of the utility could be affected by the task. Our result remains to be discussed since the observed concavity of the behavior is not the one we would have expected. Indeed, usually, people tend to prefer averages than extreme ('50-50' than '100-0' for example), but in our task, the utility in the force and the rating measurements is concave (100-0 is preferred, whoever the receiver is). One potential explanation of this unexpected concavity is what Andreoni called *impure altruism* (Andreoni, 1990). He proposed that *giving* produces a 'warm-glow' effect (a positive feeling triggered by the act of giving) and therefore utility is positive for the giver when she allocates money to NGO for example. Usually, convexity is observed when an individual has to share money with another individual whilst in our study, subject shared money with an NGO they choose before running the tasks. Thus, a post-hoc interpretation could be that subjects were more sensitive to high amounts of money whoever the receiver was and that inequity was not penalizing since it was for the *best* (as a moral concern).

We also showed that the force task was triggering faster answers but noisier ones. On the other hand, the choice task seemed more complicated to implement and less direct to get access to subjective values. Thus, each task has its advantages and disadvantages that one should take into account when choosing one of them to investigate subjective valuation processes.

In the next studies, I used rating tasks to access subjective values of participants since it is simple, precise and efficient to reveal subjective values. I will use binary choice tasks not to infer subjective values but to invest decision-making processes and to check whether ratings were predictive of choices.

## 2. Investigation of the Brain Valuation System Local Field Potentials

### A. Introduction

We have seen that the Brain Valuation System, composed of the vmPFC, the PCC, the hippocampus and the ventral Striatum, identified mostly with human fMRI studies has several interesting properties. Indeed, it is subjective and parametric: the more someone likes an object, the more the system is going to be activated. It is generic: it codes the subjective value of items independently of their category. It is automatic: it codes subjective value even if participants are engaged in a task which does not imply evaluating the likeability of the object. Another property is its dependency to baseline: the highest the baseline is, the more likely it is that a rating will be high. Finally, it codes confidence on judgments on top of the subjective value coding.

Those properties are not mutually independent. Indeed, since value coding is automatic, the vmPFC might integrate values of contextual features (because of its generality) in its baseline, which in turn will affect value judgments, since they depend on the baseline. Moreover, it confidence is the value of the response, confidence coding is a consequence of the other properties; the vmPFC automatically aggregates the value of many things, including stimulus and response.

First, it is important to note that there are some discrepancies between those results and the ones observed in animal studies. Indeed, one of the critical areas for value coding in monkey electrophysiology is the OFC, which is laterally located compared to the vmPFC and which is rarely found to correlate with subjective value in humans. In monkeys, we obviously cannot access subjective value with reports[18] thus it is classical to use either learning tasks in which monkeys need to associate a cue to a reward with a specific probability, either choice tasks in which stimuli are also associated to specific rewards and/or probabilities. Thus, without a direct access to the 'real' subjective value, it is tricky to investigate electrophysiology of the vmPFC and OFC in animals.

Nevertheless, the baseline dependency property has been replicated in both humans and monkeys vmPFC (Abitbol et al., 2015). Genericity is consensual in the vmPFC but not in the OFC. Indeed, several types of neuronal responses have been observed in Camillo Padoa Schioppa's studies in which he showed that some neurons were coding the value of reward independently from their category and other were coding the value of items from specific category. Thus, what we would observe in electrophysiological recordings in human is not trivial to predict. Automaticity and quadratic coding are also questions that have not been formally assessed in an electrophysiological study of the vmPFC or the OFC. However, there is some supporting evidence of these properties in the literature. Indeed, automaticity can be seen in Pavlovian paradigms without choices (Clark et al., 2012) and quadratic coding has been observed, as in the mentioned work of Padoa-Schioppa in which chosen value cells actually codes values in a quadratic form.

We will directly investigate the validity of all these properties in

---

[18] Even if pavlovian reactions such as lipping can be measured to get an approximate estimate (see Abitbol et al., 2015).

the following study.

In this second experiment, we had the opportunity to investigate intracranial electroencephalographic signals in 36 epileptic patients while performing a rating and a choice task on food items. Fourteen of them also completed a longer version of the task which allowed testing for the generic and automatic coding of subjective value. Those patients had the surgery for clinical purposes but we considered the recorded signals as healthy since there was no epileptic seizure during the included time periods. We had access to around 4000 cortical and hippocampal recording sites which allowed us running both region of interest analysis (ROI) and whole brain analysis similarly to fMRI studies. We used high gamma frequency as a proxy for local population coding and raised the following questions:

- What are the brain regions coding value in high gamma activity?
- Are vmPFC's partners (we did not have sufficient electrodes in PCC and no electrode in ventral Striatum) – OFC and Hippocampus – coding subjective value? What are there specificities? What are the frequency bands involved?
- Are the vmPFC properties replicated in the iEEG signal?

## B. Article

# Testing the core properties of the Brain Valuation System: an Intracranial EEG investigation.

A. Lopez-Persem(1,2), M. Petton(3,4), K. Lehongre(2,5), C. Adam(2,6), S. Rheims(3,7,8), P. Kahane(9,12,13) P. Domenech(10,11), J. Bastin(12,13), M. Pessiglione(1,2)

(1) Motivation, Brain and Behavior team, Institut du Cerveau et de la Moelle épinière (ICM), Paris, France

(2) ICM, INSERM UMRS 1127, CNRS UMR 7225, Université Pierre et Marie Curie UPMC-Paris 6, Paris, France

(3) Lyon Neuroscience Research Center, INSERM U1028, CNRS UMR5292, Brain Dynamics and Cognition Team, F-69500 Lyon-Bron, France

(4) Université Claude Bernard Lyon 1, France

(5) CENIR, ICM, UPMC/INSERM UMR S 1127 CNRS 7225, Hôpital Pitié-Salpêtrière, Paris, France

(6) Epilepsy Unit, Groupe Hospitalier Pitié-Salpêtrière, France

(7) Department of Functional Neurology and Epileptology, Hospices Civils de Lyon and Lyon 1 University, Lyon, France

(8) Epilepsy Institute, Lyon, France

(9) Neurology Department, Grenoble Alpes Hospital, F-38000 Grenoble, France

(10) Behavior, Emotion, and Basal Ganglia team – UPMC – INSERM UMRS U1127 – CNRS UMR7225

(11) CHU Henri Mondor, DHU Pepsy, Service de neurochirurgie fonctionnelle, Créteil, France

(12) University of Grenoble Alpes, F-38000 Grenoble, France

(13) INSERM U1216, F-38000 Grenoble, France

## *Abstract*

Estimating the value of alternative options is a key process in decision making. fMRI studies have identified a brain system, with the ventromedial prefrontal cortex (vmPFC) as a central component, which encodes values in a subjective, automatic and generic manner. Here, we examined the properties of this brain valuation system using a different recording technique, intracranial electroencephalography (iEEG), which offers direct access to the electrophysiological activity supposed to underpin hemodynamic responses, with much better temporal resolution. We recorded iEEG signals in 36 epileptic patients while they performed judgment tasks that involved rating different kinds of items (food, face, painting) and different features (age, likeability). The likeability ratings were significant predictors of choice and response time in a preference task where patients had to select their favorite items within pairs. In addition, both age and likeability ratings predicted second-order confidence judgments, through a quadratic (U-shaped) relationship. We identified several brain regions in which high-gamma activity (50-150 Hz) positively correlated with likeability ratings, including not only the vmPFC but also the lateral orbitofrontal cortex (lOFC) and the hippocampus. Focusing on these three regions, we characterized the dynamics of value coding across time and frequency bands. Then we used multivariate decoding across recording sites and frequency bands to assess the functional properties of valuation in the vmPFC. All properties derived from fMRI results were replicated: likeability ratings could be decoded in pre-stimulus activity (anticipation of value judgment), for both food and non-food items (generality), during both explicit and distractive tasks (automaticity), with both linear and quadratic functions (aggregation of value and confidence). Overall, our findings provide a bridge between two literatures on the neural underpinnings of subjective values: one based on electrophysiological studies in non-human primates and one based on fMRI studies in humans.

## *Introduction*

Estimating the value of alternative options is a key process in decision-making. Subjective values can be inferred from choices between options presented in pairs, but also more directly elicited using likeability or desirability ratings of options presented one by one. During the last decade, neuroeconomic studies have delineated a set of brain regions reflecting subjective values, termed brain valuation system (BVS, Rangel et al., 2008). Meta-analysis of fMRI studies have designated the ventromedial prefrontal cortex (vmPFC), the ventral striatum (VS) and the posterior cingulate cortex (PCC) as key components of the BVS, with the occasional addition of hippocampus and amygdala (Bartra et al., 2013; Clithero and Rangel, 2014; J. Peters and Büchel, 2010).

Most fMRI studies using choice and rating tasks have focused on the vmPFC and identified core properties that may affect value-based judgment and choice. First, the vmPFC can assign values to different categories of objects, such as food, money, trinkets, faces, paintings, houses, charities etc. (Chib et al., 2009; Hare et al., 2010; Lebreton et al., 2009), in accordance with the idea of a common neural currency (Levy and Glimcher, 2012). Second, the vmPFC encodes subjective value in an automatic manner, meaning even if subjects are engaged in a distractive task (Lebreton et al., 2009), or passively viewing choice options (Levy et al., 2011). Third, the value signal expressed in the vmPFC depends on pre-stimulus baseline activity, which itself depends on the pleasantness of the external or internal context (Abitbol et al., 2015; San Galli et al., 2016). Fourth, the vmpFC not only encodes option or decision value, but also confidence in the judgment or choice (De Martino et al., 2013; Lebreton et al., 2015).

The other BVS regions appeared to share at least some of these properties but also exhibited some differences. For instance, the ventral striatum might be more concerned with linking value to behavioral activation in motivation paradigms (Knutson et al., 2001; Schmidt et al., 2012), or with signaling prediction errors following outcomes in learning paradigms (O'Doherty et al., 2003; Pessiglione et al., 2006). The PCC might be more automatic – signaling the same values whether or not they are used for making a choice (Grueschow et al., 2015). The hippocampus might be particularly involved when valuation requires imagining what the outcome would be like (Lebreton et al., 2013; Peters and Büchel, 2010). Finally, other regions have repeatedly been found to correlate negatively with subjective values, such as the anterior insula and dorsal anterior cingulate cortex (Knutson et al., 2014; Pessiglione and Delgado, 2015).

However, most of these studies investigated BVS properties using fMRI, a technique that allows spatial localization at the macro-scale but not investigation of temporal dynamics. Interestingly, other techniques such as single-cell recordings in monkeys have yielded somewhat different conclusions. Neurons encoding subjective values have been found not only in the vmPFC (Abitbol et al., 2015; Bouret and Richmond, 2010; Strait et al., 2014) but also in many other brain regions, notably the lateral orbitofrontal cortex (lOFC, Padoa-Schioppa, 2007; Wallis, 2011). In an effort to bridge across techniques and species, we tested here BVS properties using intracranial electro-encephalography (iEEG), which gives access to local field potentials in deep structures of the human brain, with high temporal resolution.

Such recordings can be obtained from patients with resistant epilepsia who are implanted with intracranial electrodes for up to two weeks before surgery. This period provides a unique window into iEEG dynamics during performance of cognitive tasks in humans. Such iEEG recordings have revealed that high-frequency neural activity is involved in many cognitive tasks, with the high-gamma band (50-150Hz) as a good indicator for the participation of local neural ensembles (Fries, 2005; Ray et al., 2008; Lachaux et al., 2012). Indeed, power in the high-gamma band was shown to better correlate with spiking activity than the other frequency ranges (Ray et al., 2008; Ray and Maunsell, 2011).

In this study, we recorded iEEG signals in 36 epileptic patients while they performed rating tasks. From our large dataset, including a total of 4273 recording sites, we identified some brain regions encoding subjective values, notably the vmPFC, the lOFC and the hippocampus. We then characterized the dynamics of value representation in these regions, and validated the four properties of the vmPFC valuation signal that were identified in fMRI studies.

## *Results*

### Behavior: choice, confidence and response time are functions of ratings

All 36 patients (37.9 ± 10.7 years old, 21 females, see demographical details in Table 1) performed a series of rating and choice tasks (Figure 1A-C). A subset of patients (n=22) performed a short version of the behavioral tasks containing food items only. In a first block they had to rate the likeability of all food items presented one by one. In a second block they had to choose between items presented in pairs, according to their preference. Another subset of patients (n=14) performed a long version of the tasks with three additional components. First, two other categories of items were included: faces and paintings. Second, a distractive task was inserted where patients rated the age of faces and paintings. Third, a confidence rating task was added on top of the first-order (age or likeability) rating of food, face and painting items. The choice task was identical to that of the first version, with items blocked by category (food, face, painting).

We first analyzed the tasks included in the short version that was common to all 36 patients (red dots in Figure 1E-F). The aim was to test whether likeability ratings could predict choice behavior. Logistic regression against choice rate (in a left versus right frame) showed a significant effect of decision value (left minus right item rating) at the group level ($\beta_{Food}=0.42\pm0.07$, $t(34)=6.4$, $p=2.10^{-7}$). Linear regression against response time (RT) showed a significant effect of the unsigned decision value ($\beta_{Food}=-72.8\pm14.0$, $t(35)=-5.19$, $p=9.10^{-6}$), which can be considered as a proxy for choice difficulty.

We then analyzed the tasks that were specific to the long version, pooling non-food items (faces and paintings) to increase statistical power (black dots, Figure 1C-F). We found again that signed decision value predicted choice rate ($\beta_{Non\text{-}Food}=0.37\pm0.03$, $t(13)=12.8$, $p=6.10^{-8}$) and that unsigned decision value predicted RT ($\beta_{Non\text{-}Food}=0.37\pm19.8$, $t(13)=-6.32$, $p=6.10^{-5}$). We also verified that decision value was a significant predictor of choice in all patients for all categories (all $p<0.05$).

Next we confirmed a quadratic relationship between first-order (age and likeability) and confidence ratings ($\beta_{quad}=0.21\pm0.03$, $t(13)=6.36$, $p=2.10^{-5}$). This effect was significant for both types of items in both rating tasks ($\beta_{quad/food/like}=0.38\pm0.08$, $t(13)=4.79$, $p=4.10^{-4}$; $\beta_{quad/non\text{-}food/like}=0.30\pm0.06$, $t(13)=4.92$, $p=3.10^{-4}$; $\beta_{quad/non\text{-}food/age}=0.16\pm0.04$, $t(13)=4.62$, $p=5.10^{-4}$). Note that a linear term was included in the polynomial fit and was only significant for confidence in age rating ($\beta_{quad/non\text{-}food/age}=-0.14\pm0.05$, $t(13)=-2.81$, $p=0.01$), probably reflecting the fact that some ages were easier to rate than others. On the contrary, confidence in likeability ratings followed a pure (U-shaped) quadratic relationship.

Finally, we checked the quadratic link (inverted U-shape) between response time (from item display to first button press) and first-order judgment ($\beta_{quad}=-0.30\pm0.06$, $t(13)=-4.96$, $p=2.10^{-4}$). This quadratic link was significant for both types of items and for both judgments ($\beta_{quad/food/like}=-0.33\pm0.01$, $t(13)=-2.93$, $p=0.012$; $\beta_{quad/non\text{-}food/like}=-0.34\pm0.09$, $t(13)=-3.66$, $p=3.10^{-3}$; $\beta_{quad/non\text{-}food/age}=-0.48\pm0.02$, $t(13)=-2.53$, $p=0.025$). The linear term included in polynomial fit was not significant (all $p>0.17$).

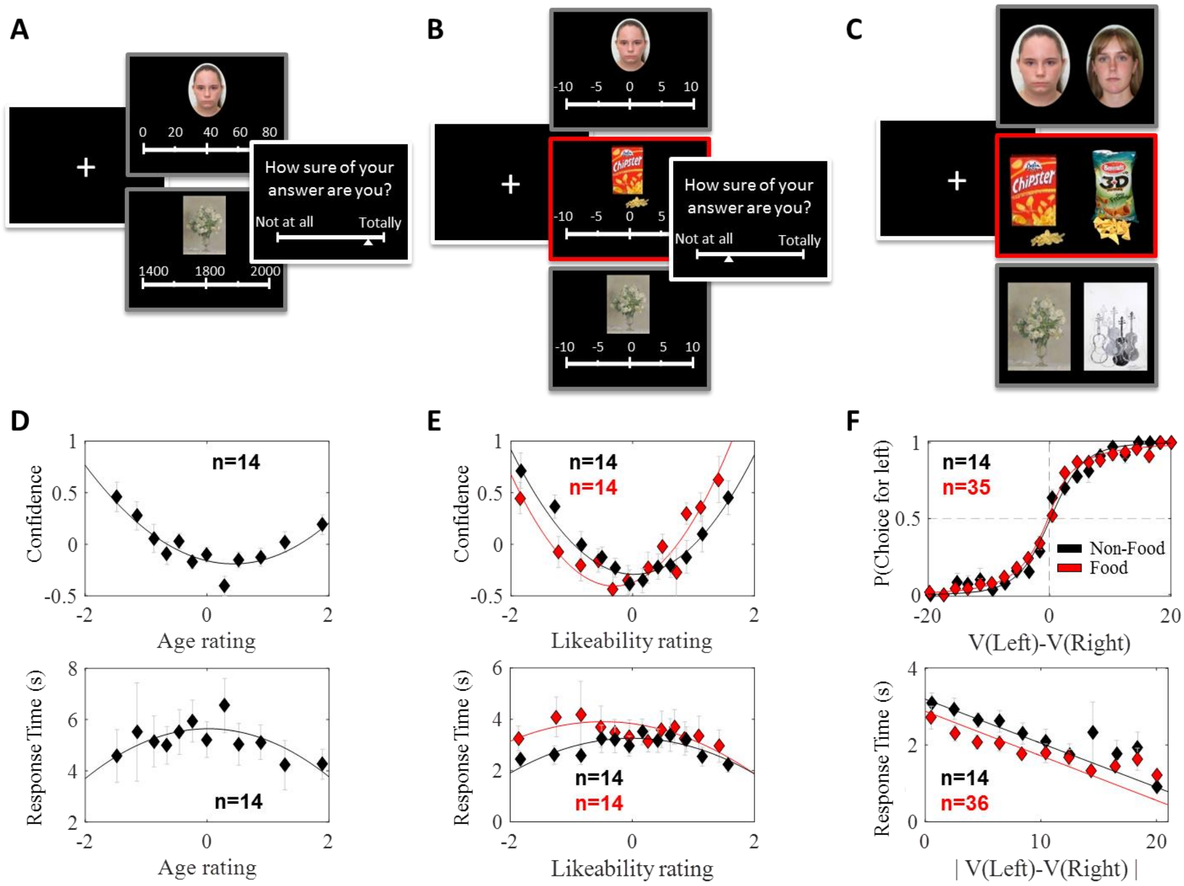


Figure 1 – Behavioral tasks and results

**Top.** Successive screens shown in example trials. In the age rating task **(A)** of non-food items (faces or paintings, framed in dark grey) patients provided judgments using analog scales between 0 and 80 years for the age of faces and between 1400 and 2000 years for the date of paintings. Confidence in the first-order rating was reported on a continuous scale going from ‘not at all confident’ to ‘totally confident’. In the likeability rating task **(B)**, patients indicated on an analog scale how much they liked the item from -10 to 10. The item could be food (red frame) or non-food (faces and paintings, dark grey). In the choice task **(C)**, two items presented in the rating tasks were displayed on the screen and patients had to select the one they preferred. **Bottom**. Behavioral results. In the rating tasks, the quadratic link was positive between confidence and age or likeability rating (**D-E**, top), and negative between RT and age or likeability rating (**D-E**, bottom), for both food (red dots) and non-food (black dots) items. In the choice tasks, likeability ratings (option values V) predicted both choice rate (**F**, top) and RT (**F**, bottom), for both food (red dots) and non-food (black dots) items. n=X indicates the number of patients tested for each task Diamonds dots indicate binned data averaged across patients. Error bars indicate inter-subject S.E.M. Lines corresponds to polynomial fit for D and E, logistic fit for top F and linear fit for bottom F.

When including squared rating and RT in a same regression model meant to explain confidence rating, we found that both regressors were significant predictors of confidence ($\beta_{rating^2}$=0.22+/-0.04, t(13)=6.05, p=4.10$^{-5}$; $\beta_{RT}$=-0.10+/-0.03, t(13)=-3.18, p=7.10$^{-3}$), but that squared ratings were better predictors than RT (t(13)=2.6, p=0.02). This suggests that the confidence expressed in ratings was not a direct readout of RT, but rather a readout of some uncertainty that also affected RT, together with other factors.

## Intracranial EEG: value coding in high-gamma frequency, a pseudo whole-brain analysis

Our data set included a total of 4273 recording contacts. Contacts with bad quality signal were removed and bipolar montages were computed for each pair of adjacent contacts. Dipoles were then labeled according to a re-parcellation of the AAL atlas ensuring that the different regions of interests (ROIs) had comparable size (see Methods and Table S1). Among the remaining 3440 bipolar recordings in our data set that can be seen on Figure 2 (top), 3194 sites could be labeled as belonging to one AAL area. Among the 115 covered areas, 77 had at least 9 bipolar recordings and were included in the 'pseudo whole-brain' analysis (Figure 2, bottom).

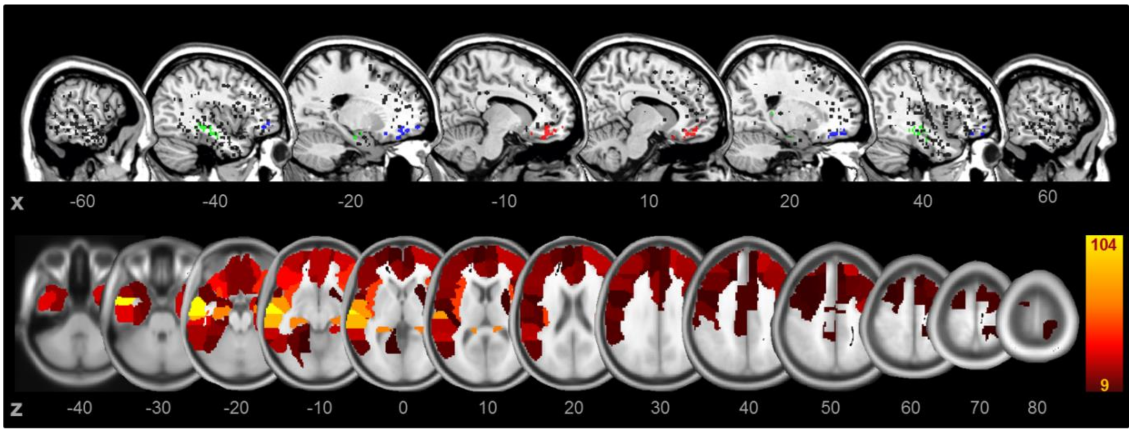


Figure 2 – Anatomical locations of the bipolar recording sites in the whole data set

**Top**. Sagittal slices of a brain template which represents the approximated locations of the dipoles recorded in 36 epileptic patients. Each dot represents one dipole. Red dots are the ones labelled as belonging to the vmPFC ROI, blue ones to the OFC and green ones to the hippocampus. **Bottom**. Axial slices of a brain template that represents all AAL areas including at least 9 dipoles (dark red). Light yellow indicates the maximum number of dipoles that was recorded in a same area (104). x and z coordinates refer to the MNI atlas.

For this analysis we focused on high-gamma band activity (50-150Hz), since it was found to correlate with both spiking and fMRI activity (Logothetis et al., 2001; Ray et al., 2008). High-gamma power was extracted from each dipole and each time point and regressed against subjective values. We took likeability ratings obtained for food items as the main proxies for subjective values, because they were collected in all our patients. For each brain region of the restructured AAL atlas, we averaged the time courses of regression estimates across all recorded dipoles. We then tested the significance of the regression in a fixed effect analysis (pooling dipoles across patients), using permutation tests at the cluster level ($p<0.05$) with a false discovery rate (FDR) procedure. We will refer to this analysis as the 'whole-brain t-value' analysis. We also ran a 'whole-brain percentage' analysis in which we computed the percentage of dipoles having a significant cluster positively correlated with subjective values.

We found 38 significant ROI in the whole-brain t-value analysis and 23 ROI containing more than 33% of electrodes coding value (arbitrary threshold) in the 'whole-brain percentage' analysis, with 17 areas in common including the vmPFC (Supplementary Table 2). Among them, four ROI were bilateral: two of them were labeled as 'OFC' (central and lateral), two as 'hippocampus', and two as 'fusiform anterior'.

Those pseudo whole-brain analyses are consistent with observations that subjective values are represented in a large set of brain regions and raise the question of how specific the BVS is. One interesting result is that the hippocampus and OFC appeared to strongly express subjective values (with higher t-values and percentages than vmPFC). Given this result and the literature reviewed in the introduction we defined the vmPFC, the hippocampus and the lOFC as our main regions of interest.

### Comparison of value coding between ROI and frequencies

Among all dipoles, 73 (belonging to 26 patients, with 22 dipoles in 8 patients who performed the long version of behavioral tasks) were anatomically located in the vmPFC ROI, which was defined as the gyrus rectus plus the fronto-medial part of orbitofrontal cortex bilaterally (in red on Figure 3A and table S1). The OFC ROI was defined as the bilateral central and laterals parts of the orbitofrontal cortex (AAL labels: Frontal Superior orbital and Frontal Middle Orbital, respectively) and comprised 160 dipoles (in blue on Figure 3A and Table S1). The hippocampus ROI included 104 dipoles on the left hemisphere and 102 on the right, yielding to a total of 206 dipoles (in green on figure 3A and Table S1).

Since ROI were selected, among the priors derived from literature, based on whole-brain analysis of high-gamma activity, we tested whether other frequency bands could also be involved in value coding. We first performed a time-frequency analysis on the evoked response of each dipole and averaged power across all dipoles belonging to each ROI (Figure 3B). We observed a global increase in the high-gamma frequency power in the three ROIs after stimulus onset. In addition, there was in

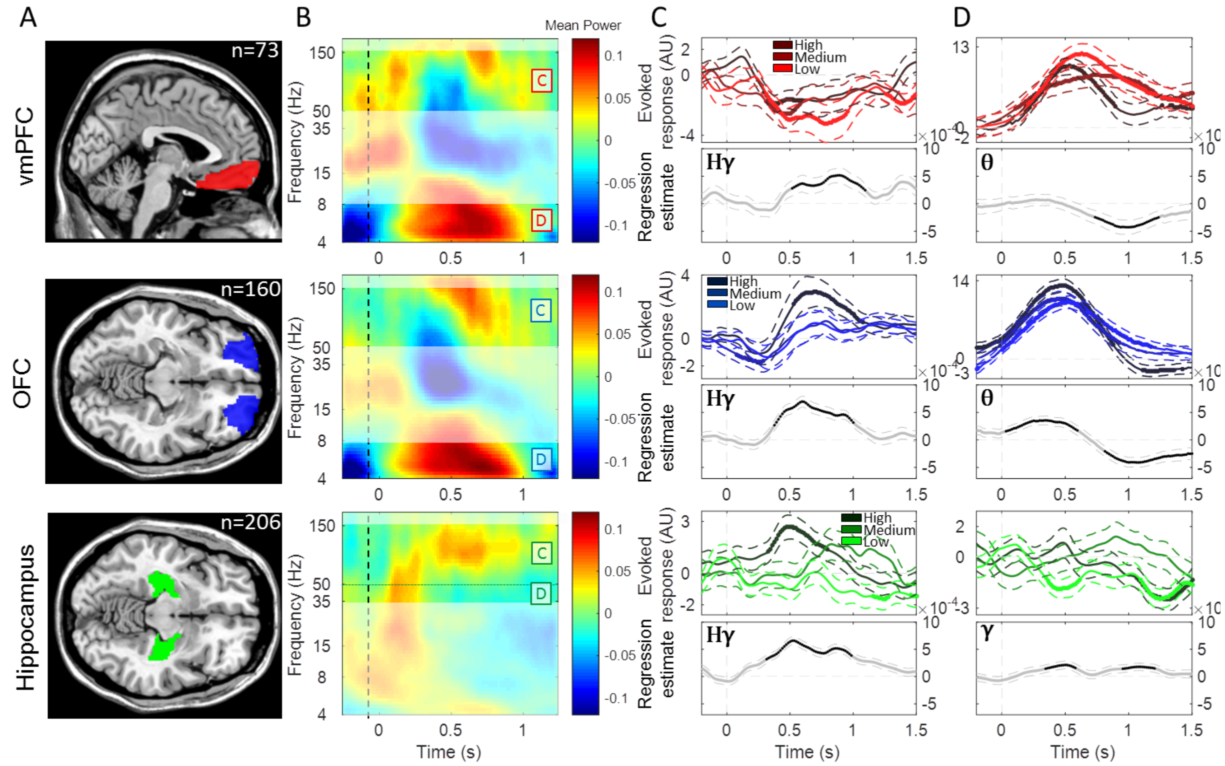


Figure 3 – Time-frequency investigation of the vmPFC, OFC and hippocampus

**A.** Anatomical localization of the vmPFC (red, first row), OFC (blue, second row) and hippocampus (green, last row). All dipoles located in those areas were included in the ROI analysis. B. Time-frequency analysis of the evoked response following stimulus onset (dashed vertical line). Hotter colors indicate higher power. Insets 'C' and 'D' label the frequency bands (not shadowed) that are investigated in panels C and D. **C.** Time courses of high-gamma (Hγ) activity splitted into tertiles of likeability ratings (top), or regressed against likeability ratings (bottom), locked on the stimulus onset (dashed vertical lines). Black points indicate significance of regression estimates. **D**. Same as C except that the analysis was focused on theta activity (θ) for the vmPFC and lOFC and on gamma activity (γ) for the hippocampus.

increase in alpha-theta power (4-15Hz), and a decrease in beta power (15-35 Hz) for both vmPFC and lOFC ROI. In contrast, we observed an early increase in the beta band (15-35Hz), followed by a later decrease, in the hippocampus.

Then, we looked for any significant correlation with likeability ratings in the five frequency envelopes defined as Theta (4-7Hz), Alpha (8-15Hz), Beta (15-35Hz), Gamma (35-50Hz) and High-gamma (50-150Hz). We found significant clusters in the high-gamma band for all three areas (vmPFC: $t(72)=108.7$ $p<1.10^{-4}$, 520-1010 ms after stimulus onset; lOFC: ($t(159)=234.2$, $p<1.10^{-4}$, 360-1090 ms; hippocampus: $t(205)=126.7$, $p<1.10^{-4}$, 310-990 ms), which confirmed the result obtained in the whole brain analysis (Figure 3C, Table 2). In order to visualize the time courses of high-gamma evoked responses we splitted the trials into tertiles of subjective values. Figure 3C). Interestingly, we observed a global decrease in the vmPFC, which was less pronounced for high than for low values, similarly to what is often observed in fMRI studies. The lOFC and hippocampus responses were on the contrary characterized by a global increase, which was more pronounced for higher values.

Finally, we found significant clusters in other frequency bands, notably in the gamma range for all ROIs, but also in the theta range for the vmPFC, in the alpha range for the hippocampus, and in virtually all bands for the OFC. Note that some of these correlations were negative (see Figure 3D and table 2).

## Core properties of value coding in the vmPFC

In the following analyses we tested whether the core properties of the BVS identified with fMRI (anticipation, generality, automaticity, quadratic coding) were also observable in iEEG activity, and how they would unfold over time. We focused on the vmPFC, because it has been the main ROI explored in fMRI studies, and on the high-gamma band, because it has been suggested to co-vary with hemodynamic activity. Our approach was to start with food likeability ratings (for which we had a larger dataset) and then to extend regression analyses to other categories (non-food items), other judgment (age and not likeability rating) and other type of coding (quadratic and not linear).

We extended the regression of high-gamma power against food likeability ratings to the entire trial, with time locking either to stimulus onset or to motor response (first button press made to move the cursor along the rating scale). We observed significant clusters not only 0.52-1.09 s after stimulus onset (as shown in Figure 3C) but also around the motor response (Figure 4A), between -265 and 100 ms ($t(72)=63.03$, $p=0.02$) and 340 and 670 ms ($t(72)=53.02$, $p=0.04$). This result suggests that value representation is maintained until the rating is provided.

In order to assess the contribution of other frequency bands in encoding value we ran a multivariate regression on the iEEG time series, each regressor corresponding to one dipole and one frequency (High-Gamma, Gamma, Beta, Alpha, Theta). Then, we computed the PRESS statistic (predicted residual error sum of squared, Myers and Montgomery, 2002; Trujillo-Ortiz et al., 2006) for each time point, following a cross-validation procedure (see methods). Significance was tested at the group level (across subjects), using permutation tests. We found a significant cluster between 860 and 1000 ms ($t(26)=34.03$, $p<5.10^{-3}$, corrected for multiple comparisons), overlapping with that seen in the linear regression on high-gamma activity locked to stimulus onset (Supplemental Figure 1). We also found significant clusters before and after the first button press (around -600 ms: $t(26)=9.30$, $p=5.10^{-3}$ and 430ms: ($t(26)=9.20$, $p=5.10^{-3}$). Thus, multivariate decoding using the full frequency spectrum confirmed the links between vmPFC iEEG activity and subjective values that was observed in the high-gamma band. We next tested each property of value coding in the vmPFC, focusing on high-gamma power.

### *Anticipation (baseline activity predicts value judgments)*

To assess whether vmPFC baseline activity would influence subjective values we tested the correlation between food likeability rating and high-gamma power in the -200 to 0 ms window aligned to stimulus onset, but it failed to reach significance ($t(72)=1.91$, $p=0.059$). However, when we separated the dorsal and the ventral part of the vmPFC according to the AAL atlas (frontal medial area and gyrus rectus, respectively), we found a significant effect in the dorsal sub-region ($t(43)=2.23$, $p=0.03$). Thus, significant baseline effect was observed in a region that was close to the vmPFC activation cluster that has been reported in fMRI studies. We also tested each vmPFC dipole and found that 43% were significant predictors of subjective values (between -200 and 0ms, all $p<1.10^{-3}$). Note that in the multivariate decoding using the full frequency spectrum and the entire vmPFC ROI (Supplemental Figure 1), there was indeed a significant cluster in baseline activity

($t(26)=19.44$, $p<5.10^{-3}$).

In order to examine whether the same code was shared between time windows (pre-stimulus activity and evoked response), we trained the multivariate value decoder on the baseline (-200 to 0 ms) and tested it on each time point of the entire trial (Figure 4B). We found that likeability ratings were decodable, with the weights obtained from the baseline, between 750 ms and 1200 ms ($t(26)=272.8$, $p<1.10^{-4}$).

### *Generality (evoked activity reflects value judgments for different sorts of items)*

To invest whether the vmPFC would encode subjective values across categories we extended the linear regression on high-gamma power to the likeability of non-food items (Figure 4C), which were obtained in the subset of patients who performed the long version of behavioral tasks. There was no significant cluster at the group level even if a trend could be observed around 800 ms ($t(21)=1.85$, $p=0.078$). When testing each dipole separately, we found that 55% were presenting at least one significant cluster, with 33% being significant for both food and non-food items, 33% only for food items and 33% only for non-food items. We could not identify any topographical organization in the location of significant dipoles. This result prevented us from drawing any conclusion, so we investigated generality through multivariate decoding.

To assess whether the value of food and non-food items would be represented with the same code in the vmPFC, we trained the multivariate decoder of food likeability ratings at all time points, and tested it on non-food likeability ratings  at all time points, in order to obtain a 'temporal generalization matrix' (King and Dehaene, 2014). From this matrix, we determined for each patient a training time window in which cross-decoding was maximal. Then we generated the decoding time-series, and tested for significant effects at the group level (Figure 4D). We found significant clusters from 310 to 810 ms after stimulus onset ($t(7)=46.05$, $p=5.10^{-3}$) and from 920 to 1230 ms after the stimulus onset ($t(7)=63.93$, $p<5.10^{-3}$). Interestingly, there was also a significant cluster around stimulus onset (from -250 ms to 250 ms, $t(7)=145.6$, $p<1.10^{-4}$), which is replicating the property that baseline activity is predictive of value judgment. These results therefore suggest that that a same code is used, not only in different time windows, but also for different categories of goods, in accordance with the idea of a common neural currency.

### *Automaticity (evoked activity reflects value judgments during non value-related tasks)*

To check whether the vmPFC would also reflect subjective values during a distractive task, we extended the linear regression of high-gamma power against the likeability of non-food items to the iEEG time series recorded during age rating (Figure 4E). We found a trend in a 860-1000 ms, which failed to survive correction for multiple comparisons (t(21)=35.73, p=0.09). When investigating separately faces and paintings, we found that this trend was driven by a significant automatic coding of face values (from 770 to 1000 ms, t(21)=76.0, $p=6.10^{-3}$). When investigating the regression results at the single dipole level, we found that 52% presented a significant cluster, with 33% during both age and likeability ratings, 33% only for age and 33% only for likeability. Again there was no topographical organization to be seen in the location of significant dipoles, so we explored the automaticity using multivariate decoding.

The multivariate decoder was trained on non-food values during likeability rating, and tested on the same values but during age rating (Figure 4F). As for generality we identified training time windows showing maximal cross-decoding at the patient level, and tested significance of decoding performance at the group level. We found a significant cluster between 170 ms and 700 ms ((t(7)=54.4, $p<1.10^{-4}$). This result suggests that a similar code was used across different tasks, whether they required an explicit value judgment (likeability rating) or not (age rating).

