



HAL
open science

Combiner les apprentissages motivés et associatifs

Maxime Carrere

► **To cite this version:**

Maxime Carrere. Combiner les apprentissages motivés et associatifs. Bio-informatique [q-bio.QM]. Université de Bordeaux, 2016. Français. NNT : 2016BORD0191 . tel-01399886

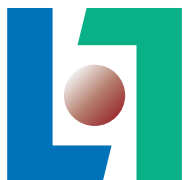
HAL Id: tel-01399886

<https://theses.hal.science/tel-01399886>

Submitted on 21 Nov 2016

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.



THÈSE

PRÉSENTÉE À

L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE DE MATHÉMATIQUES ET
D'INFORMATIQUE

par **Maxime Carrere**

POUR OBTENIR LE GRADE DE

DOCTEUR

SPÉCIALITÉ : INFORMATIQUE

Combiner les apprentissages motivés et associatifs

Date de soutenance : 11 octobre 2016

Devant la commission d'examen composée de :

Frédéric ALEXANDRE	Directeur de recherche, Inria	Directeur
Peter DOMINEY	Directeur de recherche, SBRI	Rapporteur
Mehdi KHAMASSI	Chargé de recherche HDR, ISIR	Rapporteur
Pierre-Yves OUDEYER	Directeur de recherche, Inria	Examineur
Aline DESMEDT	Maître de conférence HDR, Neurocentre Magendie	Examineur
Julien VITAY	Associate Professor, Chemnitz University of Technology	Examineur

Résumé Pour pouvoir être autonomes dans un environnement complexe, les humains comme les systèmes artificiels doivent posséder un apprentissage souple et capable de s'adapter au changement. Dans cette thèse, nous nous intéressons à comment cette autonomie peut être obtenue par interactions entre les différents systèmes d'apprentissage de notre cerveau. Pour cela, nous modélisons dans une approche inspirée de la biologie le comportement de certaines des parties du cerveau impliquées dans les apprentissages répondant et opérant, et observons comment leurs interactions permettent un apprentissage flexible dans des tâches impliquant des changements comme l'extinction et le reversal.

Title Combining associative and motivated learning.

Abstract In a complex environment, humans and artificial systems need a flexible learning system to adapt themselves to situations which can change. In this thesis, we study how autonomy can be the result of interactions between the different learning systems of our brain. In particular, in a biologically-inspired approach, we model different parts of the brain involved in respondent and operant conditioning, and show how their interactions can promote flexible learning in tasks in which situation can change, like extinction or reversal.

Keywords Learning, Biologically-inspired, Neuroscience, Amygdala, Cortex, Striatum, Neuromodulation.

Mots-clés Apprentissage, Bio-inspiré, Neuroscience, Amygdale, Cortex, Striatum, Neuromodulation.

Laboratoire d'accueil Laboratoire Bordelais de Recherche en Informatique (LaBRI). 351, cours de la Libération, Talence, France

Remerciements

Dans un but d'égalité, les remerciements suivants ont été mélangés suivant un ordre aléatoire par l'usage de `numpy.randint`. Ainsi, l'ordre des remerciements n'est, en aucun cas, un ordre d'importance!

Je tiens à remercier particulièrement Nicolas Rougier, pour son aide dans l'utilisation et l'évolution de Dana, pour son implémentation du modèle des basal ganglia que j'ai repris, ainsi que pour m'avoir fait participer au projet Pinokio, à Pint of Science, et j'en passe.

Un grand merci à Fabien pour me faire prendre une pause quand je ne le ferais pas de moi-même, ainsi que sa corrections des différents bogues des bibliothèques dont je dépends.

Je tiens à remercier Meropi pour son travail sur le modèle des basal ganglia, sans lequel je n'aurais pas eu de base sur la deuxième partie de ma thèse. Ainsi que pour m'avoir fait pratiquer mon anglais régulièrement!

Je tiens à remercier les enseignants de l'Enseirb-Matmeca, pour m'avoir permis de donner des TDs en parallèle de ma thèse durant ces trois années. Avec un remerciement tout particulier à toutes les personnes qui m'ont aider à préparer les différents TDs la première année, Denis, David, Aymerick et Georges.

Un grand merci aux différents étudiants avec qui j'ai pu collaborer au cours de cette thèse, Nicolas, Florian et Amélie pour Virtual Enaction, Pramod pour la dopamine, ainsi qu'à ceux qu'avec lesquels je n'ai fais qu'échanger.

Un grand merci aux différents invités avec qui j'ai pu discuté, confronté, et faire évoluer mes points de vue, en particulier Jessica et Davide!

Un grand merci à Thierry, parce qu'il aime bien qu'on s'occupe de lui. Ainsi que pour m'avoir laissé m'amuser sur Virtual Enaction!

Un grand merci à toute ma famille pour son aide et ses conseils, que ce soit avant, pendant, ou après la thèse.

Un grand merci à Ikram pour son aide, et son implication pour organiser les différents évènements!

Un grand merci à Xavier pour fouetter les stagiaires et éviter qu'ils fassent trop de bruit!

Un grand merci aux autres doctorant, en particulier Quentin, Grégoire, Maria, pour les différents échanges, repas, activités que l'on a pu partager.

Je tiens à remercier André, que j'ai trahis en n'utilisant finalement pas de spikes.

Je tiens à remercier -et à m'excuser- à tous ceux que j'aurais par erreur oublié dans cette page ou la précédente.

Je tiens à remercier Charlotte, en tant que fellow doctorante! Ainsi que pour son implications pour organiser les différents évènements, ainsi que Pint of Science.

Un grand merci à tout les amis, ex-enseirbien où non, pour me laisser me changer les idées quand j'en ai besoin!

Je tiens à remercier Nicolas et Chrystel, pour l'aide qu'ils m'ont apporté, ainsi que celle qu'ils apportent à toute l'équipe, pour toutes les démarches au niveau Inria, et pour toute leur aide dans l'organisation, la gestion, les démarches ou le remboursement des missions. Ce, même dans les cas compliqués comme le mien, impliquant plusieurs laboratoires, et des démarches partagées entre ceux-ci.

Un grand merci à la marmotte numéro deux pour son soutien, que ce soit à base de chocolat ou non!

J'ai quand même triché sur le caractère aléatoire de l'ordre de mes remerciements afin de garder Frédéric pour la fin. Un grand merci à Frédéric Alexandre, mon directeur de thèse. J'ai malheureusement bien trop de raisons de te remercier pour pouvoir toutes les lister ici, par conséquent un grand merci pour tout! Si je devais quand même essayer de les lister, je crois que je commencerais par les heures de discussion sur comment marche le cerveau, devant un tableau remplis de flèches et de boîtes.

Table des matières

Table des matières	vii
Introduction	1
1 Etude du comportement	5
1.1 Les paradigmes du conditionnement	5
1.1.1 Le conditionnement répondant	6
1.1.2 Le conditionnement opérant	9
1.2 Les règles du conditionnement	11
1.2.1 L'erreur de prédiction	12
1.2.2 Apprentissage dépendant du CS	13
2 Etude de la biologie	17
2.1 Le cortex sensoriel et l'hippocampe : des entrées élaborées	18
2.2 L'amygdale, lieu de l'association CS-US	21
2.3 Le conditionnement opérant, liée au conditionnement répondant	24
2.4 Le rôle de la neuromodulation	27
3 Problématique et outils utilisés	33
3.1 Problématique	33
3.2 Des modèles sous contraintes biologiques	34
3.2.1 Modèle de neurones utilisé	35
3.2.2 Apprentissage Hebbien	36
3.3 Principales bibliothèques et logiciels.	36
3.3.1 DANA	37
3.3.2 Nengo	38
3.3.3 Virtual Enaction	39
4 Le rôle des différents circuits de l'amygdale dans l'acquisition de la réponse pavlovienne	43
4.1 Les différents circuits de l'amygdale	43
4.2 Description du modèle	45
4.2.1 Règles d'apprentissage	48
4.2.2 Le rôle de la neuromodulation	50

4.2.3	Protocole expérimental	51
4.3	Résultats	52
4.3.1	Extinction-renewal	53
4.3.2	Pairing et unpairing	55
4.4	Discussion	58
4.4.1	L'apprentissage complémentaire de LA et BAf	59
4.4.2	L'extinction apprise par BAe	60
4.4.3	Un rôle simplifié de CeLOn.	62
4.4.4	Impact d'ACh sur l'apprentissage	63
4.4.5	Relations avec les paradigmes de conditionnement	64
4.4.6	Conclusion	65
5	L'apprentissage pavlovien de la boucle cortico-striatale limbique guide l'apprentissage opérant.	67
5.1	Introduction	67
5.2	Modèle préexistant	68
5.3	Le calcul de l'erreur de prédiction, via la neuromodulation	71
5.4	Expérience de reversal	73
5.5	Conclusion	74
6	Le rôle de la noradrénaline dans le compromis exploration-exploitation.	79
6.1	Introduction	79
6.2	Le système noradrenergique	81
6.3	Modéliser l'exploration	82
6.4	Expériences et résultats	84
6.5	Conclusion	86
	Conclusion	89
	Bibliographie	97

Introduction

Aristote le dit, “le tout est plus que la somme des parties”.

En 2016, année où j’écris cette thèse, un programme, AlphaGo, bat le champion du monde de Go Lee Sedol, par un score de 4 à 1, et bat avec 99.8% de certitude les programmes existants jouant au Go.

Si nous examinons AlphaGo d’un peu plus près, nous pouvons nous s’apercevoir que celui-ci est basé sur un apprentissage par réseaux de neurones [Silver *et al.*, 2016]. Comparé aux échecs, le jeu de Go présente deux difficultés majeures. Premièrement, la combinatoire, i.e le nombre de coups possibles à partir d’une position, est beaucoup plus importante, rendant très compliquée une exploration des différentes possibilités. Deuxièmement, il est difficile d’évaluer la valeur d’une position. Les joueurs professionnels parlent “d’instinct”, “de sens esthétique”, pour évaluer une position.

AlphaGo résout ce problème par la combinaison de différents apprentissages, via des réseaux de neurones calculant des valeurs différentes. Un réseau de neurones apprend ainsi à calculer la valeur des différentes positions, et un deuxième à définir quels sont les coups intéressants dans une situation donnée. Grâce à ces deux apprentissages, qui répondent précisément aux deux difficultés majeures du jeu de Go par rapport aux échecs, et de leur utilisation par un algorithme de décision classique en théorie des jeux, de type MonteCarlo, AlphaGo est capable d’être plus performant que l’humain. AlphaGo montre ainsi l’intérêt de combiner différents apprentissages pour la prise de décision. Il illustre ainsi dans une certaine mesure la citation d’Aristote, en permettant au tout, à la combinaison de différents apprentissages, d’être plus efficace que la somme des parties, que ces apprentissages et de décisions, pris séparément.

Cependant, si AlphaGo montre l’intérêt de combiner différents apprentissages, ceux-ci sont spécifiques au jeu de Go, et les valeurs qu’ils calculent ne sont pas adaptées à d’autres types de problèmes. Est-il possible de combiner des apprentissages moins spécifiques, pour obtenir un apprentissage global plus général, et adaptés à davantage de problèmes? Si nous regardons l’humain, certes moins performant au Go, notre point fort, et dans une moindre mesure celui des animaux par rapport aux machines, est notre capacité d’adaptation, notre autonomie, notre flexibilité à apprendre des tâches différentes. Je sais jouer à la fois aux échecs, aux dames, et au Go. Si pour chacun des jeux, je suis moins performant qu’un programme dédié exclusivement à ce jeu, je suis

capable d'apprendre à jouer, et acquérir une performance toute relative, dans ces trois jeux. Et je serai probablement capable d'apprendre à jouer à un nouveau jeu, qui n'a pas encore été inventé moment où j'écris cette affirmation. Étudier la façon dont l'humain combine les apprentissages qu'il possède, et créer des modèles de prises de décision basés sur son fonctionnement, peut ainsi nous permettre de créer des systèmes plus autonomes, plus flexibles que ceux qui existent actuellement, capables de s'adapter à différents types de tâches.

