



**HAL**  
open science

## Cerveau bayésien : peut-on modéliser l'émotion ?

Hugo Bottemanne

► **To cite this version:**

Hugo Bottemanne. Cerveau bayésien : peut-on modéliser l'émotion ?. *L'Encéphale*, 2021, 47 (1), pp.58-63. 10.1016/j.encep.2020.04.022 . halshs-03565384

**HAL Id: halshs-03565384**

**<https://shs.hal.science/halshs-03565384v1>**

Submitted on 13 Feb 2023

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

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



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

# Cerveau bayésien : peut-on modéliser l'émotion ?

---

*Bayesian brain: can we model emotion?*

**Hugo Bottemanne<sup>1,2</sup>**

<sup>1</sup> Service de psychiatrie adulte, groupe hospitalier de la Pitié-Salpêtrière, AP-HP, Paris.

<sup>2</sup> Equipe Control-Interoception-Attention, Institut du Cerveau et de la Moelle, Sorbonne-Université.

Email : [hugo.bottemanne@gmail.com](mailto:hugo.bottemanne@gmail.com)

## Résumé

La modélisation computationnelle permet de construire des modèles mathématiques simulant les mécanismes de perception, de prise de décision et de mise à jour des croyances. Ces modèles représentent mathématiquement ces processus complexes de traitement de l'information en combinant une distribution de probabilité antérieure, une fonction de vraisemblance et un ensemble de paramètres et d'hyperparamètres. Leur utilisation a popularisé la conception d'un système nerveux fonctionnant comme une machine prédictive, ou « cerveau bayésien ». Appliqués à la psychiatrie, ces modèles offrent des explications mécanistiques des dysfonctionnements retrouvés dans les troubles psychiatriques. Malgré les preuves de l'influence de l'émotion sur les processus cognitifs et son implication dans les troubles psychiatriques, peu de modèles computationnels proposent des représentations mathématiques de l'émotion ou intègrent des facteurs émotionnels dans leurs paramètres de modélisation. Nous présentons ici quelques hypothèses computationnelles pour la modélisation de paramètres affectifs et nous suggérons que les modèles bayésiens des troubles psychiatriques bénéficieraient de ces paramètres de modélisation.

**Mots-clés :** Cerveau bayésien ; Emotion ; Croyance ; Codage prédictif ; Inférence active ; Perception ; Prise de décision ; Neurosciences computationnelles ; Mise à jour des croyances.

## Abstract

Computational modelling builds mathematical models of cognitive phenomena to simulate patterns of perception, decision making, and belief updating. These models mathematically represent the information processing by combining an anterior probability distribution, a likelihood function and a set of parameters and hyperparameters. Their use popularized the conception of a nervous system functioning as a predictive machine, or "bayesian brain". Applied to psychiatry, these models seek to explain how psychiatric dysfunction may emerge mechanistically. Despite the significance of emotions for cognitive phenomena and for psychiatric disorders, few computational models offer mathematical representations of emotion or incorporate emotional factors into their modeling parameters. We present here some computational hypotheses for the modeling of affective parameters, and we suggest that computational psychiatry would benefit from these modeling parameters.

**Keywords :** Bayesian brain ; Emotion ; Belief ; Predictive coding ; Active inference ; Perception ; Decision-making ; Computational neurosciences ; Belief updating.

## Introduction

Bien qu'il existe une grande diversité d'approches théoriques, les émotions sont généralement définies comme des états affectifs avec une réalisation physique et psychologique susceptibles de se produire chez un organisme au cours de son interaction avec l'environnement [1].

Elles sont souvent caractérisées par une dimension qualitative, la valence, permettant de définir leur nature positive ou négative, et une dimension quantitative, l'intensité (« *arousal* »), désignant la force du ressenti subjectif ou de la manifestation physiologique associée [2]. De nombreux travaux en sciences cognitives ont montré que ces dimensions de valence et d'intensité influencent la manière dont nous traitons l'information, impactant nos processus perceptifs, décisionnels et d'apprentissage [3, 4]. Autrement dit, nous « percevons », nous « agissons » et nous « apprenons » en fonction de ce que nous ressentons. Mais comment l'émotion influence-t-elle ces mécanismes complexes ?