### *Quadratic coding (evoked activity reflects confidence judgments on ratings)*

To test whether the vmPFC would also reflect confidence in judgments, we regressed high-gamma against squared rating, pooling age and likeability ratings collected in the long version of behavioral tasks, in order to improve statistical power (Figure 4G). We did not use confidence judgments since we wanted to compare the regression results to those obtained with the squared ratings of food likeability, for which we do not have confidence rating. Interestingly, we did not find any significant cluster (all p>0.05) when the signal was locked on stimulus onset, but we did find several significant clusters when locked on the first button press (from -340 to 140 ms, t(21)=82.9, p=0.014 and from 470 to 1010 ms: t(21)=99.8, p=0.006). We found similar results with squared food ratings: no significant clusters when locking on stimulus onset but two significant clusters when locking on first button press (from -1000 ms to -370 ms: t(72)=173.6, p<1.10-4 and from -10 to 250 ms: t(72)=58.3, p=0.03) (Supplementary figure 2A, middle panel).

We conducted a supplementary analysis to invest whether linear and quadratic coding were dissociated in time. High-gamma power was extracted either from the time window showing significant linear correlation with likeability rating (500 to 900 ms after stimulus onset), or from the time window showing significant quadratic correlation with ratings (-300 to 100 ms around button press). Then the signal was binned according to ratings and fitted with a second-order polynomial function to test the significance of linear and quadratic components (Supplementary figure 2). We found that in the first window (following stimulus onset-, only the linear correlation with likeability rating was significant (linear: t(72)=3.3, $p=2.10^{-3}$; quadratic: t(72)=1.50, p=0.14), whereas in the second window (around button press), both linear and quadratic terms were significant for likeability rating (linear: t(72)=3.3, $p=2.10^{-3}$; quadratic: t(72)=3.4 ,p=0.02), and only the quadratic term for age rating (linear: t(21)=0.39, p=0.70; quadratic: t(21)=2.6, p=0.02) (Supplementary figure 2B). This confirms that linear coding is only present for values (not age), in both early and late windows (stimulus and response), whereas quadratic coding is present for all ratings (both age and value), only in a late window (close to motor response).

We also ran the polynomial regression at the dipole level (with the signal locked on first button press) and found that 90% of dipoles were significantly reflecting squared judgment, with 47% coding both squared age and squared likeability, 32% only squared likeability and 21% only squared age.

In order to invest whether the code was shared between stimulus value and response confidence, we trained a multivariate decoder on food and non-food item value during the likeability rating task, and tested it on the squared value of the same items in the same task. We found a significant cluster from -200 ms to 640 ms around stimulus onset (t(25)=95.6, $p<1.10^{-4}$), and from -730 to -250 ms before the first button press (t(25)=49.3, p=0.03). Note that an additional cluster was found between 440 and 800 ms after the first button press on the edge of significance (t(25)=45.8, p=0.050). These results suggest that confidence in the judgment is coded in the same way (with the same pattern of weights on dipoles and frequencies) as the value of the stimulus. This is in line with the previous properties if confidence is considered as a value judgment on just another type of event (the response).

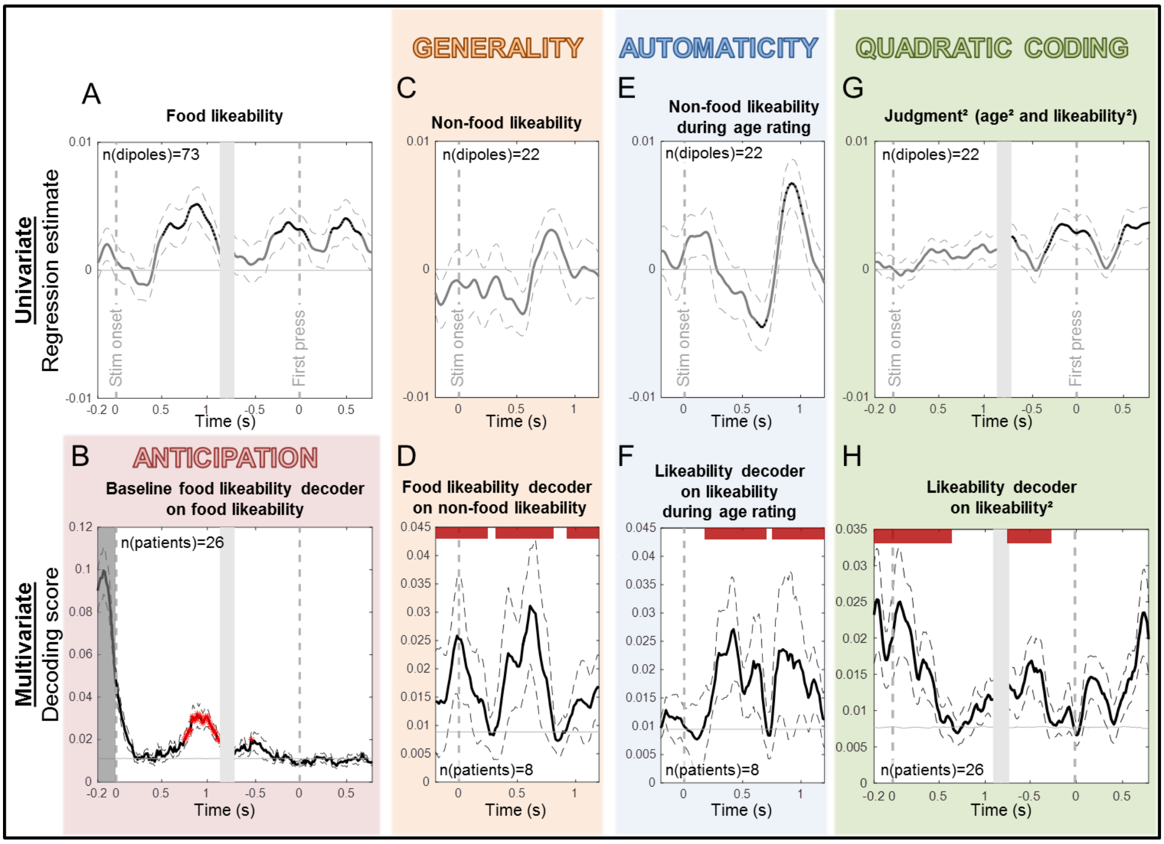


Figure 4 – Univariate & Multivariate investigation of vmPFC properties

**A.** Mean time course of regression estimates for high-gamma vmPFC activity against food likeability, locked either on stimulus onset (left) or on first button press. (right). **Anticipation property**. **B**. Mean time course of likeability decoding score from the multivariate decoder trained on the pre-stimulus time window (in grey) and tested on every time point during the food likeability rating task (different window, same items, same task). **Generality property. C**. Mean time course of regression estimates for high-gamma vmPFC activity against non-food (faces and paintings) likeability, locked on stimulus onset. **D.** Mean time course of likeability decoding score from the multivariate decoder trained on food items and tested on non-food items during the likeability rating task (different items, same task). **Automaticity property**. **E.** Mean time course of regression estimates for high-gamma vmPFC activity against non-food items (faces and paintings) likeability during the age rating task, locked on stimulus onset. **F.** Mean time course of non-food likeability decoding score from the multivariate decoder trained on during likeability rating and tested during age rating (same items, different task). **Quadratic coding property. G.** Mean time course of regression estimates for high-gamma vmPFC activity against squared ratings of both age and likeability, pooling all types of items (food, faces, paintings), locked on either stimulus onset (left) or first button press (right). **H.** Mean time course of decoding score from the multivariate decoder trained on likeability ratings (of both food and non-food items) during the likeability rating task (same items, same task, different variables). Regression estimates and decoding scores were tested at the group level, across dipoles or patients, respectively. Dashed lines indicate inter-dipole or inter-patient SEM. Black dots (top) or red parts (bottom) indicate significant clusters.

Finally, all multivariate analyses described in the previous sections were repeated with inclusion of all the dipoles across the three main ROI (vmPFC, lOFC, hippocampus), in an effort to improve multivariate decoding. We could not improve the accuracy of decoding, which suggests that the lOFC and hippocampus provides no additional information to what could be observed in the vmPFC.

## *Discussion*

Using a large dataset of iEEG signals recorded in 36 epileptic patients during judgment tasks, we were able to provide information about the BVS regarding 1) anatomical localization over the whole-brain, 2) contribution of the different frequency bands, 3) functional properties of value correlates. In the following, we successively discuss these three aspects of our study.

Using univariate and multivariate analysis on a large dataset of electrophysiological recordings in epileptic patients, we were able to replicate the fMRI established properties of the vmPFC in subjective value coding. Indeed, we found that the baseline activity was predictive of the likeability rating which is about to be provided by patients and of the evoked activity after the stimulus onset. Moreover, we showed that the vmPFC code for value of food items was similar to the one used for subjective value of non-food items, suggesting a 'common neural currency' property of this area. We also showed that the automaticity property proposed by several previous studies could also be observed in the high-gamma activity, with a coding of subjective value found during an age evaluation task. We were also able to find a quadratic coding of judgment locked on the first button press in the high-gamma, which is suggesting a coding of confidence linked to the response.

### Delineating the Brain Valuation System

Using direct regression of high-gamma activity against the likeability ratings assigned to food items, we found a large set of brain regions that was significantly reflecting subjective values. Some of these significant ROI were part of the standard BVS defined from meta-analyses of fMRI studies (e.g., vmPFC, hippocampus), some corresponded to areas where single cells were found to encode subjective values (e.g., lOFC) and some were less classically associated with valuation (e.g., fusiform gyrus). The fact that some standard results were replicated is reassuring regarding the correspondence between hemodynamic and electrophysiological signals (Logothetis et al., 2001). Also, the fact that value correlates were seen in the lOFC fills a gap between human fMRI studies and monkey single-cell recordings. It shows that linear coding of subjective values can be observed using iEEG in the human brain, as was already shown using spiking activity in monkeys (Padoa-Schioppa and Assad, 2006), and contrary to what was suggested using fMRI in humans (Howard et al., 2015). Thus, the high-gamma band, which is often considered as a robust correlate of local neural activation (Jerbi et al., 2009; Lachaux et al., 2012), seems to provide an interesting link between human and animal studies on subjective valuation.

However, the rather long list of brain regions that was significantly correlating with subjective values questions the specificity of the BVS identified with fMRI. It could also mean that the statistical threshold applied here was less stringent than what is typically used in fMRI. It remains unclear whether the additional regions play a role in the construction of subjective values, as was suggested for the hippocampus (Barron et al., 2013; Lebreton et al., 2013; Peters and Büchel, 2010), or whether they are influenced by subjective values in top-down fashion. This could be the case of the fusiform gyrus, if we may speculate that the value of visual items might direct spatial attention and improve perceptual processing of these items.

We also note that some classical BVS regions were not in the list of significant ROI. One obvious reason is that they were not covered in our sample of recording sites, as for the posterior cingulate cortex and the ventral striatum. More generally, the unequal sampling of the AAL atlas ROIs induces differences in the statistical power of the group-level analyses used to test regression. This is why we used the term ‘pseudo whole-brain’ and why the comparison between regions must be taken with caution. In particular, there were fewer dipoles in the vmPFC (n=73) compared to the lOFC (n=160) and hippocampus (n=206), which is sufficient to explain why p-values were lower. Another possibility for the absence of significant regression is that values may not be reflected in the high-gamma range.

### Dynamics of value coding across frequency bands

We then examined the dynamics of value coding across frequency bands, in the significant ROI overlapping with the literature on the neural implementation of subjective values. We acknowledge that our selection was rather arbitrary, and we may investigate other ROI in the prefrontal, cingulate or temporal cortex in the future.

We observed that subjective value was represented in the high-gamma range for all three regions, as expected from the pseudo whole-brain analysis. Interestingly, the temporal dynamics was different across the three regions, although significant correlation with subjective values was found in a similar time window (0.5-1s post-stimulus). High-gamma activity decreased in the vmPFC, this decrease being less pronounced for high values. This is typically what is observed in the hemodynamic response analyzed in fMRI studies (for an example see Abitbol et al., 2015). On the contrary, high-gamma activity rather increased in the lOFC and hippocampus, this increase being more pronounced for higher values. Thus, despite a similar positive correlation with subjective values, the underlying dynamics seemed different, which might have important functional repercussions.

When exploring lower frequencies, we found similar positive correlations with subjective values in the gamma range for all three regions at similar time points. There were also significant negative correlations in the theta/alpha band in all three regions, around 1 s post-stimulus. This is in line with the idea that lower frequencies such as theta and alpha rhythms are often coupled to high-gamma oscillations during cognitive processes (Jerbi et al., 2009). In the beta band, the only significant correlation was seen in the OFC, with a negative sign. This accords well with a recent study in monkeys (Zhang et al., 2016) reporting that reward anticipation was reflected positively in the high-gamma band (50-100Hz), but negatively in the beta band (15-29Hz). This mix of positive and negative value coding in the gamma and beta bands could be the reason why linear correlation with value in the lOFC is not captured with fMRI.

### Functional properties of value coding

All core properties of value coding were replicated in the vmPFC. These results seemed to be driven by the high-gamma band, but robust significant results were only obtained using multivariate decoding across dipoles and frequencies.

First, we found that subjective values were somehow predicted by pre-stimulus activity, particularly in the dorsal part of the vmPFC. This result is in line with what have been recently shown in both humans and monkeys (Abitbol et al., 2015). We provide here additional evidence that the code is shared between baseline activity and evoked response. Interestingly, our result does not support the mechanistic account of baseline effect proposed by Abitbol et al., 2015. Indeed, they suggested that a shift in baseline activity would persist into post-stimulus activity and impact the absolute peak of the vmPFC response, which in turn would impact the subjective value assigned to the stimulus. In the present analysis, with the multivariate decoder trained on baseline activity, such a mechanistic account would predict that the decoding score should progressively decrease with time. However, we observed that decoding score first decreases to chance level and then increases to become significant again in the time window representing stimulus value. Thus, the anticipatory effect of baseline must be generated by another mechanism, which remains to be elucidated.

Second, we found that the value of both food and non-food items was represented in the vmPFC. This is in line with several fMRI studies showing that likeability of faces and paintings was correlated with vmPFC activity (Lebreton et al., 2009), but also other kinds of items such as trinkets and money (Chib et al., 2009), consistent with the notion of a 'common neural currency' (Levy and Glimcher, 2012). This notion is critical for the ability to compare the values and choose between items belonging to seemingly incommensurable categories. Furthermore, we were able to decode non-food item values using the multivariate pattern optimized to discriminate food item values was, which provides another argument to the idea of 'common neural currency'. Contrary to some studies that suggested an anatomical segregation for value correlates of abstract and concrete rewards (e.g. ,Sescousse et al., 2010, 2013), we could not identify here any topographical organization across categories.

Third, we found that likeability ratings were expressed in the vmPFC even during the age rating task. This is replicating the result previously obtained with fMRI (Lebreton et al., 2009), and confirms that valuation in the vmPFC is a somewhat automatic process. By this we do not mean that the valuation process is irrepressible, it has been indeed shown that value representation can be extinguished or blurred by other tasks (Grabenhorst and Rolls, 2008; Grueschow et al., 2015; Plassmann et al., 2007, 2010), but just that valuation does not need an explicit value judgment of choice task to be triggered. Such automatic valuation might explain a series of ‘halo effects’ reported in the literature, where the values of distractive contextual features contaminate the judgment of target stimuli. Moreover, our multivariate analysis suggests that the code for subjective value might be shared across explicit and distractive tasks.

Four, we found that ratings were also encoded in a quadratic manner in the vmPFC. When ratings are value judgments, this quadratic coding occurs on top of linear coding. Linear coding was observed in time windows time-locked to both stimulus onset and motor response, whereas quadratic coding appeared more locked to motor response. This is in line with the quadratic coding representing the confidence in first-order rating, as was suggested in a previous fMRI study (Lebreton et al., 2015). It also accords with previous report that the vmPFC reflects confidence in choice, on top of decision value (De Martino et al., 2013). Moreover, the multivariate analysis showed that the same pattern was used to represent value and confidence. This result supports the idea that confidence is equivalent to a value judgment on the response, which may be considered just another category beyond food and non-food stimuli. Indeed, confidence is defined as the accuracy of the response (first-order judgment) and participants surely prefer to be accurate than inaccurate. In our computational model (Lebreton et al., 2015), confidence is precisely the variable that is maximized when selecting a particular rating, which fits with the intuition that subjects give the response in which they are most confident. Confidence itself might be a generic and automatic process, meaning that it would be generated similarly for different tasks and without the need for explicit confidence rating, a possibility that has not been examine in the present study.

### Limitations

To our knowledge, this is the first time that BVS dynamics and properties are investigated through intra-EEG recordings. These data have been collected in epileptic patients, who were implanted with deep electrodes for a clinical purpose. Yet we interpret those data as if they were collected in healthy subjects, making the assumption that epileptic activity did not distort the neural implementation of subjective value. This assumption might be questionable, even if most results are in accordance with previous investigations in healthy human and non-human primates.

### Conclusion

In summary, using rare intracerebral data in humans, we were able to replicate what was known on the BVS but also to provide new insights on its anatomical delineation, temporal dynamics and functional properties. Notably, we suggest that the lOFC should be included in the BVS, that the anticipatory property of baseline is not due to temporal auto-correlation, and that value representations share a same code across time windows, stimuli, tasks and events (stimulus and response). Further analyses should focus on functional connectivity between BVS regions to better characterize the circulation of information that underpins how subjective values are constructed and used to guide behavior.

## Methods

### Patients and recordings

All the patients were suffering from drug-resistant partial epilepsy and gave an informed consent. iEEG data were acquired from three different centers. Seven patients were recorded at the Epilepsy Department of the Grenoble University Hospital, eighteen patients in Lyon, in the Epilepsy Department of the Neurological Hospital and eleven in Paris, in the Pitié-Salpétrière hospital. In Lyon and Grenoble, patients underwent intracerebral recordings by means of stereotactically implanted (Lachaux et al., 2003) multilead depth electrodes (sEEG). For each patient, 12 to 18 semi-rigid electrodes were implanted depending on the patient; each electrode had a diameter of 0.8 mm and comprised 6 to 18 leads of 2 mm, 1.5 mm apart (Dixi, Besançon, France), depending on the target region. The electrode contacts were identified on each individual stereotactic scheme, and then anatomically localized using the proportional atlas of (Talairach and Tournoux, 1988) after a linear-scale adjustment used to correct for size differences between the patient's brain and the brain in the Talairach atlas. Neuronal recordings were conducted using an audio–video-EEG monitoring system (Micromed, Treviso, Italy), which allowed the simultaneous recording of 128 depth-EEG channels sampled at 512 Hz [0.1–200 Hz bandwidth]. One of the contact sites in the white matter was chosen as a reference.

In Paris, the implantation of electrodes and the participation of patients to cognitive tasks received approval from local ethic committee (CPP Paris VI, INSERM C11-16). All patients provided informed and written consent to participate in the study. The electrodes (AdTech®, Wisconsin) consist of 4-12 platinum contact electrodes with a diameter of

1 mm and length of 1.6 mm, with nickel-chromium wiring. Neural recordings were conducted with Neuralynx (ATLAS, Neuralynx®, Inc., Bozeman, MO). Spatial localizations were determined on the basis of postimplant computed tomography scans coregistered with preimplant 1.5T MRI scans.) Placement of the electrodes was determined by clinical criteria (Fried et al., 1999). The reference electrode was defined as the one with the less activity, if possible one in the white matter. Signal was band pass filtered between 0.1 Hz and 1000Hz. Localization of electrodes has been recovered automatically using the Epiloc toolbox developed by the STIM (stim.cricm.upmc.fr) engineering platform in the ICM (Institut du Cerveau et de la Moelle Epinière, (http://icm-institute.org/en/cenir-stim-stereotaxy-core-facility-techniques-images-models-2/)(Perez-Garcia et al., 2015).

Before analysis, all signals were re-referenced to their nearest neighbor on the same electrode, yielding a bipolar montage.

## Experimental tasks

### *Short version: Localizer*

The short version of the task, composed of one likeability rating task and one binary choice task among food items, was only ran in Grenoble and Lyon. All food stimuli were displayed using Presentation (Neurobehavioral Systems, Albany, CA) on a 17 CRT monitor at 85 Hz. The timing of stimulus delivery was controlled via a TTL pulse that was sent by the stimulation PC to the EEG acquisition PC to synchronize all acquisition systems. The experiment consisted of the rating task, in which patients had to give an estimate of how much they would like to eat the food item presented on the screen. The answer was given with the right hand; the cursor could be moved by pressing left and right arrows on the keyboard and then validated by pressing the space bar. Note that one patient did not complete the choice task but we included the rating task in the analysis.

### *Long version*

In Paris, the task was programmed on a PC using the Cogent 2000 (Wellcome Department of Imaging Neuroscience, London, UK) library of Matlab functions for stimulus presentation. All trials started with a fixation cross lasting for 1500 ± 500 ms.

The long version of the task consisted in three steps. The "Age Rating & Confidence task" was composed of 60 trials divided in two randomized blocks (one with faces and the other with paintings) across patients. An image appeared on the screen and patients had to rate how old the presented stimulus was, on a 21-step scale that was adapted to the category (face or painting). After validation of the age judgment, a sentence reminding participants of their rating appeared ("You gave a rating of *X*"), together with another 100-step (almost continuous) rating scale, on which they were asked to indicate how confident they were about their first-order rating ("How confident are you?", between "Not at all", and "Totally").

Second, in the "Pleasantness Rating & Confidence task", patients were asked to indicate on a 21-step scale graduated from -10 to 10 how pleasant the presented stimulus was. They rated in three different blocks three types of stimulus: faces, paintings and food items. Each likeability rating was followed by a confidence judgment on the rating, presented in the same way as during the age rating task.

Finally, patients completed three blocks of a forced binary choice task among the previous stimuli. After the fixation cross, two pictures belonging to the same category (food, faces or paintings) were displayed on each side of the screen. Patients had to choose the one they preferred by pressing the left or right arrow of the keyboard.

### Behavioral analysis

Unless otherwise specified, all dependent variables (raw, *z*-scored or binned behavioral measures and regression estimates) were computed at the subject level and tested for significance at the group level (random effect analysis) using two-tailed paired *t*-tests. All regressions were performed on *z*-scored independent and dependent variables. Data distribution was assumed to be normal but this was not formally tested. All statistical analyses were performed with Matlab Statistical Toolbox (Matlab R2015a, The MathWorks, Inc., USA).

### Electrophysiological analysis

iEEG signals were evaluated with the software package for electrophysiological analysis (ELAN-Pack) developed in the laboratory (INSERM U1028, Lyon, France) and Matlab algorithms. For each single trial, bipolar derivations were computed between adjacent electrode contacts to suppress contributions from non-local assemblies and assure that the bipolar sEEG signals could be considered as originating from a cortical volume centered within two contacts (Jerbi et al., 2009). Data were visually inspected in order to remove artifacts such as the 50Hz noise from the standard electrical line current.

### *AAL restructuration*

The AAL (Automated Anatomical Labeling) Atlas was used to label each electrodes located in the MNI space. However, in order to get brain regions with similar size to compute statistics among electrodes, the brain regions larger than 2000 voxels where separated in two brain regions according to the larger axis (until the volume of each new brain region is inferior to 2000 voxels). The brain medial regions smaller than 1000 voxels were combined (as the rectus or the frontal medial orbital brain area). The area left with less than 1000 voxels were the bilateral Amygdala (220 voxels), the bilateral Heschl (225 voxels) and the bilateral pallidum (280 (Left) and 293 (right) voxels). The new atlas was composed of 115 areas (Table S1), with an average volume of 1771 +/- 124 voxels. Only the 77 regions with at least 9 recording sites were retained for statistical analyses. On the 3440 initial recording dipoles, 3194 recording sites were folding into one of these areas. Coordinates of recording sites were computed as the MNI coordinates average of the two contacts used for each dipole.

### *High-gamma analysis*

To determine the time course of gamma band amplitude, continuous iEEG signals were first bandpass filtered in multiple successive 10 Hz wide frequency bands (e.g., 11 bands from [50–60 Hz] to [140–150 Hz]) using a zero phase shift noncausal finite impulse filter with 0.5 Hz roll-off. Next, for each bandpass filtered signal we computed the envelope using standard Hilbert transform. The obtained envelope has a sampling rate of 64 Hz (i.e., one time sample every 15.625 ms). Again, for each band this envelope signal (i.e., time varying amplitude) was divided by its mean across the entire recording session and multiplied by 100. This yields instantaneous envelope values expressed in percentage (%) of the mean. Finally, the envelope signals computed for each consecutive frequency bands (e.g., 11 bands of 10 Hz intervals between 50 and 150 Hz) were averaged together to provide a single time series (the high-gamma band envelope) across the entire session. By construction, the mean value of that time series across the recording session is equal to 100. Note that computing the Hilbert envelopes in 10 Hz sub-bands and normalizing them individually before averaging over the broadband interval allows us to counteract a bias toward the lower frequencies of the interval induced by the 1/f drop-off in amplitude. Finally, the obtained time series were smoothed on a sliding window of 250ms to get rid of potential artefacts.

### General Linear Model

Gamma envelops of each bipolar contact were epoched on each trial at two time locking (stimulus onset: -500 to 1500 ms, and first press in the rating task: -1500 to 1500 ms). Each gamma enveloped and each time point was regressed with a variable of interest to obtain a regression estimate per time point and per contact.

### Whole brain analysis

For each brain area included in the whole brain analysis, a T-value was computed across all contacts belonging to each brain area for each time point, independently of the patient (fixed effect) to allow the inclusion of area with a small number of contacts. Significance of each brain area was assessed by permutation tests on clusters with a corrected p=0.05.

### Time-frequency analysis

Time-frequency analyses were carried out using the FieldTrip toolbox for MATLAB. Spectral powers were estimated using a ''multitapering'' time-frequency transform (Slepian tapers, lower frequency range: 4–32Hz, 6 cycles and 3 tapers per window, higher frequency range : 32-200Hz, fixed time windows of 240ms, 4 to 31 tapers per window). This approach uses a constant number of cycles across frequencies up to 32Hz (hence a time window whose duration decreases when frequency increases), and a fixed time window with an increasing number of tapers above 32Hz to obtain more precise power estimates by adaptively increasing smoothing at high frequencies. Envelopes of theta, alpha, beta and gamma bands were extracted as the high-gamma frequency except that step of frequency was 1 for theta and alpha and 5 for beta and gamma. The signal was smoothed as the high-gamma only in the gamma band.

### Baseline analysis

To assess baseline significance, we tested the significance of the 14 time points (from -200 to 0ms before the stimulus onset), if this group of time points was significantly above zero with p<0.001, the considered electrode was labelled as significant.

### Decoding analysis

Three types of decoding analysis have been conducted: intra-task decoding, baseline decoding and inter-task decoding. For all of them, a GLM was run at each time point on the ratings provided by the patients with a regressor matrix comprising one column per electrode and per frequency band (number of dipoles x number of frequency bands). What we call the 'decoder' is actually the betas weight obtains from the multivariate regression.

In the intra-task decoding, a PRESS statistic was computed on each time point (Myers and Montgomery, 2002; Trujillo-Ortiz et al., 2006 - predicted residual error sum of squared). It is a way to compute the accuracy of a multivariate regression with cross-validation. To calculate the PRESS statistic, an observation i is selected. Then the regression model is fitted to the remaining n-1 observations and is use to predict the withheld observation y(i) (the predicted value is ye(i)). The prediction error (P(i)=y(i)-ye(i)) is repeated for each observation i and the PRESS statistic is defined as the sum of squares of the n P residuals. The PRESS significance was assessed by running the same analysis 200 times with permuted order of values in order to compute a threshold value of PRESS through permutation. Note that the lower the PRESS statistic is and the better the prediction. The sign of the PRESS is reversed in the text and on the figure for easiness of understanding.

Then, for all ratings we wanted to decode in the inter-task and baseline decoding, we computed the predicted values from the decoder of interest at each time point and computed the coefficient of determination between the predicted ratings and the effective ratings. Finally, we did this for each time point and obtain a temporal-generalization matrix per patient per decoding analysis.

In the baseline decoding analysis, we averaged the coefficients of determination from -200 ms to 1200 ms obtained with the decoder applied between -200 ms and 0 ms before the stimulus onset of the food rating among each patient. We then averaged at the group level the time series of decoding. We ran the same analysis 200 times with permuted order of values in order to compute a threshold value of decoding through permutation.

In the inter-task decoding analysis, for each patient, we identified the time point were the decoding score was maximal and extracted the decoding scores of the time series corresponding to the best training time. This method allowed us to get rid of time variability between tasks (i.e. if value is encoded at 500 ms in task A while in task B the value is coded at 530 ms, this analysis would allow detecting the use of the same code). Extracted decoding scores time series were again averaged across patients. To assess significance of decoding at the group level, we computed T-values of clusters with a score higher than the score obtained with random predictions for the 200 permutation tests. Then we tested whether the clusters obtained with our data had a T-value higher than the 95% of the T-values obtained with permutations.

## *References*

Abitbol, R., Lebreton, M., Hollard, G., Richmond, B.J., Bouret, S., Pessiglione, M., 2015. Neural Mechanisms Underlying Contextual Dependency of Subjective Values: Converging Evidence from Monkeys and Humans. J. Neurosci. 35, 2308–2320. doi:10.1523/JNEUROSCI.1878-14.2015

Barron, H.C., Dolan, R.J., Behrens, T.E.J., 2013. Online evaluation of novel choices by simultaneous representation of multiple memories. Nat. Neurosci. 16, 1492–1498. doi:10.1038/nn.3515

Bartra, O., McGuire, J.T., Kable, J.W., 2013. The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. NeuroImage 76, 412–427. doi:10.1016/j.neuroimage.2013.02.063

Bouret, S., Richmond, B.J., 2010. Ventromedial and Orbital Prefrontal Neurons Differentially Encode Internally and Externally Driven Motivational Values in Monkeys. J. Neurosci. 30, 8591–8601. doi:10.1523/JNEUROSCI.0049-10.2010

Chib, V.S., Rangel, A., Shimojo, S., O'Doherty, J.P., 2009. Evidence for a Common Representation of Decision Values for Dissimilar Goods in Human Ventromedial Prefrontal Cortex. J. Neurosci. 29, 12315–12320. doi:10.1523/JNEUROSCI.2575-09.2009

Clithero, J.A., Rangel, A., 2014. Informatic parcellation of the network involved in the computation of subjective value. Soc. Cogn. Affect. Neurosci. 9, 1289–1302. doi:10.1093/scan/nst106

De Martino, B., Fleming, S.M., Garrett, N., Dolan, R.J., 2013. Confidence in value-based choice. Nat. Neurosci. 16, 105–110. doi:10.1038/nn.3279

Grabenhorst, F., Rolls, E.T., 2008. Selective attention to affective value alters how the brain processes taste stimuli. Eur. J. Neurosci. 27, 723–729. doi:10.1111/j.1460-9568.2008.06033.x

Grueschow, M., Polania, R., Hare, T.A., Ruff, C.C., 2015. Automatic versus Choice-Dependent Value Representations in the Human Brain. Neuron. doi:10.1016/j.neuron.2014.12.054

Hare, T.A., Camerer, C.F., Knoepfle, D.T., O'Doherty, J.P., Rangel, A., 2010. Value Computations in Ventral Medial Prefrontal Cortex during Charitable Decision Making Incorporate Input from Regions Involved in Social Cognition. J. Neurosci. 30, 583–590. doi:10.1523/JNEUROSCI.4089-09.2010

Howard, J.D., Gottfried, J.A., Tobler, P.N., Kahnt, T., 2015. Identity-specific coding of future rewards in the human orbitofrontal cortex. Proc. Natl. Acad. Sci. 112, 5195–5200. doi:10.1073/pnas.1503550112

Jerbi, K., Ossandón, T., Hamamé, C.M., Senova, S., Dalal, S.S., Jung, J., Minotti, L., Bertrand, O., Berthoz, A., Kahane, P., Lachaux, J.-P., 2009. Task-related gamma-band dynamics from an intracerebral perspective: Review and implications for surface EEG and MEG. Hum. Brain Mapp. 30, 1758–1771. doi:10.1002/hbm.20750

King, J.-R., Dehaene, S., 2014. Characterizing the dynamics of mental representations: the temporal generalization method. Trends Cogn. Sci. 18, 203–210. doi:10.1016/j.tics.2014.01.002

Knutson, B., Fong, G.W., Adams, C.M., Varner, J.L., Hommer, D., 2001. Dissociation of reward anticipation and outcome with event-related fMRI. Neuroreport 12, 3683–3687.

Knutson, B., Katovich, K., Suri, G., 2014. Inferring affect from fMRI data. Trends Cogn. Sci. 18, 422–428. doi:10.1016/j.tics.2014.04.006

Lachaux, J.-P., Axmacher, N., Mormann, F., Halgren, E., Crone, N.E., 2012. High-frequency neural activity and human cognition: Past, present and possible future of intracranial EEG research. Prog. Neurobiol. 98, 279–301. doi:10.1016/j.pneurobio.2012.06.008

Lachaux, J.P., Rudrauf, D., Kahane, P., 2003. Intracranial EEG and human brain mapping. J. Physiol.-Paris 97, 613–628. doi:10.1016/j.jphysparis.2004.01.018

Lebreton, M., Abitbol, R., Daunizeau, J., Pessiglione, M., 2015. Automatic integration of confidence in the brain valuation signal. Nat. Neurosci. 18, 1159–1167. doi:10.1038/nn.4064

Lebreton, M., Bertoux, M., Boutet, C., Lehericy, S., Dubois, B., Fossati, P., Pessiglione, M., 2013. A Critical Role for the Hippocampus in the Valuation of Imagined Outcomes. PLoS Biol. 11, e1001684. doi:10.1371/journal.pbio.1001684

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.

doi:10.1016/j.neuron.2009.09.040

Levy, D.J., Glimcher, P.W., 2012. The root of all value: a neural common currency for choice. Curr. Opin. Neurobiol. doi:10.1016/j.conb.2012.06.001

Levy, I., Lazzaro, S.C., Rutledge, R.B., Glimcher, P.W., 2011. Choice from Non-Choice: Predicting Consumer Preferences from Blood Oxygenation Level-Dependent Signals Obtained during Passive Viewing. J. Neurosci. 31, 118–125. doi:10.1523/JNEUROSCI.3214-10.2011

Logothetis, N.K., Pauls, J., Augath, M., Trinath, T., Oeltermann, A., 2001. Neurophysiological investigation of the basis of the fMRI signal. Nature 412, 150–157. doi:10.1038/35084005

Myers, R.H., Montgomery, D.C., 2002. Response surface methodology: Process and product improvement with designed experiments. John Wiley & Sons, New York.

O'Doherty, J.P., Dayan, P., Friston, K., Critchley, H., Dolan, R.J., 2003. Temporal difference models and reward-related learning in the human brain. Neuron 38, 329–337.

Padoa-Schioppa, C., 2007. Orbitofrontal Cortex and the Computation of Economic Value. Ann. N. Y. Acad. Sci. 1121, 232–253. doi:10.1196/annals.1401.011

Padoa-Schioppa, C., Assad, J.A., 2006. Neurons in the orbitofrontal cortex encode economic value. Nature 441, 223–226. doi:10.1038/nature04676

Perez-Garcia, F., Lehongre, K., Bardinet, E., Jannin, P., Navarro, V., Hasboun, D., Vidal, S.F., 2015. Automatic segmentation of depth electrodes implanted in epileptic patients: a modular tool adaptable to multicentric protocols. Epilepsia 56, 227.

Pessiglione, M., Delgado, M.R., 2015. The good, the bad and the brain: neural correlates of appetitive and aversive values underlying decision making. Curr. Opin. Behav. Sci. 5, 78–84. doi:10.1016/j.cobeha.2015.08.006

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R.J., Frith, C.D., 2006. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature 442, 1042–1045. doi:10.1038/nature05051

Peters, J., Büchel, C., 2010. Neural representations of subjective reward value. Behav. Brain Res. 213, 135–141. doi:10.1016/j.bbr.2010.04.031

Peters, J., Büchel, C., 2010. Episodic Future Thinking Reduces Reward Delay Discounting through an Enhancement of Prefrontal-Mediotemporal Interactions. Neuron 66, 138–148. doi:10.1016/j.neuron.2010.03.026

Plassmann, H., O'Doherty, J., Rangel, A., 2007. Orbitofrontal Cortex Encodes Willingness to Pay in Everyday Economic Transactions. J. Neurosci. 27, 9984–9988. doi:10.1523/JNEUROSCI.2131-07.2007

Plassmann, H., O'Doherty, J.P., Rangel, A., 2010. Appetitive and Aversive Goal Values Are Encoded in the Medial Orbitofrontal Cortex at the Time of Decision Making. J. Neurosci. 30, 10799–10808. doi:10.1523/JNEUROSCI.0788-10.2010

Rangel, A., Camerer, C., Montague, P.R., 2008. A framework for studying the neurobiology of value-based decision making. Nat. Rev. Neurosci. 9, 545–556. doi:10.1038/nrn2357

Ray, S., Crone, N.E., Niebur, E., Franaszczuk, P.J., Hsiao, S.S., 2008. Neural Correlates of High-Gamma Oscillations (60-200 Hz) in Macaque Local Field Potentials and Their Potential Implications in Electrocorticography. J. Neurosci. 28, 11526–11536. doi:10.1523/JNEUROSCI.2848-08.2008

Ray, S., Maunsell, J.H.R., 2011. Different Origins of Gamma Rhythm and High-Gamma Activity in Macaque Visual Cortex. PLoS Biol. 9, e1000610. doi:10.1371/journal.pbio.1000610

San Galli, A., Varazzani, C., Abitbol, R., Pessiglione, M., Bouret, S., 2016. Primate ventromedial prefrontal cortex neurons continuously encode the willingness to engage in reward directed behavior. (No. biorxiv;062315v1).