Une définition de l'autonomie, qui est un des grands centres d'intérêt de l'IA et de la robotique actuellement, est la capacité à se définir par soi-même ses propres buts, en fonction de ses besoins et à apprendre les actions nécessaires pour parvenir à ces buts. Si notre cerveau possède différentes régions, et différents mécanismes d'apprentissage dans ces régions, deux apprentissages sont particulièrement nécessaires à l'autonomie. Le premier est l'apprentissage du conditionnement opérant, que j'appellerais par la suite apprentissage opérant ou apprentissage motivé, l'apprentissage des actions nécessaires à l'obtention d'un but. Cependant, pour pouvoir obtenir choisir et obtenir notre but, il nous faut pouvoir arriver à prédire quels buts peuvent être obtenus dans une situation. L'apprentissage observé lors du conditionnement répondant, que j'appellerais par la suite apprentissage répondant, ou apprentissage pavlovien, est l'apprentissage passif des contingences entre les stimuli que nous percevons et les récompenses, buts potentiels, ou les punitions, éviter une punition étant un but en soi. C'est un apprentissage associatif, qui apprend à associer stimuli et buts. Fait intéressant, ces deux apprentissages sont les deux apprentissages, appliqués au cas précis du jeu de GO, utilisés par AlphaGo. Son premier réseau de neurone apprend en effet à associer une valeur aux positions, indépendamment du choix effectué, et est ainsi proche de l'apprentissage pavlovien considéré ici. Son second réseau apprend les actions pertinentes à effectuer dans une situation donnée, et est ainsi similaire à un apprentissage opérant.

Combiner les apprentissages associatif et motivé permet ainsi de définir quels sont les buts atteignables dans une situation, et quelles sont les actions à faire pour atteindre ses buts. Nous espérons ainsi dans cette thèse, par l'étude et la modélisation de ces apprentissages et de leurs interactions dans la biologie, pouvoir apporter des réponses sur comment créer des systèmes informatiques plus autonomes, et plus flexibles, face aux différentes situations qu'ils rencontrent. Afin de présenter l'état de l'art de ce qui est connu de ces différents apprentissages, et de leurs interactions, dans le vivant, je présenterai deux approches différentes de l'étude de ces mécanismes. La première est l'approche comportementale, c'est-à-dire l'observation et la modélisation des apprentissages associatifs et motivés au travers des comportements observés chez les humains et les animaux lors du conditionnement à différentes tâches. Puis, motivé par les limites de cette approche, je présenterai dans une étude de la

biologie un état de l'art des mécanismes et régions impliqués dans les apprentissages associatif et motivé dans le cerveau. Cet état de l'art nous permettra alors de distinguer plus précisément, dans l'étude des interactions entre apprentissages motivés et opérant, quels mécanismes cérébraux participant à la flexibilité et l'adaptabilité de ces différents apprentissages nous reproduirons dans des modèles informatiques.

Cette démarche se situe ainsi à mi-chemin entre neuroscience et informatique. Du point de vue des neurosciences, les modèles réalisés peuvent contribuer à améliorer la connaissance du domaine, en mettant en avant certains mécanismes non reproduits ou explicités au préalable, et en proposant des prédictions nouvelles en découlant. Du point de vue des sciences informatiques, ces modèles mettent en avant des principes d'apprentissages pouvant permettre d'améliorer l'apprentissage de systèmes apprenants, algorithmes ou robots.

Chapitre 1

Etude du comportement

Dans cette section, je présente les principes d'apprentissages que nous cherchons à comprendre et modéliser en abordant un point de vue comportemental. En quelque sorte, l'objectif de cette section est de répondre à la question "quoi", de présenter quels sont les problèmes sur lesquels on étudie l'apprentissage, et quels sont les comportements observés, résultant de l'apprentissage. Egalement, cette approche permet d'introduire l'intérêt d'une autre approche, basée sur la biologie, qui permet d'étudier les mécanismes précis derrière les différents apprentissages, ainsi que comment sont combinés ces différents apprentissages. L'approche par l'étude de la biologie, décrite dans la section suivante, répond ainsi davantage à la question "comment".

La première partie de cette section est consacrée à la présentation des différentes tâches (que nous appellerons paradigmes) qui permettent de tester les apprentissages associatifs et motivés ainsi que des éléments mettant en évidence une relation entre ces différents apprentissages.

La deuxième partie étudie l'approche de modélisation "comportementale", qui cherche à reproduire précisément les comportements observés des animaux dans les paradigmes, l'effet de l'apprentissage, plutôt que de reproduire la façon dont l'apprentissage est implémenté dans notre cerveau. Notamment, l'étude de ces règles d'apprentissage permet de mettre en évidence différents principes et stratégies présents dans notre cerveau, sans toutefois expliquer au travers de quels mécanismes ces stratégies sont mises en oeuvre, que seule une approche basée sur la modélisation du fonctionnement du cerveau pourra permettre.

1.1 Les paradigmes du conditionnement

Il existe de nombreux paradigmes utilisés pour observer le comportement des animaux, et les fonctions spécifiques des différentes zones du cerveau. La plupart de ces paradigmes restent simples, à la fois parce qu'il est difficile de mettre en place et de faire apprendre aux animaux des problèmes complexes

(beaucoup de problèmes que nous jugerions simples nécessitent plusieurs mois d'apprentissage pour des animaux), et parce que des problèmes “simples” demandent moins de fonctions différentes afin de les résoudre, et permettent ainsi d'attribuer plus facilement des fonctions spécifiques aux différentes aires cérébrales impliquées dans leurs résolutions.

Dans cette partie, je présente plusieurs paradigmes classiques du conditionnement sur lesquels nous avons testé nos modèles, et les comportements observés des animaux face à ces paradigmes. Généralement, ces paradigmes peuvent être regroupés en deux catégories, paradigmes du conditionnement répondant, i.e. pour lesquels l'animal n'a pas à effectuer d'action pour obtenir de récompense, et apprend de manière passive à prédire quelle récompense arrivera, et paradigmes du conditionnement opérant, durant lesquels l'animal doit effectuer une ou plusieurs actions pour réussir l'essai et obtenir sa récompense.

1.1.1 Le conditionnement répondant

Le paradigme du conditionnement répondant le plus étudié est probablement le conditionnement pavlovien [Rescorla, 1988].

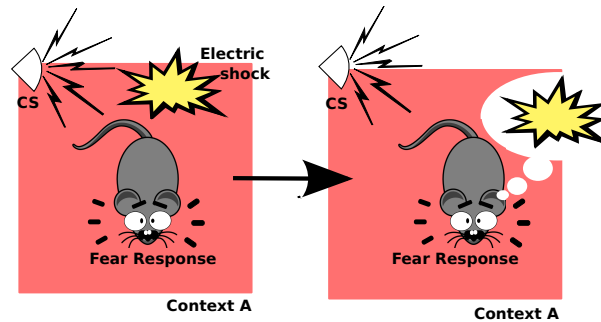


FIGURE 1.1 – Acquisition de la réponse pavlovienne. Dans une phase de conditionnement, le CS, stimulus initialement neutre, est présenté de manière répétitive associé à l'arrivée de l'US (une récompense ou une punition). Suite au conditionnement, l'arrivée du CS seul déclenche la réponse réflexe (telle que la salivation si la récompense se mange, ou une réponse de peur dans le cas d'une punition) normalement associée à l'arrivée de l'US.

Un stimulus initialement neutre (le CS, “conditioned stimulus”), précède l'arrivée d'une récompense (l'US, “unconditioned stimulus”), appétitive (nourriture, boisson) ou aversif (choc électrique). Après apprentissage, le CS provoque une réponse réflexe, ou réponse pavlovienne de l'animal, qui peut être une réponse motrice, comme le “freezing” (“gel” de l'animal, qui demeure crispé, immobile) pour l'aversif, ou comme la salivation pour l'appétitif. La réponse pavlovienne peut également être une réponse attentionnelle ou hormonale, comme la libération de neuromodulateurs tels que la dopamine ou l'acétylcho-

line. L'animal parvient ainsi à acquérir une association entre le CS et l'arrivée de la récompense, association observable par la réponse pavlovienne au moment de l'arrivée du CS.

Une fois cette association établie, elle est robuste, et n'est pas oubliée ou remplacée facilement. Deux paradigmes, l'extinction et le "renewal" (ou rétablissement spontané) démontrent cette robustesse. Après acquisition de la réponse pavlovienne, la présentation de manière répétée du CS seul, sans l'US, provoque une diminution de la réponse pavlovienne, phénomène appelé extinction. Cependant, cette extinction n'est pas un oubli de l'association CS-US apprise lors du conditionnement, comme le montre le renewal, durant lequel la réponse pavlovienne peut être spontanément rétablie, comme par exemple si l'animal est remis dans le contexte dans lequel il avait appris l'association pavlovienne [Herry *et al.*, 2008] :

Extinction et renewal :

CS + Ctxt1 -> US

CS + Ctxt2 -> noUS

CS + Ctxt1 -> rétablissement de la réponse pavlovienne.

Le blocking montre que l'association formée lors de l'acquisition n'est pas facilement remplacée. Lors du blocking, l'animal apprend à associer un premier CS, CS1, à l'arrivée de l'US. Un nouveau CS, CS2, est ensuite présenté en même temps que CS1, toujours suivis de l'US. Puis l'animal est testé uniquement avec CS2, et ne montre pas de réponse pavlovienne. Le fait que CS1 permettait déjà de prédire l'arrivée de l'US bloque donc l'apprentissage de CS2, évitant ainsi l'apprentissage d'une association superflue. La seconde association n'a pas non plus remplacée l'association déjà existante.

Blocking :

CS1 -> US

CS1 + CS2 -> US

CS1 -> réponse pavlovienne

CS2 -> absence de réponse pavlovienne

L'inhibition conditionnée est un paradigme dans lequel un premier CS, CS1, est associé à l'US, puis est présenté dans une seconde phase en présence d'un autre CS, CS2, en l'absence de récompense. Après apprentissage de cette seconde phase, l'animal ne prédit plus l'US si CS1 et CS2 sont présentés simultanément, mais le prédit toutefois si CS1 est présenté seul. L'association pavlovienne n'est pas oubliée, mais simplement inhibée par la présence d'un autre CS, qui acquiert une valeur inhibitrice.

Inhibition conditionnée :

CS1 -> US

CS1 + CS2 -> noUS

CS1 -> réponse pavlovienne

CS1 + CS2 -> absence de réponse pavlovienne

L'influence du contexte sur le conditionnement a également été très étudiée. Dans certains paradigmes, comme l'extinction contextuelle (extinction dans un contexte spécifique de l'association CS-US apprise au préalable), le terme contexte désigne un lieu, une certaine pièce dans laquelle l'animal est placé durant l'apprentissage de l'extinction. Plus généralement le contexte peut être compris comme l'ensemble des informations reçues par l'animal dans une situation donnée (informations visuelles, odeurs, sons, touché, etc), extéroceptives et intéroceptives (niveau de faim, état du corps, etc). Le contexte peut, comme le CS, être associé à l'US. L'animal apprend dans ce cas que la pièce, la situation dans laquelle il se trouve, est prédictrice d'US, au lieu d'attribuer l'arrivée de l'US à un CS particulier, à un élément particulier de la pièce. Les paradigmes de pairing et d'unpairing permettent en particulier de déterminer quelle est l'association, CS-US ou contexte-US, apprise par l'animal sous différentes conditions. Le pairing consiste en présentations successives du CS suivi de l'US, dans un contexte fixe, jusqu'à apprentissage. Puis la réaction de l'animal est observée quand uniquement le CS, ou uniquement le contexte, lui est présenté. L'animal réagit alors presque exclusivement au CS seul, mais pas au contexte seul, montrant ainsi qu'en temps normal, l'apprentissage est biaisé vers le CS saillant. Lors de l'unpairing, l'animal est également testé dans un contexte fixe, avec des présentations de CS suivis d'US, mais un délai aléatoire est inséré entre le CS et l'US. Lorsque l'animal est testé avec uniquement le CS ou uniquement le contexte, la réponse de l'animal est biaisée vers le contexte, contrairement au pairing.

Pairing :

CS1 + Ctxt -> US

CS1 -> réponse pavlovienne

Ctxt -> absence de réponse pavlovienne

Unpairing :

CS1 + Ctxt -> délai aléatoire -> US

CS1 -> absence de réponse pavlovienne

Ctxt -> réponse pavlovienne

Dans ces paradigmes, une distinction est généralement faite selon que l'US est, ou non, séparé dans le temps du CS. Si les présentations du CS et l'US sont contiguës, ou se recouvrent, on parle de "delay conditioning". Si au contraire

il existe un intervalle de temps entre la fin de la présentation du CS, et l'arrivée de l'US, on parle de "trace conditioning", i.e. le conditionnement s'effectue à partir de la trace en mémoire du CS.