Pour répondre à cette question, le recours à la modélisation computationnelle est une technique prometteuse. Les formulations bayésiennes permettent de représenter mathématiquement ces mécanismes en combinant une distribution de probabilité antérieure, une fonction de vraisemblance et un ensemble de paramètres et d'hyperparamètres. Bien que leur application mathématique utilise de nombreux cadres de calcul, elles s'inspirent généralement du théorème introduit de manière posthume par le révérend Thomas Bayes [5].

Le théorème de Bayes offre une technique « optimale » pour « combiner » les connaissances d'un sujet avec les nouvelles données observées afin de calculer la vraisemblance de l'observation, et pour « modifier » ses connaissances en fonction des nouvelles informations observées. Si un sujet connaît A et fait une observation B, l'équation propose ainsi de combiner une probabilité antérieure  $P(A)$  représentant la connaissance initiale et une fonction de vraisemblance  $P(B/A)$  représentant la probabilité de B sachant A, pour former une probabilité postérieure  $P(A/B)$  désignant la connaissance A révisée par B :

$$P(A/B) = \frac{P(A) * P(B/A)}{P(A) * P(B/A) + P(\bar{A}) * P(B/\bar{A})}$$

Une hypothèse de plus en plus populaire en sciences cognitives, la théorie du « cerveau bayésien », consiste à appliquer les principes du bayésianisme à la manière dont notre cerveau traite l'information sensorielle et génère des croyances à propos du monde [6]. Cette théorie suppose que le cerveau utilise les informations sensorielles qu'il reçoit pour encoder des croyances, appelées « *priors* », à propos de son environnement. Il peut ensuite utiliser ces *priors* pour élaborer des inférences sur les causes et les caractéristiques des entrées sensorielles ainsi que sur les valeurs et les conséquences des actions que le sujet peut entreprendre [7]. Celles-ci fonctionnent comme des filtres bayésiens en « pondérant » les entrées sensorielles et « sélectionnant » les sorties motrices au travers d'un ensemble d'inférences perceptives (comment est l'environnement dans lequel le sujet évolue) et d'inférences actives (comment devrait être l'environnement après que le sujet a réalisé une action) [Figure 1].

La nature dynamique et ambiguë du monde contraint le cerveau à adapter continuellement ses modèles pour conserver son efficacité prédictive. Le décalage que le cerveau détecte entre les modèles et les entrées sensorielles constitue une erreur de prédiction qui lui permet de corriger son modèle statistique du monde [8]. La combinaison de ces cycles d'inférences (c'est-à-dire la déduction des causes des entrées sensorielles à partir des attentes) et de mise à jour (c'est-à-dire de révision des attentes en fonction des inférences) permet au sujet de décoder la complexité des signaux qu'il perçoit et d'encoder la probabilité statistique qu'ils se produisent.

En pratique, ce modèle a fait évoluer notre compréhension de la manière dont notre cerveau traite les informations sensorielles ou apprend à partir des régularités statistiques dans son environnement. Dans le domaine de la perception, la modélisation bayésienne a permis de comprendre que le rôle des neurones du cortex visuel primaire ne se réduit pas à extraire les caractéristiques sensorielles capturées dans l'environnement, mais que leur fonction est de

réaliser des inférences sur les entrées sensorielles en utilisant des *priors* permettant de prédire quel type d'information sensorielle est la plus probable [9]. Dans le domaine de l'apprentissage, la modélisation bayésienne a montré que, pour tirer des informations des régularités statistiques issues d'un environnement, le cerveau utilise des algorithmes bayésiens qui codent la confiance pour les entrées sensorielles et utilise cette confiance pour moduler son apprentissage probabiliste [10]. Pourtant, malgré l'impact des émotions sur les processus perceptifs et cognitifs, les modélisations bayésiennes ont souvent sacrifié les facteurs émotionnels. Leur négligence a entraîné une évolution dissociée des neurosciences computationnelles et affectives, délaissant tout un champ de notre activité cognitive [11].

Un regain d'intérêt récent a entraîné l'émergence de nouveaux modèles proposant des représentations mathématiques de l'émotion ou enrichis par des facteurs émotionnels [12, 13]. Ces modèles se développent à partir des fondements théoriques des neurosciences affectives et exploitent la puissance d'analyse du cadre bayésien pour formaliser mathématiquement des dimensions complexes comme la « valence » ou « l'intensité » de l'émotion [14]. Nous présentons ici quelques hypothèses pour la modélisation de l'émotion et l'intérêt d'associer ces nouveaux paramètres computationnels dans les modèles bayésiens des troubles psychiatriques.