Schmidt, L., Lebreton, M., Cléry-Melin, M.-L., Daunizeau, J., Pessiglione, M., 2012. Neural Mechanisms Underlying Motivation of Mental Versus Physical Effort. PLoS Biol. 10, e1001266. doi:10.1371/journal.pbio.1001266

Sescousse, G., Caldú, X., Segura, B., Dreher, J.-C., 2013. Processing of primary and secondary rewards: A quantitative meta-analysis and review of human functional neuroimaging studies. Neurosci. Biobehav. Rev. 37, 681–696. doi:10.1016/j.neubiorev.2013.02.002

Sescousse, G., Redoute, J., Dreher, J.-C., 2010. The Architecture of Reward Value Coding in the Human Orbitofrontal Cortex. J. Neurosci. 30, 13095–13104. doi:10.1523/JNEUROSCI.3501-10.2010

Strait, C.E., Blanchard, T.C., Hayden, B.Y., 2014. Reward Value

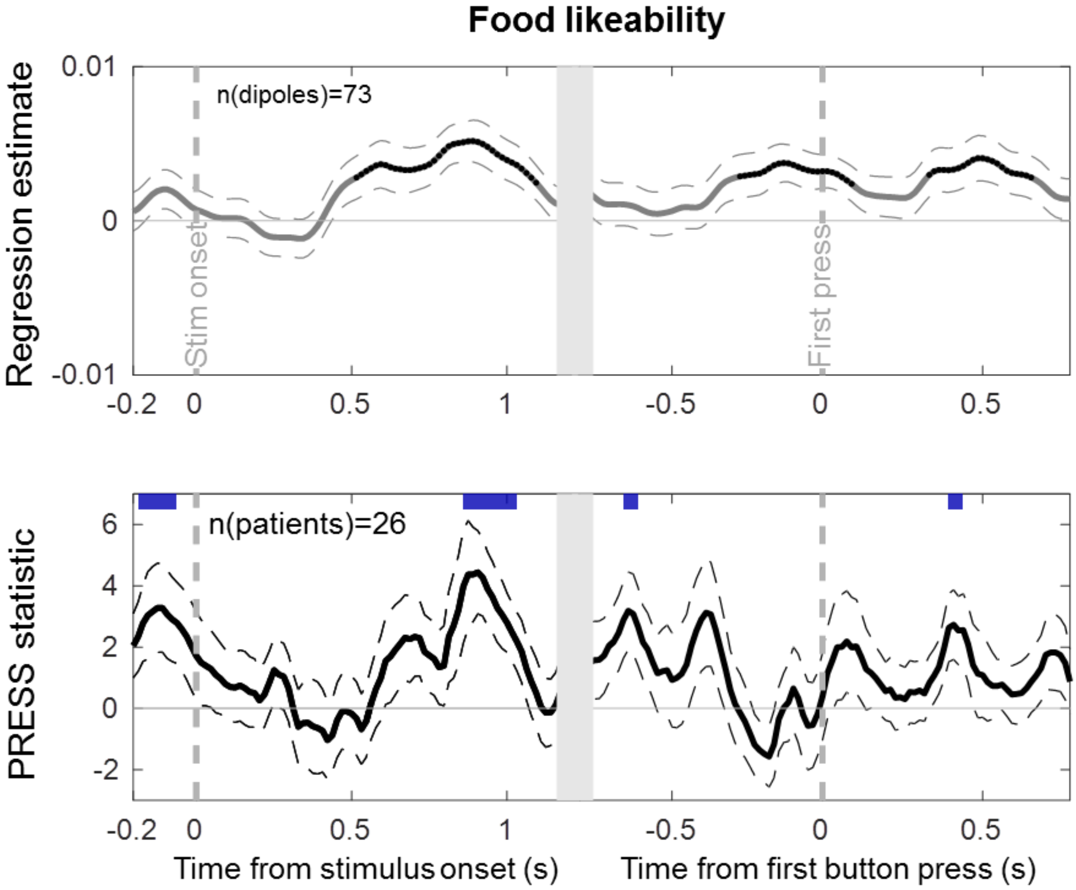
Comparison via Mutual Inhibition in Ventromedial Prefrontal Cortex. Neuron 82, 1357–1366. doi:10.1016/j.neuron.2014.04.032

Talairach, J., Tournoux, P., 1988. Co-planar stereotaxic atlas of the human brain. 3-Dimensional proportional system: an approach to cerebral imaging.

Trujillo-Ortiz, A., Hernandez-Walls, R., Barba-Rojo, K., Castro-Perez, A., 2006. press:Prediction error sum of squares. A MATLAB file.

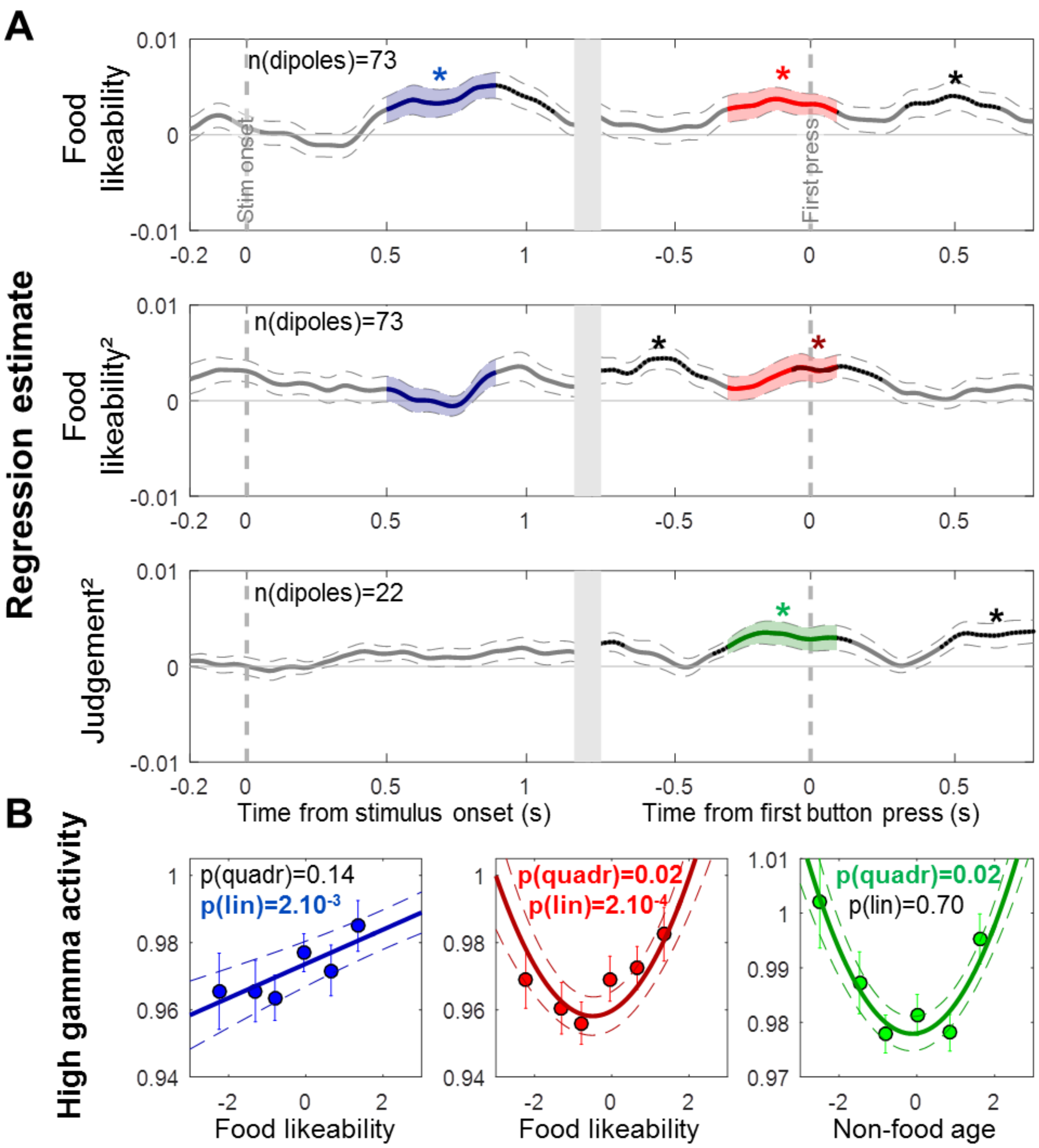
Wallis, J.D., 2011. Cross-species studies of orbitofrontal cortex and value-based decision-making. Nat. Neurosci. 15, 13–19. doi:10.1038/nn.2956

Zhang, Y., Pan, X., Wang, R., Sakagami, M., 2016. Functional connectivity between prefrontal cortex and striatum estimated by phase locking value. Cogn. Neurodyn. 10, 245–254. doi:10.1007/s11571-016-9376-2



**Supplementary Figure 1 – Univariate & Multivariate decoding of food likeability in the vmPFC.**

**Top**. Mean time course of regression estimates for high-gamma vmPFC activity against food likeability, locked either on stimulus onset (left) or on first button press. (right). **Bottom**. The multivariate decoder is a matrix where each regressor corresponds to the power in a specific frequency band recorded on a specific dipole. Mean PRESS statistic (decoding performance) is plotted as a function of time around stimulus onset first button press. Dashed lines indicate inter-subject S.E.M. Blue bars indicate significant clusters (corrected for multiple comparisons using permutation tests).



**Supplementary Figure 2 – Linear and quadratic coding of likeability and age.**

**A**: Time courses of the regression estimates of food likeability (top), squared food likeability (middle) and squared judgments (age and likeability of food and non-food) (bottom) locked on the stimulus onset (left) and on the first button press (right). In blue and red are depicted the first and second time windows on which signal has been extracted and averaged to investigate the shape of the quadratic regression presented in B. In green is depicted the same time window as in red but in patients who completed the long version of the task. Stars indicate significance of the regression. **B**: High-gamma activity splitted according to bins of food likeability (left and middle) and to bins of age rating (right) in the first time window (blue) and in the second time window (red and green). p indicates p-values for the quadratic term (quadr) and for the linear term (lin). Font bold indicates significance (p<0.05).

# Tables

### Table 1: Demographical data

| Task version | Recording center | ID | Sex | Age (years) | Epilepsy onset (age) | Suspected Epileptic focus | IQ | Hand laterality | Number of recording sites | Number of electrodes |
|---|---|---|---|---|---|---|---|---|---|---|
| Complete | Grenoble | G1 | M | 37 | 29 | L OFC | NA | NA | 165 | 13 |
| Short | | G2 | F | 31 | 24 | L temporal | 100 | R | 156 | 12 |
| Short | | G3 | M | 47 | 38 | R temporal | 96 | R | 328 | 18 |
| Short | | G4 | F | 34 | 26 | R OFC | 76 | R | 208 | 16 |
| Short & Complete | | G5 | F | 47 | 19 | Bi-fronto-temporal | 120 | R | 203 | 16 |
| Short (only rating) | | G6 | M | 57 | 37 | L ant temporal | 104 | NA | 187 | 12 |
| Complete | | G7 | F | 39 | 28 | L temporal | 76 | R | 229 | 17 |
| Short | Lyon | L1 | F | 49 | 39 | R temporal | NA | R | 149 | 14 |
| Short | | L2 | M | 21 | 14 | L fronto-temporal | NA | R | 168 | 15 |
| Short | | L3 | F | 43 | 2 | L temporale | NA | R | 129 | 13 |
| Short | | L4 | F | 40 | 15 | NA | NA | L | 109 | 11 |
| Short | | L5 | F | 38 | 28 | L temporal | NA | R | 98 | 9 |
| Short | | L6 | M | 22 | NA | NA | NA | L | 133 | 11 |
| Short | | L7 | F | 48 | 19 | R temporal | NA | R | 149 | 13 |
| Short | | L8 | M | 41 | NA | L temporal | NA | L | 50 | 5 |
| Short | | L9 | M | 30 | NA | R fronto-temporal | NA | R | 149 | 13 |
| Short | | L10 | F | 21 | 4 | L frontal / insula | NA | R | 115 | 11 |
| Short | | L11 | M | 26 | 21 | NA | NA | R | 132 | 12 |
| Short | | L12 | F | 43 | 40 | L temporal | NA | L | 126 | 13 |
| Short | | L13 | F | 26 | 4 | L ant temporal | NA | R | 127 | 13 |
| Short | | L14 | F | 31 | NA | NA | NA | R | 121 | 12 |
| Short | | L15 | F | 45 | 21 | L temporal /insula | NA | R | 110 | 10 |
| Short | | L16 | M | 44 | NA | NA | NA | R | 114 | 11 |
| Short | | L17 | M | 48 | NA | NA | NA | R | 153 | 14 |
| Short | | L18 | M | 51 | 11 | NA | NA | R | 128 | 12 |
| Complete | Paris | P1 | M | 22 | 7 | L temporal | 79 | R | 78 | 9 |
| Complete | | P2 | F | 55 | 36 | R temporal | 96 | R | 94 | 11 |
| Complete | | P3 | F | 60 | 32 | R ant temporal | 89 | L | 57 | 8 |
| Complete | | P4 | F | 30 | 16 | Dorso-medial frontal | 117 | R | 66 | 10 |
| Complete | | P5 | M | 29 | 20 | R posterior temporal | NA | | 60 | 8 |
| Complete | | P6 | F | 22 | 8 | Dorso-medial frontal | NA | R | 58 | 10 |
| Complete | | P7 | F | 34 | 9 | L Insula / Amygdala | 110 | R | 68 | 8 |
| Complete | | P8 | F | 39 | 23 | L ant temporal | 106 | R | 62 | 7 |
| Complete (without faces) | | P9 | M | 34 | 12 | R ant temporal | NA | R | 78 | 10 |
| Complete | | P10 | M | 37 | 24 | R ant temporal | NA | R | 50 | 7 |
| Complete | | P11 | F | 45 | 25 | R temporal | NA | R | 78 | 9 |

IQ: Intellectual Quotient; NA: Missing Data, OFC: Orbito-Frontal Cortex; R: Right, L: Left, Bi: Bilateral; M: Male; F: Female; ant: anterior

**Table 2: Regression estimates for food ratings in all frequencies and ROI.**

| | | vmPFC | lOFC | | Hippocampus | |
|---|---|---|---|---|---|---|
| θ<br>Theta<br>4-7 Hz | Cluster begin (s) | 0.74 | 0.08 | 0.81 | | |
| | Cluster end (s) | 1.23 | 0.6 | 1.5 | | |
| | T-value | -92.45 | 130.77 | -157.44 | | |
| | p-value | 0.01 | 0 | 0 | | |
| α<br>Alpha<br>8-15 Hz | Cluster begin (s) | | -0.2 | 0.7 | 0,92 | |
| | Cluster end (s) | | 0.41 | 1.16 | 1.4 | |
| | T-value | | 131.8 | -87.31 | -75.8 | |
| | p-value | | 0.01 | 0.03 | 0.03 | |
| β<br>Beta<br>15-35 Hz | Cluster begin (s) | | -0.2 | 0.86 | | |
| | Cluster end (s) | | 0.17 | 1.11 | | |
| | T-value | | 64.2 | 46.3 | | |
| | p-value | | 0 | 0.01 | | |
| γ<br>Gamma<br>35-50 Hz | Cluster begin (s) | 0.36 | 0.49 | | 0.35 | 0.95 |
| | Cluster end (s) | 1.23 | 1.09 | | 0.58 | 1.2 |
| | T-value | 175.26 | 106.53 | | 41.06 | 40.08 |
| | p-value | 0 | 0 | | 0.03 | 0.04 |
| Hγ<br>High-Gamma<br>50-150 Hz | Cluster begin (s) | 0.52 | 0.36 | | 0.31 | |
| | Cluster end (s) | 1.09 | 1.08 | | 0.99 | |
| | T-value | 108.69 | 234.24 | | 126.78 | |
| | p-value | 0 | 0 | | 0 | |

Each column in ROI represents one cluster. Displayed clusters are the one with duration longer than 30 ms and surviving cluster correction for multiple comparisons. P-values are assessed through permutation tests (500). p=0 corresponds to $p<1.10^{-4}$. s: second

# *Supplementary tables*

## Table S1: AAL re-parcellation

| AAL (restructured) Label | Number of sites | AAL (restructured) Label | Number of sites | AAL (restructured) Label | Number of sites |
|---|---|---|---|---|---|
| Temporal Mid L ant/ant | 148 | Frontal Mid L ant/inf | 31 | Angular L | 8 |
| Temporal Mid L ant/post | 119 | Frontal Inf Oper L | 31 | Thalamus R | 8 |
| Temporal Sup R ant | 117 | **Frontal Sup Orb R** | **30** | Occipital Mid R sup | 8 |
| Temporal Sup L ant | 111 | Amygdala L | 30 | Temporal Mid R post/post | 7 |
| **Hippocampus L** | **104** | **Rectus** | **29** | Postcentral R inf/ant | 7 |
| **Hippocampus R** | **102** | Rolandic Oper R | 26 | Putamen L | 7 |
| Temporal Inf R ant | 92 | Cingulum Ant R | 26 | Parietal Inf R | 7 |
| Temporal Mid R ant/ant | 84 | SupraMarginal R | 25 | Angular R | 6 |
| Insula L | 82 | Frontal Inf Tri R post | 25 | Cingulum Mid R post | 6 |
| Insula R | 82 | Frontal Mid L post/inf | 24 | Postcentral R inf/post | 6 |
| Temporal Mid R ant/post | 80 | ParaHippocampal L | 23 | Parietal Inf L post | 5 |
| Frontal Inf Orb R | 75 | Frontal Mid R post/sup | 23 | Supp Motor Area R post | 4 |
| Frontal Inf Orb L | 65 | Heschl R | 23 | Cingulum Post | 4 |
| Temporal Inf L ant | 57 | Postcentral L inf | 23 | Cuneus R | 4 |
| **Frontal Mid Orb R** | **53** | Temporal Pole Sup L | 22 | Precuneus R inf | 3 |
| Temporal Sup R post | 51 | Fusiform L post | 22 | Frontal Sup Medial L sup | 3 |
| Temporal Pole Mid R | 51 | Frontal Sup R post/ant | 21 | Lingual R ant | 3 |
| ParaHippocampal R | 49 | Frontal Sup L ant | 21 | Parietal Sup L ant | 3 |
| Frontal Sup R ant | 49 | Temporal Mid L post/post | 21 | Precentral R sup | 3 |
| Temporal Mid L post/ant | 48 | Temporal Mid R post/ant | 18 | Precentral L sup | 3 |
| Frontal Mid R ant inf | 47 | Rolandic Oper L | 18 | Olfactory | 2 |
| **Frontal Mid Orb L** | **45** | Frontal Sup Medial R inf | 17 | Caudate | 2 |
| **Frontal Med Orb** | **44** | SupraMarginal L | 16 | Cerebelum L | 2 |
| Precentral R inf | 43 | Cingulum Mid R ant | 15 | Calcarine R | 2 |
| Temporal Sup L post | 42 | Cingulum Mid L | 15 | Pallidum R | 2 |
| Temporal Inf L post | 40 | Supp Motor Area R ant | 14 | Cuneus L | 2 |
| Frontal Mid R ant/sup | 40 | Temporal Inf R post | 14 | Occipital Sup L | 2 |
| Frontal Inf Tri R ant | 39 | Frontal Sup R post/post | 13 | Caudate R | 2 |
| Frontal Mid L ant/sup | 37 | Frontal Sup Medial L inf | 12 | Fusiform R post | 2 |
| Frontal Mid R post/inf | 37 | Frontal Sup L post | 12 | Paracentral Lobule | 2 |
| Fusiform R ant | 36 | Putamen R | 12 | Parietal Sup R ant | 2 |
| Frontal Inf Oper R | 35 | Precentral L inf | 12 | Precuneus L inf | 2 |
| Temporal Pole Mid L | 34 | Frontal Sup Medial R sup | 10 | Occipital Inf L | 1 |
| Frontal Inf Tri L post | 34 | Lingual L ant | 10 | Supp Motor Area L ant | 1 |
| Cingulum Ant L | 34 | Heschl L | 10 | Pallidum L | 1 |
| Fusiform L ant | 33 | Amygdala R | 9 | Occipital Sup R | 1 |
| Temporal Pole Sup R | 33 | Frontal Mid L post/sup | 9 | Occipital Mid L ant | 1 |
| Frontal Inf Tri L ant | 33 | Postcentral R sup | 9 | Precuneus L sup | 1 |
| **Frontal Sup Orb L** | **32** |  |  | NA | 235 |

Inf: inferior; Sup: Superior; Ant: Anterior; Post: Posterior; L: Left; R: Right; NA: Not Attributed; Mid: Middle; Med: Median; Tri: Triangularis; Supp: Supplemantary; Orb: Orbital; Oper: Opercular.

Grey: Areas not included in the whole brain analysis (less than 9 sites). Red: vmPFC; blue: lOFC; green: Hippocampus.

**Table S2: Whole brain T-values and Percentage of electrode coding food value in high-gamma**

| AAL Label | Maximum T-value | Percentage of significant electrodes |
|---|---|---|
| **Hippocampus L** | **291.54** | **36** |
| **Hippocampus R** | **245.32** | **40** |
| Frontal Sup R post/post | 141.95 | 15 |
| **Frontal Mid Orb L** | **133.16** | **44** |
| Frontal Sup R ant | 129.83 | 34 |
| Temporal Inf L ant | 114.25 | 28 |
| Frontal Mid R ant/inf | 110.39 | 24 |
| ParaHippocampal L | 107.39 | 47 |
| **Frontal Sup Orb R** | **107.29** | **37** |
| Temporal Mid L post/ant | 99.46 | 17 |
| Fusiform L ant | 97.33 | 39 |
| Fusiform R ant | 79.98 | 44 |
| **Frontal Mid Orb R** | **77.49** | **34** |
| **Frontal Med Orb** | **77.18** | **36** |
| Frontal Inf Tri L ant | 76.97 | 36 |
| Frontal Mid L post/sup | 75.10 | 44 |
| ParaHippocampal R | 69.13 | 27 |
| Fusiform L post | 62.63 | 41 |
| **Frontal Sup Orb L** | **60.05** | **41** |
| Temporal Pole Sup R | 50.58 | 39 |
| Precentral L inf | 49.63 | 50 |
| SupraMarginal R | 49.44 | 44 |
| Cingulum Ant L | 43.29 | 24 |
| Frontal Inf Oper R | 39.18 | 26 |
| Frontal Sup Medial L inf | 37.76 | 25 |
| **Rectus** | **33.35** | **28** |
| Temporal Pole Mid L | 31.29 | 26 |
| Temporal Mid L ant/ant | 31.14 | 24 |
| Temporal Mid R ant/post | 30.71 | 28 |
| Cingulum Ant R | 30.38 | 23 |
| Temporal Inf R ant | 30.20 | 27 |
| Frontal Inf Tri R post | 29.74 | 24 |
| Temporal Inf L post | 27.30 | 25 |
| Temporal Mid R post/ant | 25.12 | 33 |
| Cingulum Mid L | 21.33 | 33 |
| Frontal Inf Oper L | 13.95 | 26 |
| Insula L | 13.26 | 23 |
| Rolandic Oper R | 10.52 | 27 |

Same abbreviations as Table S1. Areas are ordered according to the cluster T-value of regression estimate with food value. Percentages in orange indicate areas with a percentage higher than 33%.

Red: vmPFC; blue: lOFC; green: Hippocampus. Yellow: common areas between the two analysis.

## C. Discussion

Using a large dataset of 36 epileptic patients including around 4000 recording sites we were able to conduct a pseudo whole brain analysis on the Local Field Potentials recorded while patients were rating the likeability of food items. A subset of 14 patients also participated to a longer version of the task in which they had to rate the likeability and to evaluate the age of non-food items (faces and paintings). Each judgment was followed by a confidence rating. First, we showed that likeability ratings were good predictors of choices and reaction time. Moreover, confidence and reaction time during the rating tasks were found to be *quadratically* linked to the judgments, suggesting that the ratings provided by the patients could be used as an appropriate proxy to study subjective value. Second, we analyzed the regression estimates of the high gamma signal against subjective value on all the brain areas covered by our dataset and we found that all the areas of the Brain Valuation System available in this dataset were significantly revealed by this analysis (ventral and dorsal vmPFC and bilateral hippocampus). Moreover, we found that the four regions forming the OFC (bilateral central and lateral OFC) were also significantly revealed by this analysis. Other regions such as temporal areas and fusiform gyrus were found but we focused our investigation on the BVS and the OFC. In these regions, we found that the time-frequency analysis of the evoked response in the OFC and in the vmPFC were really similar compared to the one observed in the hippocampus. However, the precise investigation of the dynamics of various frequencies showed that the vmPFC was coding subjective value through a decrease of high gamma and gamma activity and a negative coding in the low frequency theta (4-7 Hz) while the OFC was coding value through an increase of high gamma, gamma and beta activity and through a negative correlation in theta and alpha bands. In the hippocampus, the gamma and high gamma activities were found involved in subjective value coding. This precise dynamics sheds light on specific differences between the vmPFC and the OFC and shows that there is parametric coding of subjective value in the human OFC which can be investigated through LFP.

According to Fries (2015), one way to interpret the rhythms involved in cognition is to consider that high frequencies such as gamma and high gamma reflect bottom-up processes and stimuli processing while intermediate frequencies such as alpha and beta would control the influences of high gamma band through top-down mechanisms. Moreover, it has been proposed that the theta rhythm could have a role in attentional sampling. Given those assumptions, phase coherence between those frequencies could be investigated in our dataset to better understand how subjective value is represented in our regions of interest.

Finally, we investigated the properties of the Brain Valuation System in the vmPFC through both univariate and multivariate regression analysis which showed that all properties were replicable. **Anticipation** effect was found in the vmPFC high gamma activity but seemed to rely on a broader range of frequencies. We also showed through the dynamics of the decoding analysis that this property was beyond a simple additive effect of the baseline activity on the evoked response. The second property – **generality** – was not found significant in the high gamma frequency (only a trend, driven by faces) but was shown to be significant in the multivariate analysis. This indicates that other frequencies might also participate to this property. The third one – **automaticity** – was found to be significant in both univariate in high gamma and in multivariate analysis and since it was assessed on non-food items it also supports indirectly the generality property. Finally, we found a **quadratic coding** of judgment in the high gamma analysis locked on the beginning of the response, which could let us suppose that confidence could be encoded in relation to the response more than to the stimulus onset. However, the multivariate analysis showed that the quadratic coding could be found around (and before) the stimulus onset, which might mean that baseline could influence both value and confidence reports. Thus, other phenomenon linked to confidence coding in other frequencies remain to be further investigated. In conclusion, our results confirmed the properties of the Brain Valuation System in the vmPFC which have been originally established in fMRI and bring new insights into the dynamics of subjective value representation.

This huge dataset was a priceless opportunity to investigate the Brain Valuation System. The presented paper here is still incomplete and further analysis will be needed to precise the results. Particularly, we do not test the validity of the properties in the hippocampus and in the OFC. For example, it would be interesting to see whether we can find a parametric coding of non-food items in the OFC and to see whether we can decode their values with the decoder trained on food. A failure of decoding could indicate that the code is different (and confirmed what has been found in multivariate pattern analysis in fMRI) despite the parametric coding if it exists (with different frequencies or recording sites involved for example). Moreover, we could try to improve the food value decoder by including the OFC and the hippocampus activity and investigate to what extent they improve it and through which manner. In other words, the multivariate decoder provides a ‘code’ (a combination of regression estimate for each frequency band and each electrode per patient) which has not been investigated in our study despite the fact that it could be highly informative. For example, I conducted a preliminary analysis on this *code* to investigate whether a pattern could be observed at various time points across subjects. I computed the average weights for each frequency band across patients on the time windows which were providing a significant decoding at the group level (baseline: -200 to 0 ms and peak: 860 to 1000 ms after the stimulus onset, in blue on Figure 37). I also included a 'control' window in which the decoding was not significant (100-500ms, in black) and I found that the pattern was really similar between the baseline and the peak code (as indicated by the significant decoding) with positive weights for beta, gamma and high gamma frequency while a totally different pattern appeared for the control window. This could indicate that it is high frequencies in a broader range than the high-gamma which are involved

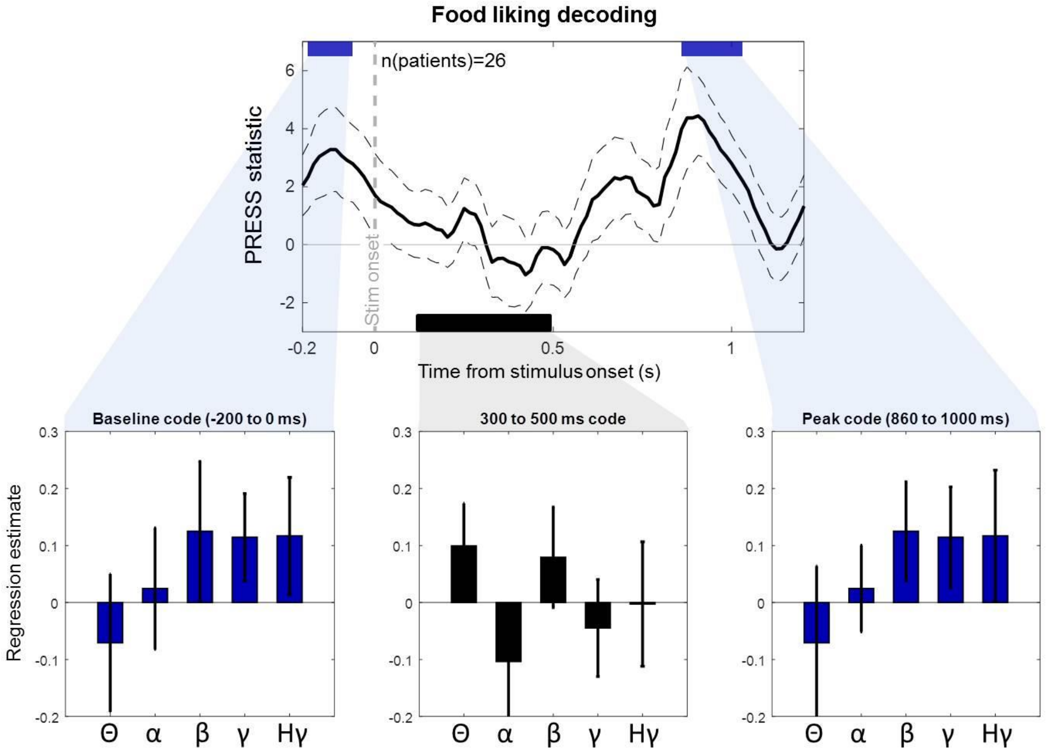


Figure 37 – Preliminary result on the 'code' underlying food value representation in the vmPFC

Top: Food value decoder: Averaged PRESS statistic (decoding performance) across patients in function of time after stimulus onset. Dashed lines indicate SEM. Rectangles indicate the time window of regression estimate investigation. Blue is indicating significance of decoding. Bottom: Averaged regression estimate of time points and dipoles accross patients for five frequencies and three time windows. Error bars indicate SEM.

in the baseline effect. Nevertheless, none of this regression estimate is significant, thus we cannot draw any conclusion from this example and further analysis could be performed to investigate this observation. I simply presented here as an overview of what could be done to invest the code underlying subjective value processing in the Brain Valuation System.

A multiple number of complementary analyses could also be run on this immense dataset. For example, pseudo whole-brain analysis on other frequencies or even on the iERP could bring precisions on the involved networks in valuation processes. Phase-locking analysis could also help to understand coupling between frequencies such as theta and high gamma in the vmPFC. Moreover, connectivity analysis would help understand interactions between the hippocampus and the vmPFC or between the OFC and the vmPFC. Our dataset had the advantage to comprise simultaneous recordings in several brain regions. As an illustration, for each patient, we created couples of electrodes between our three areas of interest. This pairing resulted in 86 pairs between the OFC and the vmPFC (19 patients), 54 pairs between OFC and hippocampus (12 patients) and 29 pairs between the vmPFC and the hippocampus (12 patients). Thus, we have a sufficient statistical power to perform such analysis and I believe this dataset will bring enlightening answers on the underlying mechanisms of subjective valuation processes.

To conclude, thanks to a rare dataset, we replicated what has been established in fMRI on the Brain Valuation System and brought new results on its dynamics. Having an access to the LFP of the Human Brain Valuation System opened a window on interesting results which nevertheless still need to be further investigated.

# 3. fMRI investigation of the Brain Valuation System during decision-making

## A. Introduction

Would you prefer eating French fries or a chocolate bar? How our brain decides between such options is a general issue in decision neuroscience. A central question is how the subjective values of such options are compared. We saw that the Brain Valuation System is critical for subjective value coding, with the vmPFC as a key component. During binary choices (between two options A and B), the BVS is assumed to signal a decision value, i.e. the difference between the two option values ($V_A – V_B$). However, what labels A and B actually represent has remained totally unclear. This issue is crucial since the brain regions involved in the valuation and decision process need to identify which option is favored by the decision value signal in order to make the appropriate selection.

In this last fMRI experiment, we supposed that prior information about which options will be faced by subjects before making a decision could be used to set up a default policy (reference option). We therefore suggested the hypothesis that decision value should be computed as the difference between default and alternative options ($V_{def} – V_{alt}$). Prior preferences were defined at the category level (savory versus sweet food for example) while choices were made at the item level (French fries versus chocolate bar). At the behavioral level, we should see that prior preferences induce a bias that leads participants to choose the default option more often (and faster) than its value would predict. At the neural level, we will investigate how the decision value is encoded in the BVS evoked response.

## B. Article

# How prior preferences determine decision-making frames and biases in the human brain

Alizée Lopez-Persem(1, 2), Philippe Domenech(2, 3, 4), Mathias Pessiglione(1, 2)

(1) Motivation, Brain and Behavior (MBB) lab, Centre de NeuroImagerie de Recherche (Cenir), Institut du Cerveau et de la Moelle épinière (ICM), Groupe Hospitalier Pitié-Salpêtrière, 75013 Paris, France
(2) Inserm U1127, CNRS U 7225, Université Pierre et Marie Curie (UPMC – Paris 6)
(3)CHU Henri Mondor, DHU Pe-PSY, Service de Neurochirurgie Fonctionnelle, Créteil, France
(4) Behavior, Emotion & Basal Ganglia (BEBG) lab, Institut du Cerveau et de la Moelle épinière (ICM), Groupe Hospitalier Pitié-Salpêtrière, 75013 Paris, France

*In revision in e-Life*

## *Abstract*

Understanding how option values are compared when making a choice is a key objective for decision neuroscience. In natural situations, agents may have a priori on their preferences that create default policies and shape the neural comparison process. We asked participants to make choices between items belonging to different categories (e.g., jazz vs. rock music). Behavioral data confirmed that items taken from the preferred category were chosen more often and more rapidly, which qualified them as default options. FMRI data showed that baseline activity in classical brain valuation regions, such as the ventromedial Prefrontal Cortex (vmPFC), reflected the strength of prior preferences. In addition, evoked activity in the same regions scaled with the default option value, irrespective of the eventual choice. We therefore suggest that in the brain valuation system, choices are framed as comparisons between default and alternative options, which might save some resource but induce a decision bias.

## *Keywords*

Decision-making; neuroeconomics; default bias; fMRI; brain valuation system; ventromedial prefrontal cortex

## *Introduction*

Standard decision theory assumes that when faced with a choice, individuals first assign subjective values to each option, and then compare these values in order to select the best option (Samuelson, 1938; Von Neumann and Morgenstern, 1947). Understanding the neural mechanisms governing this valuation/selection process has become a central aim in the field of decision neuroscience. A large set of fMRI evidence points to the ventro-medial Prefrontal Cortex (vmPFC) as a key player in the valuation process (Bartra et al., 2013; Clithero and Rangel, 2014). Neural activity in the vmPFC reflects subjective values, either measured with likeability ratings or inferred from binary choices (Kable and Glimcher, 2009; Rangel and Hare, 2010). In accordance with the idea of a common neural currency (Levy and Glimcher, 2012), the vmPFC was found to encode the subjective value of many kinds of goods, such as food, money, trinkets, faces, paintings, charities, etc. (Chib et al., 2009; Hare et al., 2010; Lebreton et al., 2009; Plassmann et al., 2007). Such value coding was observed not only during choice but also in the absence of choice, during passive viewing of items presented in the attentional focus or when performing a distractive task on these items (Lebreton et al., 2009; Levy et al., 2011; Abitbol et al., 2015).

During binary choices, it has been repeatedly shown that vmPFC activity correlates with the relative value of the two options under consideration ($V_A–V_B$). However, the framing of such decision value signal, i.e. what A and B actually represent, remains an unresolved issue. This question is of importance because the brain regions downstream in the decision process cannot operate the appropriate selection without knowing which option is favored by the relative value signal. In particular, the post-decisional frame that has often been reported (Boorman et al., 2013; Hare et al., 2011; Hunt et al., 2012) provides a decision value signal between chosen and unchosen options ($V_{ch}$-$V_{unch}$) that cannot be used for making the selection. A spatial frame, based on the location of options (e.g., $V_{left}$-$V_{right}$), has been suggested but not supported by much experimental evidence regarding the vmPFC valuation signal (Palminteri et al., 2009; Wunderlich et al., 2009; Skvortsova et al., 2014). A more promising suggestion is the attentional frame (Krajbich et al., 2010), in which the decision value signal encoded in the vmPFC depends on which option is attended to ($V_{att}$-$V_{unatt}$). Such framing provided a good account for vmPFC activity in a choice task where fixation patterns were imposed to subjects, and correctly predicted several features of spontaneous choice behavior by imposing a discount weight on the unattended option value (Krajbich et al., 2010; Lim et al., 2011). Notably, the attentional frame predicts that more fixated options should be more frequently chosen, which might explain why vmPFC activity has been found to correlate with $V_{ch}$-$V_{unch}$ in other studies. However, the attentional model assumes that visual exploration is random, which might be true in artificial laboratory tasks where subjects have no information about upcoming options, but not in natural situations where prior knowledge might play a role.