Ces différents paradigmes du conditionnement répondant mettent en évidence un apprentissage associatif pavlovien qui n'est pas uniquement la formation d'une simple association CS-US, mais une association adaptée à la situation, et qui peut être modifiée, inhibée, ou récupérée spontanément, lorsque la situation l'exige, pour permettre une plus grande adaptabilité et flexibilité dans le comportement.

Nous regardons dans cette thèse l'intérêt de l'apprentissage pavlovien mis en relation avec le conditionnement opérant. Cependant, l'apprentissage pavlovien est également, par lui même, intéressant pour l'animal, en permettant à l'animal de se préparer en avance à l'arrivée de phénomènes appétitif ou aversif, et lui permettant de rendre saillants des stimuli neutres, pour les détecter plus rapidement.

1.1.2 Le conditionnement opérant

Contrairement au conditionnement répondant, pour lequel la récompense ou punition arrive quel que soit le choix de l'animal, le conditionnement opérant, ou conditionnement instrumental, nécessite d'effectuer une action, en fonction éventuellement de plusieurs CS présentés, pour obtenir l'US. Dans les protocoles expérimentaux, l'action demandée peut être d'actionner un levier, ou d'emprunter une porte, notamment chez le rat. Chez le singe (ou l'homme), l'action demandée est souvent la sélection d'une cible, soit en la fixant du regard, soit en la sélectionnant au moyen d'une manette, de touches ou d'un écran tactile.

Le conditionnement opérant ne se résume cependant pas à l'apprentissage d'une simple association CS-action, mais est au contraire lié à l'apprentissage pavlovien, comme illustré par exemple lors du transfert pavlovien-instrumental (PIT). Le PIT débute par l'apprentissage d'une relation CS-US, par exemple une lumière suivie de l'arrivée de nourriture. Dans un second temps, en l'absence de la lumière, l'animal est entraîné à appuyer sur un levier pour obtenir de la nourriture (action-US). On observe alors que si la lumière est présentée en présence du levier, l'animal va davantage se mettre à actionner le levier qu'en présence du levier seul, montrant que l'apprentissage CS-US d'une part, et action-US d'autre part, a mené à l'apprentissage d'une relation CS-action (ou CS-action-US).

Dans le paradigme de dévaluation, un animal apprend dans un premier temps l'action à effectuer pour recevoir un US de valence positive, puis dans un second temps l'US reçu est dévalué (sa valeur est rendue aversive ou nulle, par exemple en mettant l'animal à satiété au préalable, ou en mélangeant l'US avec un US aversif). On observe alors qu'un animal davantage entraîné

durant la première phase, aura plus de difficulté à apprendre la dévaluation, et à arrêter de faire l'action associée au CS [Dezfouli et Balleine, 2012]. Ceci va dans le sens d'un apprentissage d'abord des associations CS-US et action-US, dans lesquels on a encore accès à la valeur de l'US, permettant donc de réagir plus rapidement lorsque sa valeur change. Plus l'animal est surentraîné, plus l'association CS-action est renforcée, entraînant la formation d'une habitude. La dévaluation de la valeur du but n'a alors plus d'effet sur le choix de l'action. Le choix est alors moins basé sur la valeur de l'US, mais plus lié à l'habitude quand présenté face au CS.

Un paradigme similaire à la dévaluation, que nous étudierons, est le reversal, dans lequel une règle précédemment établie est "renversée", modifiée. C'est le cas par exemple si stimulus précédemment récompensé devient neutre ou aversif, et un nouveau précédemment neutre, ou prédicteur d'une récompense aversive, devient récompensé. Le changement de règle imposé à l'animal lors du reversal peut être intra ou extra dimensionnel. Intra-dimensionnel désigne les cas où la dimension sur laquelle porte la règle reste la même après le reversal, par exemple si l'animal devait choisir des indices rouges avant le reversal pour obtenir l'US, et doit après le reversal sélectionner des indices bleus. Extra-dimensionnel désigne les cas où la dimension testée par la règle change, par exemple si l'animal doit maintenant sélectionner les indices ayant une certaine forme, ou ceux à une certaine position. Un paradigme classique souvent utilisé pour tester le reversal extra-dimensionnel chez les humains ou les singes est le Wisconsin card sorting test (WCST), dans lequel sont présentées plusieurs cartes de base, et l'animal doit associer une carte qui lui est présentée avec l'une des cartes de base, en fonction soit de sa couleur, soit de sa valeur ou de sa forme. La règle suivant laquelle effectuer les associations est inconnue du participant, qui reçoit uniquement un retour après chaque association formée l'informant si celle-ci était bonne ou pas. De plus la règle change au cours du temps, le participant doit alors détecter ce changement, et apprendre la nouvelle règle par exploration.

Les paradigmes comme le WCST, qui contiennent à un moment donné une incertitude sur l'action optimale à mener dans une situation donnée, mettent en avant une autre problématique liée au conditionnement opérant qui est le compromis entre exploration et exploitation. Si par exemple, dans une certaine situation, je sais qu'effectuer une action a 50% de chance de m'apporter une récompense, dois-je rester sur cette action, et donc exploiter la règle connue, ou au contraire, explorer d'autres actions, en espérant en trouver une avec une probabilité de récompense plus importante? Un paradigme souvent utilisé pour mettre en évidence ce compromis est la tâche du bandit manchot ("k-armed bandit task"), dans laquelle les animaux ont plusieurs choix de leviers à actionner, chaque levier ayant sa propre distribution de probabilité concernant l'arrivée de la récompense. Il faut donc arriver à trouver et à persévérer sur le meilleur choix, tout en explorant suffisamment chacun des choix pour choisir

celui avec la probabilité de gain la plus élevée. De plus, dans un environnement changeant, les probabilités liées aux différents choix peuvent changer, ce qui nécessite constamment un choix entre exploitation et exploration.

Ainsi, que ce soit pour les paradigmes répondants, dans lesquels l'US est associé à un élément prédicteur au cours de l'apprentissage, ou pour les paradigmes opérants, dans lesquels une action doit être effectuée pour atteindre la cible désirée, les problèmes à résoudre sont relativement simples, mais nécessitent des mécanismes d'apprentissages complexes pour pouvoir s'adapter aux différentes situations possibles, dans un monde incertain et changeant. De plus, si ces deux types de conditionnement sont souvent présentés comme séparés, les mécanismes sous-jacents ne sont pas indépendants, comme le montre notamment le PIT, dans lequel l'apprentissage d'associations pavloviennes guide l'opérant. Connaître la valeur d'un CS par l'apprentissage pavlovien permet d'en faire une cible de l'action, au même titre que l'US auquel il est associé. Comment peut-on alors modéliser cet apprentissage ?

1.2 Les règles du conditionnement

Une des premières approches utilisées dans l'étude des mécanismes d'apprentissages impliqués dans le conditionnement est une approche comportementale, qui se concentre sur la création de règles d'apprentissage reproduisant le comportement des animaux dans les différents paradigmes d'apprentissages. Cette approche vise donc à reproduire directement les résultats de l'apprentissage, visibles dans le comportement de l'animal, plutôt que les mécanismes cérébraux précis menant à ces comportements. Plus une règle parvient à rendre compte de comportements différents observés chez les animaux, plus elle est intéressante, car plus elle synthétise les principes utilisés par les mécanismes d'apprentissage du cerveau. Dans cette partie, sans en faire une revue exhaustive, ce qui n'est pas l'objet de cette thèse, je présenterai certaines des règles d'apprentissages de la littérature dont les principes ont été repris, modélisés ou discutés, au cours de ma thèse.

Ces règles d'apprentissages ont l'intérêt de mettre en avant les principes sous-jacents à l'apprentissage chez les animaux, elles ne reproduisent pas les mécanismes fins du cerveau, ce qui en particulier entraîne des difficultés à étendre ces règles à de nouveaux paradigmes et comportements. Ces règles d'apprentissages sont comme des fonctions, qui vont d'un ensemble de départ extrêmement grand (toutes les situations possibles) à un ensemble d'arrivée également immense (tous les comportement possibles chez les animaux). Quelle est la chance de trouver une fonction suffisamment proche du fonctionnement des animaux, à partir de quelques exemples, parmi toutes les fonctions pouvant reproduire ces exemples ? En particulier, si un nouveau critère (un nouveau comportement observé dans un nouveau paradigme) est ajouté, rien ne nous

garantit que la règle d'apprentissage trouvée au préalable va vérifier ce nouveau critère, et qu'il ne faudra pas totalement la modifier. Cette difficulté à adapter les règles existantes à de nouveaux comportements entraîne la création de règles de plus en plus complexes, respectant de plus en plus de critères, et donc de plus en plus difficiles à étendre à de nouveaux paradigmes. Ceci est une importante limitation de cette approche purement fonctionnelle.

1.2.1 L'erreur de prédiction

Parmi les premiers modèles d'apprentissage associatif arrivant à reproduire certains des comportements effectués par les animaux sont ceux utilisant un terme d'erreur pour guider l'apprentissage. Ces modèles proposent ainsi d'associer à chaque stimulus S une force d'association V_S , qui décrit à quel point S est un bon prédicteur de l'US, et dont la valeur évolue suivant sa performance pour prédire l'arrivée de l'US. Certains des premiers modèles, comme [Bush et Mosteller \[1951\]](#), proposent un terme d'erreur propre à chaque CS :

$$\Delta V_S = \alpha_S \beta (\lambda - V_S)$$

Où ΔV_S est la variation de V_S à chaque essai, λ l'US reçu à cet essai, α_S représente l'associabilité de l'indice, ici sa saillance lors de l'essai, et β le coefficient d'apprentissage modulant la vitesse à laquelle l'association est formée ou oubliée. Intuitivement, le terme d'erreur $\lambda - V_S$ permet d'acquérir progressivement l'association, car il est positif si V_S ne prédit pas suffisamment la récompense, amenant donc à une augmentation de V_S et donc de la prédiction de récompense au prochain essai. Lorsque V_S prédit totalement l'US, l'erreur est nulle, et l'apprentissage s'arrête.

Cependant, ce modèle fait l'hypothèse que chaque CS apprend en fonction de son erreur de prédiction. En particulier, appliqué au paradigme du blocking tel que décrit dans la partie précédente, ce modèle prédit que le second CS, CS B, quand présenté simultanément avec le CS A ayant déjà acquis l'association, formerait également une association CS B-US, comme son erreur de prédiction, indépendante de celle de CS A, serait élevée. Au contraire, il est observé que les différents CS sont en compétition pour prédire l'US, et lors du blocking, un CS déjà bon prédicteur bloque l'apprentissage des autres CS déjà présents. La règle de Rescorla-Wagner rend compte de ce phénomène [[Rescorla et Wagner, 1972](#)] avec une erreur similaire à celle précédemment utilisée. Dans cette règle, l'apprentissage de l'association propre à chaque CS dépend de l'erreur de prédiction de l'ensemble des CS :

$$\Delta V_S = \alpha_S \beta (\lambda - \sum V_S)$$

Avec $\sum V_S$ la somme des forces d'associations pour chacun des CS présents lors de l'essai. Ainsi, cette règle apprend une association pour chaque CS V_S à

partir de l'erreur de prédiction dans un essai. La prédiction de l'animal dans un essai est elle-même obtenue en sommant les prédictions obtenues à partir des différents CS visibles.

La règle de Rescorla-Wagner permet de rendre compte du phénomène de blocking, car si un ou plusieurs CS prédisent complètement l'US, l'erreur est nulle *quel que soit le CS*, et donc aucun nouveau CS ne peut apprendre. De même, elle explique également l'inhibition conditionnée, en permettant qu'une association négative du CS inhibiteur compense une association positive dans le calcul de l'erreur. Cette règle prédit donc que nous apprenons à partir de notre erreur de prédiction, nous évitant ainsi d'apprendre des associations superflues quand d'autres CS prédisent déjà suffisamment, ainsi que permettant aux différentes prédictions d'éventuellement se compenser quand la situation l'exige. Elle est de plus particulièrement intéressante, car une telle erreur de prédiction a été retrouvée dans de nombreuses parties du cerveau, notamment dans celles impliquées dans l'apprentissage pavlovien [Lee *et al.*, 2013]. En revanche, elle utilise la même erreur de prédiction quel que soit le CS : si deux CS de valeur initiale identique sont présentés en même temps, suivis d'un US, chacun des CS prédit à la fin de l'apprentissage 50% de l'US. Or en réalité, s'il y a bien une compétition, quel CS acquiert une valeur dépend aussi de l'historique, ou de la nature, du CS. Un exemple est l'inhibition latente, dans laquelle un CSA est dans un premier temps présenté seul, non suivi d'US, pendant plusieurs essais. CSA est ensuite couplé à un CSB, et suivi d'un US. La règle de Rescorla-Wagner prédit alors donc que chacun des CS va apprendre à parts égales à prédire l'US. Or en réalité, le CSB va acquérir la quasi-totalité de l'association, l'apprentissage est donc biaisé vers le stimulus nouveau, au détriment du stimulus familier. Plus généralement, ce phénomène montre que l'apprentissage d'association se fait non seulement par rapport à l'US, et de l'erreur de prédiction, mais également en fonction d'indices relatifs aux CS.