## **L'émotion est-elle un « facteur » de l'inférence ?**

Si j'apprécie un arôme de fleur d'oranger en traversant un jardin florentin baigné de soleil, l'émotion positive qui m'envahit peut-elle influencer ma perception du sourire de la passante croisée peu après, mon seuil de décision pour l'achat de la sculpture coûteuse chez l'antiquaire du coin de la rue et mes croyances sur mes capacités de négociation en italien ? De nombreux

travaux suggèrent que l'émotion modifie le traitement perceptif, la prise de décision et le raisonnement [15, 16], agrégés en un système de « rationalité affective » [17].

Pour comprendre le rôle de l'émotion sur ce traitement complexe de l'information, on peut la représenter comme un « facteur » influençant la manière dont le cerveau réalise des inférences et met à jour ses croyances. Dans la théorie du cerveau bayésien, cela suppose que l'émotion peut moduler les signaux montants ou les signaux descendants, c'est-à-dire agir sur les erreurs de prédiction ou les attentes prédictives codées par notre cerveau.

En faveur de l'hypothèse d'une influence sur les signaux montants, plusieurs travaux expérimentaux ont montré que le contexte émotionnel facilite le traitement des erreurs de prédiction dans le cortex [18]. Les caractéristiques émotionnelles associées à un stimulus augmentent généralement sa « saillance », c'est-à-dire son importance en terme attentionnel ou épistémique, et induisent un effet de cadrage sur ce stimulus [20]. Par exemple, la peur associée à la perception d'une menace augmente considérablement la saillance du stimulus de menace et le codage de l'erreur de prédiction aversive au niveau du striatum ventral [19]. L'amplification du poids des erreurs de prédiction et de la rapidité de leur traitement cortical favorise les apprentissages associés à des contextes émotionnels [21]. Ainsi, lorsque l'on visualise un stimulus émotionnel juste avant un signal prédictif de récompense au cours d'une tâche d'apprentissage associatif, la vitesse d'apprentissage de l'association stimulus-récompense est accélérée. De plus, cet effet est renforcé lorsque la valence émotionnelle est congruente avec le type d'information impliqué [22].

Mais l'émotion pourrait aussi perturber les prédictions descendantes en influençant les croyances antérieures impliquées dans l'inférence. Plusieurs travaux ont montré qu'une émotion positive comme la joie favorise des attentes prédictives d'événements appétitifs [23], tandis qu'une

émotion négative comme la tristesse favorise des attentes prédictives aversives [24]. En perturbant les attentes prédictives, l'émotion pourrait agir comme un filtre sur la perception du monde, le colorant d'une tonalité affective congruente à l'état émotionnel. Ainsi lorsque l'on demande à des sujets de déterminer rapidement si des individus sont équipés d'armes ou feu ou d'objets neutres à partir de signaux ambigus, une émotion congruente comme la colère augmente la probabilité que des objets neutres soient identifiés à tort comme des objets liés à la violence [25]. Cette influence est déterminante dans de nombreuses situations de prise de décision au cours desquelles le cerveau doit calculer la « valeur » des options de décision en faisant des prédictions sur la récompense et le coût associé aux résultats des actions potentielles [26]. Dans le cadre du « *delay discounting* », correspondant à la capacité à différer une petite récompense immédiate pour recevoir une plus grande récompense future, plusieurs travaux ont montré que les émotions positives favorisent les décisions menant à des récompenses différées et plus importantes tandis que les émotions négatives entraînent un biais décisionnel pour des récompenses immédiates et plus réduites [27].

Traduit en langage computationnel, l'émotion pourrait ainsi constituer un hyperparamètre impactant la « confiance » dans les erreurs de prédiction, c'est-à-dire la saillance d'un signal, ou encore influençant le type d'attentes prédictives utilisées par le sujet pour percevoir et comprendre la nature de son environnement. Pourtant, cette représentation dualiste dissociant facteurs affectifs et cognitifs comme s'il s'agissait de deux entités computationnelles divergentes semble insuffisante pour proposer une modélisation pertinente de l'émotion.

## **L'émotion est-elle un « produit » de la confiance épistémique ?**



Plutôt que de la considérer comme une force extrinsèque modulant les processus cognitifs, on peut aussi se représenter l'émotion comme une caractéristique « endogène » de l'inférence et de la mise à jour des croyances.