Here, we hypothesize that the framing of the decision value encoded in the vmPFC is imposed by prior preferences. In other words, vmPFC activity should scale positively with the value of the option that is preferred a priori, which we call the default option, and negatively with that of the alternative ($V_{def}$-$V_{alt}$). This hypothesis is compatible with the observation that vmPFC activity correlates with $V_{ch}$-$V_{unch}$, since choices usually follows on prior preferences. Yet, the interpretation is fundamentally different, as $V_{def}$-$V_{alt}$ is a pre-decisional value signal susceptible to drive option selection. Our hypothesis builds on the literature about optimal foraging, which argues that stay/switch choice is the natural case of decision-making (Stephens and Krebs, 1986). In this framework, staying on a same patch is the default option against which all alternatives must be compared. Several studies investigated such stay/switch decisions and implicated the dorsal anterior cingulate cortex in promoting a shift away from the default option (Hayden et al., 2011; Kolling et al., 2012; Kvitsiani et al., 2013), while others induced default policies by manipulating prior probabilities of being correct (Boorman et al., 2013; Fleming et al., 2010; Mulder et al., 2012; Scheibe et al., 2010). Although experimental manipulations vary across these studies, the default option is always defined as the option that would be selected in the absence of further information processing about its value relative to the alternatives. This definition provides objective criteria to identify the default option in a choice set: it should be selected faster and more frequently than the alternatives.

Therefore, our hypothesis implies that prior preferences should (1) induce a bias in favor of the default option, and (2) determine the frame of the value comparison process. The purpose of the present study was to examine how these two constraints would shape the brain valuation signal. To do so, we exploited the hierarchical structure of preferences: individuals have global preferences between categories of goods that can be locally reversed when comparing particular items. For instance, someone may prefer pop to jazz music in general, but nonetheless pick Keith Jarrett if the only other option is Britney Spears. In a binary choice, the prior preference at the category level thus designates a default option (i.e., the item belonging to the preferred category), but the option values still need to be compared at the item level in order to reach a final decision.

We conducted an fMRI experiment where participants made binary choices between items belonging to different categories. Preferences between categories were inferred from likeability ratings that were collected for every item before the scanning session. In the following analyses, we first establish the presence of a bias toward the default option in both choice and response time, above and beyond the prior preference between categories. Using computational modeling, we provide evidence that the default bias is best accounted for by a shift in the starting point of a drift diffusion process, which is proportional to the prior preference between categories. Then, we show that the default bias is unrelated to gaze fixation pattern, precluding an attentional framing. Finally, we uncover two effects of prior preference in fMRI data: (1) vmPFC baseline activity reflects the a priori shift in favor of the default option, and (2) vmPFC evoked response represents the value of the default option, irrespective of the eventual choice.

## *Results*

### **Behavior**

Prior to the scanning session, participants (n=24) rated the likeability of items belonging to three different domains (food, music, magazines). Each domain included four categories of 36 items (see methods). At that time, participants were unaware of these categories. This is because the presentation of items for likeability ratings was blocked by domain but not by categories, which were randomly intermixed. During the scanning session, subjects performed series of choices between two items (Figure 1), knowing that one choice in each domain would be randomly selected at the end of the experiment and that they would stay in the lab for another 15 minutes to enjoy their reward (listening to the selected CD, eating the selected food and reading the selected magazine). Trials were blocked in series of nine choices between items belonging to the same two categories within a same domain. The two categories were announced at the beginning of the block, such that subjects could form a prior preference (although they were not explicitly asked to do so). We quantified this prior preference as the difference between mean likeability ratings (across all items within each of the two categories), which is hereafter denoted as $DV_{CAT}$. In most cases (84±3% on average), preferences inferred from mean ratings matched the preferences between categories that subjects directly expressed in post-scanning debriefing tasks. Moreover, the confidence in these choices between categories, which subjects provided on an analog rating scale during debriefing, was significantly correlated to $DV_{CAT}$ ($r=0.44\pm0.06$, $t(23)=7.88$, $p=5.10^{-8}$). These explicit measures taken after the scanning session therefore validate our quantification of implicit preferences between categories. In the following, we analyze choices and response times to assess the presence of a bias in favor of the default option (i.e., the item belonging to the preferred category).

We fitted a simple logistic regression model including a constant, the default option value, denoted $V_{IT}(def)$, and the alternative option value, denoted $V_{IT}(alt)$, to choices expressed in the 'default vs. alternative' frame. Regression coefficient estimates showed that the two option values were equally contributive to the choice ($V_{IT}(def)$: $\beta=0.060\pm0.005$, $t(23)=11.90$, $p=3.10^{-11}$; $V_{IT}(alt)$: $\beta=-0.060\pm0.004$, $t(23)=-14.21$, $p=7.10^{-13}$). Crucially, the constant was significantly positive ($\beta=0.68\pm0.13$, $t(23)=5.40$, $p=2.10^{-5}$), bringing evidence for a bias toward the default option. This constant was significantly reduced when including $DV_{CAT}$ in the regression model ($\beta=0.31\pm0.16$, $t(23)=1.94$, $p=0.06$), with the effect of $DV_{CAT}$ itself being significant ($\beta=0.021\pm0.006$, $t(23)=3.53$, $p=2.10^{-3}$), which established a direct link between prior preference and default bias. We also introduced past choices (coded 1 vs. -1 when default option was chosen vs. unchosen) in the regression model but they yielded no significant effect on choice rate. Consistently, the constant estimate was not different when restricting the logistic regression to the first choice in a block ($\beta_{first}=0.58\pm0.49$, $\beta_{all}=0.68\pm0.13$, difference: $t(23)=0.97$ $p=0.44$), confirming that the default bias was not resulting from the history of past choices. To illustrate this result (Figure 2A), we plotted the choice rate, P(def), as a function of the decision value, $DV_{IT}=V_{IT}(def)-V_{IT}(alt)$. This plot shows that even when the two options have the same value

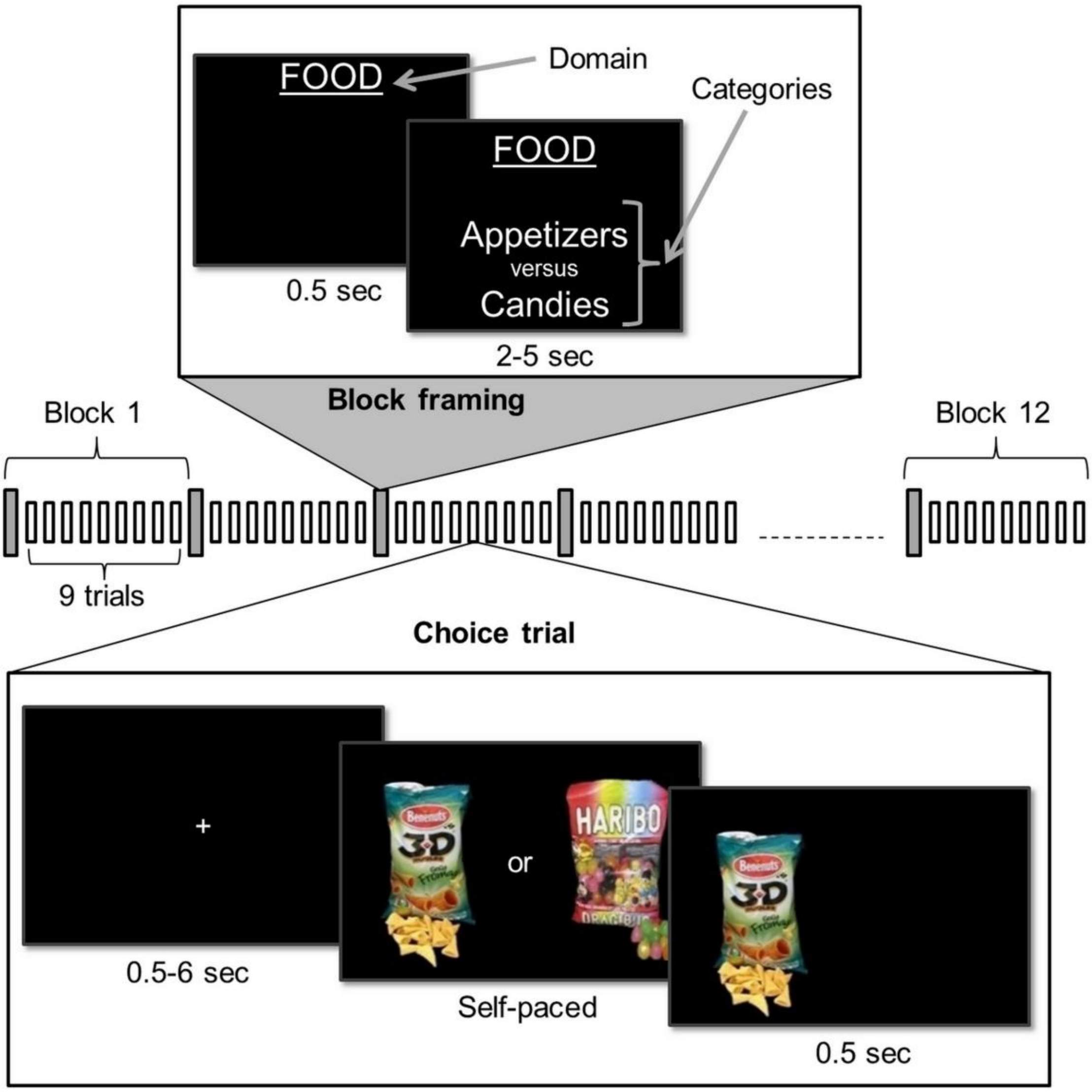


Figure 1. Choice task

Participants performed the choice task inside the MRI scanner. The task was composed of four 12-block sessions. During a block, subjects first saw an instruction screen indicating the reward domain (e.g., food) and the two categories from which choice options were drawn. Then, they had to make a series of 9 binary choices, each confronting the two categories with two novel items. Choice was self-paced and feedback on chosen option was provided to the subject.

($DV_{IT}$=0 on the x-axis), the choice rate is not at chance level (50% on the y-axis), which would denote indifference, but shifted toward the default option (by 15.7±1.7% on average). Thus, these results provide behavioral evidence for a 'choice bias' occurring on top of the decision value ($DV_{IT}$), i.e. above and beyond what could be predicted by the difference in likeability rating.

To account for choice response time (RT), we fitted a general linear model (GLM) including the main effects and the interaction of two factors: the unsigned decision value ($|DV_{IT}|$) and the choice type (default vs. alternative). As typically reported, we found a significant effect of unsigned decision value ($t(23)=-6.8$, $p=6.10^{-6}$), indicating that choices were longer when option values were closer. We also found a significant effect of choice type ($t(23)=-5.47$, $p=1.10^{-5}$), indicating that subjects were faster to pick the default option than the alternative. There was no significant interaction between the two factors ($t(23)=0.59$, $p=0.56$). Thus the 'RT bias' corresponds to the difference between intercepts for a null decision value (Figure 2B). This RT bias means that subjects were significantly faster when choosing the default (by 357±50 ms on average), irrespective of the decision value.

To assess whether the choice and RT biases could arise from the same underlying computation, we tested their correlation across blocks (Figure 2C). This is possible in our design because each block corresponds to a confrontation between two given categories, some being very close and others far apart in terms of mean likeability (i.e., they vary in terms of $DV_{CAT}$). We fitted a regression model to each block in order to extract choice and RT biases for each pair of categories. Correlation across blocks was estimated at the subject level and then tested against the null hypothesis at the group level. We found a significant correlation between the two biases ($r=0.24\pm0.06$, $t(23)=3.78$, $p=1.10^{-3}$), suggesting a common underlying mechanism, which we further characterized using computational modeling.

### Computational modeling

To account for both choice and RT distributions, we employed an analytical approximation to the Drift Diffusion Model (DDM). The DDM assumes that choices result from a sequential sampling process, through which a decision variable accumulates evidence until it reaches a boundary (Ratcliff, 1978; Ratcliff and McKoon, 2008). DDMs were originally developed to explain perceptual decisions but they have already been successfully applied to economic (value-based) decisions (Gold and Shadlen, 2007; Basten et al., 2010; Krajbich et al., 2010, 2012). In our DDM, the boundaries corresponded to the default and alternative choices, and the mean of the drift rate was the signed decision value, $DV_{IT}$ (inset in Figure 2D). A priori, the choice and RT biases could arise from a change in the drift rate or from a shift in the starting point, S. The latter possibility is more consistent with the negative correlation that was observed between choice bias and RT (Figure 2D) and tested at the group level (r=-0.68±0.08, t(23)=-8.06, p=4.10$^{-8}$). Indeed, in the DDM framework, a bias in the starting point has less impact on choices when the decision process lasts longer (Brunton et al., 2013).

To formally disentangle between these possibilities, we compared a DDM where the starting point is fixed at zero and the drift rate equal to $DV_{IT}$ (null model) to six alternative DDMs where either the starting point or the drift rate is allowed to change across subjects, and for some of them across blocks. The first three models ('start family') test the hypothesis of a shift in the starting point. The shift was captured with a single free parameter in model 1 ('1 free S'), with one free parameter per block in model 2 ('12 free S'), or as a free parameter scaled by $DV_{CAT}$ in model 3 ('S=a*$DV_{CAT}$'). Thus, the starting point was respectively considered constant across blocks (but possibly different from zero), freely adjusted to each block, or proportional to the prior preference. The last three models (drift family) test the hypothesis of a gain in the drift rate, which in any case was proportional to $DV_{IT}$. The shift was captured with a single additional parameter to $DV_{IT}$ in model 4, with one additional parameter per block in model 5, and with an additional term scaled by $DV_{CAT}$ in model 6.

We first conducted a family model comparison to examine the possibilities that the choice and RT biases were due to a shift in the starting point (models 1-3) or a change in the drift rate (models 4-6), relative to the null model (Figure 2E, top). The most plausible mechanism was the shift in the starting point (start family: exceedance probability, xp=0.997). Then, we compared the three models within this family (Figure 2E, bottom) and found evidence in favor of model 3 (xp=0.920), suggesting that the starting point varied across blocks proportionally to prior preferences. We verified this conclusion by testing the correlation across blocks between the posterior means of the 12-free-S model and the prior preference $DV_{CAT}$ (Figure 2F). The correlation was significant at the group level (r=0.35±0.07, t(23)=5.12, p=3.10$^{-5}$), strengthening the idea that prior preference was imposing a shift in the starting point that resulted in both choice and RT biases. Thus, the correlation observed between choice and RT biases was driven by variations in $DV_{CAT}$ across blocks, the two biases trending to zero when $DV_{CAT}$ was close to null.

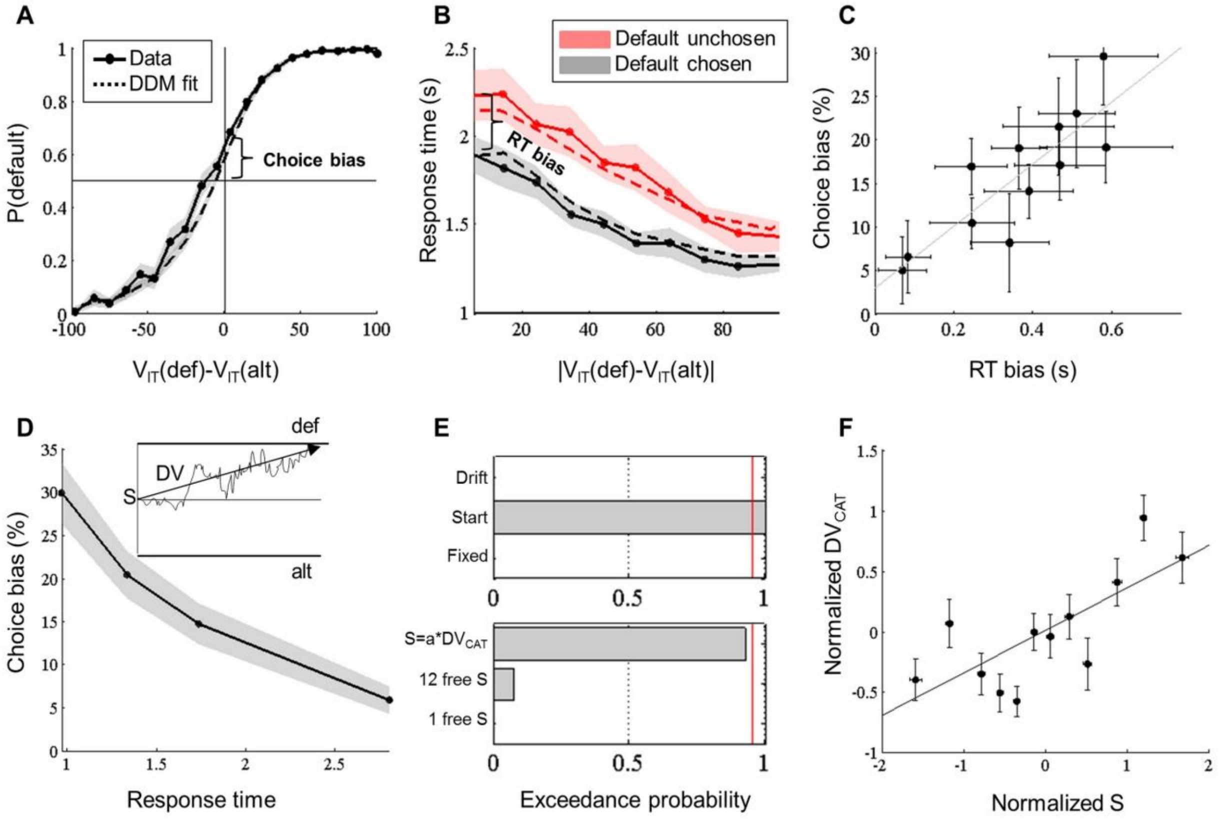


**Figure 2. Behavioral results (MRI experiment)**

(A) Probability of choosing the default option, P(def), plotted as a function of decision value, $DV_{IT}$, divided into 20 bins. Values correspond to likeability ratings given by the subject prior to scanning. Both probabilities observed in choice data (solid line) and simulated from the fitted Drift Diffusion Model (dashed line) are shown. Choice bias was defined as the difference between the observed probability for a null decision value and the expected equiprobability (50%). (B) Choice response time (RT) plotted as a function of the absolute decision value, $|DV_{IT}|$ divided into 10 bins, separately for trials in which the default option was chosen (black) and unchosen (red). Both RT observed in behavioral data (solid line) and simulated from the fitted Drift Diffusion Model (dashed line) are shown. RT bias was defined as the difference between the intercepts observed for the two types of choice. (C) Correlation of choice and RT biases across blocks. (D) Choice bias plotted as a function of response time, divided into 4 bins. Inset illustrates the Drift Diffusion model (adapted from (Voss et al., 2013), with S the starting point, DV the mean drift rate and def / alt the thresholds for choosing default / alternative options. Choice bias was larger for shorter RT, suggesting that it could arise from a bias in the starting point. (E) Family model comparison between different theoretical accounts of choice and RT biases. Top: the null model (‘Fixed’) is compared to models in which either the starting point (‘Start’) or the drift rate (‘Drift’) is allowed to favor the default option. Bottom: the model with a single free starting point (‘1 free s’) is compared to models in which the starting point is varied across blocks, either in proportion to the value difference between categories ‘s=a*$DV_{CAT}$’ or as a set of 12 independent parameters (‘12 free s’). Red line corresponds to 95% exceedance probability. (F) Correlation across blocks between $DV_{CAT}$ and starting point s (from fitting the ’12-free-s’ model). This suggests that the starting point is adjusted in each block to the average value difference between the two confronted categories. Shaded areas and error bars represent ± inter-subject SEM.

### Eye-tracking

The fact that decision bias was best explained by shifting the starting point made less likely an interpretation in terms of attentional dynamics. This is because in previous studies, gaze fixation pattern was found to affect the drift rate and not the starting point (Krajbich et al., 2010). We nevertheless investigated the possibility that the effect of prior preferences on choice and RT biases could be mediated by the pattern of gaze fixation. This possibility would imply that subjects pay more attention to the default option than to the alternative, which we examined using eye-tracking measurements.

Another group of participants (n=23) performed the same series of rating and choice tasks, while their gaze position on the screen was recorded using an eye-tracking device. All the behavioral results described in the previous section were replicated (Figure 3A and 3B), with a significant bias in both choice (15.5±1.7%, $t(22)=5.12$, $p=4.10^{-4}$) and RT (341±42 ms, $t(22)=-6.69$, $p=1.10^{-6}$), and a significant correlation between the two ($r=0.22\pm0.06$, $t(22)=3.87$, $p=8.10^{-4}$).

We also replicated a number of results predicted by the attentional Drift Diffusion Model (aDDM), in which a parameter $\theta$ down-weights the value of the unattended item in the decision value, hence in the drift rate (Krajbich et al., 2010; Krajbich and Rangel, 2011; Lim et al., 2011; Krajbich et al., 2012). As predicted by the aDDM, we notably observed that choice probability was higher for the item fixated last ($t(22)=-11.68$, $p=7.10^{-11}$), and for the most fixated item during the decision process ($t(22)=-4.71$, $p=1.10^{-4}$), irrespective of decision value. These results confirm that fixation pattern had the expected effects on choice. However, none of these effects could account for the bias toward the default option that was observed in our task.

To test the link between prior preference and gaze fixation pattern, we compared the duration of fixation for the default and alternative options, separately for trials in which the default was chosen and unchosen. We found that the default option was fixated longer when it was chosen (difference: 81±11 ms, $t(22)=7.52$, $p=2.10^{-7}$). Conversely, it was the alternative option that was fixated longer when the default option was not chosen (difference: 41±17 ms, $t(22)=2.41$, $p=2.10^{-2}$). Consistently, the ANOVA conducted on fixation time revealed a significant main effect of the 'chosen vs. unchosen' (factor $F(1,88)=4.03$, $p=0.048$), but no main effect of the 'default vs. alternative' factor ($F(1,88)=0.43$, $p=0.41$). The interaction was significant ($F(1,88)=17.45$, $p=1.10^{-4}$), reflecting the fact that the default option was more frequently

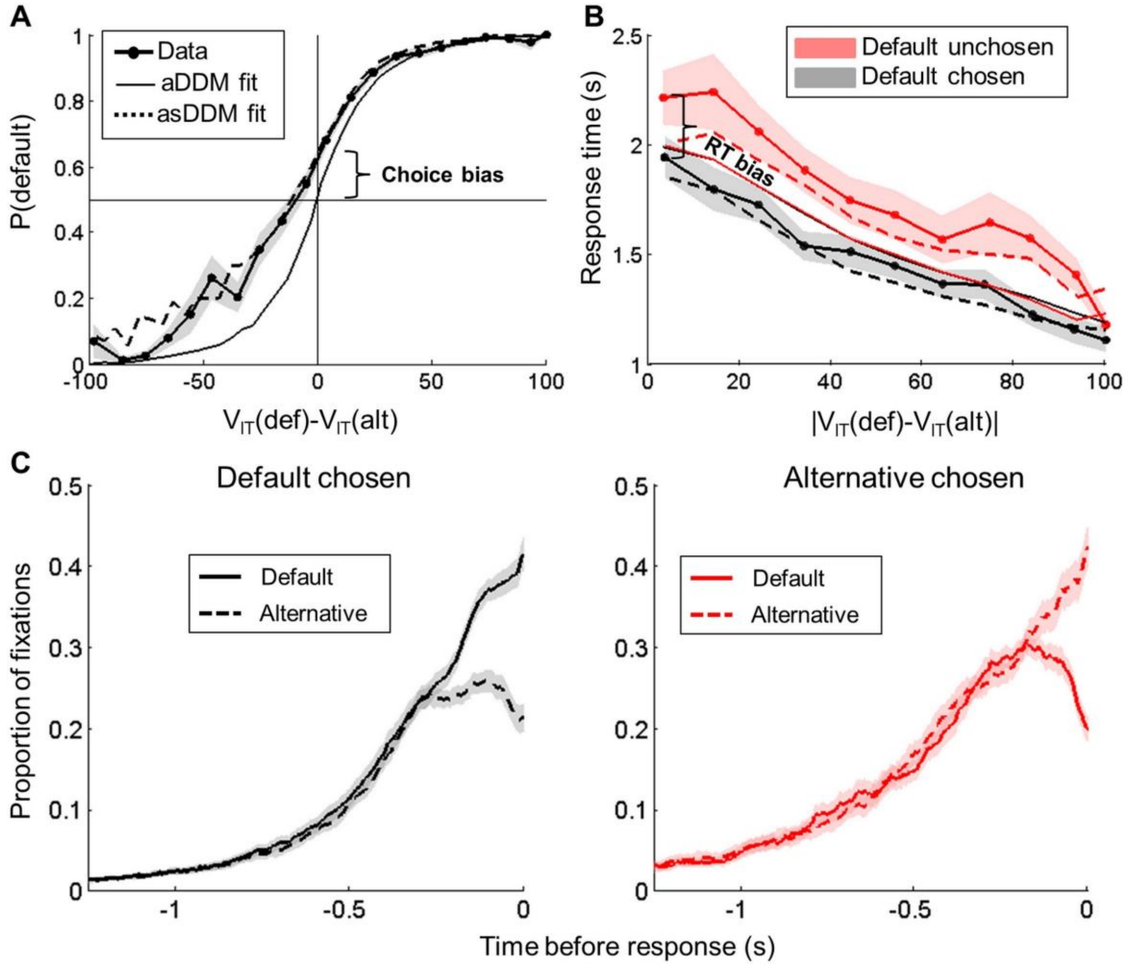


**Figure 3. Behavioral results (eye-tracking experiment)**

(A) Probability of choosing the default option plotted as a function of decision value $DV_{IT}$. The three curves correspond to probabilities actually observed in choice data (lines with circles) and simulated from either the fitted attentional Drift Diffusion Model (aDDM, solid lines) or the same model fitted with a starting point proportional to prior preference $DV_{CAT}$ (asDDM, dashed line). (B) Choice response time (RT) plotted as a function of the absolute decision value $|DV_{IT}|$, separately for trials in which the default option was chosen (left) and unchosen (right). The different curves correspond to RT observed in behavioral data (lines with circles) and simulated from either the fitted aDDM (solid line with circles) or asDDM (dashed line). Note that the aDDM alone cannot reproduce choice and RT biases. (C) Proportion of fixations (number of trials over all trials) to the default and alternative options at each time point when default is chosen (left) or unchosen (right). Curves are time-locked to choice (button press). They do not add up to one because at a given time point in a given trial, subjects may fixate none of the two options.Shaded areas are ± inter-subject SEM.

chosen, and with larger decision values. Thus, fixation duration was indeed predictive of choice, but was not influenced by prior preference. To control for the dynamics of the decision process, we computed the proportion of fixations for each option at each time point. The time courses locked to stimulus onset revealed a clear preference for looking at the left option during the first 250 ms, and then at the right option during the next 250 ms, but no significant preference for the default option at the beginning of the trial. The time courses locked on the response confirmed the fixation bias toward the chosen option, with a similar pattern whether default or alternative option was chosen (Figure 3C). This result was further confirmed by a model comparison showing that fixation duration for each option was better explained (xp=0.999) by a GLM including the unsigned decision value and the choice (chosen vs. unchosen option) than by GLMs including an additional regressor that indicated the prior preference (default vs. alternative option).

Finally, we compared four variants of the DDM to contrast how fixation pattern and prior preference influence the decision process. The first was the null model, with a starting point S fixed at zero and a weighting factor $\theta$ fixed at one. The second was the sDDM selected as the best model in the first experiment, with S proportional to $DV_{CAT}$ and $\theta$ still fixed at one. The third was the standard aDDM, with S fixed at zero and a freely fitted $\theta$. The fourth was termed asDDM and included both S proportional to $DV_{CAT}$ and fitted $\theta$. The most plausible model was the asDDM (xp=0.95), with the weight on $DV_{CAT}$ significantly above zero (0.02±0.003, t(22)=7.85, $p=1.10^{-6}$), and a $\theta$ significantly below one ($\theta$=0.94±0.03, t(22)=2.14 p=0.04). The fits of choice and RT are illustrated for the aDDM and asDDM (Figure 3A and B). Although using the fixation pattern (with $\theta$) improved the fit, only the prior preference (with S) could explain the decision bias toward the default option. We reached similar conclusions when the advantage for the attended option was additively included in the drift rate, on top of decision value (as in Cavanagh et al., 2014). In fact, gaze fixation pattern failed to produce the default bias simply because the default option was no more looked at than the alternative option.

### fMRI

Our behavioral results establish that prior preferences exert a bias on choices, which in the DDM framework was best explained by a proportional shift in the starting point. We analyzed fMRI data first to examine whether the bias toward the default option could be observed in baseline neural activity, second to assess whether prior preference could frame the comparison between option values that might be implemented in the evoked neural response.

### Baseline activity

To examine whether the prior preference ($DV_{CAT}$) was encoded in baseline activity, we fitted a GLM (GLM0, see methods) convolved with a finite impulse response (FIR) function to fMRI data. GLM0 contained an indicator delta function for option display that was parametrically modulated by the option values, $V_{IT}$(def) and $V_{IT}$(alt). In the following we analyze regression coefficient estimates obtained for the indicator function on volumes acquired before and after option display. The contrast performed at the individual level weighted block-specific indicator functions by z-scored $DV_{CAT}$. Group-level statistical test ($p<0.005$, uncorrected) performed on this contrast for the volume acquired two seconds before option display revealed activity scaling with $DV_{CAT}$ in the vmPFC, ventral striatum and left hippocampus (Figure 4A, left panel), which are regions classically identified as parts of the brain valuation system (e.g., (Lebreton et al., 2009). The ventral striatum and left hippocampus were the only regions that survived cluster-level family-wise error (FWE) correction at the whole-brain level. The vmPFC cluster only survived small-volume correction ($p=6.10^{-3}$) within a ROI based on independent criterion – a sphere centered on the peak of the cluster that positively reflected value in a previous meta-analysis (Bartra et al., 2013). In order to illustrate the time course of this effect in the vmPFC, we simply averaged BOLD activity levels (coefficient estimates for indicator functions) in high and low $DV_{CAT}$ blocks separated with a median split (figure 4A, right panel). The difference between high and low $DV_{CAT}$ appeared to be maintained during the decision process and to progressively vanish at the end of the trial.

### Evoked response

Our key hypothesis was that the decision value signal is framed by the prior preference, as a comparison between default and alternative options. This opposition of default vs. alternative options partially overlaps with that of chosen vs. unchosen options (here, in 77.8 ± 1.04% of the choices), the latter contrast being classically used to localize value comparison signals. We started the analysis by replicating this classical approach with a standard GLM, before dissociating the two possible frames with a more exhaustive GLM.

The first GLM only contained $V_{IT}$(ch) and $V_{IT}$(unch) as parametric modulators of a categorical regressor (delta function) indicating option display (GLM1, figure 4 – figure supplement 1A, left), all convolved with a canonical hemodynamic response function (HRF). As expected, the classical contrast $V_{IT}$(ch)-$V_{IT}$(unch) revealed significant correlation in brain valuation regions such as the vmPFC, ventral striatum and posterior cingulate cortex. Yet, this pattern of activation was not very specific of the brain valuation system, as it also included other brain

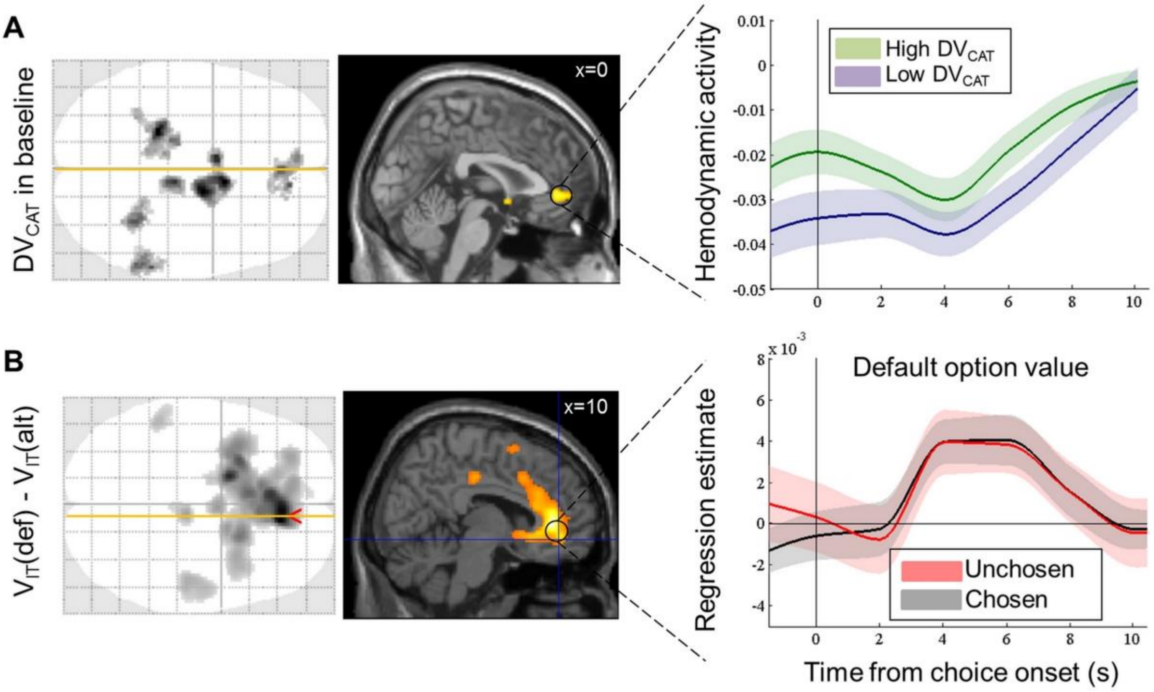


**Figure 4. Neural correlates of the bias and framing effects of prior preference**

(A) Bias in value coding and vmPFC baseline. Left: axial glass brain and sagittal slice of statistical maps relating to the prior preference ($DV_{CAT}$), one volume (2 sec) before the display of choice options. Areas shown in black (on glass brain) and yellow (on sagittal slice) showed significant positive group-level random effect (one-sample t-test, $p<0.005$ uncorrected for display purposes, minimum extent: 100 voxels). Right: time course of peri-stimulus fMRI activity in the vmPFC region, shown separately for high (green) and low (purple) $DV_{CAT}$. Black vertical line (time 0) indicates the onset of choice options. (B) Frame of value coding in vmPFC response. Left: axial brain glass and sagittal slice of statistical maps relating the canonical hemodynamic response to the decision value ($DV_{IT}$), at the onset of choice option (same threshold as in A) (one-sample t-test, $p<0.005$ uncorrected for display purposes, minimum extent: 100 voxels). Right: time courses of the regression estimate (beta) obtained in the vmPFC region for the default option value when it was chosen (black) or unchosen (red). Shaded areas are ± inter-subject SEM.

regions such as the intra-parietal lobules. The opposite contrast, $V_{IT}$(unch)-$V_{IT}$(ch), yielded significant correlation in the dorsal anterior cingulate cortex (dACC) and bilateral anterior insula, which are classically associated with choice difficulty or with the value of foregone options (Kolling et al., 2012; Shenhav et al., 2014), as well as in the inferior frontal gyrus and middle occipital gyrus. The equivalent GLM containing $V_{IT}$(def) and $V_{IT}$(alt) as parametric modulators yielded similar results (GLM1', figure 4 – figure supplement 1A), which is expected due to the shared variance between the chosen and default option values.

In order to disambiguate between representations of pre-choice values, $V_{IT}$(def) and $V_{IT}$(alt), and post-choice values, $V_{IT}$(ch) and $V_{IT}$(unch), we built a second GLM (GLM2, figure 4 – figure supplement 1B, left) that included two value regressors for each choice type (default chosen versus unchosen). These value regressors were parametric modulators of the categorical regressor indicating option display. The GLM also contained a regressor modeling the choice type itself, to dissociate value coding from option selection, and a boxcar function parametrically modulated by $DV_{CAT}$, to account for tonic effects of prior preference. This GLM allows computing both the $V_{IT}$(ch)-$V_{IT}$(unch) and the $V_{IT}$(def)-$V_{IT}$(alt) contrasts on the canonical evoked response.

Critically, we found significant activation (surviving cluster-level, whole-brain FWE correction) in the vmPFC and ventral striatum with the $V_{IT}$(def)-$V_{IT}$(alt) contrast (Figure 4B, left) but not with the $V_{IT}$(ch)-$V_{IT}$(unch) contrast (Figure 4 – figure supplement 1B, right). The neural response implementing the $V_{IT}$(def)-$V_{IT}$(alt) comparison was specific to the brain valuation system, since no other regions than vmPFC and ventral striatum passed the corrected threshold (Table S1 - Supplementary file 1). In particular, activity in the parietal or temporo-parietal cortex followed the $V_{IT}$(ch)-$V_{IT}$(unch) contrast as in the classical GLM. The opposite contrast, $V_{IT}$(unch)-$V_{IT}$(ch), again activated the dorsal anterior cingulate, anterior insula and middle occipital gyrus (Figure 4 – figure supplement 1B, right). The latter activation might relate to the fact that visual inspection of choice options was longer when choice was more difficult. No brain region was significantly associated with the $V_{IT}$(alt)-$V_{IT}$(def) contrast.