1.2.2 Apprentissage dépendant du CS

Parmi les modèles qui mettent en avant un apprentissage dépendant d'indices relatifs aux CS, deux modèles ont particulièrement eu une influence, les modèles de Pearce-Hall et de Mackintosh. Ces deux modèles mettent en avant deux principes différents, principes en apparence contradictoires.

Le modèle de Mackintosh [Mackintosh, 1975] est basé sur l'observation que si à un moment donné, un CS prédit davantage l'US que les autres, ce dernier va apprendre davantage, ce dont les règles basées exclusivement sur une erreur de prédiction ne peuvent rendre compte. Intuitivement, le modèle de Mackintosh peut être traduit par "les meilleurs prédicteurs apprennent plus". Pour ce faire, il utilise une règle similaire au modèle de Bush51, mais avec une associabilité α_A^n variant selon que le CS A est bon prédicteur ou non à l'essai n :

$$\Delta V_A = \alpha_A^n \beta (\lambda^n - V_A)$$

avec β le coefficient d'apprentissage. α_A^n dépend donc de la prédiction de CS A par rapport aux prédictions de tous les autres CS présents lors de l'essai n :

$$\begin{aligned}\Delta\alpha_A^n &< 0 \text{ if } |\lambda^n - V_A^{n-1}| > |\lambda^n - V_X^{n-1}| \\ \Delta\alpha_A^n &> 0 \text{ if } |\lambda^n - V_A^{n-1}| < |\lambda^n - V_X^{n-1}|\end{aligned}$$

avec λ^n la récompense obtenue à l'essai n , V_A^{n-1} l'associabilité du CS A à l'essai précédent n , et V_X^{n-1} la somme des associabilité des CS différents de A. Ces deux équations regardent donc si l'erreur de prédiction basée sur le CS A est plus faible que celle basée sur l'ensemble des autres stimuli, et menant dans ce cas à une plus grande associabilité, et donc à un apprentissage plus rapide du meilleur prédicteur. Cependant, cette version du modèle utilise une erreur de prédiction similaire à celle de Bush51, et est donc incapable d'expliquer des phénomènes comme le blocking ou l'inhibition conditionnée. Ceci a par la suite été corrigé dans une version étendue, [Schmajuk et Moore, 1985], dans laquelle la variation d'association ΔV_S contient un terme prenant en compte la somme des prédictions de tous les CS présents lors de l'essai.

Au contraire, la règle de Pearce-Hall [Pearce et Hall, 1980] est basée sur l'idée que "les stimuli incertains, i.e qui ont été présents dans des essais dans lesquels le modèle a commis une erreur de prédiction, apprennent plus". Ceci permet notamment de rendre compte de phénomènes comme l'inhibition latente, que le modèle de Mackintosh ne peut expliquer. De manière similaire au modèle de Mackintosh, cette règle modifie l'associabilité α_s^n d'un stimulus A dans la règle d'apprentissage :

$$\alpha_A^n = \gamma|\lambda^n - \sum V_{NET}^{n-1}| + (1 - \gamma)\alpha_A^{n-1}$$

γ est un paramètre définissant à quel point l'associabilité dépend des événements se déroulant à l'essai n par rapport aux événements s'étant déroulés aux essais précédents. $\sum V_{NET}^{n-1}$ est un terme correspondant à la prédiction totale du CS lors de l'essai n :

$$\sum V_{NET}^{n-1} = \sum V^{n-1} - \sum \bar{V}^{n-1}$$

avec \bar{V}^{n-1} la valeur de l'association CS-no-US, autrement dit à quel point le stimulus permet de prédire l'absence d'US. Cette association, CS-no-US permet notamment de rendre compte de phénomènes comme l'extinction ou l'inhibition conditionnée, en permettant à certains CS d'apprendre une valeur d'extinction, permettant d'inhiber l'association CS-US apprise à partir de ces mêmes, ou d'autres, CS, sans toutefois l'oublier.

Les modèles de Mackintosh et de Pearce-Hall sont en apparence basés sur des principes contradictoires, expliquant respectivement pourquoi un stimulus ayant déjà une valeur associative forte va avoir tendance à accaparer l'apprentissage, et pourquoi l'inhibition latente et les phénomènes de pré-exposition

modulent l'apprentissage. [Le Pelley \[2004a\]](#) propose un modèle "hybride", conciliant les deux précédents modèles, permettant d'expliquer les expériences dont rend compte chacun des modèles. Dans son modèle, la variation d'association ΔV_A pour un CS A dépend à la fois de l'associabilité de Mackintosh α_A , et de l'associabilité de Pearce-Hall σ_A , telles que décrites plus haut. Si l'erreur de prédiction R est positive, i.e l'apprentissage est excitateur :

$$\Delta V_A = \alpha_A \sigma_A \beta_E (1 - V_A + \overline{V_A}) * R$$

Si R est négative, le modèle apprend à inhiber une acquisition déjà acquise :

$$\Delta V_A = \alpha_A \sigma_A \beta_I (1 + V_A - \overline{V_A}) * R$$

Avec β_E et β_I les coefficients d'apprentissage pour l'apprentissage excitateur et inhibiteur. De manière intéressante, [[Le Pelley, 2004a](#)] impose une priorité dans la façon dont les différentes associabilités affectent l'apprentissage, en bornant l'associabilité de Mackintosh α_S entre 1 et 0.05, tandis que celle de Pearce-Hall σ_S est bornée entre 1 et 0.5. Ainsi le coefficient de Mackintosh peut potentiellement presque supprimer l'apprentissage, quand la règle est déjà prédite par un autre CS, comme par exemple lors du blocking. En revanche, l'associabilité de Pearce-Hall permet d'accélérer l'apprentissage des stimuli incertains par rapport à ceux auxquels on est habitué, mais sans toutefois supprimer leur apprentissage. Ce modèle prouve en particulier que les différentes règles vues précédemment, en apparence contradictoires peuvent coexister dans le cerveau, mais n'explique pas quels mécanismes précis font émerger ces deux associabilités et impactent l'apprentissage. En particulier, la différence d'influence entre les deux associabilités dans la règle précédente suggère que ces deux associabilités influent sur l'apprentissage via des moyens, des mécanismes différents.

Dans cette partie, nous voyons donc que les problèmes auxquels nous nous intéressons dans cette thèse sont des problèmes dont l'expression est relativement simple, mais impliquant des règles d'apprentissage complexes qui mettent en application différentes stratégies pour permettre un apprentissage à la fois robuste et souple, adapté à la situation. Les comportements des animaux peuvent être dans une certaine mesure reproduits par différentes règles d'apprentissages qui permettent de bien mettre en avant certaines des principales stratégies utilisées dans l'apprentissage. Cependant, en raison de la complexité combinatoire du problème, étudier le fonctionnement au niveau de la biologie nous est nécessaire pour aller plus loin, et pour comprendre quels sont les mécanismes exacts impliqués dans les apprentissages associatifs, motivés, et dans leurs interactions. Un autre avantage d'une approche basée sur la compréhension et la modélisation des mécanismes biologiques impliqués dans l'apprentissage est qu'en incluant davantage de contraintes sur le fonctionnement du cerveau, elle est plus à même de généraliser et de prédire le comportement dans de nouvelles circonstances. La comparaison des deux approches peut

cependant permettre de dégager les avantages et inconvénients de chacune, et en particulier de mettre en évidence ce que peut apporter un modèle, certes plus complexe, basé sur les stratégies mises en oeuvre dans notre cerveau. Il est donc temps de nous intéresser au “comment”, comment les mécanismes des apprentissages associatifs et motivés sont implémentés biologiquement dans le cerveau, et comment nous pouvons les modéliser.

Chapitre 2

Etude de la biologie

Cette section est consacrée à l’approche bio-inspirée, à l’étude des mécanismes de la biologie. Mon objectif n’est pas ici de décrire de manière extensive chaque partie du cerveau, mais davantage de donner une idée générale du fonctionnement des apprentissages associatifs et motivés, des parties du cerveau impliquées, et des travaux de modélisations bio-inspirés déjà effectués. Les parties qui ont été étudiées plus en détails dans cette thèse seront par la suite décrites plus précisément, afin d’apporter les informations justifiant nos choix d’implémentation, dans les parties qui leur sont consacrées.

La figure 3.2 représente les principales régions et flots d’informations que je présenterai dans cette partie, et qui jouent un rôle majeur dans l’apprentissage motivé et associatif. Le cortex sensoriel en bleu, et l’hippocampe en vert, ont des apprentissages respectivement perceptifs et épisodiques, et sont présentés dans cette partie en qualité de sources d’informations en entrée, et en interaction, avec les régions responsables des apprentissages associatifs et motivés. L’amygdale, en rouge sur le schéma, est impliquée dans l’apprentissage pavlovien, et de la réponse de peur. Le cortex frontal et les ganglions de la base sont impliqués dans les apprentissages opérants et répondant au niveau des différentes boucles d’apprentissages. Ce schéma met également en évidence les interactions possibles entre les apprentissages opérants et motivés : interactions directes via les connexions directes entre les deux parties, ou indirectes, via des neuromodulateurs comme la dopamine, la noradrénaline ou l’acétylcholine. Les principaux neuromodulateurs seront également décrits dans cette partie, en tant que système d’interactions entre les différents types d’apprentissage du cerveau. Ne sont pas décrits dans cette figure les interactions des apprentissages répondants et opérants vers le cortex et l’hippocampe, interactions directes ou indirectes via la neuromodulation.

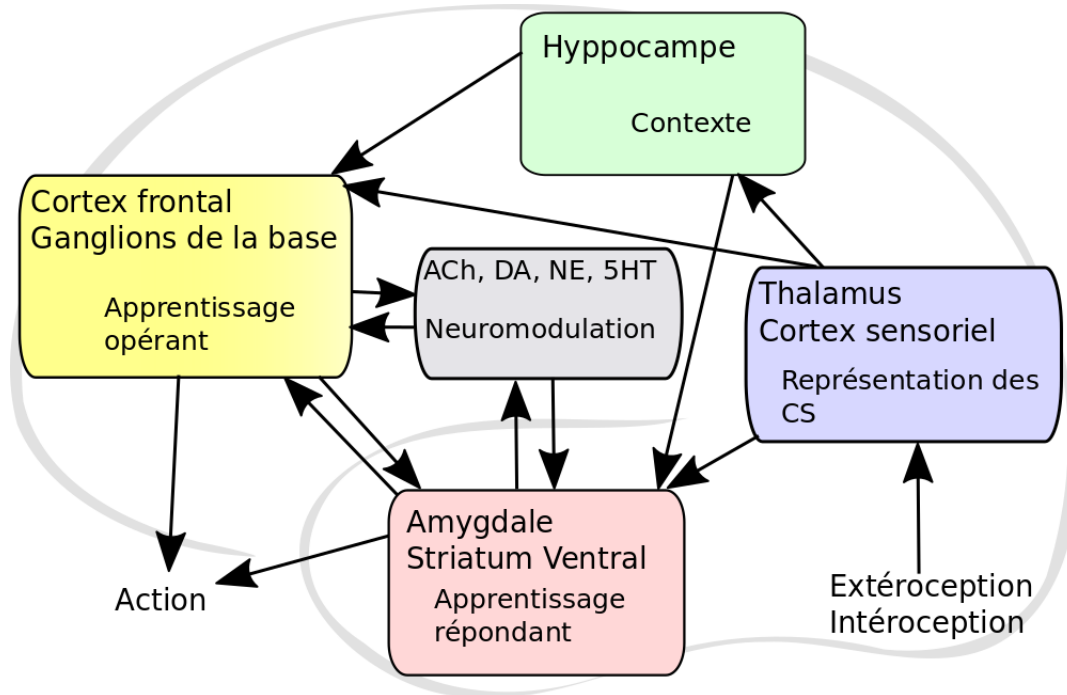


FIGURE 2.1 – Les principaux flux d'information considérés dans l'étude des apprentissages associatifs et motivés. Le cortex sensoriel (en bleu) apprend des représentations de CS complexes sous forme de catégorie. L'hippocampe (en vert) apprend des informations contextuelles. Ces deux régions fournissent ces informations en entrées des régions limbiques comme l'amygdale ou le striatum ventral, impliquées dans l'apprentissage pavlovien (en rouge), et du cortex frontal et des ganglions de la base (en jaune), impliqués dans l'apprentissage opérant, dirigé vers un but, ou dans la formation d'habitudes. Ces deux régions sont respectivement responsables de l'action (réponse pavlovienne ou action motivée). Les différents neuromodulateurs (en gris) interagissent réciproquement avec les différentes régions décrites ici. En rapport avec le sujet de cette thèse, nous nous intéresserons plus particulièrement à leur relations avec les apprentissages motivés et associatifs.