L'une des perspectives héritée de la théorie computationnelle de l'émotion affirme que les émotions dépendent non seulement des croyances à propos du monde mais aussi des états motivationnels qui y sont associés. Cette théorie représente l'émotion sous la forme d'une dyade croyance-désir reflétant le degré de confiance pour un état de l'environnement, c'est-à-dire la croyance sur la vraisemblance statistique d'un événement, et d'appétence pour cet état, c'est-à-dire le désir qu'il se produise [28]. Par exemple, si je me baigne dans un lagon paradisiaque de l'Océan Indien et que j'aperçois un aileron grisâtre graviter autour de moi, la peur qui me saisit est corrélée à ma « croyance » dans la probabilité d'un événement dangereux et au « désir » qu'il ne se produise pas. Mon émotion se manifeste en fait après la mise à jour de mon attente prédictive à propos de la probabilité d'un événement aversif que je redoute. Je ressens de la peur si je « crois » qu'un requin est présent et que je ne « souhaite » pas être dévoré.

Dans notre exemple, mon émotion émerge après la mise à jour de la vraisemblance d'un événement aversif (ici, la présence d'un requin) compte tenu des preuves à ma disposition (un aileron grisâtre), en lien avec une cascade hiérarchique de croyances associées à l'événement aversif : je crois qu'un requin a un aileron, je crois qu'il y a des requins dans l'Océan Indien, je crois que les requins ont des écailles grisâtres. La valence de l'événement est déterminée par mon état motivationnel vis-à-vis de l'événement, formalisé comme un autre niveau hiérarchique de croyances prédictives sur les conséquences de l'événement [29]. Si je crois qu'un requin est dangereux, que se faire dévorer la jambe est douloureux et que je souhaite conserver ma jambe

pour continuer à faire des marathons, alors l'événement se voit attribuer une valence négative dans la dyade « croyance-désir ».

Cette conception permet de différencier des émotions factives comme la joie et la tristesse se rapportant à des événements pour lesquelles on possède un haut degré de certitude parce qu'ils se sont déjà produits ou parce que l'état de l'environnement est associé à une faible stochasticité, et des émotions épistémiques comme la peur ou l'espoir, caractérisées par un faible degré de certitude sur l'état de l'environnement [30]. Par exemple, si j'apprends que l'un de mes amis vient de subir un grave accident de voiture et que je n'ai pas d'informations sur son état de santé, mon incertitude à propos de l'événement est associée à une émotion de peur qui persistera jusqu'à ce que je réduise cette incertitude. Si j'assiste au contraire à l'enterrement de cet ami, alors j'ai une certitude élevée dans la probabilité qu'il soit mort tout en ayant le désir que ce ne soit pas le cas : l'émotion de tristesse qui m'envahit est le reflet de cet important degré de certitude pour un événement que je ne souhaite pas selon la dyade croyance-désir.

Dans cette perspective, l'émotion reflète le décalage entre les attentes prédictives et les changements dans l'environnement, mais constitue surtout un produit de la « confiance épistémique » et de « l'attribution motivationnelle » associées aux signaux perçus. Elle n'est plus un simple facteur qui influence la cognition et qu'on pourrait modéliser sous la forme d'un paramètre, mais elle est endogène à la manière dont le cerveau traite l'information.

## **L'émotion est-elle un « reflet » de la transition d'incertitude ?**

D'autres représentations théoriques présentent l'émotion comme un artefact phénoménologique de la transition d'incertitude de l'environnement [14].

Selon la théorie de l'inférence active, le degré d'incertitude du monde est représenté comme la « précision », c'est-à-dire notre confiance, dans les prédictions sur les conséquences sensorimotrices de nos actions [31]. Nous avons vu que cette théorie stipule que notre cerveau encode un modèle statistique du monde constitué par l'ensemble des croyances à propos de l'état actuel et futur de son environnement. Lorsque je confirme mes croyances antérieures, je réalise que ma prédiction est valide, ce qui indique que mon modèle statistique du monde est cohérent : je diminue ainsi indirectement la perception de l'incertitude de mon environnement. Lorsqu'au contraire mes croyances antérieures sont contredites par mes entrées sensorielles, je comprends que mon modèle statistique du monde n'est plus assez valide pour le prédire correctement et j'augmente indirectement la perception de l'incertitude de mon environnement. Autrement dit, plus mes croyances antérieures sont confirmées par mes actions et ma perception, plus les conséquences de mes actions sont prévisibles, plus le monde me semble contrôlable. Plus mes croyances antérieures sont infirmées par mes actions et ma perception, plus l'incertitude du monde augmente, plus je risque d'être impuissant à contrôler mon environnement.