As we realized that the evoked response might have contaminated baseline vmPFC activity at the next trial, we estimated another version of GLM2, with the four parametric modulators - $V_{IT}$(def_chosen), $V_{IT}$(alt_unchosen), $V_{IT}$(def_unchosen), $V_{IT}$(alt_chosen) - replaced by the same values but for the options presented in the previous trial (the first trial of each block was discarded). None of these parametric modulators had a significant effect at the time points preceding option display ($t=-2$ and $t=0$ s from stimulus onset). Thus, the values of the options presented in the previous trial did not affect baseline vmPFC activity, beyond the variance that they shared with $DV_{CAT}$.

Thus, the neural decision value encoded in the vmPFC seemed to be expressed in the pre-choice frame (default minus alternative), rather than in the post-choice frame (chosen minus unchosen). However, inspecting the regression coefficient estimates obtained for $V_{IT}$(def) and $V_{IT}$(alt) separately suggested that the contrast was driven by $V_{IT}$(def), as no significant effect was observed with $V_{IT}$(alt) alone. To verify that $V_{IT}$(def) was similarly encoded in the vmPFC irrespective of the eventual choice, we fitted a FIR version of GLM2, and extracted regression estimates from the same independent vmPFC ROI as used previously (Bartra et al., 2013). We found that regression estimates for $V_{IT}$(def) were significant for both choice types (default chosen: $\beta= 6.10^{-3}\pm2.10^{-3}$, $t=3.23$, $p=3.10^{-3}$; default unchosen: $\beta=6.10^{-3}\pm3.10^{-3}$, $t=2.67$, $p=0.01$), four seconds after option display. For illustration purposes, we have plotted the time course of regression estimates extracted from the vmPFC cluster associated with $V_{IT}$(def)-$V_{IT}$(alt) in our main analysis (Figure 4B, left).

To further challenge our conclusion regarding the encoding of decision value in the vmPFC, we compared variants of GLM2 that included different parametric regressors locked to option display. The two-by-two model space tested the possibilities of (1) pre-choice (default minus alternative) versus post-choice (chosen minus unchosen) framing for value coding and (2) best option value ($V_{IT}$(def) or $V_{IT}$(Ch)) versus differential value coding ($DV_{IT}$). Bayesian model selection indicated that the pre-choice family was more plausible (xp=0.92), and that within this family, $V_{IT}$(def) alone better accounted for vmPFC activity than $DV_{IT}$ (xp=0.94). By comparison, the post-choice family best accounted for dACC activity (xp=0.99), with a clear advantage for the decision value $DV_{IT}$ over $V_{IT}$(Ch) alone (xp=0.9996).

At first, we were surprised that the vmPFC (or any other brain region) seemed to hold no representation of $V_{IT}$(alt) despite the fact that, behaviorally, $V_{IT}$(alt) impacted choices to a similar extent as $V_{IT}$(def). Then we reasoned that the vmPFC might have encoded both option values on top of the decision value, following on the idea that such valuation processes are automatically triggered when stimuli are presented (Lebreton et al., 2009; Levy et al., 2011). This would imply that vmPFC activity should correlate not only with the difference between option values but also with their sum. The net result would be a correlation with $V_{IT}$(def), since $V_{IT}$(alt) would be subtracted out, which was observed in our analysis. In order to test this hypothesis, we fitted a last variant of GLM2 that included both the difference and sum of option values in the pre-choice frame as parametric modulators of option display. The common variance, linked to the presence of $V_{IT}$(def) in both the sum and difference, was removed such that each regressor had unique variance. We examined the regression estimates extracted from the independent vmPFC ROI and found significant effects for both the difference ($\beta_{def\text{-}alt}=0.10\pm0.04$, $t(23)=2.74$, $p=0.01$) and the sum ($\beta_{def+alt}=0.14\pm0.06$, $t(23)=2.47$, $p=0.02$). In addition, model comparison showed that the GLM with orthogonalized sum and difference was a much better explanation of vmPFC response (xp=0.993) than the GLM containing only $V_{IT}$(def). This result suggests that the vmPFC evoked response signals the two option values (sum) on top of the decision value (difference). In macroscopic measures of brain activity such as fMRI, the positive and negative correlation with $V_{IT}$(alt) appear to cancel each other, but these representations might be dissociable using techniques with better spatial resolution that can access microscopic scales.

## *Discussion*

In this study, we examined how prior preference shapes the neural representation of decision value. We observed two major phenomena in vmPFC activity: 1) baseline activity was shifted in proportion to the strength of prior preference, as was the starting point in a drift diffusion model accounting for the decision bias in favor of the default option, 2) evoked activity signaled the value of the option belonging to the preferred category, suggesting that the choice was framed as a comparison between default and alternative.

Although they were not instructed to do so, subjects likely formed prior preferences at the beginning of blocks, when the two categories confronted in the upcoming series of choices were announced on the screen. Preference between two categories was inferred from likeability ratings averaged across items belonging to each category. In a vast majority of cases, this notion of preference matched the preference directly expressed by subjects in binary choices between categories made during post-scan debriefing. Moreover, the difference between mean likeability ratings ($DV_{CAT}$) was proportional to the confidence expressed in these choices between categories, in keeping with the notion that choice and confidence proceed from the same decision value (De Martino et al., 2013). These debriefing observations validate our notion of prior preference, which then served to designate the default option in the pair of items that was presented for choice.

Indeed items from the preferred category could be qualified as default options, because they were chosen faster and more frequently than their alternatives. These choice and RT biases are minimal requirements for a default option, i.e. an option that should be chosen in the absence of further information processing. Such criteria have been used in other paradigms where the goal was to maximize an objective reward, with for instances the default being defined as the pre-selected option in a perceptual decision task (Fleming et al., 2010), as the current patch in a foraging task (Hayden et al., 2011; Kolling et al., 2012; Kvitsiani et al., 2013), or as the long-term best option in a probabilistic instrumental learning task (Boorman et al., 2013). These studies reported that when the two option values were similar, subjects (both humans and monkeys) favored the default option. This phenomenon has been coined 'default bias', or 'status quo bias' in cases where the default option was defined as the pre-selected choice. Here, the same phenomenon was observed in the case of subjective preference. Importantly, the default bias was estimated once option values were matched, therefore it goes beyond what could be predicted from the difference in likeability ratings between categories. This bias could lead to preference reversals, meaning that subjects would pick the default option in spite of the alternative option having received a higher rating. Thus, the default bias denotes suboptimal decision-making, which could be compensated by the fact that following a default policy is on average less costly in terms of time or cognitive resource, than systematic unbiased comparison of option values. This phenomenon is therefore much different from the optimal use of prior information that has been observed in a variety of perceptual decision-making paradigms, subjects being biased only when tricked with invalid cues (Link and Heath, 1975; Bogacz et al., 2006; Scheibe et al., 2010; Mulder et al., 2012).

Within the drift diffusion framework, the default bias observed in choice and RT was best accounted for by a shift in the starting point. This is consistent with perceptual decision-making studies showing that prior information on probability or payoff is also incorporated in the starting point (Scheibe et al., 2010; Summerfield and Koechlin, 2010; Mulder et al., 2012). However, this is not compatible with the idea that the effect of prior preference on choice and RT biases is mediated by the pattern of gaze fixations. This idea implies that subjects pay more attention to the default option, which through the attentional DDM mechanism would favor the default choice, because the attended option has more weight than the alternative in the drift rate. In fact, subjects looked equally often at the default and alternative options in the eye-tracking experiment. Our results nonetheless confirmed that the pattern of gaze fixation does inform the prediction of choices, in a manner that is nicely captured by the attentional DDM. Thus, although the attentional DDM is perfectly compatible with our data, it could not by itself explain the default bias. The best account of choice and RT was in fact obtained with a model that cumulated the down-weighting of unattended options in the drift rate, as suggested by the attentional DDM, and the shift in the starting point that explains the default bias.

In our best model, the shift in starting point was proportional to the prior preference ($DV_{CAT}$). A striking parallel was found at the neural level, with the prior preference being reflected in the baseline activity of valuation regions including the vmPFC, ventral striatum and posterior cingulate cortex. This is in line with a previous study showing that baseline vmPFC activity is sensitive to contextual factors, both in humans and monkeys, and provide a bias in subsequent valuation processes (Abitbol et al., 2015). Other contextual manipulations were found to modulate vmPFC activity and subsequent valuation, for instance mood induction has been shown to affect reward-related vmPFC activity (Young and Nusslock, 2016). In contrast, cueing manipulation that affected perceptual decisions through a shift in starting point had no influence on vmPFC activity (Scheibe et al., 2010; Summerfield and Koechlin, 2010; Mulder et al., 2012). This dissociation suggests that the recruitment of vmPFC was not related to the general process of changing the accumulation starting point, but to biasing value-based decisions (as opposed to perceptual decisions). In fact the shift in baseline vmPFC activity was maintained throughout the decision process, and was hence added to the evoked activity, which followed a canonical hemodynamic response. As both baseline and evoked activity scaled with the value of the default, respectively at the category and item levels, they together contributed to favoring the default option over the alternative. Thus, the mechanics is analog to the DDM process, but the dynamics is somewhat different. In fact, the neural dynamics is not compatible with the vmPFC implementing the DDM, since we observed no ramping signal corresponding to an accumulation-to-bound process; neither is it compatible with the vmPFC output being sent to a distant accumulator, since the shift in starting point should not be integrated over time. Therefore, we do not suggest that the DDM used to capture behavioral patterns is literally implemented as such in the brain, just that the general logics and some key features appeared to match vmPFC activity during choices. We also note that other types of modeling would have been possible to capture behavioral effects, notably a Bayesian account where prior preference would affect the mean and perhaps the variance of a prior distribution on decision value.

The analysis of the evoked response showed that the vmPFC and ventral striatum encode the decision value in a frame that opposes the default to the alternative option. This pre-choice framing supports the idea of an anatomical separation between the valuation and selection processes, with the vmPFC being implicated in the former but not the latter. It could be a very general frame for value coding in the vmPFC, because most studies found a correlation between vmPFC activity and the value of chosen options (e.g., (Hare et al., 2011; Boorman et al., 2013), which are partially confounded with default options as we have shown here. We note that an opposite dissociation was found by Boorman and colleagues (2013), with the vmPFC encoding option values in post-choice frame, and not pre-choice frame. As decision-making dynamics was not explored in this study, it is unclear whether participants truly implemented a default strategy as defined here, which implies an anticipation of a default response, associated with shortening of response time. Accordingly, the representation of chosen option value was largely delayed in comparison to our study (peaking 10s after option display), possibly related to the necessity of storing expected values in a learning context.

Another partial confound is with choice easiness or confidence, which was also found to be integrated in vmPFC activity in addition to value (De Martino et al., 2013; Lebreton et al., 2015). The pre-choice framing could also be reconciled with the theory that the vmPFC encodes the value of the attended option, if we assume that when they have no prior information on the choice, subjects set up a default on the fly, which could be the option they just look at. By contrast, we found a post-choice framing of decision value (unchosen vs. chosen) in the dACC and anterior insula, which could be related either to choice difficulty or to the value of shifting away from the default policy, which might require cognitive control (Hare et al., 2011; Kolling et al., 2012; Shenhav et al., 2013).

A last potential issue is that the correlation with decision value ($DV_{IT}$) was driven by the default option, although the default and alternative options had the same weight on choices, and despite the two options being reflected in other regions such as dACC. Our interpretation is that both option values are encoded in the vmPFC on top of the decision value. As a result, the correlation with the alternative option value would be cancelled out, and the correlation with the default option value would be doubled, as suggested by the following equation:

$$Signal(vmPFC) \sim (V_{IT}(def) - V_{IT}(alt)) + V_{IT}(def) + V_{IT}(alt) = 2 * V_{IT}(def)$$

This interpretation is consistent with both the idea that the vmPFC automatically encodes the value of items that fall under the attentional focus (Lebreton et al., 2009; Levy et al., 2011) and the idea that the vmPFC computes a decision value whenever a choice process is engaged (Plassmann et al., 2007; Grueschow et al., 2015). It would also explain why many studies report a correlation with the chosen value alone and not the decision value, as the unchosen value would be cancelled out for the same reasons (Wunderlich et al., 2010; Kolling et al., 2012; Hunt et al., 2012). Model comparison supported this post-hoc interpretation, as including the two option values (sum) on top of the decision value (difference) provided the best account of vmPFC activity during choice. Other techniques than fMRI, with better spatial resolution, would be needed to investigate whether the different value representations rely on different populations on neurons.

In conclusion, our findings show that when decision-makers have a prior preference, the brain valuation system is configured so as to compare default and alternative options, with prior and novel information being encoded in baseline and evoked activity, respectively. Such framing could have been selected to solve natural decision problems, with the advantage of saving time and/or cognitive resource, and the disadvantage of biasing choice toward the default policy. How the valuation system adapts to artificial economic choices, in which two novel options present themselves simultaneously, still needs to be investigated. One may speculate that the brain would start by defining a default option, and then proceed to the comparison as usual. If this is correct, identifying the trial-wise and/or subject-wise default policy might be essential for understanding how the brain makes value-based decisions. However, we only have a proof of concept here, the generality of the 'default vs. alternative framing' remains to be established. Further research is also required to specify the contribution of the different brain regions that are involved in the valuation and selection processes, notably the dACC. The present results suggest that the vmPFC provides a decision value, which is also represented in the ventral striatum. How such decision value is used by the brain to make a selection remains to be explained.

## *Materials and Methods*

### Participants

The study was approved by the Pitié-Salpétrière Hospital ethics committee. All subjects were recruited via e-mail within an academic database and gave informed consent before participation in the study. They were right-handed, between 20 and 32 years old, with normal vision, no history of neurological or psychiatric disease, and no contra-indication to MRI (pregnancy, claustrophobia, metallic implants). They were not informed during recruitment that they would win food items, music CD and magazines to avoid biasing the sample. In total, 24 subjects (23.8±2.8 years old, 12 females) were included in the fMRI experiment and paid a fixed amount (80€) for their participation. In the eye-tracking experiment, 24 right-handed subjects (24±3.4 years old, 11 females) were recruited following the same procedure with the same inclusion criteria. No statistical method was used to predetermine sample size, but our sample size is similar to those generally employed in the field. One subject was excluded due to technical issue with the eye-tracking device.

### Tasks

All tasks were programmed on a PC in MATLAB language, using the Psychophysics Toolbox extensions (RRID:SCR_002881, Brainard, 1997; Pelli, 1997). Subjects performed the rating task outside the scanner and the choice task during fMRI scanning. Prior to each task, they were instructed and trained on short versions (24 trials) to get familiarized with the range of items and the mode of response.

During the rating task, subjects were asked to estimate the likeability of all 432 items that they could potentially obtain at the end of the experiment. These items were blocked by reward domain: food, music and magazines. Unbeknown to subjects, each reward domain was divided into 4 categories of 36 items. The 12 categories were: appetizers, biscuits, candies, chocolate (food domain); news, comics, cultural, generalist (magazine domain); French, jazz, rock, urban (music domain). The order of presentation was randomized within each reward domain, such that categories were intermingled. The series of trials consisted of displaying pictures of the items one by one on the computer screen. A pseudo-continuous rating scale (101 points) was presented below the picture, with three reference graduations (do not like at all, neutral, like a lot). Subjects could move a cursor along the scale by pressing a key with the right index finger to go left or another key with the right middle finger to go right. The cursor was initially positioned at the middle of the rating scale. Rating was self-paced and subjects had to press a button with the left hand to validate their response and proceed to the next trial. At the beginning of each block, the reward domain was announced on a black screen.

Likeability ratings were used for pairing options in the choice task. For each domain, mean ratings were used to rank categories according to subjective preference. The most preferred categories (ranked 1 and 2) were opposed to the least preferred ones (ranked 3 and 4), making a total of 4 oppositions (1-3, 1-4, 2-3, 2-4). To generate the series of choices for each opposition, items were sorted in the order of likeability rating. Half the choices varied the difference between ratings while keeping the average constant (item ranked mean+X was paired with item ranked mean-X); the other half varied the average while keeping the difference minimal (item ranked X was paired with item ranked X-1). Thus, the mean value and relative value of choice options were orthogonalized. A total of 36 choices were generated for each inter-categorical opposition, and presented in a randomized order. The 36 choices were divided into 4 blocks of 9 trials, presented in 4 different fMRI sessions. As there were 12 possible oppositions (4 per domain), this makes a total of 432 trials, meaning that each item being presented twice.

At the beginning of each block, the domain was announced on a black screen for 0.5 second, then the two opposed categories were displayed for 2 to 5 seconds, followed by a 0.5 second fixation cross. Each block was composed of a series of 9 choices. Choice trials started with the display of the two options side by side. The side of a given category as well as the best rated option was counter-balanced across trials. Subjects were asked to indicate their preference by pressing one of two buttons, with their left or right index finger, corresponding to the left and right options. The chosen picture was framed with a white square for 0.5 seconds, followed by a black screen with fixation cross lasting for 0.5 to 6.5 seconds.

Importantly, subjects were not asked to generate a prior preference at the beginning of blocks, when categories are revealed. They were only told that contextual information would be given before each series of choices, and that would not require any response from their part. They also knew that at the end of the experiment, one trial per domain would be randomly selected and that they would be given the options chosen in these trials.

Following the scanning session, subjects had to complete a debriefing task in which they were presented the opposed categories two by two. They were asked to first select the category that they preferred and then to rate their confidence in their choice using an analog scale. Finally, they spent an additional 20 minutes in the lab to eat the food item, listen to the music album and read the magazine that they just won.

### Behavior

All analyses were performed with Matlab Statistical Toolbox (Matlab R20013b, The MathWorks, Inc., USA). Two dependent variables were recorded: choice (which option was selected) and response time (between option onset and button press). The influence of likeability ratings on these variables were assessed with logistic or linear regression models, as explained in the results section. Regression estimates were computed at the individual level and tested for significance at the group level using one-sample two-tailed t-test. Correlations between variables of interest were also computed at the individual level, using Pearson's coefficient, and similarly tested at the group level.

### Eye-tracking

In the eye-tracking experiment, gaze position was recorded with a 60Hz sampling frequency using The Eye Tribe device (http://theeyetribe.com), during each block of the choice task. There was no constraint on the head, subjects were simply asked to avoid head movement. A screen providing feedback on the eyes position was inserted in the trial sequence every time gaze was lost. The number of excluded trials due to loss of gaze position varied between 0 and 6, depending on subjects.

Fixation duration was computed for each trial and option, as the time during which gaze position was inside a square window delineated the corresponding picture on the screen. Last fixation was defined as the picture being looked at when choice was made. The proportion of fixation was calculated as the number of trials in which gaze position was on the corresponding picture at a given time point. Note that these proportions for the two options do not add up to one because gaze position can be outside the two windows.

### Modeling

We used the EZ2 analytical approximation for the drift diffusion model (Wagenmakers et al., 2007; Grasman et al., 2009) to account for the probability of choosing the default option and the response time, on a trial-by-trial basis. As proposed by (Ratcliff, 1978; Wagenmakers et al., 2007; Grasman et al., 2009), we defined the probability of choosing the default option as:

$$P(def) = \frac{\varphi(-A, S-A)}{e^{\frac{2\mu A}{\sigma^2}} - 1}$$

As proposed by EZ2 (Grasman et al., 2009), we defined the corresponding RT as:

$$RT(def\ chosen) = Tnd + \frac{(A-S) * \big(\varphi(S,A) + \varphi(0, A-S) + 2A\varphi(A-S, 0)\big)}{-\mu\varphi(A-S, A)\varphi(-A, 0)}$$

With $\varphi(x,y) = e^{\frac{-2\mu y}{\sigma^2}} - e^{\frac{-2\mu x}{\sigma^2}}$, $A$ the amplitude between boundaries, $S$ the starting point, $\mu$ the mean of the drift rate, $\sigma$ the standard deviation of the drift rate and $Tnd$ the non-decision time. To compute response time in trials where the alternative option is chosen, we replace $(\mu, S)$ by $(-\mu, A-s)$.

The free parameters $A$, $Tnd$, $\sigma$, $\mu$ and $S$ were estimated with the behavioral data. Different versions of the model were compared to account for the changes in choice and RT patterns that were induced across blocks by the variations in prior preference. In all cases, $A$, $Tnd$ and $\sigma$, were estimated for each individual but constant across blocks. In the null model, $\mu$ was proportional to the decision value (difference in likeability rating between options, $DV_{IT}$, such that $\mu = \alpha DV_{IT}$) and $S$ was set to zero. The model space (see details in the results section) explored the possibilities that $\mu$ and $s$ could differ from their initial setting ($\mu = \alpha DV_{IT} + \beta$ / s=z), vary across blocks (12 free $\alpha$ for $\mu$ / 12 free z for s) , or be informed by the prior preference (difference in mean likeability rating between categories, $DV_{CAT}$, such that $\mu = \alpha DV_{IT} + \beta DV_{CAT}$ / s= $\beta DV_{CAT}$)). In the attentional versions of the model, $\mu$ was also informed by gaze fixations, as follows:

$$\mu = \frac{(V_{IT}(def) - \theta V_{IT}(alt) * D_{def} - (V_{IT}(alt) - \theta V_{IT}(def)) * D_{alt}}{D_{def} + D_{alt}}$$

With $D_{def}$ and $D_{alt}$ the total durations of fixation for the default and the alternative options during the considered trial, and $\theta$ the weight discounting the value of the unfixated item relative to the fixated one (Krajbich et al., 2010).

All versions of the drift diffusion model were fitted separately for each individual to choices and RTs using Matlab VBA-toolbox (available at http://mbb-team.github.io/VBA-toolbox/), which implements Variational Bayesian analysis under the Laplace approximation

(Daunizeau et al., 2014). This iterative algorithm provides a free-energy approximation for the model evidence, which represents a natural trade-off between model accuracy (goodness of fit) and complexity (degrees of freedom) (Friston et al., 2007; Penny, 2012). Additionally the algorithm provides an estimate of the posterior density over the model free parameters, starting with Gaussian priors. Individual log model evidences were then taken to group-level random-effect Bayesian model selection (BMS) procedure (Penny et al., 2010). BMS provide an exceedance probability (xp) that measures how likely it is that a given model (or family of models) is more frequently implemented, relative to all the others considered in the model space, in the population from which participants were drawn (Rigoux et al., 2014; Stephan et al., 2009).

### fMRI

Functional echo-planar images (EPIs) were acquired with a T2*-weighted contrast on a 3-T magnetic resonance scanner (Siemens Trio). Interleaved 2-mm slices separated by a 1.5 mm gap and oriented along a 30° tilted plane were acquired to cover the whole brain with a repetition time of 2.01 seconds. The first five scans were discarded to allow for equilibration effects. All analyses were performed using statistical parametric mapping (SPM8, RRID:SCR_007037) environment (Wellcome Trust Center for NeuroImaging, London, UK). Structural T1-weighted images were coregistered with the mean EPI, segmented, and normalized to the standard Montreal Neurological Institute (MNI) T1 template. Normalized T1-images were averaged across subjects to localize group-level functional activations by superimposition. During preprocessing, EPIs were spatially realigned, normalized (using the same transformation as for structural images), and smoothed with an 8-mm full-width at half-maximum Gaussian kernel.

We used four general linear models (GLMs) to explain pre-processed time-series at the individual level.

The first model (GLM0) was built for whole-brain search of voxels encoding prior preference in baseline activity. It was composed of a finite impulse response function (FIR) that included 7 time points per trial, from one TR (-2.01s) before to five TR (10.05s) after choice onset. The different blocks were modeled in separate regressors, each being parametrically modulated by the two option values (default and alternative). For each time point, we computed a contrast that weighted all trials of a given block by the corresponding prior preference ($DV_{CAT}$). Four subjects were excluded from this analysis because they presented at least one block without sufficient variance to estimate the parametric regression coefficients.

The second model (GLM1) included a stick function capturing option display (only one event per trial), parametrically modulated by the two option values (chosen and unchosen). The three regressors were convolved with a canonical hemodynamic response function.

The third model (GLM2) included two categorical regressors: a boxcar function over blocks and the same stick function as in GLM1. The boxcar function was parametrically modulated by $DV_{CAT}$, to account for tonic effects of prior preference. The stick function was parametrically modulated by 5 variables: chosen option (default or alternative), $V_{IT}$(def) when default chosen, $V_{IT}$(alt) when default chosen, $V_{IT}$(def) when default unchosen and $V_{IT}$(alt) when default unchosen. This allowed computing orthogonal contrasts for the decision value in the pre-choice (default vs. alternative) and post-choice (chosen vs. unchosen) frames. The regressors were convolved with a canonical HRF to localize brain regions where the evoked response reflected the decision value. In a subsequent analysis the same regressors were convolved with the same FIR as used for GLM0, in order to examine the dynamics of value coding in regions of interest (ROI).

The fourth model (GLM3) was equivalent to GLM1 except that the stick function modeling option display was modulated by the sum and difference of option values, in the pre-choice frame (default vs. alternative). Common variance between the two parametric regressors was removed such that they could explain a unique variance in the BOLD signal.

Motion artifacts were corrected in all GLMs by adding subject-specific realignment parameters as covariates of no interest. Regression coefficients were estimated at the individual level and then taken to group-level random-effect analysis using one-sample two-tailed t-test. In ROI analyses they were extracted from spheres of 6mm radius positioned on coordinates defined independently from the present dataset: for the vmPFC we took the peak coordinates [-2 40 -8] from a meta-analysis of value representation (Bartra et al., 2013), and for the dACC we took the peak coordinate [-6 24 34] of a negative correlation with chosen option value (Boorman et al., 2013).

Four variants of GLM2 were also compared to better characterize value coding in the two ROI. The four regressors modeling option values were replaced by a single regressor: 1) default option value, 2) pre-choice decision value (default minus alternative), 3) chosen option value, 4) post-choice decision value (chosen minus unchosen). This was meant to assess whether value representation concerned only one option or the difference between the two, and whether it was expressed in a pre-choice or post-choice frame. All models were fitted to individual time-series extracted from vmPFC and dACC ROI, so as to compute group-level exceedance probabilities, following a BMS procedure similar to that used for behavioral data analysis.

## *Author contributions*

A.L.-P., P.D. and M.P. conceived and designed the experiment. A.L.-P. collected the data and conducted the analysis. P.D and M.P. provided expertise and feedback on data analysis. A.L.-P., P.D. and M.P. wrote the manuscript. M.P. supervised the research.

## *Acknowledgements*

We are grateful to the PRISME platform for help in behavioral and eye-tracking data collection and to the CENIR staff for assistance in MRI data acquisition. The study was funded by a Starting Grant for the European Research Council (ERC-BioMotiv). This work also benefited from the program "Investissements d'avenir" (ANR-10-IAIHU-06). A.L.-P. received a PhD fellowship from the Direction Générale de l'Armement and from the LabEx Bio-Psy. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. We wish to thank Antonio Rangel for suggestions on data analysis and Lionel Rigoux for guidance on computational modeling.

## *Competing financial interests*

The authors declare no competing financial interests.

## *References*

Abitbol, R., Lebreton, M., Hollard, G., Richmond, B.J., Bouret, S., Pessiglione, M., 2015. Neural Mechanisms Underlying Contextual Dependency of Subjective Values: Converging Evidence from Monkeys and Humans. J. Neurosci. 35, 2308–2320. doi:10.1523/JNEUROSCI.1878-14.2015

Bartra, O., McGuire, J.T., Kable, J.W., 2013. The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. NeuroImage 76, 412–427. doi:10.1016/j.neuroimage.2013.02.063

Basten, U., Biele, G., Heekeren, H.R., Fiebach, C.J., 2010. How the brain integrates costs and benefits during decision making. Proc. Natl. Acad. Sci. 107, 21767–21772. doi:10.1073/pnas.0908104107

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., Cohen, J.D., 2006. The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. Psychol. Rev. 113, 700–765.

Boorman, E.D., Rushworth, M.F., Behrens, T.E., 2013. Ventromedial Prefrontal and Anterior Cingulate Cortex Adopt Choice and Default Reference Frames during Sequential Multi-Alternative Choice. J. Neurosci. 33, 2242–2253. doi:10.1523/JNEUROSCI.3022-12.2013

Brainard, D.H., 1997. The psychophysics toolbox. Spat. Vis. 10, 433–436.

Brunton, B.W., Botvinick, M.M., Brody, C.D., 2013. Rats and Humans Can Optimally Accumulate Evidence for Decision-Making. Science 340, 95–98. doi:10.1126/science.1233912

Cavanagh, J.F., Wiecki, T.V., Kochar, A., Frank, M.J., 2014. Eye Tracking and Pupillometry are Indicators of Dissociable Latent Decision Processes. J. Exp. Psychol. Gen. 143, 1476–1488. doi:10.1037/a0035813

Chib, V.S., Rangel, A., Shimojo, S., O'Doherty, J.P., 2009. Evidence for a Common Representation of Decision Values for Dissimilar Goods in Human Ventromedial Prefrontal Cortex. J. Neurosci. 29, 12315–12320. doi:10.1523/JNEUROSCI.2575-09.2009

Clithero, J.A., Rangel, A., 2014. Informatic parcellation of the network involved in the computation of subjective value. Soc. Cogn. Affect. Neurosci. 9, 1289–1302. doi:10.1093/scan/nst106

Daunizeau, J., Adam, V., Rigoux, L., 2014. VBA: A Probabilistic Treatment of Nonlinear Models for Neurobiological and Behavioural Data. PLoS Comput. Biol. 10, e1003441. doi:10.1371/journal.pcbi.1003441

De Martino, B., Fleming, S.M., Garrett, N., Dolan, R.J., 2013. Confidence in value-based choice. Nat. Neurosci. 16, 105–110. doi:10.1038/nn.3279

Fleming, S.M., Thomas, C.L., Dolan, R.J., 2010. Overcoming status quo bias in the human brain. Proc. Natl. Acad. Sci. 107, 6005–6009. doi:10.1073/pnas.0910380107

Friston, K., Mattout, J., Trujillo-Barreto, N., Ashburner, J., Penny, W., 2007. Variational free energy and the Laplace approximation. NeuroImage 34, 220–34. doi:10.1016/j.neuroimage.2006.08.035

Gold, J.I., Shadlen, M.N., 2007. The Neural Basis of Decision Making. Annu. Rev. Neurosci. 30, 535–574. doi:10.1146/annurev.neuro.29.051605.113038

Grasman, R.P.P.P., Wagenmakers, E.-J., van der Maas, H.L.J., 2009. On the mean and variance of response times under the diffusion model with an application to parameter estimation. J. Math. Psychol. 53, 55–68. doi:10.1016/j.jmp.2009.01.006

Grueschow, M., Polania, R., Hare, T.A., Ruff, C.C., 2015. Automatic versus Choice-Dependent Value Representations in the Human Brain. Neuron. doi:10.1016/j.neuron.2014.12.054

Hare, T.A., Camerer, C.F., Knoepfle, D.T., O'Doherty, J.P., Rangel, A., 2010. Value Computations in Ventral Medial Prefrontal Cortex during Charitable Decision Making Incorporate Input from Regions Involved in Social Cognition. J. Neurosci. 30, 583–590. doi:10.1523/JNEUROSCI.4089-09.2010

Hare, T.A., Schultz, W., Camerer, C.F., O'Doherty, J.P., Rangel, A., 2011. Transformation of stimulus value signals into motor commands during simple choice. Proc. Natl. Acad. Sci. 108, 18120–18125. doi:10.1073/pnas.1109322108

Hayden, B.Y., Pearson, J.M., Platt, M.L., 2011. Neuronal basis of sequential foraging decisions in a patchy environment. Nat. Neurosci. 14, 933–939. doi:10.1038/nn.2856

Hunt, L.T., Kolling, N., Soltani, A., Woolrich, M.W., Rushworth, M.F., Behrens, T.E., 2012. Mechanisms underlying cortical activity during value-guided choice. Nat. Neurosci. 15, 470-S3. doi:10.1038/nn.3017

Kable, J.W., Glimcher, P.W., 2009. The Neurobiology of Decision: Consensus and Controversy. Neuron 63, 733–745. doi:10.1016/j.neuron.2009.09.003

Kolling, N., Behrens, T.E.J., Mars, R.B., Rushworth, M.F.S., 2012. Neural Mechanisms of Foraging. Science 336, 95–98. doi:10.1126/science.1216930

Krajbich, I., Armel, C., Rangel, A., 2010. Visual fixations and the computation and comparison of value in simple choice. Nat. Neurosci. 13, 1292–1298. doi:10.1038/nn.2635

Krajbich, I., Lu, D., Camerer, C., Rangel, A., 2012. The Attentional Drift-Diffusion Model Extends to Simple Purchasing Decisions. Front. Psychol. 3. doi:10.3389/fpsyg.2012.00193

Krajbich, I., Rangel, A., 2011. Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. Proc. Natl. Acad. Sci. 108, 13852–13857. doi:10.1073/pnas.1101328108

Kvitsiani, D., Ranade, S., Hangya, B., Taniguchi, H., Huang, J.Z., Kepecs, A., 2013. Distinct behavioural and network correlates of two interneuron types in prefrontal cortex. Nature 498, 363–366. doi:10.1038/nature12176

Lebreton, M., Abitbol, R., Daunizeau, J., Pessiglione, M., 2015. Automatic integration of confidence in the brain valuation signal. Nat. Neurosci. 18, 1159–1167. doi:10.1038/nn.4064

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. doi:10.1016/j.neuron.2009.09.040

Levy, D.J., Glimcher, P.W., 2012. The root of all value: a neural common currency for choice. Curr. Opin. Neurobiol. doi:10.1016/j.conb.2012.06.001

Levy, I., Lazzaro, S.C., Rutledge, R.B., Glimcher, P.W., 2011. Choice from Non-Choice: Predicting Consumer Preferences from Blood Oxygenation Level-Dependent Signals Obtained during Passive Viewing. J. Neurosci. 31, 118–125. doi:10.1523/JNEUROSCI.3214-10.2011

Lim, S.-L., O'Doherty, J.P., Rangel, A., 2011. The Decision Value Computations in the vmPFC and Striatum Use a Relative Value Code That is Guided by Visual Attention. J. Neurosci. 31, 13214–13223. doi:10.1523/JNEUROSCI.1246-11.2011

Link, S.W., Heath, R.A., 1975. A sequential theory of psychological discrimination. Psychometrika 40, 77–105. doi:10.1007/BF02291481

Mulder, M.J., Wagenmakers, E.-J., Ratcliff, R., Boekel, W., Forstmann, B.U., 2012. Bias in the Brain: A Diffusion Model Analysis of Prior Probability and Potential Payoff. J. Neurosci. 32, 2335–2343. doi:10.1523/JNEUROSCI.4156-11.2012

Palminteri, S., Boraud, T., Lafargue, G., Dubois, B., Pessiglione, M., 2009. Brain Hemispheres Selectively Track the Expected Value of Contralateral Options. J. Neurosci. 29, 13465–13472. doi:10.1523/JNEUROSCI.1500-09.2009

Pelli, D.G., 1997. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. Spat. Vis. 10, 437–442.

Penny, W.D., 2012. Comparing dynamic causal models using AIC, BIC and free energy. NeuroImage 59, 319–30. doi:10.1016/j.neuroimage.2011.07.039

Penny, W.D., Stephan, K.E., Daunizeau, J., Rosa, M.J., Friston, K.J., Schofield, T.M., Leff, A.P., 2010. Comparing families of dynamic causal models. PLoS Comput. Biol. 6, e1000709. doi:10.1371/journal.pcbi.1000709

Plassmann, H., O'Doherty, J., Rangel, A., 2007. Orbitofrontal Cortex

Encodes Willingness to Pay in Everyday Economic Transactions. J. Neurosci. 27, 9984–9988. doi:10.1523/JNEUROSCI.2131-07.2007

Rangel, A., Hare, T., 2010. Neural computations associated with goal-directed choice. Curr. Opin. Neurobiol. 20, 262–270. doi:10.1016/j.conb.2010.03.001

Ratcliff, R., 1978. A theory of memory retrieval. Psychol. Rev. 85, 59–108. doi:10.1037/0033-295X.85.2.59

Ratcliff, R., McKoon, G., 2008. The diffusion decision model: theory and data for two-choice decision tasks. Neural Comput. 20, 873–922.

Rigoux, L., Stephan, K.E., Friston, K.J., Daunizeau, J., 2014. Bayesian model selection for group studies - Revisited. NeuroImage 84, 971–85. doi:10.1016/j.neuroimage.2013.08.065

Samuelson, P.A., 1938. The numerical representation of ordered classifications and the concept of utility. Rev. Econ. Stud. 6, 65–70.

Scheibe, C., Ullsperger, M., Sommer, W., Heekeren, H.R., 2010. Effects of Parametrical and Trial-to-Trial Variation in Prior Probability Processing Revealed by Simultaneous Electroencephalogram/Functional Magnetic Resonance Imaging. J. Neurosci. 30, 16709–16717. doi:10.1523/JNEUROSCI.3949-09.2010

Shenhav, A., Botvinick, M.M., Cohen, J.D., 2013. The Expected Value of Control: An Integrative Theory of Anterior Cingulate Cortex Function. Neuron 79, 217–240. doi:10.1016/j.neuron.2013.07.007

Shenhav, A., Straccia, M.A., Cohen, J.D., Botvinick, M.M., 2014. Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value. Nat. Neurosci. 17, 1249–1254. doi:10.1038/nn.3771

Skvortsova, V., Palminteri, S., Pessiglione, M., 2014. Learning To Minimize Efforts versus Maximizing Rewards: Computational Principles and Neural Correlates. J. Neurosci. 34, 15621–15630. doi:10.1523/JNEUROSCI.1350-14.2014

Stephan, K.E., Penny, W.D., Daunizeau, J., Moran, R.J., Friston, K.J., 2009. Bayesian model selection for group studies. NeuroImage 46, 1004–17. doi:10.1016/j.neuroimage.2009.03.025

Stephens, D.W., Krebs, J.R., 1986. Foraging theory. Princeton University Press.