2.1 Le cortex sensoriel et l'hippocampe : des entrées élaborées

En plus des CS basiques (tons élémentaires, orientations visuelles), unimodaux, envoyées aux systèmes associatifs et motivés via le thalamus [Li, 1996; LeDoux, 2000; Bachevalier *et al.*, 1997], des représentations plus sophistiquées sont également fournies aux apprentissages répondant et opérant, principalement via le cortex sensoriel et l'hippocampe.

Le rôle principalement attribué au cortex sensoriel est l'apprentissage de catégories à partir des informations perceptuelles élémentaires, présentes dans les aires primaires [Thorpe et Fabre-Thorpe, 2001]. Chaque modalité sensorielle (ouïe, vision, toucher, odeur, goût) est ainsi constituée d'une hiérarchie d'aires corticales, présentant un gradient dans les catégories apprises : plus

l'aire corticale est élevée dans la hiérarchie, plus elle va encoder des concepts abstraits [Bussey et Saksida, 2002]. Ainsi si les aires primaires encodent des informations telles que les orientations visuelles pour le cortex visuel, ou des fréquences élémentaires pour le cortex auditif, les aires plus élevées peuvent encoder des représentations plus élaborées, telles que des visages [Puce *et al.*, 1996; Haxby *et al.*, 2001] ou des outils [Chao *et al.*, 1999]. Ces différentes modalités, et l'interoception, sont intégrées au niveau du cortex perirhinal, permettant ainsi des représentations multimodales [Burwell *et al.*, 1995]. Les catégories ainsi présentes dans le cortex sensoriel sont envoyées à l'amygdale [LeDoux, 2000], à l'hippocampe via l'entorhinal, et au cortex frontal [Miller et Vogt, 1984].

De très nombreux modèles et théories ont été développés sur le cortex sensoriel, en particulier pour le cortex visuel (voir [Swindale, 1996] pour une revue extensive des premiers modèles). Si un apprentissage non-supervisé a été proposé par de nombreux modèles, basé sur un apprentissage statistique de régularité [Rumelhart *et al.*, 1986; Becker, 1992; Olshausen et Field, 1996], celui-ci n'est pas cohérent avec le fait que les catégories apprises ne sont pas uniquement statistiques, mais sont celles en particulier pertinentes pour le comportement [Sigala et Logothetis, 2002; Baker *et al.*, 2002; Schoups *et al.*, 2001]. L'apprentissage supervisé, et notamment la rétro-propagation d'un terme d'erreur, n'est pas non plus plausible biologiquement, car il nécessiterait à la fois la propagation d'une erreur spécifique des couches les plus élevées aux couches les plus primaires [Zipser et Rumelhart, 1990], ce qui n'est pas observé, et la présence d'un "professeur" qui indiquerait "la bonne" activation de catégories pertinentes.

En revanche, l'apprentissage dans le cortex semble intégrer une composante attentionnelle, qui est nécessaire à l'apprentissage de catégories [Ahissar et Hochstein, 1993; Schoups *et al.*, 2001]. Certains modèles [Roelfsema et van Ooyen, 2005; Pauli et O'Reilly, 2008; Carrere et Alexandre, 2013] reprennent ainsi un signal attentionnel global pour apprendre les représentations pertinentes. Dans ces modèles, l'arrivée d'une erreur dans une tâche provoque l'envoi d'un signal attentionnel proportionnel à l'erreur, qui impacte de manière globale les neurones apprenant des représentations, leur permettant d'apprendre davantage dans ces circonstances. L'apprentissage proposé est d'une certaine façon l'apprentissage statistique non-supervisé, mais uniquement dans les situations nécessitant de l'attention, permettant la représentation de catégories pertinentes à ces situations. En particulier, dans [Carrere et Alexandre, 2013], nous avons proposé que l'apprentissage d'associations CS-US guide l'apprentissage attentionnel du cortex, illustrant ainsi que si le cortex sensoriel est une des entrées de l'apprentissage associatif, les interactions peuvent être réciproques. Ainsi, le cortex sensoriel peut créer et fournir aux systèmes associatifs et motivés des représentations, uni-modales ou non, à la fois complexes, comme des objets ou des sons composés, et pertinentes, ceux auxquels on a prêté attention

par le passé.

Si l'apprentissage du cortex est un apprentissage de généralités, de catégories formées à partir de nombreux exemples différents, l'hippocampe présente au contraire un apprentissage de cas particuliers. En condensant sous forme d'épisode un ensemble d'information de nature variée, en provenance de l'ensemble des zones du cortex sensoriel, l'hippocampe apprend ainsi un épisode, un contexte tel que défini dans les expériences de la partie précédente, et le stocke sous forme de cas particulier. Différentes régions de l'hippocampe sont impliquées dans cet apprentissage, le cortex entorhinal, DG, CA1 et CA3.

L'hippocampe reçoit des informations multimodales de l'ensemble des régions du cortex sensoriel via le cortex perirhinal, puis le cortex entorhinal [Amaral et Witter, 1995], qui sont orthogonalisées dans DG [McNaughton et Nadel, 1990; Andersen P., 2007]. Le processus d'orthogonalisation, ou "séparation de motifs", permet d'obtenir un ensemble d'activités dans DG très distinctes à partir des ensembles d'activités du cortex entorhinal, permettant ainsi de séparer des exemples relativement proches les uns des autres. Cette théorie est en partie soutenue par le très grand nombre de neurones de DG comparé à celui de l'entorhinal (respectivement plus d'un million et 200 000 chez le rat, Andersen P. [2007]). Les informations perceptuelles, une fois orthogonalisées dans DG, sont passées à CA3, qui les stocke au moyen de neurones possédant des connexions récurrentes et peu denses [Amaral et Witter, 1995; Treves et Rolls, 1994]. Cet apprentissage a pour caractéristique d'être rapide, et de permettre une mémorisation sur le long terme [Bliss et Collingridge, 1993; Shapiro et Eichenbaum, 1999]. Les informations stockées dans CA3 peuvent ensuite être retrouvées au moyen de "complétion de motif", i.e trouver l'exemple le plus similaire à celui présenté, et sont ensuite transmises et reconstruites par CA1, qui constitue la sortie de l'hippocampe [O'Reilly et Rudy, 2001], et projette notamment sur l'amygdale, le cortex entorhinal et le cortex préfrontal [Andersen P., 2007; Phelps, 2004; Wang *et al.*, 2016]. L'hippocampe possède également une influence indirecte sur certains de ces systèmes, notamment l'amygdale, via son impact sur la libération de neuromodulateurs tels que l'acétylcholine [Calandrea *et al.*, 2006a].

De manière similaire au cortex visuel, de nombreux modèles ont été effectués du fonctionnement de l'hippocampe, et en particulier du rôle de CA3 en tant que mémoire auto-associative, c'est à dire qui va être capable d'apprendre à associer des épisodes avec eux même, et de faire de la complétion de motif [Hebb, 1950; Marr, 1971; McNaughton et Morris, 1987; McNaughton et Nadel, 1990; Rendeiro *et al.*, 2014]. L'aspect séquentiel de l'apprentissage de l'hippocampe, sa capacité à apprendre un épisode formé de la séquence des événements récemment arrivés, a également été mis en avant par plusieurs travaux de modélisation centrés sur CA3 [Tulving, 2002; Levy, 1989, 1996; Wallenstein et Hasselmo, 1997]. Cependant, les modèles récents mettent en évidence la complexité des rôles relatifs des différentes régions de l'hippocampe, impli-

quant davantage de subtilité dans leur fonctions exactes dans l'apprentissage épisodique [Lisman, 1999; Kassab et Alexandre, 2015], ainsi que la complexité des interactions entre l'hippocampe et ses afférents comme le cortex sensoriel [O'Reilly et Rudy, 2001].

De manière là aussi similaire au cortex sensoriel, nous présentons particulièrement dans cette thèse l'hippocampe en tant qu'entrée des régions impliquées dans l'apprentissage opérant et répondant, mais ces interactions sont en réalité réciproques. Par exemple une des régions impliquées dans l'apprentissage répondant, l'amygdale, peut moduler l'apprentissage de l'hippocampe [Huff et Rudy, 2004; McGaugh, 2004], de manière directe en influençant les activités de l'entorhinal [Richter-Levin, 2004], et le transfert des informations du perirhinal à l'entorhinal [Paz et Paré, 2013], ou de manière indirecte en projetant sur les centres émetteurs d'acétylcholine nécessaires à l'apprentissage hippocampique [Desmedt *et al.*, 2015]. Cette influence permet ainsi de provoquer un apprentissage de contextes pertinents, comme par exemple des épisodes marquants émotionnellement, pour lesquels l'amygdale a exhibé une forte réponse de peur. Également, l'hippocampe a un rôle reconnu dans l'apprentissage du cortex sensoriel, en permettant une consolidation sur le long terme des représentations du cortex durant le sommeil [Marshall et Born, 2007].

2.2 L'amygdale, lieu de l'association CS-US

Différentes régions semblent être impliquées dans le conditionnement répondant, et semblent en particulier dépendre de la réponse demandée. Par exemple, le cervelet a un rôle majeur dans certains cas particuliers d'associations pavloviennes liés au conditionnement du clignement de l'oeil suite à la réception d'un US aversif, comme une émission d'air dirigée vers l'oeil, si l'intervalle de temps entre le CS et l'US est très bref [Steinmetz, 2000]. Pour des intervalles de temps plus long, ou pour des US plus forts, comme lors du conditionnement à la peur, l'amygdale a été montrée comme lieu de l'association CS-US [LeDoux, 2003].

L'amygdale est composée de plusieurs noyaux, qui peuvent être regroupés sous deux catégories, les structures similaires aux structures corticales et celles similaires aux structures striatales. Les structures similaires au cortex sont au nombre de deux, le noyau latéral de l'amygdale (LA), et le noyau basal de l'amygdale (BA), lui même subdivisé en un noyau basolatéral (BLA) et basomédial (BM). Ces structures sont essentiellement composées de neurones excitateurs, avec une minorité d'interneurones inhibiteur [Carlsen, 1988]. Au contraire, le noyau central de l'amygdale (CeA) est une région fortement similaire en structure et connectivité au striatum ventral (partie des ganglions de la base dont nous parlerons plus loin), et en particulier au noyau accumbens et au pallidum [Cassell *et al.*, 1999]. Il est ainsi majoritairement composé de neu-

rones inhibiteurs Gabaergiques, et peut être séparé en deux noyaux, le noyau latéral (CeL) et le noyau médial (CeM), dont les fonctions sont distinctes [Ciocchi *et al.*, 2010]. Trois populations de neurones inhibiteurs GABAergiques sont également intercalés entre ces différents noyaux, logiquement appelées cellules intercalées (ITC, pour “intercalated cells”) [Busti *et al.*, 2011]. Ces populations sont positionnées respectivement adjacent à LA, entre LA/BLA et CeL, et entre BL/BM et CeM.