Ce cadre théorique suppose que l'émotion positive et l'émotion négative sont respectivement associées à la confirmation et l'infirmité des croyances antérieures [14]. Lorsque mes entrées sensorielles confirment mes croyances antérieures, la prédictibilité de mon environnement augmente et je ressens une émotion positive, tandis que lorsque mes entrées sensorielles infirment mes croyances antérieures, la prédictibilité de mon environnement diminue et je ressens une émotion négative [32]. L'émotion est alors le « reflet » de la transition d'incertitude à propos de l'environnement. Dans l'exemple du requin dans l'Océan Indien, je ressens de la surprise lorsque le signal sensoriel (aileron grisâtre) vient contrarier mon attente prédictive, puis de la peur si la présence de ce signal augmente l'incertitude sur les conséquences possibles de

mes actions (je ne sais pas si je pourrais nager assez vite pour éviter de me faire dévorer). À l'inverse, si j'aperçois le même requin derrière la vitre d'un parc aquatique, je ne ressens ni surprise (je m'attendais à voir un requin dans ce contexte) ni peur (j'ai une grande confiance dans la prédictibilité de mon environnement).

Dans ce cadre, les émotions sont ainsi générées par le cerveau au fur et à mesure de l'évolution de son modèle statistique du monde. Elles formalisent alors l'évolution de la prévisibilité de l'environnement, représentée mathématiquement comme la précision des prédictions sur les conséquences sensorimotrices de l'action [33]. L'émotion perd son efficacité causale pour devenir l'écho du mécanisme de l'inférence et de la mise à jour des croyances. Elle ne constitue plus un « facteur » de la transition d'incertitude, mais un « marqueur » de cette transition. Elle n'est plus ce qui influence le mécanisme computationnel, mais un reflet phénoménologique de l'inférence hiérarchique.

## **Modéliser l'émotion pour améliorer la psychiatrie computationnelle ?**

Le champ en plein essor de la psychiatrie computationnelle propose une approche mécanistique des troubles psychiatriques. Il suggère que les symptômes cliniques reflètent des défauts sous-jacents de traitement de l'information qu'on peut modéliser sous la forme d'une perturbation de la distribution de probabilité antérieure, des erreurs de prédiction ou de la précision de ces paramètres. Considérant ces troubles non plus sous l'angle de la sémiologie descriptive mais au prisme d'une dimension mécanistique, la psychiatrie computationnelle propose de disséquer les perturbations du traitement de l'information associées à ces symptômes pour comprendre leur étiologie neurobiologique et étendre les perspectives thérapeutiques.

Concernant les troubles de l'humeur, les modèles d'inférence hiérarchique représentent généralement l'humeur comme une strate supérieure de paramètres appelés « hyperpriors » qui fournissent des prédictions sur les paramètres de niveau inférieur [Figure 2]. Si l'émotion reflète la violation des attentes prédictives sur les conséquences sensorimotrices des actions, alors l'humeur peut être modélisée comme la précision associée à l'émotion, c'est-à-dire la confiance dans la prédictibilité du monde [33]. Dans ce cadre, la dépression, modélisée comme une humeur basse insensible au contexte, survient lorsque le sujet est certain qu'il expérimentera un environnement incertain. Lorsqu'il est persuadé que ses actions n'auront aucune influence pour réduire cette incertitude, son vécu émotionnel est marqué par une oscillation d'affects de valence négative comme l'angoisse et la tristesse et il se retrouve incapable d'adopter les stratégies comportementales nécessaires à la formation de nouvelles inférences [34]. La mise à jour des croyances s'effectue alors à partir de paramètres perturbés ne capturant pas bien la relation entre les attentes prédictives et les états de l'environnement, formant des boucles d'attentes prédictives et de sensorialité inadaptées qui se renforcent progressivement au cours du temps. D'autres modèles utilisant des cadres de calcul alternatifs et associant facteurs émotionnels et paramètres computationnels ont permis de formuler des hypothèses sur la capacité de l'humeur à moduler la perception des stimuli, l'humeur positive conduisant à considérer les événements comme plus positifs et l'humeur négative à les considérer comme plus négatifs qu'ils ne le sont objectivement [35]. Ces modèles expliquent aussi les biais dans la prise de décision retrouvés chez les patients souffrant de trouble dépressif en explorant la manière dont l'humeur influence le poids relatif attribué aux options de choix [36].