Summerfield, C., Koechlin, E., 2010. Economic Value Biases Uncertain Perceptual Choices in the Parietal and Prefrontal Cortices. Front. Hum. Neurosci. 4. doi:10.3389/fnhum.2010.00208

Von Neumann, J., Morgenstern, O., 1947. Theory of games and economic behavior.

Voss, A., Nagler, M., Lerche, V., 2013. Diffusion Models in Experimental Psychology: A Practical Introduction. Exp. Psychol. Former. Z. Für Exp. Psychol. 60, 385–402. doi:10.1027/1618-3169/a000218

Wagenmakers, E.-J., van der Maas, H.L.J., Grasman, R.P.P.P., 2007. An EZ-diffusion model for response time and accuracy. Psychon.

Bull. Rev. 14, 3–22.

Wunderlich, K., Rangel, A., O'Doherty, J.P., 2010. Economic choices can be made using only stimulus values. Proc. Natl. Acad. Sci. 107, 15005–15010. doi:10.1073/pnas.1002258107

Wunderlich, K., Rangel, A., O'Doherty, J.P., 2009. Neural computations underlying action-based decision making in the human brain. Proc. Natl. Acad. Sci. 106, 17199–17204.

Young, C.B., Nusslock, R., 2016. Positive mood enhances reward-related neural activity. Soc. Cogn. Affect. Neurosci. 11, 934–944. doi:10.1093/scan/nsw012

## *Supplementary figure*

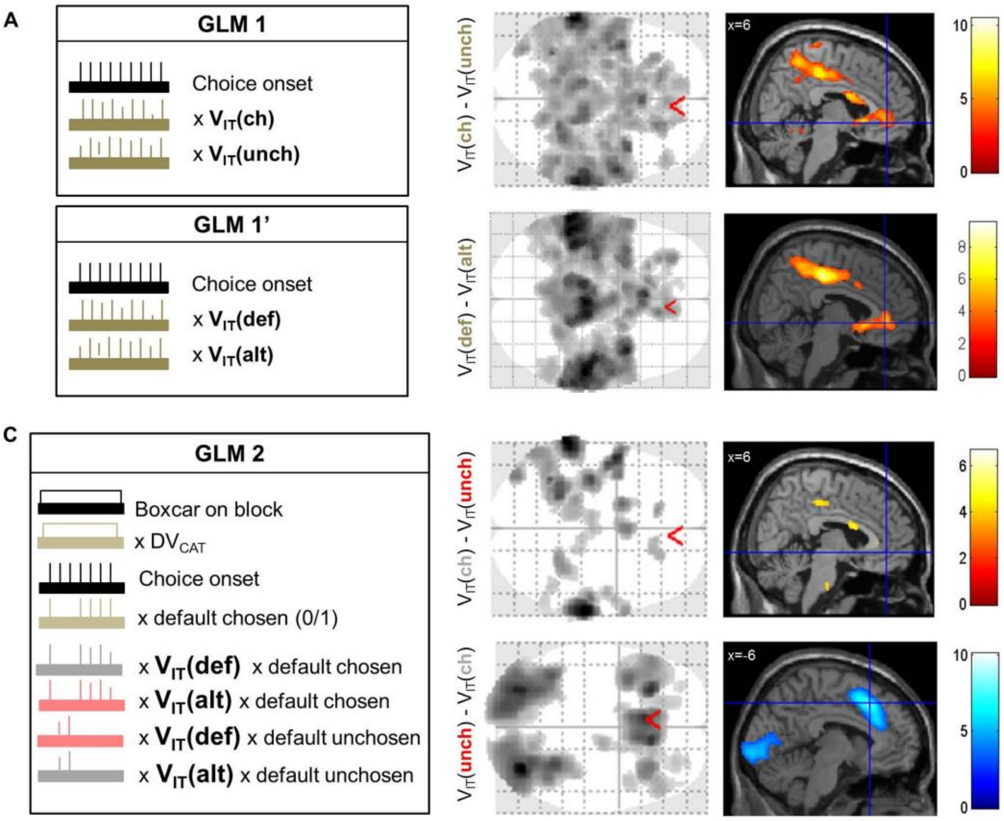


Figure 4 – figure supplement 1. Dissociation of neural value representations expressed in the pre-choice vs. post-choice frame.

(A) Simple contrasts between chosen and unchosen value (GLM1) or between default and alternative options (GLM1′). The box on the left illustrates GLM1 and GLM1′, with a delta function indicating the onset of choice options (black), modulated by two parametric regressors representing the option values (brown). The maps on the right show in black (on glass brain) and with color code (on sagittal slice) clusters that significantly reflected post-choice (top) and pre-choice (bottom) decision value at the group level (one-sample t-test, $p<0.005$ uncorrected for display purposes, minimum extent: 100 voxels). In both cases significant clusters were found in the vmPFC and ventral striatum, among many other areas, with global maxima in the intra-parietal lobules. (B) Dissociation between decision value coding in the pre-choice and post-choice frames using GLM2. The box on the left illustrates GLM2, with a boxcar and delta functions (black) indicating the duration of blocks and the onset of choice options, respectively. The boxcar function was parametrically modulated by the prior preference, whereas the delta function was parametrically modulated by choice type and the four interactions between choice type and option values, i.e. default and alternative option values when default is chosen and when it is not chosen. The maps on the right show significant clusters (using same tests and thresholds as in A) in the contrasts between grey and red regressors, which correspond to chosen versus unchosen values. These contrasts show differences in parameter estimates for the value of chosen options (irrespective of whether they are default or alternative), relative to unchosen options. Compared to the result shown in A, the chosen minus unchosen option value contrast still yielded significant activation outside the brain valuation system, but not in the vmPFC or ventral striatum (with identical statistical threshold). The orthogonal contrast, between default and alternative options values (irrespective of whether they are chosen or not), did result in activation of the vmPFC and ventral striatum, as shown in Figure 4B and Table S1 (Supplementary file 1).

## *Supplementary file 1*

**Table S1 – Activation list for decision value coding in the pre-choice and post-choice frames (GLM 3).**

Regions are listed that survived voxel-based thresholding of $p<0.001$ uncorrected, and whole-brain cluster-level FWE correction. [x, y, z] coordinates refer to the Montreal Neurological Institute (MNI) space.

$V_{IT}(def)-V_{IT}(alt)$

| **Cluster** | **Cluster size** | **peak side** | **x** | **y** | **z** | **Z score** | **T value** | **Cluster p value** |
|---|---|---|---|---|---|---|---|---|
| vmPFC | 1262 | R | 2 | 40 | 10 | 4.85 | 6.51 | $9.10^{-10}$ |
| Ventral striatum | 675 | L | -18 | 4 | 10 | 4.74 | 6.28 | $3.10^{-6}$ |

$V_{IT}(ch)-V_{IT}(unch)$

| **Cluster** | **Cluster size** | **peak side** | **x** | **y** | **z** | **Z score** | **T value** | **Cluster p value** |
|---|---|---|---|---|---|---|---|---|
| Rolandic operculum | 342 | L | -56 | 2 | 8 | 4.92 | 6.66 | 0.004 |
| Inferior parietal lobule | 367 | R | 64 | -24 | 40 | 4.71 | 6.21 | 0.003 |
|  | 368 | L | -66 | -38 | 36 | 4.56 | 5.90 | 0.003 |
| Angular gyrus | 208 | R | -24 | -44 | 28 | 4.52 | 5.82 | 0.035 |

$V_{IT}(unch)-V_{IT}(ch)$

| **Cluster** | **Cluster size** | **peak side** | **x** | **y** | **z** | **Z score** | **T value** | **Cluster p value** |
|---|---|---|---|---|---|---|---|---|
| Middle occipital gyrus | 4901 | R | 32 | -80 | 18 | 6.18 | 10.08 | $<1.10^{-11}$ |
| dACC | 2202 | R | 10 | 26 | 32 | 5.74 | 9.66 | $2.10^{-11}$ |
| Anterior insula | 1774 | L | -28 | 26 | 4 | 5.83 | 8.99 | $6.10^{-10}$ |
| Middle occipital gyrus | 5289 | L | -28 | -84 | 20 | 5.74 | 8.73 | $6.10^{-4}$ |
| Anterior insula | 876 | R | 30 | 26 | -2 | 4.84 | 6.48 | $4.10^{-6}$ |

## C. Discussion

In this second experiment, we used fMRI to investigate the neural substrate of decision-making. We found that prior preferences for categories of objects such as savory food compared to sweet food induced biases in the behavior toward the option of the preferred category, and thus defined a default option. Interestingly, the strength of preference for a category was modulating the baseline activity of the Brain Valuation System and the vmPFC evoked activity was correlating with the decision value in the frame of the default policy.

The general validity of the 'default frame' followed by the vmPFC signal remains to be established and one major limitation we could address on this study is the impact of the experimental design on our study. We did not present it in the discussion of the paper but we have reasons to believe that this bias could be observed without any instruction on categories as we did in our experiment. Two unpublished results are supporting this idea. First, before running the experiment, we ran what we call a 'pilot' study in order to set up the experiment. Initially, participants were asked to make choices between items belonging to either two different domains (such as a choice between eating a chocolate bar and listening to Michael Jackson) either one unique domain (such as a choice between reading 'Le Monde' or 'Le Figaro'). Surprisingly we could not find evidence for a default response in the inter-domain condition, but when investigating further the choices in the intra-domain condition, we realized that evidence for a default response could be found in specific choices such as 'politically lefty newspaper against righty newspaper' or even 'savory food against sweet food'. These examples of categories were defined post-hoc and allowed us redefining our design: proposed choices needed to be intra-domain but inter-category to evoke a default response. I will come back on this constraint of the design later but here what we need to recall is that default response evidence was found in the behavior without any instruction. Moreover, a second unpublished result supports this idea. In a dataset assessing cognitive dissonance from another team in the lab, subjects had first to rate travel destinations and then to make choice between them. There was no instruction about categories in this task. When creating post-hoc categories for travel destinations such as 'Hot/Cold' or 'Urban/Wild', I could find evidence for a default response in the behavior, with an effect in both choice and reaction time (Figure 38). Those arguments are obviously not a proof of concept but they support the hypothesis that defining a default option might be something 'natural', and also that it is possible that the vmPFC is using

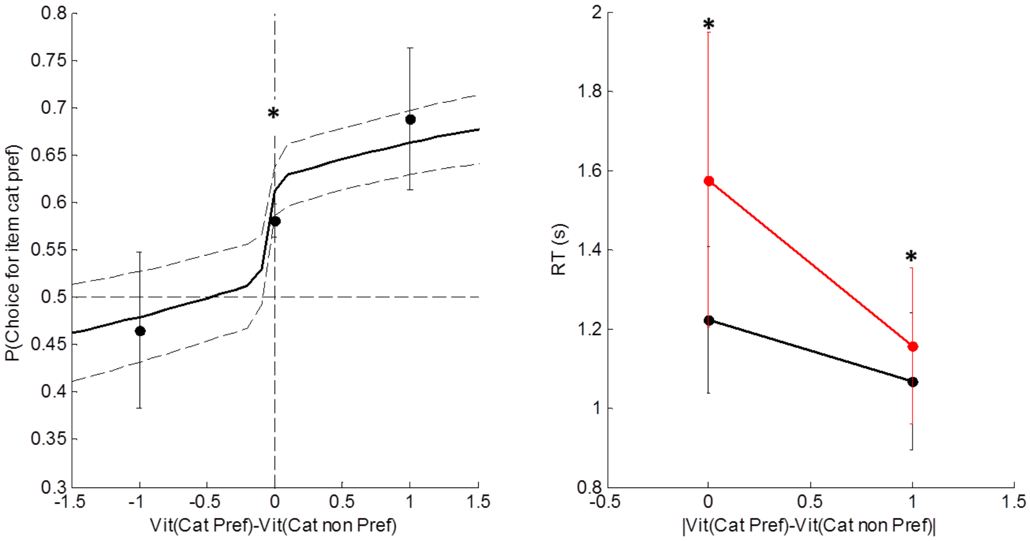


Figure 38 – Example of default bias observed in a study without framing.

**Left:** Proportion of choices for an item (travel destination) according to the decision value following the frame of preference for categories such as 'hot vs cold' countries or 'urban vs wild' destination. The proportion of choice for the item belonging to the preferred category is significantly different from 0.5 when the decision value is 0. **Right:** The bias is also seen in the reaction time, with the reaction time for a choice in favor of an item belonging to the non preferred category depicted in red. Stars indicate significance. The Range of the difference value is small because they were selected to be so (cognitive dissonance protocole).

the default frame of decision value coding without any instruction. However, this needs to be experimentally assessed and proven.

The second point we do not discuss in the paper is the formalization of a default. As mentioned, our design triggered evidence for the existence of default options in the context of choices which are intra-domain and inter-category but not inter-domain. This constraint is raising limits that worth being discussed. What is a default option? What makes it emerging? These questions are more about the representation of options and concepts than value representation but testing how the hierarchical organization of conceptual knowledge would impact and define a default option is a parallel but thrilling question which might deserve to be assessed in future studies.

Finally, this study is bringing new insights on how the brain implements decision-making but the story is still incomplete since we did not address the question of the comparison process itself in this experiment.

# Chapter 3: General discussion

With the three studies performed during my PhD, we were able to 1) validate the tasks used in the field to measure subjective values by showing that they are able to reveal the same hidden utility function in the behavior; 2) precise the role of the Brain Valuation System and the OFC, by replicating some fundamental properties of the vmPFC in electrophysiological data and by bringing some insights on the neural dynamics of this network; 3) propose a general framework to investigate decision-making processes by showing that default policy could shape the neural signal of the vmPFC.

In the following section, I will first discuss methodological aspects of these studies, then theoretical implications which can be drawn from them and I will finish by discussing some limitations of this work.

## 1. Methodological approach

### A. On the advantages of using model fitting & model comparisons.

In the first and the third studies, we used model fitting and model comparison to respectively identify what was the underlying utility function of the behavior and investigate the cognitive processes underlying the biases observed in the behavior. This methodology allowed us revealing hidden parameters of the behavior in both studies and for the third study to develop hypothesis on the potential neural mechanisms involved. In both studies, we used models as tools to further investigate the processes of interest even if we do not claim that the brain is actually implementing those exact models: we use them as '*as if*' mechanisms and as a way to explore in greater details the behavior. Indeed, if we take the example of the Drift Diffusion Model, we saw in the introduction that it might not be directly implemented by neurons (see page 104 and (Latimer et al., 2015)), and that it is consequently probable that an 'accumulation-to-bound process' does not actually take place in the context of subjective valuation. However, in our third study, we used the Drift Diffusion Model to fit our data and we found that the observed bias was best accounted for by a shift in the starting point. This result led us to look for a shift in the baseline activity in the brain, as suggested by the biological equivalent of a starting point. Thanks to this insight, we found that the Brain Valuation System baseline activity was actually sensitive to prior preference strength and it helped us to better understand the origin of the biases observed in the behavior.

Moreover, still in the first and the third study, we used model comparisons to bring answers to usually unsolved problems. For example, generally, we cannot conclude anything from an absence of significant difference between two conditions[19]. However, with a model comparison, we were allowed to conclude that there was significantly no difference in the utility function revealed by the three measures since the family model 'same utility function' won the model comparison against 'different utility functions'. The other example comes from the third study in which we supposed that we were observing only one value ($V_{IT}$(def)) in the vmPFC potentially because of its automatic property which would let us suppose that it codes both values of the presented items but also the decision value such as the signal would follow:

$$\text{Signal}_{vmPFC} \sim V_{IT}(\text{def}) + V_{IT}(\text{alt}) + [V_{IT}(\text{def}) - V_{IT}(\text{def})]$$

Which is an equivalent of:

---

[19] Since we can only reject the null hypothesis or accept the alternative but not the opposite.

$\text{Signal}_{\text{vmPFC}} \sim 2^* V_{\text{IT}}(\text{def})$

Thus, at the populational level with a spatial and temporal resolution such as the one of fMRI, one would only observe a correlation with $V_{\text{IT}}(\text{def})$. Thanks to a Bayesian model comparison applied to our fMRI data, we were able to show that the vmPFC signal was better explained by the sum and the difference of the option values instead of the single value $V_{\text{IT}}(\text{def})$. This is providing an additional evidence in favor of our claim.

Given those examples, I would conclude that it is not only useful to use model comparison in neuroscience of decision-making but it might be considered as necessary in situations in which one would need to disentangle between several mechanistic accounts leading to similar observations at the group level.

## B. LFP & BOLD: Similarities & differences

Classically, in humans, if one is interested in neural dynamics, she would use EEG or MEG to investigate neural activity but with the cost of a low spatial resolution. On the other hand, if one is interested in neural events supposed to occur in deep brain structures, fMRI would be the neuroimaging method which would be selected, but it comes also with a loss of temporal precision, thus dynamics of processes cannot be properly investigated. In animals, electrophysiology recordings allow investigating the dynamics of neural processes in a precise location but the technic is highly invasive since it implies implanting electrodes inside the brain. Moreover, this last technic is also constrained to one or few brain regions simultaneously recorded.

During my PhD, I used fMRI to investigate neural decision processes but I also had the opportunity to use electrophysiological recordings (iEEG) in humans in tasks in which the BOLD response is well established. Indeed, rating tasks on food, faces and paintings usually trigger BOLD responses in the Brain Valuation System (Lebreton et al., 2009). Thus, the results obtained in the second study with iEEG can be (methodologically) compared to those classically observed in fMRI.

The first interesting observation on the high gamma activity recorded in the vmPFC in our study is that the evoked response is negative and value is coded through an inverse proportional decrease: the lowest the value is, the greatest the decrease in high gamma will be. In fMRI studies, we also observe this decrease in the BOLD signal (Figure 39) but it is still unexplained. The vmPFC is also part of the Default Mode Network, a network activated during resting state and deactivated by executive tasks. Thus, one could speculate on this link to understand the exact dynamics of the vmPFC. However, even if this deserves to be investigated, it is not the topic of this manuscript and a

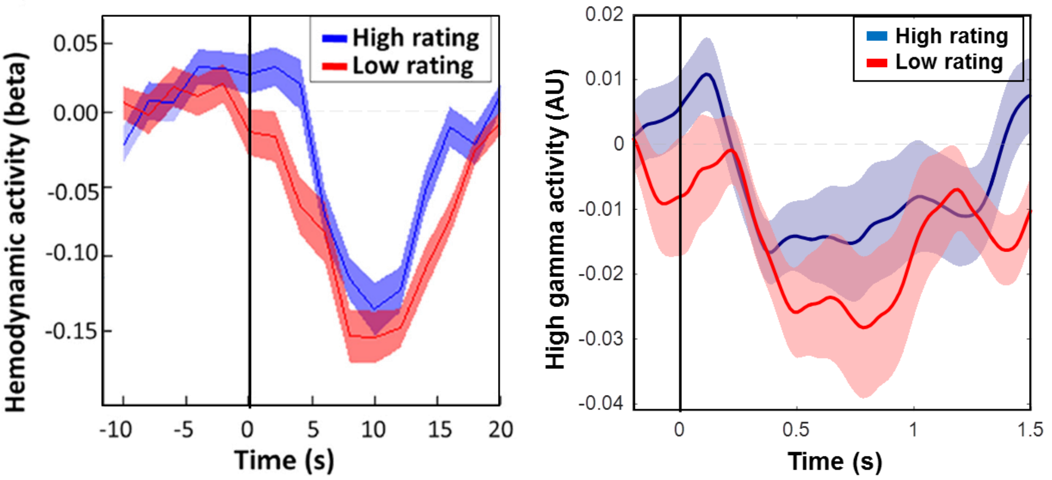


Figure 39 – vmPFC evoked activity in fMRI and iEEG.

Left. The vmPFC evoked BOLD response is characteriezd by a decrease in the signal in fMRI studies (Figure adapted from Abitbol et al., 2015). Right. The same pattern of response was observed in the high gamma vmPFC response (adapted from the figure 3C of the iEEG study). Blue and red respectively indicate high and low likeability ratings.

whole study would be necessary to bring some insights to this question. Here, the point is that we observe the same dynamics in BOLD and in high gamma activity in the vmPFC.

As we saw in the introduction (page 75), a classical discrepancy between animal and human studies in the field is the OFC involvement in value-based processes: while it is rarely found to parametrically code value in humans, it is less clear in animals. However the main difference might not be due to the species given the established homologies (Mackey and Petrides, 2010) but potentially to the neuroimaging techniques. Indeed, in our second study, we found a strong involvement of the OFC in subjective value coding in high gamma (but also in the other frequencies), which is uncommon in humans (as suggested by the fMRI meta-analysis done by Bartra et al., (2013) in which the OFC does not belong to the human Brain Valuation System).

High-frequency local field potential (LFP) signal have been shown to correlate with the evoked BOLD response (Logothetis et al., 2001). In our study, we found high gamma activity correlation with value in the OFC so we should logically observe it in an equivalent fMRI study, but it is not the case. In the vmPFC, we found an involvement of the theta phase in value coding, as in the OFC, while we could not find any involvement of intermediate frequencies such as in the OFC. Thus, one would rightfully ask what are the interactions between alpha and beta bands with the BOLD signal. One interesting study investigated the link between the different frequencies and the BOLD signal amplitude and timing (Magri et al., 2012). They wanted to test the predictions of a theory developed by Kilner et al., (2005) in which the authors suggest that an increase in low frequencies such as alpha would lead to a reduction in the BOLD signal. Magri et al., (2012) tested and confirmed this prediction. They also went further by showing that the interaction between the gamma and alpha bands was reflected in the amplitude of the BOLD signal while between gamma and beta, it was reflected in the latency of the BOLD response. Thus, given these results, we can suppose that the interaction between value coding in alpha and high gamma frequencies may affect the BOLD signal observed in fMRI and could explain why we do not classically observe an OFC response to subjective value in fMRI but more often in electrophysiological data. This is obviously speculations which would need to be properly tested but I believe that this kind of LFP result is highly informative to decipher the neural code of subjective valuation.

In the following sections, I will discuss the theoretical implications of the work conducted during my PhD.

## 2. Theoretical implications

### A. On the behavior

#### *Internal context & instability of preferences*

As presented in the first chapter, economists have spent centuries trying to formalize theories such as the 'Expected Utility Theory'(Von Neumann and Morgenstern, 1947) or the 'Prospect Theory' (Kahneman and Tversky, 1979) in order to explain biases in choices and instability of preferences. In our first behavioral study, we showed that subjective values elicited through different measures could be considered as equivalent. However, we found that the accuracy to predict choices never exceeded 80%, which is i) in line with what is usually observed in the literature (see Lebreton et al.,( 2009) for an example in which they found a maximal prediction of 75%); ii) confirming that choices can easily reverse. We saw that many factors can induce those reversals but one which is interesting here is '**internal context'** such as the ongoing brain activity. Indeed, in both studies, we found an effect of the vmPFC baseline activity on the provided ratings (iEEG study) and on choices (fMRI study). In the first case, we did not manipulate the external environment but still, we were able to observe a dependency between ongoing activity and provided responses, as observed in perceptual choices (Hesselmann et al., 2008) or subjective ratings (Abitbol et al., 2015). Such a dependency shown without any involvement of a measurable factor (like music listening or trial number as in Abitbol et al., (2015) makes this result strongly reliable and I believe that ongoing activity is one main factor of preference instability.

In the second study, we saw that choices could be biased by prior information (i.e. by informing subjects on which type of items they would face). We found that the behavioral effect was strong and easily replicable (in the Eye-tracking experiment and in a pilot experiment not mentioned in the text) and that it was not induced by instructions since the only thing which was told to the subject regarding the framing was '*Information will be given at the beginning of each block, you won't have to do anything with it*' and they were never asked to '*prepare an answer*'. Thus, this manipulation is a good example on how context (here context is prior knowledge about categories) can bias choices and induce irrational choices (i.e. choosing the default even if the alternative is better). Furthermore, in line with the previous results cited on the influence of baseline activity on the behavior, we found that prior information value was affecting the vmPFC baseline activity and that it could be a good candidate to explain the observed behavioral biases in our experiment.

### *Natural choices & default options*

This '*natural*' bias toward the default allows us raising the question of the 'natural choice'. Indeed, it has been suggested (Stephens and Krebs, 1986) that a natural choice would follow a 'switch/stay' frame with the 'stay' option defined as the default option in order to (for example) avoiding giving away a prey already captured which would have not been exploited yet. Our finding adds supportive evidence for this idea since we showed that the vmPFC is following this frame. An implication of this idea is that choices usually presented in the lab (without priors) would prevent the formation of any default option and rule out any naturalness in the process. In this point of view, investigating the brain activity in such a situation would in the best case show how the brain *solves the problem* (i.e. computes a decision value without prior to frame it) but not how it is naturally processing a choice. We believe that the attentional frame *attended/unattended* (Lim et al., 2011) could potentially be a solution to this frame problem in un-natural situations in which the default would be set up on the fly, according to the fixation pattern. We see here how critical the idea of a default option is for neural computations but it still needs to be further investigated since the definition of the 'default' (i.e. the label assignment) is not trivial. For example, there are challenging situations in which knowing what is the default is not easy such as intertemporal choices or dietary choices. Indeed, in situations in which control needs to be implemented to choose the long-term best option (delayed or healthy option), knowing which one is the default is not intuitive. In such cases, the default option is defined either as the immediate/tasty option because a default option should not be costly to be selected; either as the delayed/healthy option because it is the one which would be in average more rewarding than the other. A previous study by Boorman et al.,

(2013) actually used an experimental design in which a default option was defined as the long-term best option since more rewarding than the others in average. They showed behavioral evidence for the default option (it was chosen more often) but they found opposite result in the vmPFC and in the dACC compared to our study. For now, we cannot provide any satisfying explanation of this discrepancy. If control is needed to override the default response, we could speculate on the fact that the long term best option would need some control to be chosen and consequently the natural framing effect could be annihilated. However, more experiments are needed to establish the generality of our result. Moreover, some experiments are currently conducted in the lab in order to investigate 'natural choices' situations and the link with control.

## B. Properties of the Brain Valuation System

As mentioned in the literature review and in the iEEG study, the Brain Valuation System responds proportionally to subjective value in a generic and automatic manner, but it is also dependent on its own baseline activity and it codes confidence on top of the subjective value.

### *Replication & new insights on the valuation process*

In our both neuroimaging studies, we were able to replicate the **genericity** property with faces and painting in the iEEG study and with music and magazines in the fMRI study together with food items. This result is in line with the 'common neural currency' theory (Levy and Glimcher, 2012) and it provides a new replication which make this property strongly reliable. Moreover, we saw that this property partially relies on the high gamma frequency, and that the same code is used by the vmPFC to represent values of various categories of items.

We were able to replicate the **automaticity** property (i.e. any stimulus that comes under attention is assigned a value) directly in the iEEG study by assessing value coding during an age evaluation task but also in the fMRI study in which value of categories influenced the baseline activity when presented to the subjects. This property has strong implications and can provide interesting insights for the underlying mechanisms of choices based on subjective values. Indeed, we can suppose here that the vmPFC is processing values as the visual cortex is processing visual features. This analogy posits the vmPFC as an 'evaluator' and support the idea of a distinction between the valuation and the choice selection process.

Moreover, we saw that the **baseline dependency** was also true in both studies, with a higher baseline linked to higher likeability ratings. Our studies also showed that this baseline dependency was true in two situations: with or without experimental factors influencing it like in the fMRI (display of the category labels) or in the iEEG study respectively. Moreover, while the fMRI result about the baseline effect supports the mechanistic account of the baseline effect proposed by Abitbol et al., 2015 (see page 92 and Figure 29), it is not the case for the iEEG study. Indeed, they supposed that a shift in baseline activity would persist into post-stimulus activity and impact the absolute peak of the vmPFC activity which in turn would impact the subjective value rating given to a specific stimulus. In the multivariate analysis we conducted on the vmPFC signal with the decoder trained on the baseline, such a mechanistic account would predict that the decoding score would progressively decrease in function of time. However, we observed that the decoding score first decreases and then increases at the time to which subjective value is coded. This is suggesting that another mechanism might underlie the anticipation effect and it raises interesting questions for further research in this direction.

Finally, in the iEEG study, we were also able to replicate the quadratic coding property (the fMRI study did not allowed us to test it). This last property were shown to be locked on the beginning of the response in the high gamma frequency, suggesting that local neuronal populations are indeed representing squared judgments in relation to response. However, we do not claim here that the vmPFC is computing itself the confidence level in the response given the fact that some studies found neural correlates of confidence in other areas such as rlPFC (Fleming and Dolan, 2012; negative correlate). Then, two points of view could be compared: either confidence is a subjective value per se and it is consequently represented in the vmPFC; either correlates of confidence in the vmPFC signal could actually be interpreted as a value assigned to the confidence level itself: we usually like to be confident and we often try to reduce uncertainty. Moreover, if we combine the properties of the vmPFC and try to define it as an brain area assigning a *hedonic value* to *every kind* of objects, even *internal* states, without an *explicit request* in order to *orient* and *energize* the behavior, then - if confidence is the probability that an answer is correct (a subjective value can be assigned to this probability) - it should be represented in the vmPFC (in the two perspectives). Consequently, it might interact with the value of the object at stake. This is obviously speculative and experiment(s) would be needed to further investigate how confidence is represented in the brain and whether the vmPFC is representing confidence or an affective value linked to confidence.

### *The role of the vmPFC in decision-making*

Many studies showed that the vmPFC is coding the value of the chosen option during a binary choice (Boorman et al., 2013; Hunt et al., 2012). We also replicated this result in our fMRI study but we went further by showing that in our study this result was largely due to an overlap between the default option and the chosen option and that the vmPFC was coding the value of the default option even if it was not chosen. Thus, we provide a new theory on how the decision value could be computed in the brain. Furthermore, we propose that it could explain many results observed in past studies such as a correlation of the vmPFC BOLD signal with the chosen option value or even with the attended option value given that the attended option could be defined as the default if no any other information, as I mentioned in the last section.

Given the results obtained in the fMRI experiment, I added a block of choices in the design of the iEEG experiment in which savoury food is opposed to sweet food. I simply told the patients before the task that on each trial they would have to choose between a savory item and a sweet item without any further information. On four patients (out of five screened), I could see the bias in choices and response time (see Figure 40 for an example). These results are encouraging and show that the effect is quite robust. The iEEG data are currently being analyzed for 4 dipoles in the vmPFC among those patients. We will try to collect more data on this block to properly address the dynamics of the default policy.

In a study investigating the dynamics of value-guided choices (Hunt et al., 2012), they tested the predictions of a biophysically plausible network model similar to Wang's model presented page 106 and found that experimental observations were matching those predictions. Indeed, they found that in low frequency bands (2-8Hz) vmPFC signal was first correlated with the value of both options and then the difference between the chosen and the unchosen option. First, we see here that as them, we found that subjective value could be reflected in low frequencies such as theta in our iEEG study. Then, given the fact that we claim that the vmPFC activity is not following the post-decisional frame ‘Chosen-Unchosen’, it could be interesting to test our paradigm with MEG (or iEEG) and see whether the theta band power follows the choice or the default frame. It would allow infirming or confirming the plausibility of the proposed model. For now, we only have some preliminary results regarding the vmPFC activity during the binary choices in our iEEG data set. Given that we did not have any framing or inter-categorical choices in those data, we separated choices in two conditions: choices for the option which had the highest rating and choices for the lowest (congruent and incongruent choices). As in our fMRI study, choices which are interesting here are the incongruent ones. When we investigated the high gamma activity related to those choices, we found that the unchosen option value was positively correlated to the high gamma activity while the chosen was negatively correlated. This is obviously a preliminary result and we did not yet

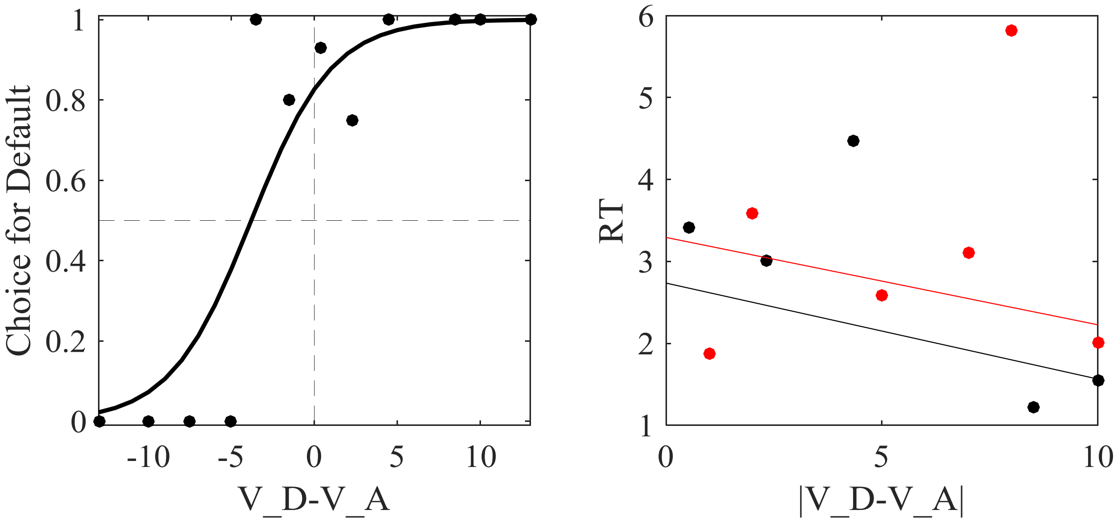


Figure 40 – Example of one epileptic patient for choices between savoury and sweet food.

Left: Proportion of choices for the default in function of the signed difference Vdefault-Valternative. (V_D: Vdefault; V_A: Valternative). Right: Response time according to the unsigned difference. Red dots are bins for response time when the default is not chosen. Dots are data and line are logistic fit for choices and linear fit for response time.

investigated how the theta band was involved in value coding but it let us suppose that the vmPFC activity is not following the post-choice frame and it could be an argument for the selection being implemented elsewhere in the brain.

So far we have focused on the healthy functioning of the vmPFC, but another way to address our questions is to investigate the behavior when the brain region is damaged. It has been shown that choices are noisier in patients with lesions to the vmPFC (Fellows, 2007) but interestingly, there is no difference in the response time (Henri-Bhargava et al., 2012) between vmPFC patients and control patients. This last result might suggest that the vmPFC is more linked to the valuation process than to the selection process: if the vmPFC is implementing both valuation and selection, decision time would be affected in one way or another (faster because of quick indifferent choices or slower because of undecision triggered by less distinguishable values). On the other hand, if vmPFC is only assigning values to options, values would be unstable but the selection of option would still rely on those values and the decision time would not be globally affected. However, it would be interesting to test the predictions we can formulate with our 'default' task: we could expect that the default bias, if implemented in the vmPFC, would disappear in such patients. However, given that response time does not seem to be affected by vmPFC lesions, it would imply that all response time would be in the range of the one observed for choices toward the alternative (under the assumption that having a default increases speed for choices <u>toward the default</u> instead of a decreasing speed for choices <u>toward the alternative</u> or a mixture of the two phenomena). Thus, it could be interesting to investigate the behavior of this kind of patients to further address the 'default hypothesis'.

To conclude on this part, we propose here that the vmPFC is an ‘**evaluator**’ and that it could provide a decision value in a default frame. Indeed, in our observations, we lack evidences supporting the idea that the vmPFC is implementing the option selection itself as suggested by the previous mentioned study.

## C. Components of the Brain Valuation System

A limitation of our approach is the tendency to focus on the vmPFC alone. However, as we saw in the first chapter and in my studies, the vmPFC is part of a network mainly composed of the ventral striatum and the posterior cingulate cortex. It is also strongly connected to the hippocampus and to its close OFC neighbor. Therefore, it appears as a central hub of subjective value processing and it makes it of high interest for us but how it is distinguishable from its partners is what I am going to address in the following section.

### *OFC & vmPFC*

We found a strong parametric coding of value in the high gamma activity of the OFC in our iEEG study. This result is in direct contradiction with what has been established in fMRI studies suggesting that OFC codes value in an identity-specific representation (Howard et al., 2015) (i.e. in a non-parametric manner). As suggested in the discussion of the methods, we could suppose that OFC activity linearly linked to value is not seen in fMRI because of the involvement of lower frequencies in value coding. But this is speculative and this question should be directly assessed. Thus, our study brings new results regarding the OFC and an inclusion of the OFC in the Brain Valuation System could be considered given the similarities observed in the iEEG study between the vmPFC and the OFC. As mentioned in the discussion of this article, more analysis would be needed to better characterize differences between those two regions.

### *Ventral striatum & vmPFC*

In our studies, we did not focus on the striatum ventral despite the fact that it is a central part of the Brain Valuation System. In our fMRI study, the ventral striatum seemed to have a similar activity to the vmPFC in both the baseline ongoing and evoked activity. Hover, our GLM of interest[20] failed to reveal a significant difference in this area despite a trend in the same direction as in the vmPFC. Given this negative result, we did not focus on it. In the iEEG study, the technical reason is because it was not covered by our iEEG dataset. Indeed, it is rare that the striatum is the epileptic focus thus it is not targeted by neurosurgeons when patients are implanted. Thus, we were not able to investigate the dynamics of this area.