LA reçoit des entrées du thalamus et de l'ensemble du cortex sensoriel, et est responsable de l'apprentissage de l'association entre le CS sensoriel présent dans le contexte et l'US pour le conditionnement pavlovien [LeDoux, 2000, 2007]. Cette acquisition est robuste, et n'est pas oubliée même lorsque l'animal prédit correctement l'extinction [Maren, 2005; Repa *et al.*, 2001]. L'association présente dans LA est transmise à la fois à BA Lee *et al.* [2013] et à CeL [Ciocchi *et al.*, 2010]. BA est la région de l'amygdale qui s'est le plus étendue lors de l'évolution récente [Cardinal *et al.*, 2002]. Elle intègre à la fois des informations venant de LA, de l'hippocampe, notamment pour le conditionnement au contexte spatial ou temporel [Goossens et Maren, 2001; Biedenkapp et Rudy, 2009; Eichenbaum *et al.*, 2012; Carretero-Guillén *et al.*, 2013a], de la partie prélimbique (PL) [Corcoran et Quirk, 2007] et de la partie infralimbique (IL) du cortex préfrontal médial (mPFC), en particulier pour l'apprentissage de l'extinction [Sierra-Mercado *et al.*, 2011; Wallis, 2007].

BLA semble posséder des populations hétérogènes, avec des neurones calculant différentes valeurs relatives à l'arrivée de l'US. Belova *et al.* [2015] observe en particulier à la fois des neurones réagissant aux récompenses positives, des neurones réagissant aux punitions, et des neurones réagissant indépendamment de la valence de l'US. Herry *et al.* [2008] montre également la présence de neurones encodant l'extinction dans BLA, inhibant les neurones prédisant l'US négatif pour permettre l'extinction. [Schoenbaum *et al.*, 1999] observe également des neurones de BLA susceptibles de renverser leur sélectivité lors d'un changement de règle (ie. devenant sélectif pour le nouveau stimulus récompensé, et ne réagissant plus à l'ancien) plutôt que d'être inhibé par une autre population prédisant la nouvelle règle. BLA projette également des projections sur le cortex Orbitofrontal (OFC) [Corbit *et al.*, 2003], sur PL [Sotres-Bayon *et al.*, 2012], ainsi que sur le striatum ventral [Setlow *et al.*, 2003]. Schoenbaum *et al.* [2003] montre notamment que BLA est nécessaire pour apporter à l'OFC des informations sur la nature de l'US. Ainsi, si LA est plutôt associé à l'acquisition d'une association CS-US robuste, BLA semble davantage liée au calcul de différentes valeurs et associations liées à l'US, permettant plus de souplesse dans le fonctionnement, comme par exemple dans le cas de l'extinction, ou du reversal. De plus BLA possède des échanges réciproques avec différentes régions impliquées dans l'apprentissage motivé telles que OFC et PL, permettant ainsi de donner des informations liées à l'US à l'apprentissage motivé.

Au contraire, CeA est souvent considéré comme le pôle moteur de l'amygdale, déclenchant la réponse pavlovienne de peur, et reçoit des entrées de LA et de BA, éventuellement au travers des ITCs, sur CeL [Ciocchi et al. \[2010\]](#) ainsi que directement sur CeM [\[Lee et al., 2013; Amano et al., 2010\]](#). CeL possède de manière similaire à BLA des neurones réagissant spécifiquement à l'US (CeLOn), ainsi que des neurones réagissant dans ces situations d'extinction (CeLOff) [Ciocchi et al. \[2010\]](#). CeLOff projette sur CeM, inhibant ainsi la réponse de peur, tandis que CeLOn la favorise par l'inhibition de CeLOff.

La réponse de peur, comme le comportement de "freezing", est déclenchée par les projections de CeA sur le periaqueducal grey (PAG) [\[Herry et Johansen, 2014\]](#). Un autre aspect de la réponse pavlovienne déclenchée par CeA est la réponse hormonale, impliquant neuromodulateurs et hormones neuropeptidiques tels que l'acétylcholine ou la dopamine [\[Pessoa, 2010; Lee et al., 2013\]](#), qui influencent différents facteurs communs à l'ensemble du cerveau, tels que le niveau d'attention, le stress, la nouveauté de la situation présente.

L'apprentissage de l'amygdale, de manière similaire à la règle de Rescorla-Wagner, semble dépendre d'une erreur de prédiction présente suite à l'arrivée de l'US. La dopamine, neuromodulateur associé à l'erreur de prédiction, est nécessaire à l'apprentissage de l'amygdale [\[Lamont et Kokkinidis, 1998; Greba et al., 2001\]](#), en permettant la désinhibition des circuits de l'amygdale par son action sur les interneurons [\[Bissiere et al., 2003\]](#) et en excitant directement les neurones pyramidaux excitateurs de LA et BA [\[Lamont et Kokkinidis, 1998\]](#). [Lee et al. \[2013\]](#) propose ainsi que la dopamine guide l'apprentissage par l'élévation du ratio signal-sur-bruit dans LA et BA. Un autre signal contenant l'erreur de prédiction, pour les US aversifs est glutamatergique, provenant d'un retour du PAG sur l'amygdale [\[Herry et Johansen, 2014\]](#).

Différents modèles de l'amygdale ont été développés, et peuvent être regroupés en deux approches : la modélisation de populations spécifiques de l'amygdale, ou l'intégration d'une version plus simple de l'amygdale dans un réseau plus complexe. Certains modèles s'attachent à reproduire précisément les populations spécifiques de certains noyaux de l'amygdale, sans toutefois modéliser les différents circuits auxquels ces populations appartiennent. Par exemple, [Li et al. \[2009\]](#) modélise précisément les interactions entre interneurons inhibiteurs et neurones excitateurs dans LA, dans des paradigmes d'acquisition et d'extinction. Cependant, l'extinction est apprise, au lieu d'être une inhibition dynamique, ce qui est plus plausible au vu du renewal. Au contraire, [Vlachos et al. \[2011\]](#) parvient à reproduire le comportement des neurones de peur et d'extinction exhibés par [\[Herry et al., 2008\]](#) dans BLA par deux populations en inhibition mutuelle, recevant respectivement le CS et le contexte. [Kim et al. \[2013\]](#) modélise également l'aspect temporel des interactions entre neurones de LA lors du "trace conditionning", lorsqu'un délai est inséré entre la fin du CS et le début de l'arrivée de l'US.

Au contraire, une autre approche est la modélisation de l'amygdale dans

une vue plus globale du fonctionnement du cerveau. [Moustafa *et al.*, 2013] modélise l’acquisition de l’extinction contextuelle par les interactions entre l’amygdale et l’OFC, l’extinction étant uniquement apprise dans l’OFC, qui inhibe la réponse d’une amygdale, sans toutefois représenter le rôle de BLA dans l’acquisition de l’extinction (qui n’est en effet pas nécessaire dans certaines conditions Laurent *et al.* [2008]). [Pauli *et al.*, 2011] modélise l’interaction entre l’amygdale et un système de prise de décision, dans lequel l’amygdale est une simple couche de neurone, apportant une prédiction de l’US au réseau, correspondant ainsi au fonctionnement de populations spécifiques à BLA. [Meeter *et al.*, 2005] propose un modèle simple des interactions entre l’hippocampe et un système d’acquisition de la réponse pavlovienne correspondant aux fonctions de l’amygdale, dans lequel l’information contextuelle apprise par l’hippocampe est passée en entrée d’une couche représentant l’amygdale, au même titre qu’un CS.

Ainsi, l’amygdale apprend une réponse pavlovienne basée sur une diversité d’informations (CS, contexte, informations venant des régions impliquées dans l’apprentissage motivé), et ce au niveau de différentes populations, noyaux et circuits. Différents modèles de l’amygdale ont été proposés, et s’attachent à la description soit de populations spécifiques à l’amygdale, soit de la place de l’amygdale dans différents circuits du cerveau. A notre connaissance, aucun ne modélise les différents circuits internes à l’amygdale, pour en différencier les contributions relatives à l’apprentissage associatif. En rapport avec cette thèse, l’amygdale est en interaction réciproque avec des régions du cerveau impliquées dans l’apprentissage motivé, comme IL et PL, lui fournissant le contexte d’extinction [Sierra-Mercado *et al.*, 2011; Wallis, 2007] ou une mémoire de travail nécessaire à une réponse rapide [Do-Monte *et al.*, 2015].

2.3 Le conditionnement opérant, liée au conditionnement ré pondant

L’apprentissage motivé prend essentiellement place dans les interactions entre le cortex frontal, les ganglions de la base et le thalamus, sous forme d’un ensemble de circuits correspondant à des boucles et impliqués dans la motivation, la valeur émotionnelle, l’élaboration de plan et le mouvement [Alexander et Crutcher, 1990; Maia *et al.*, 2008]. Si le nombre exact de boucles varie suivant les auteurs, une division commune est la séparation en trois boucles, limbique (ou affective, motivationnelle), associative, et motrice [Di Martino *et al.*, 2008]. Chacune de ces boucles comprend plusieurs chemins, direct, indirect, et hyperdirect [Mink, 1996]. Le chemin direct, ou “go pathway”, est formé par les connexions du cortex au Striatum (STR), puis de STR au Globus pallidum interne (Gpi), et finalement de Gpi au Thalamus, et déclenche la réponse motrice. Au contraire les chemins indirect, Cortex-STR-Globus pallidus externe

(Gpe)-Gpi-Thalamus et hyperdirect, Cortex-noyau subthalamique (STN)-Gpi-Thalamus ont tous deux un rôle dans l'inhibition de la réponse motrice [Jahfari *et al.*, 2011]. Cependant, dans une approche fonctionnelle, je m'intéresse dans cette section aux différentes informations codées par les différentes boucles, plus qu'à la manière dont la décision est déclenchée par les boucles.

La partie frontale de la boucle limbique inclut le cortex Orbitofrontal (OFC) et le cortex antérieur cingulaire (Acc). OFC est une région en forte interaction avec l'amygdale, l'hippocampe, le cortex sensoriel, et est impliqué dans le calcul de la récompense [Holland et Gallagher, 2004; Roesch et Olson, 2004; Rolls, 2004b]. L'OFC présente une organisation topographique selon le type de récompense, suivant un gradient antéro-postérieur des récompenses les plus abstraites (argent) aux plus concrètes (nourriture, boisson) [Kringelbach et Rolls, 2004]. L'OFC est souvent décrit comme encodant des états émotionnels à partir des stimuli sensoriels [Rolls, 2004a], à partir desquels les décisions peuvent être prises. Les projections de l'amygdale sur l'OFC permettent en particulier de fournir des informations relatives à l'US pour la prise de décision [Holland et Gallagher, 2004]. De manière similaire, Acc est connecté à l'amygdale, au cortex sensoriel, et à l'hippocampe, et projette sur le striatum [Stevens *et al.*, 2011]. Acc est notamment impliqué dans la détection ou la désambiguation de conflit [Carter *et al.*, 1998; Cardinal *et al.*, 2002] ou de nouveauté grâce aux projections de l'hippocampe [Krebs *et al.*, 2013]. L'OFC est ainsi plus lié à l'encodage de la valeur des stimuli extérieurs, tandis qu'Acc les met en relation avec l'interoception [Bechara et Naqvi, 2004]. OFC projette notamment sur le shell du noyau accumbens (NAcc), région du striatum ventral, tandis qu'Acc projette sur le core de NAcc [Cardinal *et al.*, 2002].

Le striatum ventral est impliqué dans la prédiction de l'arrivée de récompense pour le calcul de l'erreur de prédiction [Schultz *et al.*, 2000; Tanaka *et al.*, 2004]. Il est ainsi impliqué dans des fonctions pavloviennes similaires à celles de l'amygdale, fait souligné par de très fortes similitudes structurales entre l'amygdale et Nacc [Cassell *et al.*, 1999], en particulier du shell du noyau accumbens, situé dans la continuité spatiale de CeA dans le cerveau. Il a également de fortes projections sur le système dopaminergique impliqué dans l'erreur de prédiction, et en particulier l'aire tegmentale ventrale (VTA) et la partie dorsale de la substantia nigra pars compacta (SNc) [Lynd-Balta et Haber, 1994]. Si le striatum ventral est hétérogène, possédant de nombreuses régions (voir [Haber, 2011] pour une revue extensive), une dissociation est observée entre le shell et le core du noyau accumbens Penner et Mizumori [2012]; Cardinal *et al.* [2002]. Si l'amygdale projette à la fois sur le shell et le core, l'hippocampe et le cortex préfrontal ne projettent que sur le shell, tandis qu'Acc ne projette spécifiquement pas sur le shell [Haber, 2011]. De plus, le shell reçoit aussi des entrées dopaminergiques d'une région spécifique de VTA, au contraire du reste de VS qui reçoit de l'intégralité de la partie dorsale de VTA/SNc [Haber, 2011].