Concernant les troubles psychotiques, de nombreux modèles simulent la symptomatologie hallucinatoire et délirante sous la forme d'une perturbation du processus d'inférence et de mise

à jour des croyances [37]. Ils proposent par exemple que le poids excessif des erreurs de prédiction dans la mise à jour confère une « saillance » inadaptée à des signaux sans pertinence épistémique, entraînant l'émergence d'idées délirantes, ou qu'une influence trop importante des attentes prédictives dans l'inférence perturbe le processus de perception, à l'origine des hallucinations [38].

L'intégration de facteurs émotionnels dans ces modèles computationnels permettrait de mieux comprendre l'interaction entre les dimensions cognitives et émotionnelles de la symptomatologie. La phase de transition psychotique est généralement caractérisée par une difficulté à réaliser des associations causales adaptées à partir du vécu sensoriel et émotionnel, entraînant l'émergence progressive de modèles erronés, les idées délirantes, pour expliquer le monde [39]. L'émotion associée à ces expériences semble être déterminante pour la rapidité du passage de croyances incertaines vers des convictions délirantes, puis pour la force de leur enracinement dans le système doxastique de l'individu [40]. L'intégration de ces facteurs permettrait de créer des modèles de la symptomatologie psychotique plus écologiques rendant compte du vécu subjectif des patients [41]. Par ailleurs, le rôle de l'émotion dans l'émergence des croyances atypiques et des dysfonctionnements décisionnels associés aux troubles obsessionnel-compulsifs et anxieux illustre également la nécessité d'intégrer ces dimensions dans les modélisations de ces troubles [42, 43].

## **Conclusion**

Le recours à la modélisation computationnelle est une technique prometteuse pour comprendre l'interaction entre l'émotion et la cognition. Plusieurs modèles bayésiens formulent des représentations mathématiques des dimensions de valence et d'intensité de l'émotion. Plutôt

que de la considérer comme un « facteur » exogène à la cognition, ils suggèrent que l'émotion est l'une des caractéristiques endogènes de l'inférence active et de la mise à jour des croyances.

Par ailleurs, de nombreux troubles psychiatriques impliquent une perturbation de la perception, de la prise de décision et de la mise à jour des croyances en lien avec des dysfonctionnements émotionnels. L'association de paramètres computationnels classiques et de facteurs émotionnels dans les modèles de ces troubles offrirait la possibilité de formuler de nouvelles hypothèses sur les perturbations cognitives et affectives qui les caractérisent. Cette approche mécanistique enrichie d'une dimension « affective » pourrait nous permettre dans un second temps de développer de nouvelles thérapies ciblant spécifiquement ces paramètres computationnels.

Malgré ces perspectives futures, les modèles existants ne caractérisent souvent qu'en partie ce que constitue le vécu émotionnel. Modéliser l'émotion et ses perturbations représente encore un défi majeur pour la psychiatrie computationnelle.

## Figures

**Titre figure 1 :** Triade computationnelle associant perception, prise de décision et mise à jour des croyances. **Légende Figure 1 :** Le cerveau bayésien suppose la confrontation entre les informations sensorielles et des attentes prédictives sur le monde. Ces attentes prédictives influencent le traitement des entrées sensorielles et des sorties motrices au travers d'un ensemble d'inférences perceptives et actives. Le décalage entre les attentes prédictives et les entrées sensorielles forme des erreurs de prédiction qui permettent au cerveau de réviser ses attentes prédictives sur le monde.

**Titre Figure 2 :** Représentation schématique de l'interconnexion entre attentes prédictives, émotion et humeur. **Légende Figure 2 :** Certains modèles suggèrent que l'émotion n'est pas une force extrinsèque mais un reflet du mécanisme de l'inférence hiérarchique associé à une mise à jour de l'attente prédictive sur un état du monde. L'humeur constitue alors un paramètre de niveau supérieur régulant l'impact des attentes prédictives sur la prise de décision et la perception.

## Liens d'intérêt

L'auteur déclare ne pas avoir de liens d'intérêts.