However, during my PhD, I had the opportunity to get access to OCD patients[21] who were drug resistant and implanted in either the striatum or the subthalamic nucleus. We recorded brain activity of these patients while they were performing rating and choice tasks. Data are currently being analyzed but we were able to see that during task in which patients had to indicate whether they would be ready to exert a specific effort in exchange of a specific reward; reward and effort dimension were both integrated in the beta frequency of the ventral striatum (Figure 41). However, in this same patient, we were not able to find subjective value coding during neither likeability ratings nor binary choices. This preliminary result is promising since it specifies how

---

[20] The one allowing disentangling chosen and default option values.

[21] Obsessive-Compulsive Disorder

subjective value might be represented in the ventral striatum. If this result is replicated in other patients, we might be able to bring some precisions on the role of the ventral striatum in subjective valuation.

Thus, striatum ventral seems to be in close relationship with the vmPFC but subtle differences still needs to be investigated.

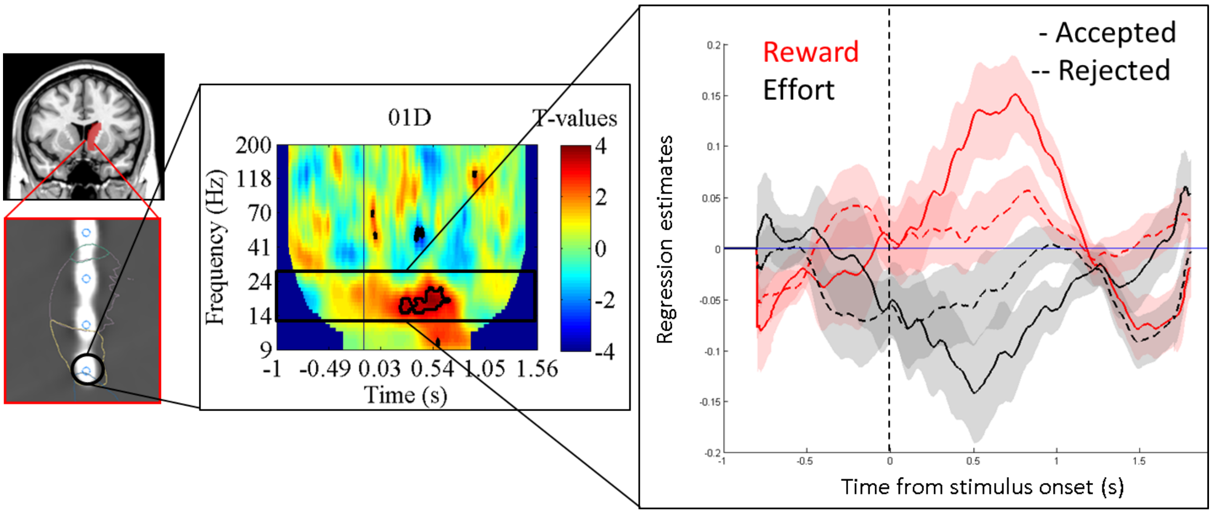


Figure 41 – Preliminary results on ventral striatal dynamics

**Left**: OCD patient is implanted with an electrode along the caudate (red). **Middle**: We regressed the time frequency of the evoked response of the ventral striatum against the utility (Effort-Reward) and found a significant cluster in the beta band around 500 ms after the stimulus onset. **Right**: When plotting the regression estimates of reward and effort value separately for accepted and rejected response, we found that reward were encoded positively and effort negatively in the ventral striatal beta band. The effect was stronger when the response was 'accepted'.

## *Posterior Cingulate Cortex & vmPFC*

For the same reasons as for the ventral striatum, we did not focus on the Posterior Cingulate Cortex (PCC) in the iEEG study. However, interestingly, we observed in the fMRI study that the PCC was responding to the post-decisional frame ($V_{chosen} – V_{unchosen}$) contrary to the vmPFC which was responding to the default frame. The PCC seemed to have a similar activity to parietal lobules (See in the fMRI article: Figure 4 – Supplemental file 1 – Panel C). This result is intriguing since it has been proposed that the PCC and parietal cortex would process sensory signals to recognize items and that its connectivity with the vmPFC would allow computing the subjective value (Polanía et al., 2015). However, our result is not supporting this claim. One potential reconciling explanation could be an attentional effect which would be confounded with the choice effect. However our data do not allow testing this hypothesis and a supplementary experiment would be needed to assess this specific question.

We saw in the last section that in my studies, the focus was done mainly on the vmPFC. This is a strong limitation when one is considering brain networks to implement complex processes such as decision-making. I will discuss this limitation among others in the following part of this chapter.

## 3. Acknowledged limits & open remarks

### A. Focus on the Brain Valuation System

The vmPFC can be considered as an *evaluator* communicating with its partners to construct values. Consequently, a *comparator* computing the choice itself must exist. In our studies, and especially in the fMRI study, we did not address this question directly even if we could have further invested the BOLD activity in dACC and dlPFC since we saw they could be candidates for the comparison process. It is still unclear how the comparison process is done in the brain (see the debates presented page 109). Thanks to our big iEEG data set, we will be able to address this question and particularly to investigate how information is communicated between those areas if they really implement the comparison process. I believe that connectivity analysis and/or phase coupling analysis would provide valuable answers or at least insights on the neural mechanisms of decision-making.

Another limitation is the strong use of our own priors on brain areas and the risk of reverse inference in our manner to interpret the data. Indeed, as an example, if I take the map of the figure 4 – supplementary figure 1 in the fMRI study in which the contrast '$V_{chosen}$-$V_{unchosen}$' is depicted, we indeed observed a part of the vmPFC but the stronger regression estimates are actually observed in the bilateral parietal lobule. Thus, without any prior, if we ask where in the brain the choice value is computed, the response would naturally be: '*mostly in the parietal lobule'*. This area has been often observed in this kind of contrast in economic choices but it is rarely discussed. However, one study investigated the contributions of the vmPFC and the posterior parietal cortex in value-based decisions (Jocham et al., 2014) and they proposed that posterior parietal cortex could be involved in fast decisions under time pressure while the vmPFC could code the value of the choices when there is no time pressure. They also proposed that the two systems work in parallel and that it could also explain why patients with lesion to the vmPFC are still able to make choices (but noisier). Thus, this example shows that instead of considering the decision process as a series of events (valuation – comparison – selection), it is much more probable that several systems work in parallel and that other brain areas can compute subjective value[22] besides the Brain Valuation System. Thus, given the numerous evidences toward the Brain Valuation System to code subjective value, we can considered ourselves to be *right* investigating it but we should not lose sight of the other networks to

---

[22] maybe in a more approximatively way given the result of Jocham et al., 2014

understand the global mechanism underlying value-based decision-making.

## B. Focus on ‘positive’ values.

In my work, I rarely mentioned ‘aversive values’. In other words, I focused on the ‘pleasure’ aspect of valuation but not on the ‘pain aspect’. I used rating scales which could allow subjects indicating that they did ‘not like’ the presented item, suggesting a negative value but it is actually a neutral value which is at stake here given that they would not suffer from winning it. Then, in situations in which one would have to evaluate how much an event/action would be painful, an interesting question is to know whether the BVS respond in the same manner than for positive values (the more painful the more deactivated). Some fMRI and electrophysiological studies report an encoding of values in the negative domain in the vmPFC and ventral striatum (Hosokawa et al., 2007; Tom et al., 2007; Plassmann et al., 2010). However, when investigating the effect of negative values on brain activity, one consistent result is the proportional activation of the anterior insula with the amount of punishments such as electric shocks (Seymour et al., 2004) or aversive pictures (Nitschke et al., 2006). Thus, it seems that there is an opponent system to the positive Brain Valuation System and it is not inconsistent with behavioral observation if we refer to the asymmetry observed between gains and losses (Kahneman and Tversky, 1979). It is also intuitive to think that two systems might exist to code the two domains given the biological limitations of the neural code (a neuron cannot fire ‘negatively’). Moreover, we could also suppose that those systems work in parallel with more or less contribution of each of them according to the domain of valuation.

## C. What is value in the brain?

### *How is value built up?*

In this whole thesis, I assessed the concept of subjective value and its neural implementation in the context of its interaction with other variables and in the context of decision-making. However, one central question around subjective value I did not assess is ***how is it constructed?*** An interesting study investigated how the value of a new item can be constructed from the value of its components (Barron et al., 2013) and showed that both vmPFC and hippocampus were involved in constructing the novel value from the value of its attributes. Thus, it seems that it is indeed these regions which allow assigning values by integrating various components and that past experience plays a critical role. However, it is unclear **how** features are integrated or even weighted in a multi-attribute option value such as a food item. We have seen that the OFC has been often found to represent value in a non-parametric manner. If so, no specific pattern of coding has been found and further experiments are needed to decipher the neural code used by the OFC to assign value to features. I believe that recent finding on the neural representation of conceptual knowledge (found to use a hexagonally symmetric code as grid cells in spatial navigation) (Constantinescu et al., 2016) brings enlightening lines of approach to investigate the neural code of subjective value, especially when items to evaluate are multi-dimensional and when it might be useful to map some representation on continuous multi-dimensional spaces.

### *Value as an internal estimate?*

I briefly mentioned in the first chapter that subjective value could be understood not only as a single value but also as an internal distribution with a mean (correlated to the subjective value report) and a variance (anti-correlated to the confidence report). In my work, I always regressed the neural signal with the subjective report without taking into account this definition (for obvious reasons of simplicity). However, integrating both confidence and likeability rating in a function which would provide an approximation of the supposed internal distribution (with a mean and a variance) would probably help characterizing much more precisely the brain activity linked to subjective value processing if value is represented through this theoretical form.

Moreover, if subjective value is an internal distribution, it would be interesting to investigate the precise link between control, valuation and confidence. Indeed, we could suppose that the more we apply control on the estimation of value, the more precise it would be and potentially the more confident we would also be (through a decrease of the variance). Thus, considering value as an internal estimate instead of a single value might also help to understand the processes at stake.

## 4. Conclusion

Inconsistencies and biases have always been observed in the human behavior. Philosophers, economists and scientists have investigated value-based decision-making for many centuries and the advent of neuroimaging techniques have recently brought new insights on the neural mechanisms underlying this process.

The present thesis investigated the neural properties of subjective values through a rare neuroimaging method and confirmed the critical role of the Brain Valuation System in the subjective value processing. Moreover, thanks to the use of a new framework to investigate decision-making, it brought behavioral and neural evidence to a general solution for the neural implementation of decisions based on subjective values. Altogether, the studies conducted during this PhD support the idea that the Brain Valuation System, and in particular the vmPFC, can be considered as an *hedonometer* used to orient and drives our decisions and actions.

The neural implementation of decision-making still remains to be deciphered but I believe this work is a small but important building block in the global understanding of the human brain fundamental functioning.

# References

Abitbol, R., Lebreton, M., Hollard, G., Richmond, B.J., Bouret, S., Pessiglione, M., 2015. Neural Mechanisms Underlying Contextual Dependency of Subjective Values: Converging Evidence from Monkeys and Humans. J. Neurosci. 35, 2308–2320. doi:10.1523/JNEUROSCI.1878-14.2015

Abler, B., Walter, H., Erk, S., Kammerer, H., Spitzer, M., 2006. Prediction error as a linear function of reward probability is coded in human nucleus accumbens. NeuroImage 31, 790–795. doi:10.1016/j.neuroimage.2006.01.001

Aizman, O., Brismar, H., Uhlén, P., Zettergren, E., Levey, A.I., Forssberg, H., Greengard, P., Aperia, A., 2000. Anatomical and physiological evidence for D1 and D2 dopamine receptor colocalization in neostriatal neurons. Nat. Neurosci. 3, 226–230.

Alexander, G.E., Crutcher, M.D., 1990. Functional architecture of basal ganglia circuits: neural substrates of parallel processing. Trends Neurosci. 13, 266–271. doi:10.1016/0166-2236(90)90107-L

Alexander, G.E., DeLong, M.R., Strick, P.L., 1986. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annu. Rev. Neurosci. 9, 357–381.

Amiez, C., Joseph, J.-P., Procyk, E., 2006. Reward encoding in the monkey anterior cingulate cortex. Cereb. Cortex N. Y. N 1991 16, 1040–1055. doi:10.1093/cercor/bhj046

Amiez, C., Joseph, J.-P., Procyk, E., 2005. Anterior cingulate error-related activity is modulated by predicted reward. Eur. J. Neurosci. 21, 3447–3452. doi:10.1111/j.1460-9568.2005.04170.x

Andreoni, J., 1990. Impure Altruism and DOnations to Public Goods : A theory of Warm-Glow Giving. Econ. J. 100, 464–477.

Arieli, A., Sterkin, A., Grinvald, A., Aertsen, A., 1996. Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. Science 273, 1868–1871.

Asch, S.E., 1951. Effects of group pressure upon the modification and distortion of judgments. Groups Leadersh. Men 222–236.

Balleine, B.W., Dickinson, A., 1998. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. Neuropharmacology 37, 407–419. doi:10.1016/S0028-3908(98)00033-1

Barch, D.M., Braver, T.S., Nystrom, L.E., Forman, S.D., Noll, D.C., Cohen, J.D., 1997. Dissociating working memory from task difficulty in human prefrontal cortex. Neuropsychologia 35, 1373–1380.

Barron, H.C., Dolan, R.J., Behrens, T.E.J., 2013. Online evaluation of novel choices by simultaneous representation of multiple memories. Nat. Neurosci. 16, 1492–1498. doi:10.1038/nn.3515

Bartra, O., McGuire, J.T., Kable, J.W., 2013. The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. NeuroImage 76, 412–427. doi:10.1016/j.neuroimage.2013.02.063

Basten, U., Biele, G., Heekeren, H.R., Fiebach, C.J., 2010. How the brain integrates costs and benefits during decision making. Proc. Natl. Acad. Sci. 107, 21767–21772. doi:10.1073/pnas.0908104107

Bechara, A., Damasio, H., Damasio, A.R., Lee, G.P., 1999. Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. J. Neurosci. 19, 5473–5481.

Becker, G.M., DeGroot, M.H., Marschak, J., 1964. Measuring utility by a single-response sequential method. Behav. Sci. 9, 226–232.

Berke, J.D., 2009. Fast oscillations in cortical-striatal networks switch frequency following rewarding events and stimulant drugs. Eur. J. Neurosci. 30, 848–859. doi:10.1111/j.1460-9568.2009.06843.x

Bernoulli, D., 1954. Exposition of a new theory on the measurement of risk. Econom. J. Econom. Soc. 23–36.

Bitzer, S., Park, H., Blankenburg, F., Kiebel, S.J., 2014. Perceptual decision making: drift-diffusion model is equivalent to a Bayesian model. Front. Hum. Neurosci. 8. doi:10.3389/fnhum.2014.00102

Blanchard, T.C., Hayden, B.Y., 2014. Neurons in Dorsal Anterior Cingulate Cortex Signal Postdecisional Variables in a Foraging Task. J. Neurosci. 34, 646–655. doi:10.1523/JNEUROSCI.3151-13.2014

Bogacz, R., Gurney, K., 2007. The basal ganglia and cortex implement optimal decision making between alternative actions. Neural Comput. 19, 442–477.

Boly, M., Balteau, E., Schnakers, C., Degueldre, C., Moonen, G., Luxen, A., Phillips, C., Peigneux, P., Maquet, P., Laureys, S., 2007. Baseline brain activity fluctuations predict somatosensory perception in humans. Proc. Natl. Acad. Sci. 104, 12187–12192.

Bonnelle, V., Veromann, K.-R., Burnett Heyes, S., Lo Sterzo, E., Manohar, S., Husain, M., 2015. Characterization of reward and effort mechanisms in apathy. J. Physiol. Paris 109, 16–26. doi:10.1016/j.jphysparis.2014.04.002

Boorman, E.D., Behrens, T.E., Rushworth, M.F., 2011. Counterfactual Choice and Learning in a Neural Network Centered on Human Lateral Frontopolar Cortex. PLoS Biol. 9, e1001093. doi:10.1371/journal.pbio.1001093

Boorman, E.D., Behrens, T.E.J., Woolrich, M.W., Rushworth, M.F.S., 2009. How Green Is the Grass on the Other Side? Frontopolar Cortex and the Evidence in Favor of Alternative Courses of Action. Neuron 62, 733–743. doi:10.1016/j.neuron.2009.05.014

Boorman, E.D., Rushworth, M.F., Behrens, T.E., 2013. Ventromedial Prefrontal and Anterior Cingulate Cortex Adopt Choice and

Default Reference Frames during Sequential Multi-Alternative Choice. J. Neurosci. 33, 2242–2253. doi:10.1523/JNEUROSCI.3022-12.2013

Botvinick, M.M., 2007. Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. Cogn. Affect. Behav. Neurosci. 7, 356–366.

Botvinick, M.M., Cohen, J.D., Carter, C.S., 2004. Conflict monitoring and anterior cingulate cortex: an update. Trends Cogn. Sci. 8, 539–546. doi:10.1016/j.tics.2004.10.003

Bouret, S., Richmond, B.J., 2010. Ventromedial and Orbital Prefrontal Neurons Differentially Encode Internally and Externally Driven Motivational Values in Monkeys. J. Neurosci. 30, 8591–8601. doi:10.1523/JNEUROSCI.0049-10.2010

Brown, S.D., Heathcote, A., 2008. The simplest complete model of choice response time: Linear ballistic accumulation. Cognit. Psychol. 57, 153–178. doi:10.1016/j.cogpsych.2007.12.002

Budescu, D.V., Wallsten, T.S., Au, W.T., 1997. On the Importance of Random Error in the Study of Probability Judgment. Part II: Applying the Stochastic Judgment Model to Detect Systematic Trends. J. Behav. Decis. Mak. 10, 173–188. doi:10.1002/(SICI)1099-0771(199709)10:3<173::AID-BDM261>3.0.CO;2-6

Busemeyer, J.R., Townsend, J.T., 1993. Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. Psychol. Rev. 100, 432–459.

Butter, C.M., McDonald, J.A., Snyder, D.R., 1969. Orality, Preference Behavior, and Reinforcement Value of Nonfood Object in Monkeys with Orbital Frontal Lesions. Science 164, 1306–1307. doi:10.1126/science.164.3885.1306

Buzsaki, G., 2004. Neuronal Oscillations in Cortical Networks. Science 304, 1926–1929. doi:10.1126/science.1099745

Carlsson, A., 1959. Detection and assay of dopamine. Pharmacol. Rev. 11, 300–304.

Catani, M., Thiebaut de Schotten, M., 2012. Atlas of human brain connections. Oxford Univ. Press, Oxford.

Chib, V.S., Rangel, A., Shimojo, S., O’Doherty, J.P., 2009. Evidence for a Common Representation of Decision Values for Dissimilar Goods in Human Ventromedial Prefrontal Cortex. J. Neurosci. 29, 12315–12320. doi:10.1523/JNEUROSCI.2575-09.2009

Clark, J.J., Hollon, N.G., Phillips, P.E., 2012. Pavlovian valuation systems in learning and decision making. Curr. Opin. Neurobiol. 22, 1054–1061. doi:10.1016/j.conb.2012.06.004

Cléry-Melin, M.-L., Schmidt, L., Lafargue, G., Baup, N., Fossati, P., Pessiglione, M., 2011. Why Don’t You Try Harder? An Investigation of Effort Production in Major Depression. PLoS ONE 6, e23178. doi:10.1371/journal.pone.0023178

Clore, G., Schwarz, N., Conway, M., 1994. Affective causes and consequences of social information processing, in: Handbook of Social Cognition, Second Edition: Volume 1: Basic Processes. pp. 323–417.

Cohen, M.X., Axmacher, N., Lenartz, D., Elger, C.E., Sturm, V., Schlaepfer, T.E., 2009a. Nuclei Accumbens Phase Synchrony Predicts Decision-Making Reversals Following Negative Feedback. J. Neurosci. 29, 7591–7598. doi:10.1523/JNEUROSCI.5335-08.2009

Cohen, M.X., Axmacher, N., Lenartz, D., Elger, C.E., Sturm, V., Schlaepfer, T.E., 2009b. Neuroelectric signatures of reward learning and decision-making in the human nucleus accumbens. Neuropsychopharmacology 34, 1649–1658.

Cohen, M.X., Axmacher, N., Lenartz, D., Elger, C.E., Sturm, V., Schlaepfer, T.E., 2009c. Good vibrations: cross-frequency coupling in the human nucleus accumbens during reward processing. J. Cogn. Neurosci. 21, 875–889.

Constantinescu, A.O., O’Reilly, J.X., Behrens, T.E., 2016. Organizing conceptual knowledge in humans with a gridlike code. Science 352, 1464–1468.

Cunningham, W.A., Johnsen, I.R., Waggoner, A.S., 2011. Orbitofrontal cortex provides cross-modal valuation of self-generated stimuli. Soc. Cogn. Affect. Neurosci. 6, 286–293. doi:10.1093/scan/nsq038

Daw, N.D., Niv, Y., Dayan, P., 2005. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat. Neurosci. 8, 1704–1711. doi:10.1038/nn1560

Daw, N.D., O’Doherty, J.P., Dayan, P., Seymour, B., Dolan, R.J., 2006. Cortical substrates for exploratory decisions in humans. Nature 441, 876–879. doi:10.1038/nature04766

Day, J.J., Carelli, R.M., 2007. The Nucleus Accumbens and Pavlovian Reward Learning. The Neuroscientist 13, 148–159. doi:10.1177/1073858406295854

Day, J.J., Wheeler, R.A., Roitman, M.F., Carelli, R.M., 2006. Nucleus accumbens neurons encode Pavlovian approach behaviors: evidence from an autoshaping paradigm. Eur. J. Neurosci. 23, 1341–1351. doi:10.1111/j.1460-9568.2006.04654.x

Dayan, P., Niv, Y., Seymour, B., D. Daw, N., 2006. The misbehavior of value and the discipline of the will. Neural Netw. 19, 1153–1160. doi:10.1016/j.neunet.2006.03.002

de Araujo, I.E., Rolls, E.T., Velazco, M.I., Margot, C., Cayeux, I., 2005. Cognitive Modulation of Olfactory Processing. Neuron 46, 671–679. doi:10.1016/j.neuron.2005.04.021

De Martino, B., Fleming, S.M., Garrett, N., Dolan, R.J., 2013. Confidence in value-based choice. Nat. Neurosci. 16, 105–110.

doi:10.1038/nn.3279

DeShazo, J.R., Fermo, G., 2002. Designing Choice Sets for Stated Preference Methods: The Effects of Complexity on Choice Consistency. J. Environ. Econ. Manag. 44, 123–143. doi:10.1006/jeem.2001.1199

Dickinson, A., Balleine, B., 1994. Motivational control of goal-directed action. Anim. Learn. Behav. 22, 1–18. doi:10.3758/BF03199951

Donnelly, N.A., Holtzman, T., Rich, P.D., Nevado-Holgado, A.J., Fernando, A.B.P., Van Dijck, G., Holzhammer, T., Paul, O., Ruther, P., Paulsen, O., Robbins, T.W., Dalley, J.W., 2014. Oscillatory Activity in the Medial Prefrontal Cortex and Nucleus Accumbens Correlates with Impulsivity and Reward Outcome. PLoS ONE 9, e111300. doi:10.1371/journal.pone.0111300

Drugowitsch, J., Moreno-Bote, R., Churchland, A.K., Shadlen, M.N., Pouget, A., 2012. The Cost of Accumulating Evidence in Perceptual Decision Making. J. Neurosci. 32, 3612–3628. doi:10.1523/JNEUROSCI.4010-11.2012

Eldar, E., Rutledge, R.B., Dolan, R.J., Niv, Y., 2016. Mood as Representation of Momentum. Trends Cogn. Sci. 20, 15–24. doi:10.1016/j.tics.2015.07.010

Elliott, R., 2000. Dissociable Functions in the Medial and Lateral Orbitofrontal Cortex: Evidence from Human Neuroimaging Studies. Cereb. Cortex 10, 308–317. doi:10.1093/cercor/10.3.308

Engel, A.K., Fries, P., Singer, W., 2001. Dynamic predictions: oscillations and synchrony in top-down processing. Nat. Rev. Neurosci. 2, 704–716. doi:10.1038/35094565

Erev, I., Wallsten, T.S., Budescu, D.V., 1994. Simultaneous over- and underconfidence: The role of error in judgment processes. Psychol. Rev. 101, 519–527. doi:10.1037/0033-295X.101.3.519

Evarts, E.V., Thach, W.T., 1969. Motor mechanisms of the CNS: cerebrocerebellar interrelations. Annu. Rev. Physiol. 31, 451–498.

Fehr, E., Fischbacher, U., 2004. Third-party punishment and social norms. Evol. Hum. Behav. 25, 63–87.

Fellows, L.K., 2007. The role of orbitofrontal cortex in decision making: a component process account. Ann. N. Y. Acad. Sci. 1121, 421–430. doi:10.1196/annals.1401.023

Festinger, L., 1943. Studies in decision: I. Decision-time, relative frequency of judgment and subjective confidence as related to physical stimulus difference. J. Exp. Psychol. 32, 291.

Figner, B., Knoch, D., Johnson, E.J., Krosch, A.R., Lisanby, S.H., Fehr, E., Weber, E.U., 2010. Lateral prefrontal cortex and self-control in intertemporal choice. Nat. Neurosci. 13, 538–539. doi:10.1038/nn.2516

Fiorillo, C.D., 2003. Discrete Coding of Reward Probability and Uncertainty by Dopamine Neurons. Science 299, 1898–1902.

doi:10.1126/science.1077349

FitzGerald, T.H.B., Seymour, B., Dolan, R.J., 2009. The Role of Human Orbitofrontal Cortex in Value Comparison for Incommensurable Objects. J. Neurosci. 29, 8388–8395. doi:10.1523/JNEUROSCI.0717-09.2009

Fleming, S.M., Dolan, R.J., 2012. The neural basis of metacognitive ability. Philos. Trans. R. Soc. B Biol. Sci. 367, 1338–1349. doi:10.1098/rstb.2011.0417

Fox, M.D., Snyder, A.Z., Zacks, J.M., Raichle, M.E., 2006. Coherent spontaneous activity accounts for trial-to-trial variability in human evoked brain responses. Nat. Neurosci. 9, 23–25. doi:10.1038/nn1616

Frank, M.J., 2005. When and when not to use your subthalamic nucleus: Lessons from a computational model of the basal ganglia. Presented at the International Workshop on Modelling Natural Action Selection, Edinburgh., pp. 53–60.

Frederick, S., Loewenstein, G., O'donoghue, T., 2002. Time discounting and time preference: A critical review. J. Econ. Lit. 40, 351–401.

Fries, P., 2015. Rhythms for Cognition: Communication through Coherence. Neuron 88, 220–235. doi:10.1016/j.neuron.2015.09.034

Fries, P., 2005. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. Trends Cogn. Sci. 9, 474–480. doi:10.1016/j.tics.2005.08.011

Gendolla, G.H.E., 2000. On the impact of mood on behavior: An integrative theory and a review. Rev. Gen. Psychol. 4, 378–408. doi:10.1037/1089-2680.4.4.378

Glimcher, P., 2014. Understanding the Hows and Whys of Decision-Making: From Expected Utility to Divisive Normalization. Cold Spring Harb. Symp. Quant. Biol. 79, 169–176. doi:10.1101/sqb.2014.79.024778

Gluth, S., Sommer, T., Rieskamp, J., Büchel, C., 2015. Effective Connectivity between Hippocampus and Ventromedial Prefrontal Cortex Controls Preferential Choices from Memory. Neuron 86, 1078–1090. doi:10.1016/j.neuron.2015.04.023

Gold, J.I., Shadlen, M.N., 2007. The Neural Basis of Decision Making. Annu. Rev. Neurosci. 30, 535–574. doi:10.1146/annurev.neuro.29.051605.113038

Goldin, P.R., McRae, K., Ramel, W., Gross, J.J., 2008. The Neural Bases of Emotion Regulation: Reappraisal and Suppression of Negative Emotion. Biol. Psychiatry 63, 577–586. doi:10.1016/j.biopsych.2007.05.031

Grabenhorst, F., Rolls, E.T., 2008. Selective attention to affective value alters how the brain processes taste stimuli. Eur. J. Neurosci. 27, 723–729. doi:10.1111/j.1460-9568.2008.06033.x

Grace, A.A., 1991. Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: a hypothesis for the etiology of schizophrenia. Neuroscience 41, 1–24.

Grasman, R., Wagenmakers, E.-J., van der Maas, H., 2007. EZ2: An extension of the EZ-diffusion model for Response Time and Accuracy.

Grasman, R.P.P.P., Wagenmakers, E.-J., van der Maas, H.L.J., 2009. On the mean and variance of response times under the diffusion model with an application to parameter estimation. J. Math. Psychol. 53, 55–68. doi:10.1016/j.jmp.2009.01.006

Greicius, M.D., Supekar, K., Menon, V., Dougherty, R.F., 2009. Resting-State Functional Connectivity Reflects Structural Connectivity in the Default Mode Network. Cereb. Cortex 19, 72–78. doi:10.1093/cercor/bhn059

Grèzes, J., Armony, J.., Rowe, J., Passingham, R.., 2003. Activations related to “mirror” and “canonical” neurones in the human brain: an fMRI study. NeuroImage 18, 928–937. doi:10.1016/S1053-8119(03)00042-9

Grueschow, M., Polania, R., Hare, T.A., Ruff, C.C., 2015. Automatic versus Choice-Dependent Value Representations in the Human Brain. Neuron. doi:10.1016/j.neuron.2014.12.054

Güth, W., Schmittberger, R., Schwarze, B., 1982. An experimental analysis of ultimatum bargaining. J. Econ. Behav. Organ. 3, 367–388. doi:10.1016/0167-2681(82)90011-7

Hare, T.A., Camerer, C.F., Rangel, A., 2009. Self-Control in Decision-Making Involves Modulation of the vmPFC Valuation System. Science 324, 646–648. doi:10.1126/science.1168450

Hare, T.A., O’Doherty, J., Camerer, C.F., Schultz, W., Rangel, A., 2008. Dissociating the Role of the Orbitofrontal Cortex and the Striatum in the Computation of Goal Values and Prediction Errors. J. Neurosci. 28, 5623–5630. doi:10.1523/JNEUROSCI.1309-08.2008

Hare, T.A., Schultz, W., Camerer, C.F., O’Doherty, J.P., Rangel, A., 2011. Transformation of stimulus value signals into motor commands during simple choice. Proc. Natl. Acad. Sci. 108, 18120–18125. doi:10.1073/pnas.1109322108

Hayden, B.Y., Heilbronner, S.R., Pearson, J.M., Platt, M.L., 2011. Surprise Signals in Anterior Cingulate Cortex: Neuronal Encoding of Unsigned Reward Prediction Errors Driving Adjustment in Behavior. J. Neurosci. 31, 4178–4187. doi:10.1523/JNEUROSCI.4652-10.2011

Hayden, B.Y., Pearson, J.M., Platt, M.L., 2011. Neuronal basis of sequential foraging decisions in a patchy environment. Nat. Neurosci. 14, 933–939. doi:10.1038/nn.2856

Hayden, B.Y., Pearson, J.M., Platt, M.L., 2009. Fictive Reward Signals in the Anterior Cingulate Cortex. Science 324, 948–950. doi:10.1126/science.1168488

Heath, R.A., 1984. Random-walk and accumulator models of psychophysical discrimination: a critical evaluation. Perception 13, 57–65.

Heeger, D.J., 1992. Normalization of cell responses in cat striate cortex. Vis. Neurosci. 9, 181–197.

Henri-Bhargava, A., Simioni, A., Fellows, L.K., 2012. Ventromedial frontal lobe damage disrupts the accuracy, but not the speed, of value-based preference judgments. Neuropsychologia 50, 1536–1542. doi:10.1016/j.neuropsychologia.2012.03.006

Hershberger, W.A., 1986. An approach through the looking-glass. Anim. Learn. Behav. 14, 443–451.

Hesselmann, G., Kell, C.A., Eger, E., Kleinschmidt, A., 2008. Spontaneous local variations in ongoing neural activity bias perceptual decisions. Proc. Natl. Acad. Sci. 105, 10984–10989.

Horst, N.K., Laubach, M., 2013. Reward-related activity in the medial prefrontal cortex is driven by consumption. Front. Neurosci. 7. doi:10.3389/fnins.2013.00056

Hosokawa, T., Kato, K., Inoue, M., Mikami, A., 2007. Neurons in the macaque orbitofrontal cortex code relative preference of both rewarding and aversive outcomes. Neurosci. Res. 57, 434–445.

Howard, J.D., Gottfried, J.A., Tobler, P.N., Kahnt, T., 2015. Identity-specific coding of future rewards in the human orbitofrontal cortex. Proc. Natl. Acad. Sci. 112, 5195–5200. doi:10.1073/pnas.1503550112

Huber, J., Payne, J.W., Puto, C., 1982. Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. J. Consum. Res. 9, 90–98.

Huettel, S.A., Song, A.W., McCarthy, G., 2004. Functional magnetic resonance imaging. Sinauer Associates Sunderland.

Huijgen, J., Samson, S., 2015. The hippocampus: A central node in a large-scale brain network for memory. Rev. Neurol. (Paris) 171, 204–216. doi:10.1016/j.neurol.2015.01.557

Hunt, L.T., Behrens, T.E., Hosokawa, T., Wallis, J.D., Kennerley, S.W., 2015. Capturing the temporal evolution of choice across prefrontal cortex. eLife 4. doi:10.7554/eLife.11945

Hunt, L.T., Kolling, N., Soltani, A., Woolrich, M.W., Rushworth, M.F., Behrens, T.E., 2012. Mechanisms underlying cortical activity during value-guided choice. Nat. Neurosci. 15, 470-S3. doi:10.1038/nn.3017

Iversen, S.D., Mishkin, M., 1970. Perseverative interference in monkeys following selective lesions of the inferior prefrontal convexity. Exp. Brain Res. 11, 376–386.

Izuma, K., Saito, D.N., Sadato, N., 2008. Processing of Social and Monetary Rewards in the Human Striatum. Neuron 58, 284–294. doi:10.1016/j.neuron.2008.03.020

Jevons, W.S., 1879. Political economy.

Jevons, W.S., 1871. The Theory of Political Economy, London: Macmillan. ed.

Jocham, G., Furlong, P.M., Kröger, I.L., Kahn, M.C., Hunt, L.T., Behrens, T.E.J., 2014. Dissociable contributions of ventromedial prefrontal and posterior parietal cortex to value-guided choice. NeuroImage 100, 498–506. doi:10.1016/j.neuroimage.2014.06.005

Jones, B., Mishkin, M., 1972. Limbic lesions and the problem of stimulus—reinforcement associations. Exp. Neurol. 36, 362–377.

Jung, J., Bayle, D., Jerbi, K., Vidal, J.R., Hénaff, M.-A., Ossandon, T., Bertrand, O., Mauguière, F., Lachaux, J.-P., 2011. Intracerebral gamma modulations reveal interaction between emotional processing and action outcome evaluation in the human orbitofrontal cortex. Int. J. Psychophysiol. 79, 64–72. doi:10.1016/j.ijpsycho.2010.09.014

Kahneman, D., Knetsch, J.L., Thaler, R.H., 1986. Fairness and the Assumptions of Economics. J. Bus. 59, S285. doi:10.1086/296367

Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decision under risk. Econom. J. Econom. Soc. 263–291.

Kalenscher, T., Lansink, C.S., Lankelma, J.V., Pennartz, C.M.A., 2010. Reward-Associated Gamma Oscillations in Ventral Striatum Are Regionally Differentiated and Modulate Local Firing Activity. J. Neurophysiol. 103, 1658–1672. doi:10.1152/jn.00432.2009

Kalenscher, T., Pennartz, C.M.A., 2008. Is a bird in the hand worth two in the future? The neuroeconomics of intertemporal decision-making. Prog. Neurobiol. 84, 284–315. doi:10.1016/j.pneurobio.2007.11.004

Kemp, J.M., Powell, T.P.S., 1971. The Connexions of the Striatum and Globus Pallidus: Synthesis and Speculation. Philos. Trans. R. Soc. Lond. B Biol. Sci. 262, 441–457. doi:10.1098/rstb.1971.0106

Kennerley, S.W., Behrens, T.E.J., Wallis, J.D., 2011. Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. Nat. Neurosci. 14, 1581–1589. doi:10.1038/nn.2961

Kennerley, S.W., Wallis, J.D., 2009. Evaluating choices by single neurons in the frontal lobe: outcome value encoded across multiple decision variables. Eur. J. Neurosci. 29, 2061–2073. doi:10.1111/j.1460-9568.2009.06743.x

Kepecs, A., Mainen, Z.F., 2012. A computational framework for the study of confidence in humans and animals. Philos. Trans. R. Soc. B Biol. Sci. 367, 1322–1337. doi:10.1098/rstb.2012.0037

Khamassi, M., Mulder, A.B., Tabuchi, E., Douchamps, V., Wiener, S.I., 2008. Anticipatory reward signals in ventral striatal neurons of behaving rats. Eur. J. Neurosci. 28, 1849–1866. doi:10.1111/j.1460-9568.2008.06480.x

Kiani, R., Shadlen, M.N., 2009. Representation of confidence associated with a decision by neurons in the parietal cortex. science 324, 759–764.

Kilner, J.M., Mattout, J., Henson, R., Friston, K.J., 2005. Hemodynamic correlates of EEG: A heuristic. NeuroImage 28, 280–286. doi:10.1016/j.neuroimage.2005.06.008

Kim, H., Adolphs, R., O'Doherty, J.P., Shimojo, S., 2007. Temporal isolation of neural processes underlying face preference decisions. Proc. Natl. Acad. Sci. 104, 18253–18258.