Ces distinctions anatomiques mettent en évidence deux fonctions séparées à l'intérieur de la boucle limbique : une fonction associative, plus pavlovienne, apprenant la valeur des états sensoriels, dans la boucle OFC/vmPFC-shell-Thal et une fonction plus motivationnelle, Acc-core-thal, mettant en relation les buts et les besoins. Dans cette vision, la première répondrait à la question "quelle est la valeur extéroceptive, dans l'absolu, liée à l'US, de la situation", tandis que la deuxième répondrait plutôt à "quelle est sa valeur par rapport à mon état interne, par rapport à mes besoins". [Cardinal et al. \[2002\]](#) met également en évidence une différence de comportement, dans laquelle le circuit impliquant le core est impliqué dans la réponse préparatoire (préparant l'arrivée d'une récompense non immédiate dans le temps, et donc nécessitant une motivation), tandis que le shell est impliqué dans une réponse consommatrice (l'arrivée de récompense est proche dans le temps, la priorité est donc de connaître sa valeur pour la consommer).

La boucle associative implique majoritairement le cortex préfrontal latéral (IPFC) et dorso-latéral (dlPFC, ou prélimbique (PL) chez les rongeurs), ainsi que les projections du cortex préfrontal médial (mPFC) et le striatum dorso-médial (dmS). Si elle possède, comme la boucle limbique, des interactions réciproques avec le système dopaminergique, elle est uniquement en interaction avec SNc, et pas VTA [[Lynd-Balta et Haber, 1994](#)]. Elle ne reçoit également pas d'entrées striatales venant de l'amygdale ou de l'hippocampe [[Haber, 2011](#)]. Cette boucle est impliquée dans l'élaboration de comportement pour atteindre un but [[Balleine et al., 2009](#)]. Des lésions de dmS empêchent l'élaboration de nouvelles stratégies, tandis que des lésions de PL gênent spécifiquement le changement de stratégie, mais pas le reversal [Ragozzino \[2007\]](#). Ainsi, au contraire de la boucle limbique, la boucle associative n'est pas impliquée dans le calcul de la valeur de stimuli, mais dans le choix de l'action et de la stratégie à effectuer pour atteindre le choix effectué par la boucle limbique. Des lésions de la boucle associative n'affectent pas la performance dans des tâches composées de simples associations stimulus-réponse [[White, 2009](#)], mais la détériore sur des paradigmes plus complexes [[Adams et al., 2001](#); [Feathers-tone et McDonald, 2005](#)]. [Tanaka et al. \[2004\]](#) montre également l'implication de cette boucle lorsque la récompense à obtenir est loin dans le futur, et son absence d'implication lorsque la récompense est immédiate. Ainsi la boucle associative est responsable de la création de stratégies complexes et flexibles à partir des buts définis par la boucle limbique, mais pas des actions élémentaires nécessaires pour atteindre ces buts.

Au contraire, la boucle motrice connecte le cortex moteur et sensorimoteur au striatum dorso-lateral (dlS), et est, comme son nom l'indique, impliquée dans l'acquisition de séquences motrices et d'associations directes stimulus-réponses [[Balleine et al., 2009](#)]. La réponse des neurones de dlS est corrélée aux comportements moteurs, et ce même tard lors de l'entraînement (surentraînement) [[Barnes et al., 2005](#); [Yin et al., 2009](#)]. Les stratégies et comportements

élaborées par la boucle associative seraient ainsi consolidés sous formes d'habitudes et contrôlés par cette boucle, tard dans l'apprentissage [Tricomi *et al.*, 2009].

Une vision synthétique des boucles cortico-striatales impliquées dans la prise de décision, telles que décrites dans cette section, seraient que la boucle est impliquée dans le choix du but, à la fois par sa valeur absolue (la valeur de l'US reçu) et par sa valeur relative (relative à nos besoins et motivations). La boucle associative sélectionne ou élabore des stratégies pour atteindre ce but, et la boucle motrice effectue les actions correspondantes. Au fur et à mesure de l'apprentissage, ces stratégies peuvent être consolidées en habitudes, dans un but d'économie des ressources corticales, et déclenchées directement par la boucle motrice. Il se dégage de cet état de l'art que la partie limbique est particulièrement en relation avec cette thèse, car en interaction réciproque avec l'amygdale, lieu de l'apprentissage associatif, et possédant des similarités structurelles et fonctionnelles.

2.4 Le rôle de la neuromodulation

La neuromodulation est un principe d'activité neuronale dans lequel une population de neurones influence le fonctionnement et l'apprentissage d'un grand nombre de régions du cerveau par la libération de transmetteur, le neuromodulateur. Comme son nom l'indique, son rôle est essentiellement modulateur, c'est à dire qu'il ne va pas générer d'activité par lui même, mais va au contraire influencer une activation existante en la facilitant, ou en l'inhibant [Friston, 2002]. Cette modulation se fait au moyen de récepteurs spécifiques aux différents neuromodulateurs sur les neurones [Dayan, 2012]. En particulier, suivant le récepteur présent sur un neurone, un même neuromodulateur pourra avoir un effet différent sur ce neurone (par exemple pour la dopamine les récepteurs D1 excitateurs et D2 inhibiteurs). Pour chaque neuromodulateur, il est nécessaire de faire la différence entre le comportement phasique (libération à faible échelle de temps, par exemple suite à l'arrivée d'un CS ou US) marquant précisément dans le temps un certain évènement, et tonique (libération soutenue sur une plus grande période de temps), qui correspond davantage à un état et à un mode général d'apprentissage [Dayan, 2012]. Si nous évoquons certains aspects phasiques des neuromodulateurs, nous nous intéresserons particulièrement à leurs effets toniques, car permettant de moduler l'importance des différents composants impliqués dans l'apprentissage, menant ainsi à des apprentissages différents.

Quatre neuromodulateurs principaux ont eu leur rôle fonctionnel mis en avant par Doya [2002], la dopamine (DA), la sérotonine (5HT), la noradrénaline (NE) et l'acétylcholine (ACh). Doya [2002] propose quatre rôles distincts pour chacun des neuromodulateurs : la dopamine représente l'erreur de prédic-

tion, la sérotonine la durée de la prédiction de récompense, la noradrénaline le caractère aléatoire de la prédiction, et l'acétylcholine la vitesse de mise à jour des représentations en mémoire. Yu et Dayan, [Yu et Dayan, 2005], ont par la suite proposé des rôles similaires pour l'acétylcholine et la noradrénaline, en rapport avec différents types d'incertitude. En particulier, le rôle d'ACh sur l'apprentissage en mémoire et son rapport à l'incertitude ne sont pas forcément contradictoires, comme par exemple discuté précédemment : certains apprentissages corticaux ou hippocampiques peuvent dépendre d'incertitudes liées à la situation pour apprendre les représentations adaptées, et ainsi potentiellement aider les systèmes associatifs et opérants.

Yu et Dayan [2005] propose ainsi un rôle attentionnel à l'acétylcholine, représentant l'incertitude attendue, i.e. lié à une règle connue, mais stochastique. C'est le cas lors de la première présentation d'un stimulus ("je ne connais pas donc je sais que je suis incertain"), et lorsque la règle apprise présente une incertitude sur le long terme, liée à un caractère aléatoire de l'arrivée de récompense dans cette règle (par exemple si l'arrivée d'US n'arrive que dans 50% des cas où le modèle choisit une certaine réponse). L'acétylcholine est libérée par plusieurs structures du basal forebrain : le noyau basal, le septum médian et la bande de Broca, et impacte une grande variété de structures corticales, hippocampiques et limbiques [Everitt et Robbins, 1997a].

- L'effet observé de l'acétylcholine sur le cortex, via le noyau basal, est attentionnel, et augmente le ratio signal-sur-bruit [Pauli et O'Reilly, 2008], permettant d'augmenter l'impact des informations venant des aires primaires par rapport au retour des structures contenant des représentations plus élaborées, facilitant ainsi la création de nouvelles catégories.
- L'effet observé de l'acétylcholine sur l'hippocampe via le septum médian et sur l'amygdale [Everitt et Robbins, 1997a], est de favoriser l'apprentissage contextuel dans l'hippocampe et la formation de règle contexte-US dans l'amygdale, au détriment de la règle CS-US [Hasselmo, 2006; Calandreau *et al.*, 2006a]. Ceci peut être mis en relation avec un rôle également facilitateur de ACh, comme l'acétylcholine n'est reçue que par la partie de l'amygdale recevant les informations contextuelles, BLA.
- L'effet observé de l'acétylcholine émise par la bande de Broca, notamment au moyen de ses projections sur des structures du cortex frontal telles qu'Acc, est la capacité à discriminer des règles à partir de condition visuelles [Everitt et Robbins, 1997a]. Ceci est donc également cohérent avec un effet facilitateur d'ACh sur les zones corticales, ici sur Acc, qui est précisément impliqué dans la gestion de conflit.

La noradrénaline (ou norépinéphrine, NE), est libérée principalement par le Locus Coeruleus (LC) [Aston-Jones et Cohen, 2005]. Le rôle phasique de la noradrénaline apparaît proche de l'acétylcholine pour prêter attention aux CS marquants : LC répond phasiquement aux CS [Rajkowski *et al.*, 1994],

via les projections de CeM sur LC lors de l'arrivée d'un CS [Bouret *et al.*, 2003]. Son rôle tonique a été associé par [Yu et Dayan, 2005] à l'incertitude inattendue, c'est à dire l'incertitude lorsqu'une règle connue est modifiée soudainement. Une augmentation rapide de NE est par exemple observée lors du reversal, lorsqu'une règle précédemment établie est soudainement modifiée [Aston-Jones *et al.*, 1997]. McClure *et al.* [2006] propose un modèle de NE tonique dans lequel l'incertitude inattendue est détectée en mesurant le conflit entre deux fenêtres de temps stockant respectivement un historique récent et à long terme des récompenses. Ainsi si l'historique à long terme est différent de l'historique récent, c'est que la règle a soudainement changé, correspondant bien à l'incertitude inattendue. Ce rôle tonique, lié à la détection de conflit, dépend en grande partie des projections du cortex préfrontal sur LC [Jodo *et al.*, 1998]. Acc a notamment été observé comme jouant un rôle important dans la libération de NE tonique, en particulier lié à la nouveauté [Krebs *et al.*, 2013].

Le niveau tonique de NE est également corrélé à un niveau d'éveil ou d'intérêt ("arousal") [Rajkowski *et al.*, 1994]. En accord avec cette vision, il a été montré que NE provoquait une augmentation de l'exploration en permettant de prêter attention à de nouveaux indices sensoriels, de manière similaire à ACh [Aston-Jones et Cohen, 2005]. Également, des animaux avec une détérioration de NE dans le circuit cortex-striatum-thalamus exhibent des comportements répétitifs [Gilzenrat *et al.*, 2002].

Plusieurs mécanismes précis, non mutuellement exclusifs, peuvent être proposés pour cette exploration :

- Tait *et al.* [2007] montre que des lésions des projections de LC sur le cortex préfrontal médial détériorent le reversal lorsque celui-ci est extra-dimensionnel, mais pas intra-dimensionnel. Ainsi, le rôle exploratoire de NE dans le cortex pourrait être de porter l'attention sur des dimensions ou représentations différentes, notamment par la facilitation de nouvelles représentations.
- Un autre effet de la noradrénaline est son effet inhibiteur sur le shell via les récepteurs alpha-adrenergiques [Nicola et Malenka, 1998], qui pourrait permettre l'exploration par l'inhibition des valeurs apprises dans le striatum des différents états corticaux.
- Un troisième mécanisme potentiel déclenché par NE est son rôle sur la synchronisation des fréquences d'activations entre les différentes régions des boucles cortico-striatales [Gilzenrat *et al.*, 2002]. Une augmentation de NE diminuerait ainsi la cohérence entre les différentes parties de la boucle, détériorant ainsi la récupération des valeurs apprises au préalable, facilitant ainsi de nouvelles décisions et un comportement plus exploratoire.