## Bibliographie

- [1] LeDoux JE. Emotional processing, but not emotions, can occur unconsciously. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions*. Oxford University Press. 1994;
- [2] Tellegen A, Watson D, Clark LA. On the dimensional and hierarchical structure of affect. *Psychol. Sci*, 1999. 10, 297–303 10.1111/1467-9280.00157
- [3] Anderson E, Siegel EH, Barrett LF. What you feel influences what you see: the role of affective feelings in resolving binocular rivalry. *J. Exp. Soc. Psychol.*, 47 (4) (2011), pp. 856-860, 10.1016/j.jesp.2011.02.009
- [4] Storbeck J, Clore G. Affective arousal as information: how affective arousal influences judgments, learning, and memory. *Soc. Personal. Psychol. Compass*, 2 (5) (2008), pp. 1824-1843
- [5] Efron, B. Bayes' Theorem in the 21<sup>st</sup> Century. *Science*. 2013;vol. 340, no 6137, p. 1177-1178
- [6] Knill DC, Pouget A. The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci*. 2004 Dec;27(12):712-9.
- [7] Friston K. The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 2010. 11, 127–138.
- [8] Mathys C, Daunizeau J, Friston KJ, Stephan KE. A bayesian foundation for individual learning under uncertainty. *Frontiers in Human Neuroscience*, 2011. 5: 1–20.
- [9] Lee TS, Mumford D. Hierarchical Bayesian inference in the visual cortex. *J Opt Soc Am A Opt Image Sci Vis*. 2003 Jul;20(7):1434-48
- [10] Meyniel F, Dehaene S. Brain networks for confidence weighting and hierarchical inference during probabilistic learning. *Proc Natl Acad Sci U S A*. 2017;114(19):E3859–E3868.
- [11] Schmidt LA, Schulkin J. Toward a computational affective neuroscience. *Brain Cogn*. 2000 Feb;42(1):95-8.
- [12] Smith R, Parr T, Friston KJ. Simulating Emotions: An Active Inference Model of Emotional State Inference and Emotion Concept Learning. *Front Psychol*. 2019;10:2844.
- [13] Lodewyckx T, Tuerlinckx F, Kuppens P, Allen N, Sheeber L. A hierarchical state space approach to affective dynamics. *J Math Psychol*. 2011;55(1):68–83.
- [14] Joffily M, Coricelli G. Emotional valence and the free-energy principle. *PLoS computational biology*, 2013. 9(6), e1003094.
- [15] DeSteno D, Petty RE, Wegener DT, Rucker DD. Beyond valence in the perception of likelihood: the role of emotion specificity. *J Pers Soc Psychol*. 2000 Mar;78(3):397-416.