Kinchla, R.A., Smyzer, F., 1967. A diffusion model of perceptual memory. Percept. Psychophys. 2, 219–229. doi:10.3758/BF03212471

Knutson, B., Fong, G.W., Adams, C.M., Varner, J.L., Hommer, D., 2001. Dissociation of reward anticipation and outcome with event-related fMRI. Neuroreport 12, 3683–3687.

Kobayashi, S., Pinto de Carvalho, O., Schultz, W., 2010. Adaptation of Reward Sensitivity in Orbitofrontal Neurons. J. Neurosci. 30, 534–544. doi:10.1523/JNEUROSCI.4009-09.2010

Kobayashi, S., Schultz, W., 2008. Influence of Reward Delays on Responses of Dopamine Neurons. J. Neurosci. 28, 7837–7846. doi:10.1523/JNEUROSCI.1600-08.2008

Kolling, N., Behrens, T., Wittmann, M., Rushworth, M., 2016. Multiple signals in anterior cingulate cortex. Curr. Opin. Neurobiol. 37, 36–43. doi:10.1016/j.conb.2015.12.007

Kolling, N., Behrens, T.E.J., Mars, R.B., Rushworth, M.F.S., 2012. Neural Mechanisms of Foraging. Science 336, 95–98. doi:10.1126/science.1216930

Koszegi, B., Rabin, M., 2006. A model of reference-dependent preferences. Q. J. Econ. 1133–1165.

Kouneiher, F., Charron, S., Koechlin, E., 2009. Motivation and cognitive control in the human prefrontal cortex. Nat. Neurosci. 12, 939–945. doi:10.1038/nn.2321

Krajbich, I., Armel, C., Rangel, A., 2010. Visual fixations and the computation and comparison of value in simple choice. Nat. Neurosci. 13, 1292–1298. doi:10.1038/nn.2635

Krajbich, I., Lu, D., Camerer, C., Rangel, A., 2012. The Attentional Drift-Diffusion Model Extends to Simple Purchasing Decisions. Front. Psychol. 3. doi:10.3389/fpsyg.2012.00193

Krajbich, I., Rangel, A., 2011. Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. Proc. Natl. Acad. Sci. 108, 13852–13857. doi:10.1073/pnas.1101328108

Kringelbach, M., 2004. The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. Prog. Neurobiol. 72, 341–372. doi:10.1016/j.pneurobio.2004.03.006

Kringelbach, M.L., 2005. The human orbitofrontal cortex: linking reward to hedonic experience. Nat. Rev. Neurosci. 6, 691–702. doi:10.1038/nrn1747

LaBerge, D., 1962. A recruitment theory of simple behavior. Psychometrika 27, 375–396.

Latimer, K.W., Yates, J.L., Meister, M.L.R., Huk, A.C., Pillow, J.W., 2015. Single-trial spike trains in parietal cortex reveal discrete steps during decision-making. Science 349, 184–187. doi:10.1126/science.aaa4056

Lebreton, M., Abitbol, R., Daunizeau, J., Pessiglione, M., 2015. Automatic integration of confidence in the brain valuation signal. Nat. Neurosci. 18, 1159–1167. doi:10.1038/nn.4064

Lebreton, M., Bertoux, M., Boutet, C., Lehericy, S., Dubois, B., Fossati, P., Pessiglione, M., 2013. A Critical Role for the Hippocampus in the Valuation of Imagined Outcomes. PLoS Biol. 11, e1001684. doi:10.1371/journal.pbio.1001684

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. doi:10.1016/j.neuron.2009.09.040

Lebreton, M., Kawa, S., Forgeot d'Arc, B., Daunizeau, J., Pessiglione, M., 2012. Your Goal Is Mine: Unraveling Mimetic Desires in the Human Brain. J. Neurosci. 32, 7146–7157. doi:10.1523/JNEUROSCI.4821-11.2012

Leech, R., Sharp, D.J., 2013. The role of the posterior cingulate cortex in cognition and disease. Brain 137, 12–32. doi:10.1093/brain/awt162

Lega, B.C., Kahana, M.J., Jaggi, J., Baltuch, G.H., Zaghloul, K., 2011. Neuronal and oscillatory activity during reward processing in the human ventral striatum: NeuroReport 1. doi:10.1097/WNR.0b013e32834b2975

Levy, D.J., Glimcher, P.W., 2012. The root of all value: a neural common currency for choice. Curr. Opin. Neurobiol. doi:10.1016/j.conb.2012.06.001

Levy, D.J., Glimcher, P.W., 2011. Comparing Apples and Oranges: Using Reward-Specific and Reward-General Subjective Value Representation in the Brain. J. Neurosci. 31, 14693–14707. doi:10.1523/JNEUROSCI.2218-11.2011

Levy, I., Lazzaro, S.C., Rutledge, R.B., Glimcher, P.W., 2011. Choice from Non-Choice: Predicting Consumer Preferences from Blood Oxygenation Level-Dependent Signals Obtained during Passive Viewing. J. Neurosci. 31, 118–125. doi:10.1523/JNEUROSCI.3214-10.2011

Li, Y., Vanni-Mercier, G., Isnard, J., Mauguière, F., Dreher, J.-C., 2016. The neural dynamics of reward value and risk coding in the human orbitofrontal cortex. Brain 139, 1295–1309. doi:10.1093/brain/awv409

Lim, S.-L., O'Doherty, J.P., Rangel, A., 2011. The Decision Value Computations in the vmPFC and Striatum Use a Relative Value Code That is Guided by Visual Attention. J. Neurosci. 31, 13214–13223. doi:10.1523/JNEUROSCI.1246-11.2011

Lin, A., Adolphs, R., Rangel, A., 2011. Social and monetary reward learning engage overlapping neural substrates. Soc. Cogn. Affect. Neurosci. 7, 274–281. doi:10.1093/scan/nsr006

Link, S.W., 2003. Confidence and random walk theory, in: Proceeding of the Nineteenth Annual Meeting of the International Society for Psychophysics. Presented at the International Society for Psychophysics, B. B. E. Borg, Stockholm, Sweden.

Link, S.W., Heath, R.A., 1975. A sequential theory of psychological discrimination. Psychometrika 40, 77–105. doi:10.1007/BF02291481

Logothetis, N.K., Pauls, J., Augath, M., Trinath, T., Oeltermann, A., 2001. Neurophysiological investigation of the basis of the fMRI signal. Nature 412, 150–157. doi:10.1038/35084005

Louie, K., Grattan, L.E., Glimcher, P.W., 2011. Reward Value-Based Gain Control: Divisive Normalization in Parietal Cortex. J. Neurosci. 31, 10627–10639. doi:10.1523/JNEUROSCI.1237-11.2011

Louie, K., Khaw, M.W., Glimcher, P.W., 2013. Normalization is a general neural mechanism for context-dependent decision making. Proc. Natl. Acad. Sci. 110, 6139–6144.

Luce, R.D., 1959. On the possible psychophysical laws. Psychol. Rev. 66, 81.

Mackey, S., Petrides, M., 2010. Quantitative demonstration of comparable architectonic areas within the ventromedial and lateral orbital frontal cortex in the human and the macaque monkey brains: Comparable cortical areas in human and macaque. Eur. J. Neurosci. 32, 1940–1950. doi:10.1111/j.1460-9568.2010.07465.x

Magri, C., Schridde, U., Murayama, Y., Panzeri, S., Logothetis, N.K., 2012. The Amplitude and Timing of the BOLD Signal Reflects the Relationship between Local Field Potential Power at Different Frequencies. J. Neurosci. 32, 1395–1407. doi:10.1523/JNEUROSCI.3985-11.2012

McClure, S.M., Berns, G.S., Montague, P.R., 2003. Temporal prediction errors in a passive learning task activate human striatum. Neuron 38, 339–346.

McDiarmid, C.G., Rilling, M.E., 1965. Reinforcement delay and reinforcement rate as determinants of schedule preference. Psychon. Sci. 2, 195–196. doi:10.3758/BF03343402

McMillen, T., Holmes, P., 2006. The dynamics of choice among multiple alternatives. J. Math. Psychol. 50, 30–57. doi:10.1016/j.jmp.2005.10.003

McNamee, D., Rangel, A., O'Doherty, J.P., 2013. Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex. Nat. Neurosci. doi:10.1038/nn.3337

Meyniel, F., Pessiglione, M., 2013. Better Get Back to Work: A Role for Motor Beta Desynchronization in Incentive Motivation. J. Neurosci. 34, 1–9. doi:10.1523/JNEUROSCI.1711-13.2014

Meyniel, F., Sergent, C., Rigoux, L., Daunizeau, J., Pessiglione, M., 2013. Neurocomputational account of how the human brain decides when to have a break. Proc. Natl. Acad. Sci. 110, 2641–2646. doi:10.1073/pnas.1211925110

Meyniel, F., Sigman, M., Mainen, Z.F., 2015. Confidence as Bayesian Probability: From Neural Origins to Behavior. Neuron 88, 78–92. doi:10.1016/j.neuron.2015.09.039

Mill, J.S., 1901. Utilitarianism. Longmans, Green and Company.

Milosavljevic, M., Malmaud, J., Huth, A., Koch, C., Rangel, A., 2010. The Drift Diffusion Model Can Account for the Accuracy and Reaction Time of Value-Based Choices Under High and Low Time Pressure. Available SSRN 1901533.

Milosavljevic, M., Navalpakkam, V., Koch, C., Rangel, A., 2012. Relative visual saliency differences induce sizable bias in consumer choice. J. Consum. Psychol. 22, 67–74. doi:10.1016/j.jcps.2011.10.002

Mitchell, J.F., Sundberg, K.A., Reynolds, J.H., 2007. Differential Attention-Dependent Response Modulation across Cell Classes in Macaque Visual Area V4. Neuron 55, 131–141. doi:10.1016/j.neuron.2007.06.018

Mogenson, G., Jones, D., Yim, C., 1980. From motivation to action: Functional interface between the limbic system and the motor system. Prog. Neurobiol. 14, 69–97. doi:10.1016/0301-0082(80)90018-0

Mora, F., Avrith, D.B., Phillips, A.G., Rolls, E.T., 1979. Effects of satiety on self-stimulation of the orbitofrontal cortex in the rhesus monkey. Neurosci. Lett. 13, 141–145. doi:10.1016/0304-3940(79)90031-4

Mukamel, R., 2005. Coupling Between Neuronal Firing, Field Potentials, and fMRI in Human Auditory Cortex. Science 309, 951–954. doi:10.1126/science.1110913

Nambu, A., Tokuno, H., Takada, M., 2002. Functional significance of the cortico–subthalamo–pallidal “hyperdirect” pathway. Neurosci. Res. 43, 111–117. doi:10.1016/S0168-0102(02)00027-5

Newsome, W.T., Britten, K.H., Movshon, J.A., 1989. Neuronal correlates of a perceptual decision. Nature 341, 52–54. doi:10.1038/341052a0

Nir, Y., Fisch, L., Mukamel, R., Gelbard-Sagiv, H., Arieli, A., Fried, I., Malach, R., 2007. Coupling between Neuronal Firing Rate, Gamma LFP, and BOLD fMRI Is Related to Interneuronal Correlations. Curr. Biol. 17, 1275–1285. doi:10.1016/j.cub.2007.06.066

Nir, Y., Mukamel, R., Dinstein, I., Privman, E., Harel, M., Fisch, L., Gelbard-Sagiv, H., Kipervasser, S., Andelman, F., Neufeld, M.Y., Kramer, U., Arieli, A., Fried, I., Malach, R., 2008. Interhemispheric correlations of slow spontaneous neuronal fluctuations revealed in human sensory cortex. Nat. Neurosci. 11, 1100–1108. doi:10.1038/nn.2177

Nitschke, J.B., Sarinopoulos, I., Mackiewicz, K.L., Schaefer, H.S., Davidson, R.J., 2006. Functional neuroanatomy of aversion and its anticipation. NeuroImage 29, 106–116. doi:10.1016/j.neuroimage.2005.06.068

O'Doherty, J., Kringelbach, M.L., Rolls, E.T., Hornak, J., Andrews, C., 2001. Abstract reward and punishment representations in the human orbitofrontal cortex. Nat. Neurosci. 4, 95–102.

O'Doherty, J.P., 2007. Lights, Camembert, Action! The Role of Human Orbitofrontal Cortex in Encoding Stimuli, Rewards, and Choices. Ann. N. Y. Acad. Sci. 1121, 254–272. doi:10.1196/annals.1401.036

Ohshiro, T., Angelaki, D.E., DeAngelis, G.C., 2011. A normalization model of multisensory integration. Nat. Neurosci. 14, 775–782. doi:10.1038/nn.2815

O'Reilly, J.X., Schuffelgen, U., Cuell, S.F., Behrens, T.E.J., Mars, R.B., Rushworth, M.F.S., 2013. Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. Proc. Natl. Acad. Sci. 110, E3660–E3669. doi:10.1073/pnas.1305373110

Ostlund, S.B., Balleine, B.W., 2007. Orbitofrontal Cortex Mediates Outcome Encoding in Pavlovian But Not Instrumental Conditioning. J. Neurosci. 27, 4819–4825. doi:10.1523/JNEUROSCI.5443-06.2007

Padoa-Schioppa, C., 2013. Neuronal Origins of Choice Variability in Economic Decisions. Neuron 80, 1322–1336. doi:10.1016/j.neuron.2013.09.013

Padoa-Schioppa, C., 2009. Range-Adapting Representation of Economic Value in the Orbitofrontal Cortex. J. Neurosci. 29, 14004–14014. doi:10.1523/JNEUROSCI.3751-09.2009

Padoa-Schioppa, C., Assad, J.A., 2008. The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. Nat. Neurosci. 11, 95–102. doi:10.1038/nn2020

Padoa-Schioppa, C., Assad, J.A., 2006. Neurons in the orbitofrontal cortex encode economic value. Nature 441, 223–226. doi:10.1038/nature04676

Padoa-Schioppa, C., Cai, X., 2011. The orbitofrontal cortex and the computation of subjective value: consolidated concepts and new perspectives: OFC and subjective value. Ann. N. Y. Acad. Sci. 1239, 130–137. doi:10.1111/j.1749-6632.2011.06262.x

Padoa-Schioppa, C., Jandolo, L., Visalberghi, E., 2006. Multi-stage mental process for economic choice in capuchins. Cognition 99, B1–B13. doi:10.1016/j.cognition.2005.04.008

Paglieri, F., 2013. The costs of delay: Waiting versus postponing in intertemporal choice. J. Exp. Anal. Behav. 99, 362–377.

Palminteri, S., Boraud, T., Lafargue, G., Dubois, B., Pessiglione, M., 2009. Brain Hemispheres Selectively Track the Expected Value of Contralateral Options. J. Neurosci. 29, 13465–13472. doi:10.1523/JNEUROSCI.1500-09.2009

Parent, A., 1990. Extrinsic connections of the basal ganglia. Trends Neurosci. 13, 254–258. doi:10.1016/0166-2236(90)90105-J

Parent, A., Hazrati, L.-N., 1995. Functional anatomy of the basal ganglia. I. The cortico-basal ganglia-thalamo-cortical loop. Brain Res. Rev. 20, 91–127. doi:10.1016/0165-0173(94)00007-C

Pavlov, I.P., 1927. Conditioned reflexes. Investig. Physiol. Act. Cereb. Cortex.

Pearson, J.M., Hayden, B.Y., Raghavachari, S., Platt, M.L., 2009. Neurons in Posterior Cingulate Cortex Signal Exploratory Decisions in a Dynamic Multioption Choice Task. Curr. Biol. 19, 1532–1537. doi:10.1016/j.cub.2009.07.048

Pessiglione, M., Schmidt, L., Draganski, B., Kalisch, R., Lau, H., Dolan, R.J., Frith, C.D., 2007. How the Brain Translates Money into Force: A Neuroimaging Study of Subliminal Motivation. Science 316, 904–906. doi:10.1126/science.1140459

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R.J., Frith, C.D., 2006. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature 442, 1042–1045. doi:10.1038/nature05051

Peters, J., Büchel, C., 2010. Neural representations of subjective reward value. Behav. Brain Res. 213, 135–141. doi:10.1016/j.bbr.2010.04.031

Peters, J., Büchel, C., 2010. Episodic Future Thinking Reduces Reward Delay Discounting through an Enhancement of Prefrontal-Mediotemporal Interactions. Neuron 66, 138–148. doi:10.1016/j.neuron.2010.03.026

Peterson, W., Birdsall, T., Fox, W., 1954. The theory of signal detectability. Trans. IRE Prof. Group Inf. Theory 4, 171–212.

Philiastides, M.G., Biele, G., Heekeren, H.R., 2010. A mechanistic account of value computation in the human brain. Proc. Natl. Acad. Sci. 107, 9430–9435. doi:10.1073/pnas.1001732107

Plassmann, H., O’Doherty, J., Rangel, A., 2007. Orbitofrontal Cortex Encodes Willingness to Pay in Everyday Economic Transactions. J. Neurosci. 27, 9984–9988. doi:10.1523/JNEUROSCI.2131-07.2007

Plassmann, H., O’Doherty, J., Shiv, B., Rangel, A., 2008. Marketing actions can modulate neural representations of experienced pleasantness. Proc. Natl. Acad. Sci. 105, 1050–1054.

Plassmann, H., O’Doherty, J.P., Rangel, A., 2010. Appetitive and Aversive Goal Values Are Encoded in the Medial Orbitofrontal Cortex at the Time of Decision Making. J. Neurosci. 30, 10799–10808. doi:10.1523/JNEUROSCI.0788-10.2010

Pleskac, T.J., Busemeyer, J.R., 2010. Two-stage dynamic signal detection: A theory of choice, decision time, and confidence. Psychol. Rev. 117, 864–901. doi:10.1037/a0019737

Polanía, R., Moisa, M., Opitz, A., Grueschow, M., Ruff, C.C., 2015. The precision of value-based choices depends causally on fronto-parietal phase coupling. Nat. Commun. 6, 8090. doi:10.1038/ncomms9090

Pouget, A., Drugowitsch, J., Kepecs, A., 2016. Confidence and certainty: distinct probabilistic quantities for different goals. Nat. Neurosci. 19, 366–374. doi:10.1038/nn.4240

Pratt, J.W., 1964. Risk Aversion in the Small and in the Large. Econometrica 32, 122. doi:10.2307/1913738

Prevost, C., Pessiglione, M., Metereau, E., Clery-Melin, M.-L., Dreher, J.-C., 2010. Separate Valuation Subsystems for Delay and Effort Decision Costs. J. Neurosci. 30, 14080–14090. doi:10.1523/JNEUROSCI.2752-10.2010

Quilodran, R., Rothé, M., Procyk, E., 2008. Behavioral Shifts and Action Valuation in the Anterior Cingulate Cortex. Neuron 57, 314–325. doi:10.1016/j.neuron.2007.11.031

Rabinowitz, N.C., Willmore, B.D.B., Schnupp, J.W.H., King, A.J., 2011. Contrast Gain Control in Auditory Cortex. Neuron 70, 1178–1191. doi:10.1016/j.neuron.2011.04.030

Rangel, A., Camerer, C., Montague, P.R., 2008. A framework for studying the neurobiology of value-based decision making. Nat. Rev. Neurosci. 9, 545–556. doi:10.1038/nrn2357

Rangel, A., Hare, T., 2010. Neural computations associated with goal-directed choice. Curr. Opin. Neurobiol. 20, 262–270. doi:10.1016/j.conb.2010.03.001

Ratcliff, R., 1978. A theory of memory retrieval. Psychol. Rev. 85, 59–108. doi:10.1037/0033-295X.85.2.59

Ratcliff, R., McKoon, G., 2008. The diffusion decision model: theory and data for two-choice decision tasks. Neural Comput. 20, 873–922.

Ray, S., Crone, N.E., Niebur, E., Franaszczuk, P.J., Hsiao, S.S., 2008a. Neural Correlates of High-Gamma Oscillations (60-200 Hz) in Macaque Local Field Potentials and Their Potential Implications in Electrocorticography. J. Neurosci. 28, 11526–11536. doi:10.1523/JNEUROSCI.2848-08.2008

Ray, S., Hsiao, S.S., Crone, N.E., Franaszczuk, P.J., Niebur, E., 2008b. Effect of Stimulus Intensity on the Spike-Local Field Potential Relationship in the Secondary Somatosensory Cortex. J. Neurosci. 28, 7334–7343. doi:10.1523/JNEUROSCI.1588-08.2008

Rescorla, R.A., Wagner, A.R., others, 1972. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. Class. Cond. II Curr. Res. Theory 2, 64–99.

Rich, E.L., Wallis, J.D., 2014. Medial-lateral Organization of the

Orbitofrontal Cortex. J. Cogn. Neurosci. 26, 1347–1362. doi:10.1162/jocn_a_00573

Roesch, M.R., 2004. Neuronal Activity Related to Reward Value and Motivation in Primate Frontal Cortex. Science 304, 307–310. doi:10.1126/science.1093223

Roitman, J.D., Shadlen, M.N., 2002. Response of Neurons in the Lateral Intraparietal Area during a Combined Visual Discrimination Reaction Time Task. J. Neurosci. 22, 9475–9489.

Rolls, E.T., Burton, M.J., Mora, F., 1980. Neurophysiological analysis of brain-stimulation reward in the monkey. Brain Res. 194.

Rolls, E.T., Sienkiewicz, Z.J., Yaxley, S., 1989. Hunger modulates the responses to gustatory stimuli of single neurons in the caudolateral orbitofrontal cortex of the macaque monkey. Eur. J. Neurosci. 1, 53–60.

Rosenkilde, C.E., Bauer, R.H., Fuster, J.M., 1981. Single cell activity in ventral prefrontal cortex of behaving monkeys. Brain Res. 209, 375–394.

Rudebeck, P.H., Behrens, T.E., Kennerley, S.W., Baxter, M.G., Buckley, M.J., Walton, M.E., Rushworth, M.F.S., 2008. Frontal Cortex Subregions Play Distinct Roles in Choices between Actions and Stimuli. J. Neurosci. 28, 13775–13785. doi:10.1523/JNEUROSCI.3541-08.2008

Rudorf, S., Hare, T.A., 2014. Interactions between Dorsolateral and Ventromedial Prefrontal Cortex Underlie Context-Dependent Stimulus Valuation in Goal-Directed Choice. J. Neurosci. 34, 15988–15996. doi:10.1523/JNEUROSCI.3192-14.2014

Ruff, C.C., Fehr, E., 2014. The neurobiology of rewards and values in social decision making. Nat. Rev. Neurosci. 15, 549–562. doi:10.1038/nrn3776

Rushworth, M.F., Kolling, N., Sallet, J., Mars, R.B., 2012. Valuation and decision-making in frontal cortex: one or many serial or parallel systems? Curr. Opin. Neurobiol. 22, 946–955. doi:10.1016/j.conb.2012.04.011

Rushworth, M.F.S., Noonan, M.P., Boorman, E.D., Walton, M.E., Behrens, T.E., 2011. Frontal Cortex and Reward-Guided Learning and Decision-Making. Neuron 70, 1054–1069. doi:10.1016/j.neuron.2011.05.014

Sadaghiani, S., Hesselmann, G., Kleinschmidt, A., 2009. Distributed and Antagonistic Contributions of Ongoing Activity Fluctuations to Auditory Stimulus Detection. J. Neurosci. 29, 13410–13417. doi:10.1523/JNEUROSCI.2592-09.2009

Sallet, J., Quilodran, R., Rothé, M., Vezoli, J., Joseph, J.-P., Procyk, E., 2007. Expectations, gains, and losses in the anterior cingulate cortex. Cogn. Affect. Behav. Neurosci. 7, 327–336.

Samuelson, P.A., 1937. A note on measurement of utility. Rev. Econ. Stud.

4, 155–161.

Sato, F., Lavallee, P., Levesque, M., Parent, A., 2000. Single-axon tracing study of neurons of the external segment of the globus pallidus in primate. J. Comp. Neurol. 417, 17–31. doi:10.1002/(SICI)1096-9861(20000131)417:1<17::AID-CNE2>3.0.CO;2-I

Savage, L., 1954. The Foundations of Statistics, New York: Wiley. ed.

Schmidt, L., Clery-Melin, M.-L., Lafargue, G., Valabregue, R., Fossati, P., Dubois, B., Pessiglione, M., 2009. Get Aroused and Be Stronger: Emotional Facilitation of Physical Effort in the Human Brain. J. Neurosci. 29, 9450–9457. doi:10.1523/JNEUROSCI.1951-09.2009

Schmidt, L., d'Arc, B.F., Lafargue, G., Galanaud, D., Czernecki, V., Grabli, D., Schupbach, M., Hartmann, A., Levy, R., Dubois, B., Pessiglione, M., 2007. Disconnecting force from money: effects of basal ganglia damage on incentive motivation. Brain 131, 1303–1310. doi:10.1093/brain/awn045

Schmidt, L., Lebreton, M., Cléry-Melin, M.-L., Daunizeau, J., Pessiglione, M., 2012. Neural Mechanisms Underlying Motivation of Mental Versus Physical Effort. PLoS Biol. 10, e1001266. doi:10.1371/journal.pbio.1001266

Schnyer, D.M., Nicholls, L., Verfaellie, M., 2005. The role of VMPC in metamemorial judgments of content retrievability. J. Cogn. Neurosci. 17, 832–846. doi:10.1162/0898929053747694

Schultz, W., 2016a. Dopamine reward prediction-error signalling: a two-component response. Nat. Rev. Neurosci.

Schultz, W., 2016b. Reward functions of the basal ganglia. J. Neural Transm. 123, 679–693. doi:10.1007/s00702-016-1510-0

Schultz, W., 2000. Multiple reward signals in the brain. Nat. Rev. Neurosci. 1, 199–207.

Schultz, W., 1998. Predictive reward signal of dopamine neurons. J. Neurophysiol. 80, 1–27.

Schultz, W., 1997. Dopamine neurons and their role in reward mechanisms. Curr. Opin. Neurobiol. 7, 191–197. doi:10.1016/S0959-4388(97)80007-4

Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. Science 275, 1593–1599.

Sescousse, G., Caldú, X., Segura, B., Dreher, J.-C., 2013. Processing of primary and secondary rewards: A quantitative meta-analysis and review of human functional neuroimaging studies. Neurosci. Biobehav. Rev. 37, 681–696. doi:10.1016/j.neubiorev.2013.02.002

Sescousse, G., Redoute, J., Dreher, J.-C., 2010. The Architecture of Reward Value Coding in the Human Orbitofrontal Cortex. J. Neurosci. 30, 13095–13104. doi:10.1523/JNEUROSCI.3501-10.2010

Seymour, B., O'Doherty, J.P., Dayan, P., Koltzenburg, M., Jones, A.K., Dolan, R.J., Friston, K.J., Frackowiak, R.S., 2004. Temporal difference models describe higher-order learning in humans.

Nature 429, 664–667. doi:10.1038/nature02581

Shenhav, A., Botvinick, M., 2015. Uncovering a Missing Link in Anterior Cingulate Research. Neuron 85, 455–457. doi:10.1016/j.neuron.2015.01.020

Shenhav, A., Botvinick, M.M., Cohen, J.D., 2013. The Expected Value of Control: An Integrative Theory of Anterior Cingulate Cortex Function. Neuron 79, 217–240. doi:10.1016/j.neuron.2013.07.007

Shenhav, A., Straccia, M.A., Cohen, J.D., Botvinick, M.M., 2014. Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value. Nat. Neurosci. 17, 1249–1254. doi:10.1038/nn.3771

Sidgwick, H., 1907. The methods of ethics. Hackett Publishing.

Sidman, M., Brady, J.V., Boren, J.J., Conrad, D.G., Schulman, A., 1955. Reward schedules and behavior maintained by intracranial self-stimulation. Science 122, 830–831.

Simmons, J.M., Ravel, S., Shidara, M., Richmond, B.J., 2007. A Comparison of Reward-Contingent Neuronal Activity in Monkey Orbitofrontal Cortex and Ventral Striatum: Guiding Actions toward Rewards. Ann. N. Y. Acad. Sci. 1121, 376–394. doi:10.1196/annals.1401.028

Simonson, I., 1989. Choice based on reasons: The case of attraction and compromise effects. J. Consum. Res. 16, 158–174.

Singer, W., Gray, C.M., 1995. Visual feature integration and the temporal correlation hypothesis. Annu. Rev. Neurosci. 18, 555–586. doi:10.1146/annurev.ne.18.030195.003011

Snodgrass, J.G., Townsend, J.T., Ashby, F.G., 1985. Stochastic Modeling of Elementary Psychological Processes. Am. J. Psychol. 98, 480. doi:10.2307/1422636

Soon, C.S., He, A.H., Bode, S., Haynes, J.-D., 2013. Predicting free choices for abstract intentions. Proc. Natl. Acad. Sci. 110, 6217–6222. doi:10.1073/pnas.1212218110

Stauffer, W.R., Lak, A., Kobayashi, S., Schultz, W., 2016. Components and characteristics of the dopamine reward utility signal: Dopamine utility signal. J. Comp. Neurol. 524, 1699–1711. doi:10.1002/cne.23880

Stauffer, W.R., Lak, A., Schultz, W., 2014. Dopamine Reward Prediction Error Responses Reflect Marginal Utility. Curr. Biol. 24, 2491–2500. doi:10.1016/j.cub.2014.08.064

Stephens, D.W., Krebs, J.R., 1986. Foraging theory. Princeton University Press.

Strait, C.E., Blanchard, T.C., Hayden, B.Y., 2014. Reward Value Comparison via Mutual Inhibition in Ventromedial Prefrontal Cortex. Neuron 82, 1357–1366. doi:10.1016/j.neuron.2014.04.032

Stuss, D.T., 2011. Functions of the Frontal Lobes: Relation to Executive Functions. J. Int. Neuropsychol. Soc. 17, 759–765.

doi:10.1017/S1355617711000695

Tanner Jr, W.P., Swets, J.A., 1954. A decision-making theory of visual detection. Psychol. Rev. 61, 401.

Thorndike, E.L., 1898. Animal intelligence: An experimental study of the associative processes in animals. Psychol. Rev. Monogr. Suppl. 2, i.

Thorpe, S.J., Rolls, E.T., Maddison, S., 1983. The orbitofrontal cortex: neuronal activity in the behaving monkey. Exp. Brain Res. 49, 93–115.

Timofeev, I., Steriade, M., 2004. Neocortical seizures: initiation, development and cessation. Neuroscience 123, 299–336.

Tobler, P.N., Fiorillo, C.D., Schultz, W., 2005. Adaptive Coding of Reward Value by Dopamine Neurons. Science 307, 1642–1645. doi:10.1126/science.1105370

Tom, S.M., Fox, C.R., Trepel, C., Poldrack, R.A., 2007. The Neural Basis of Loss Aversion in Decision-Making Under Risk. Science 315, 515–518. doi:10.1126/science.1134239

Tremblay, L., Hollerman, J.R., Schultz, W., 1998. Modifications of reward expectation-related neuronal activity during learning in primate striatum. J. Neurophysiol. 80, 964–977.

Tremblay, L., Schultz, W., 2000a. Modifications of reward expectation-related neuronal activity during learning in primate orbitofrontal cortex. J. Neurophysiol. 83, 1877–1885.

Tremblay, L., Schultz, W., 2000b. Reward-related neuronal activity during go-nogo task performance in primate orbitofrontal cortex. J. Neurophysiol. 83, 1864–1876.

Tremblay, L., Schultz, W., 1999. Relative reward preference in primate orbitofrontal cortex. Nature 398, 705–708.

Tsetsos, K., Gao, J., McClelland, J.L., Usher, M., 2012. Using Time-Varying Evidence to Test Models of Decision Dynamics: Bounded Diffusion vs. the Leaky Competing Accumulator Model. Front. Neurosci. 6. doi:10.3389/fnins.2012.00079

Tsetsos, K., Wyart, V., Shorkey, S.P., Summerfield, C., 2014. Neural mechanisms of economic commitment in the human medial prefrontal cortex. eLife 3, e03701.

Tusche, A., Bode, S., Haynes, J.-D., 2010. Neural Responses to Unattended Products Predict Later Consumer Choices. J. Neurosci. 30, 8024–8031. doi:10.1523/JNEUROSCI.0064-10.2010

Tversky, A., Kahneman, D., 1992. Advances in prospect theory: Cumulative representation of uncertainty. J. Risk Uncertain. 5, 297–323. doi:10.1007/BF00122574

Tversky, A., Kahneman, D., 1985. The framing of decisions and the psychology of choice, in: Environmental Impact Assessment, Technology Assessment, and Risk Analysis. Springer, pp. 107–129.

Tversky, A., Kahneman, D., 1981. The framing of decisions and the psychology of choice. Science 211, 453–458.

doi:10.1126/science.7455683

Tversky, A., Shafir, E., 1992. Choice under conflict: The dynamics of deferred decision. Psychol. Sci. 3, 358–361.

Usher, M., McClelland, J.L., 2001. The time course of perceptual choice: The leaky, competing accumulator model. Psychol. Rev. 108, 550–592. doi:10.1037//0033-295X.108.3.550

Valenstein, E.S., Beer, B., 1962. Reinforcing brain stimulation in competition with water reward and shock avoidance. Science 137, 1052–1054.

Van Der Meer, M.A., Redish, A.D., 2009. Low and high gamma oscillations in rat ventral striatum have distinct relationships to behavior, reward, and spiking activity on a learned spatial decision task. Front. Integr. Neurosci. 3, 9.

van der Meer, M.A.A., 2009. Covert expectation-of-reward in rat ventral striatum at decision points. Front. Integr. Neurosci. 3. doi:10.3389/neuro.07.001.2009

van Wingerden, M., Vinck, M., Lankelma, J., Pennartz, C.M.A., 2010. Theta-Band Phase Locking of Orbitofrontal Neurons during Reward Expectancy. J. Neurosci. 30, 7078–7087. doi:10.1523/JNEUROSCI.3860-09.2010

Varian, H.R., 1992. Microeconomic analysis. Norton & Company.

Vickers, D., 1970. Evidence for an accumulator model of psychophysical discrimination. Ergonomics 13, 37–58.

Von Neumann, J., Morgenstern, O., 1947. Theory of games and economic behavior.

Wagenmakers, E.-J., van der Maas, H.L.J., Grasman, R.P.P.P., 2007. An EZ-diffusion model for response time and accuracy. Psychon. Bull. Rev. 14, 3–22.

Wallis, J.D., 2011. Cross-species studies of orbitofrontal cortex and value-based decision-making. Nat. Neurosci. 15, 13–19. doi:10.1038/nn.2956

Wallis, J.D., Kennerley, S.W., 2011. Contrasting reward signals in the orbitofrontal cortex and anterior cingulate cortex: Contrasting reward signals in the orbitofrontal cortex. Ann. N. Y. Acad. Sci. 1239, 33–42. doi:10.1111/j.1749-6632.2011.06277.x

Wallis, J.D., Miller, E.K., 2003. Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. Eur. J. Neurosci. 18, 2069–2081. doi:10.1046/j.1460-9568.2003.02922.x

Wang, X.-J., 2002. Probabilistic decision making by slow reverberation in cortical circuits. Neuron 36, 955–968.

Wimmer, G.E., Shohamy, D., 2012. Preference by Association: How Memory Mechanisms in the Hippocampus Bias Decisions. Science 338, 270–273. doi:10.1126/science.1223252

Wunderlich, K., Rangel, A., O'Doherty, J.P., 2010. Economic choices can be made using only stimulus values. Proc. Natl. Acad. Sci. 107, 15005–15010. doi:10.1073/pnas.1002258107

Wunderlich, K., Rangel, A., O'Doherty, J.P., 2009. Neural computations underlying action-based decision making in the human brain. Proc. Natl. Acad. Sci. 106, 17199–17204.

Yaxley, S., Rolls, E.T., Sienkiewicz, Z.J., 1988. The responsiveness of neurons in the insular gustatory cortex of the macaque monkey is independent of hunger. Physiol. Behav. 42, 223–229.

Young, C.B., Nusslock, R., 2016. Positive mood enhances reward-related neural activity. Soc. Cogn. Affect. Neurosci. 11, 934–944. doi:10.1093/scan/nsw012

Zhang, Y., Pan, X., Wang, R., Sakagami, M., 2016. Functional connectivity between prefrontal cortex and striatum estimated by phase locking value. Cogn. Neurodyn. 10, 245–254. doi:10.1007/s11571-016-9376-2

Zylberberg, A., Barttfeld, P., Sigman, M., 2012. The construction of confidence in a perceptual decision. Front. Integr. Neurosci. 6. doi:10.3389/fnint.2012.00079

# List of abbreviations

ACC: Anterior cingulate cortex
BDM Becker–DeGroot–Marschak (Auction Bid)
BOLD: Blood-oxygen-level dependent
BVS: Brain valuation system
dACC: dorsal anterior cingulate cortex
DDM: Drift Diffusion Model
dmPFC: Dorsolateral prefrontal cortex
dmPFC: Dorsomedial prefrontal cortex
DV: Decision value
fMRI: Functional magnetic resonance imaging
iEEG: intracranial electro-encephalograpby
IFG: Inferior frontal gyrus
IPL: Inferior parietal lobule
LFP: Local field potential
LOFC: lateral orbitofrontal cortex
MEG: Magneto-encephalography
MNS: Mirror neuron system
MOFC: medial orbitofrontal cortex
OFC: orbitofrontal cortex
PMC: Premotor cortex
PCC: Posterior cingulate cortex
RT: Reaction times.
SDT: Signal detection theory
STS: Superior temporal sulcus
SV: Subjective value
tDCS: Transcranial direct current stimulation
TMS: Transcranial magnetic stimulation
vmPFC: ventromedial prefrontal cortex
VS: ventral striatum
WTP: Willingness top pay

# List of figures