- [16] Angie AD, Connelly S, Waples EP, Kligyte V. The influence of discrete emotions on judgement and decision-making: a meta-analytic review. *Cogn Emot.* 2011 Dec;25(8):1393-422. doi: 10.1080/02699931.2010.550751.
- [17] Slovic P, Finucane M, Peters E, MacGregor D. Rational actors or rational fools: implications of the affect heuristic for behavioral economics. *J. Socio. Econ.*, 31 (4) (2002), pp. 329-342, 10.1016/S1053-5357(02)00174-9
- [18] Vogel BO, Shen C, Neuhaus AH. Emotional context facilitates cortical prediction error responses. *Hum Brain Mapp.* 2015 Sep;36(9):3641-52.
- [19] Robinson OJ, Overstreet C, Charney DR, Vytal K, Grillon C. Stress increases aversive prediction error signal in the ventral striatum. *Proc Natl Acad Sci U S A.* 2013;110(10):4129–4133.
- [20] Rowe G, Hirsh JB, Anderson AK. Positive affect increases the breadth of attentional selection. *Proc. Natl. Acad. Sci. U.S.A.*, 104 (1) (2007), pp. 383-388
- [21] Watanabe N, Sakagami M, Haruno M. Reward prediction error signal enhanced by striatum-amygdala interaction explains the acceleration of probabilistic reward learning by emotion. *J. Neurosci.*, 33 (2013), pp. 4487-4493
- [22] Watanabe N, Haruno M. Effects of subconscious and conscious emotions on human cue-reward association learning. *Scientific reports*, 2015. 5, 8478. doi:10.1038/srep08478
- [23] Trevisani DP, Johnson SL, Carver CS. Positive Mood Induction and Facial Affect Recognition among Students at Risk for Mania. *Cognitive therapy and research*, 2008. 32(5), 639. doi:10.1007/s10608-007-9140-3
- [24] Handley IM, Lassiter GD, Nickell EF, Herchenroeder LM. Affect and automatic mood maintenance. *J. Exp. Soc. Psychol.* 2004. 40, 106–112 10.1016/S0022-1031(03)00086-6
- [25] Baumann J, Desteno D. Emotion guided threat detection: expecting guns where there are none. *J Pers Soc Psychol.* 2010 Oct;99(4):595-610. doi: 10.1037/a0020665.
- [26] Mitchell DG. The nexus between decision-making and emotion regulation: a review of convergent neurocognitive substrates. *Behav Brain Res.* 2011;217:215–231.
- [27] Liu L, Feng T, Chen J, Li H. The value of emotion: how does episodic prospection modulate delay discounting?. *PloS one*, 2013. 8(11), e81717. doi:10.1371/journal.pone.0081717
- [28] Green OH. The Belief-Desire Theory of Emotions. In: *The Emotions. Philosophical Studies Series.* 1992; Vol 53. Springer, Dordrecht. doi.org/10.1007/978-94-011-2552-9\_6
- [29] Reizenzein, R. Emotions as metarepresentational states of mind: Naturalizing the belief–desire theory of emotion. *Cognitive Systems Research.* 2009; Vol 10, Issue 1, Pages 6-20
- [30] Reizenzein, R. Emotional experience in the computational belief-desire theory of emotion. *Emotion Review.* 2009. 1: 214–222.
- [31] Seth AK, Friston KJ. Active interoceptive inference and the emotional brain. *Philosophical Transactions of the Royal Society B: Biological Sciences.* 2016;371, 20160007.
- [32] Friston K, Seth A, Joffily M, Barrett L. Active inference and emotion. In Andrew Fox, Regina Lapate, Alexander Shackman, and Richard Davidson (eds). *The Nature of Emotion. Fundamental Questions.* Second edition, Oxford University Press, pp. 28-33, 2018, 9780190612573

- [33] Clark JE, Watson S, Friston KJ. What is mood? A computational perspective. *Psychological medicine*, 2018. 48(14), 2277–2284. doi:10.1017/S0033291718000430
- [34] Stephan KE, Manjaly ZM, Mathys CD, et al. Allostatic Self-efficacy: A Metacognitive Theory of Dyshomeostasis-Induced Fatigue and Depression. *Front Hum Neurosci*. 2016;10:550. doi:10.3389/fnhum.2016.00550
- [35] Eldar E, Niv Y. Interaction between emotional state and learning underlies mood instability. *Nature communications*, 2015. 6, 6149. doi:10.1038/ncomms7149
- [36] Vinckier F, Rigoux L, Oudiette D, Pessiglione M. Neuro-computational account of how mood fluctuations arise and affect decision making. *Nature communications*, 2018. 9(1), 1708. doi:10.1038/s41467-018-03774-z
- [37] Krystal JH, Murray JD, Chekroud AM, et al. Computational Psychiatry and the Challenge of Schizophrenia. *Schizophr Bull*. 2017;43(3):473–475.
- [38] Fletcher PC, Frith CD. Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nat Rev Neurosci*. 2009 Jan;10(1):48-58. doi: 10.1038/nrn2536.
- [39] Warman DM, Martin JM. Jumping to conclusions and delusion proneness: the impact of emotionally salient stimuli. *J Nerv Ment Dis*. 2006 Oct;194(10):760-5.
- [40] Livingstone K, Harper S, Gillanders D. An exploration of emotion regulation in psychosis. *Clin Psychol Psychother*. 2009;16(5):418-30. doi: 10.1002/cpp.635.
- [41] Williams, D. Hierarchical Bayesian models of delusion. *Consciousness and Cognition*. 2018;61:129-147. doi: 10.1016/j.concog.2018.03.003
- [42] Sachdev PS, Malhi GS. Obsessive-compulsive behaviour: a disorder of decision-making. *Aust N Z J Psychiatry*. 2005;39:757–763.
- [43] Cornwell BR, Garrido MI, Overstreet C, Pine DS, Grillon C. The Unpredictive Brain Under Threat: A Neurocomputational Account of Anxious Hypervigilance. *Biol Psychiatry*. 2017;82(6):447–454. doi:10.1016/j.biopsych.2017.06.031