Un phénomène que nous ne regardons pas, comme nous nous concentrons sur l'effet direct des neuromodulateurs sur les systèmes responsables des apprentissages ré pondant et opérant, mais néanmoins présent, est l'interaction entre

les différentes régions productrices de neuromodulateurs. NE a par exemple un effet inhibiteur sur les centres libérant l'acétylcholine [Vizi, 1980]. Mis en perspective avec les résultats de Yu et Dayan [2005], ceci pourrait suggérer que l'acétylcholine représente l'incertitude en général, mais est inhibée lors de l'incertitude inattendue par la noradrénaline, et n'est donc au final libérée que lors de l'incertitude attendue telle que définie par Yu et Dayan [2005].

Les deux autres neuromodulateurs, la sérotonine (5HT) et la dopamine (DA), ont un rôle plus controversé. Certains modèles comme Daw *et al.* [2002] proposent que la sérotonine tonique soit impliquée dans la prédiction du taux moyen de récompense appétitive, et que la dopamine tonique ait un rôle équivalent pour les récompenses aversives, ou punitions. Des modèles plus récents incluent cependant un rôle de 5HT dans le conditionnement aversif [Cools *et al.*, 2010], bien que restreint à l'inhibition du comportement et l'évitement passive de punition (éviter une punition en ne déclenchant pas d'action), ainsi qu'un rôle de la dopamine sur le taux de récompense appétitif Niv *et al.* [2007], reproduisant ainsi l'effet de DA sur la vigueur de la réponse.

La sérotonine est libérée par la partie dorsale du noyau du raphé [Fuxe, 1965]. Différents effets de 5HT ont été observés sur la prise de décision. 5HT a notamment un effet sur la prise de risque, une diminution de sérotonine entraîne une préférence pour des choix plus risqués, et facilite le passage d'un choix moins risqué à un choix plus risqué [Long *et al.*, 2009]. Ainsi la sérotonine augmenterait la valeur subjective du choix le moins risqué dans une prise de décision. [Doya, 2002] expose aussi le rôle de 5HT lors du choix entre récompense immédiate et récompense plus importante mais arrivant après un délai. La présence de 5HT facilite alors le second choix, et est ainsi corrélée à la "patience" de l'individu. Tanaka *et al.* [2007] montre que le mécanisme sous-jacent pourrait être l'équilibre entre les différentes boucles cortico-striatales dans la prise de décision, en observant que de faibles niveaux de sérotonine sont corrélés dans VS à des prédictions pour les récompenses arrivant rapidement, et que de hauts niveaux de sérotonine sont corrélés dans DS à des prédictions pour les récompenses arrivant plus tardivement. Balasubramani *et al.* [2014] propose un modèle de 5HT reproduisant les différents effets mentionnés ci-dessus, dans lequel 5HT contrôle l'équilibre entre la valeur de l'action et le risque de celle-ci.

La dopamine est générée par l'aire tegmentale ventrale (VTA) et par la substantia nigra pars compacta (SNc) [Schultz *et al.*, 1997a]. La dopamine phasique envoyée de VTA à Nacc, l'amygdale et les aires préfrontales est corrélée à l'erreur de prédiction, pour apprendre à prédire l'arrivée de récompenses. La dopamine phasique envoyée de SNc au striatum dorsal semble plus impliquée dans l'action, comme la modification d'actions habituelles pour maximiser la récompense [Niv *et al.*, 2007; Daw *et al.*, 2005; Packard et Knowlton, 2002]. Au niveau tonique, la dopamine est associée au compromis entre exploration et exploitation [Humphries *et al.*, 2012]. Contrairement au rôle de NE dans

ce même paradigme, la dopamine impacte différentes régions, en particulier des régions plus motrices. [Humphries et al. \[2012\]](#) modélise ainsi l'effet de DA par une action sur la sortie de BG, dans une région motrice responsable de la sélection de l'action. [Berridge et Robinson \[1998\]](#) proposent également un rôle de la dopamine davantage orienté vers l'action, avec un effet sur notre motivation à choisir une récompense, mais pas sur la valeur intrinsèque que nous prêtons à cette récompense. [[Nicola et Malenka, 1998](#)] En accord avec la différence d'effet entre la dopamine émise, par VTA, dans NAcc et celle, émise par SNc, dans le striatum dorsal [Wise \[2004\]](#), il est également possible que la dopamine ait un rôle à la fois similaire à NE pour une exploration sensorielle ([Nicola et Malenka \[1998\]](#) observe un effet similaire de DA et de NE toniques dans le shell), et un rôle plus orienté vers une exploration motrice, motivationnelle, comme modélisé par [Humphries et al. \[2012\]](#), par la dopamine émise par SNc.

VTA et SNc reçoivent tous deux des projections de l'amygdale, et ont ainsi été regardés comme moyen pour l'apprentissage associatif de transmettre des valeurs à l'apprentissage motivé dans le cas du PIT [El-Amamy et Holland \[2007\]](#). [El-Amamy et Holland \[2007\]](#) observe que couper les projections entre l'amygdale et SNc empêche le PIT. Cependant, ceci n'implique pas nécessairement que le PIT passe directement par les projections entre l'amygdale et SNc. Il est également possible que couper les projections entre l'amygdale et SNc détériore le fonctionnement de SNc, et celui des boucles cortico-striatales recevant l'erreur de prédiction de SNc, qui elles-mêmes seraient impliquées dans le PIT.

Cette courte revue des différents neuromodulateurs montre que leurs effets locaux peuvent être similaires, et dépendent des récepteurs, pouvant entraîner une facilitation de l'activation, de la plasticité, ou de l'inhibition. Ces neuromodulateurs ont une influence sur les neurones d'un grand nombre de régions impliquées dans les apprentissages associatifs et motivés. Cependant, leurs effets sur le comportement et la performance dans les paradigmes diffèrent d'un neuromodulateur à un autre, majoritairement pour deux raisons. D'une part ces neuromodulateurs impactent différents ensembles de régions et d'informations, et d'autre part ces neuromodulateurs encodent différentes informations, et sont ainsi libérés dans différentes circonstances. Bien comprendre et représenter ces deux différences est essentiel pour mettre en évidence l'intérêt de la neuromodulation dans les différents apprentissage.

Le cortex sensoriel et l'hippocampe permettent donc de fournir des informations complexes, respectivement des concepts plus ou moins abstraits, et des contextes au systèmes associatifs et motivés. L'amygdale apparaît dans cet état de l'art comme le lieu majeur de l'apprentissage pavlovien, ainsi qu'également une partie de la boucle limbique incluant le shell du noyau accumbens. Les boucles cortico-striatales, et leurs interactions, sont le lieu de l'apprentissage

opérant, en utilisant les valeurs des différents buts obtenus par le système limbique pour définir des stratégies et effectuer les actions nécessaires à l'obtention du but. Les neuromodulateurs impactent l'ensemble des régions mentionnées ci-dessus, modulant leur activités et leurs modes de fonctionnement respectifs. Les interactions entre l'amygdale, la boucle limbique, et les boucles associatives et motrices, que ce soit interactions directes ou indirectes par la neuromodulation, inscrites dans une vision systémique du cerveau, semblent former le substrat biologique majeur du sujet de cette thèse.

Si ces différentes régions ont toutes fait l'objet de nombreux travaux de modélisation, certaines questions restent ouvertes, en particulier liées à la façon dont ses régions et leurs apprentissages interagissent les unes avec les autres. Appliquée à notre sujet, une des questions auxquels nous tenterons particulièrement d'apporter une réponse est liée à la modélisation de l'amygdale, et des compétitions entre les populations des différents noyaux, qui n'ont pas été modélisés jusqu'à présent. Mis en relation avec notre problématique, ceci nous permettra d'étudier et de mettre en avant certaines des stratégies permettant à l'amygdale d'apprendre en fonction de certaines de ces entrées plutôt que d'autres, de manière pertinente à la population. Une autre question ouverte est l'intérêt de la neuromodulation tonique dans cette compétition. En particulier, nous tenterons d'explicitier l'effet des neuromodulateurs dans une vision systémique, et de mettre en avant l'effet de la neuromodulation au niveau des interactions entre systèmes d'apprentissages. Nous nous intéresserons ainsi particulièrement à comment l'acétylcholine peut influencer l'apprentissage d'associations pavloviennes à partir d'information fournies par l'apprentissage motivé, et comment la noradrénaline tonique peut influencer la façon dont l'apprentissage pavlovien transmet les valeurs qu'il a apprises à l'apprentissage opérant.

Chapitre 3

Problématique et outils utilisés

3.1 Problématique

Un fait qui apparaît à la fois en étudiant les différents apprentissages au niveau du comportement, et au niveau de la biologie, est la diversité des mécanismes sous-jacents. Au niveau comportemental, cette diversité se traduit par une complexité de réponse, de comportements à reproduire, dans des paradigmes pourtant en apparence simples. Au niveau biologique, cette diversité se traduit par la diversité des structures et mécanismes impliqués, ainsi que par la diversité des informations représentées. Il nous faut donc faire un choix précis sur l'objet de cette thèse. A quelles mécanismes précis, quelles interactions spécifiques, s'intéresser pour mieux appréhender la façon dont apprentissages répondants et opérants sont combinés dans le cerveau ?

Une première réponse peut être donnée par l'étude des boucles limbiques et de l'amygdale, qui nous montre que certains paradigmes simples, comme l'association directe stimulus-réponse, ou l'association CS-US simple, peuvent être appris respectivement par la boucle motrice seule, ou par l'amygdale ou le cervelet seul. L'étude de ces paradigmes ne permet donc pas l'étude des interactions entre apprentissages associatif et motivés. Quels paradigmes précis étudier ? Comme présenté en introduction, un des intérêts de l'étude des apprentissages répondants et opérants dans le cerveau est qu'ils permettent l'autonomie, la flexibilité face aux différentes situations. Dans cette optique, il est préférable d'étudier les paradigmes mettant en évidence cette flexibilité. Deux paradigmes sont ainsi particulièrement intéressants : l'extinction et le reversal. De plus, comme vu précédemment, ces paradigmes impliquent tous deux à la fois des régions impliqués dans l'apprentissage répondant, comme l'amygdale et le striatum ventral, et des zones impliquées dans l'apprentissage opérant, l'OFC et les boucles associatives et limbiques.

De manière non contradictoire, l'acétylcholine et la norépinéphrine, représentant respectivement l'incertitude attendue et inattendue, sont deux neuro-modulateurs dont les niveaux toniques sont particulièrement impliquées dans

la capacité d'adaptation aux différentes situations. Comme vu précédemment, NE en particulier, est impliquée dans le reversal, en permettant de détecter le changement de règle, et permet une adaptation plus rapide, en augmentant l'exploration, pour permettre de plus rapidement trouver et apprendre la nouvelle règle. ACh a un effet sur l'extinction dans l'amygdale, en agissant précisément sur la partie des circuits de l'amygdale apprenant l'extinction.

Ainsi, la problématique de cette thèse, inscrite dans l'étude des interactions entre apprentissage associatifs et motivés, sera plus précisément d'étudier comment les interactions entre apprentissage opérant et répondants permettent la flexibilité du comportement face à des changements, comme ce peut être le cas lors de l'extinction, ou du reversal.

Je propose de répondre à cette question, dans une approche systémique, par l'étude de différentes interactions. Un premier modèle sera consacré à l'apprentissage pavlovien de l'amygdale, et montrera en particulier comment les différents circuits de l'amygdale apprennent l'extinction contextuelle en fonction d'informations fournies par des structures impliquées dans l'apprentissage opérant. Ce modèle s'attachera en particulier à regarder comment les différents circuits de l'amygdale permettent un apprentissage flexible, et adapté à la situation, ainsi que le rôle majeur de l'acétylcholine dans celui-ci.

Un second modèle sera dédié aux boucles limbiques, associatives et motrices, et montrera comment les interactions entre celles-ci permettent de s'adapter au reversal. Il s'intéressera également au rôle de l'habituation dans le reversal, en montrant comment la formation d'habitudes permet des décisions rapides dans les cas fortement appris, et comment la boucle limbique, effectuant un apprentissage pavlovien similaire à celui de l'amygdale, est impliquée dans le calcul et mise à jour de valeurs nécessaires au reversal de l'action.

Finalement, un troisième modèle sera plus précisément centré sur la Noradrénaline, neuromodulateur détectant précisément le changement de situation. Il s'attachera à modéliser comment NE est calculée, et ses effets sur les boucles de prise de décision du modèle précédent, permettant de réguler le compromis entre exploitation et exploration. En particulier, nous verrons que ce compromis peut se faire par la modulation des valeurs apprises par l'apprentissage pavlovien et utilisées par l'apprentissage motivé pour définir un but.

Pris ensembles, ces trois modèles permettent de mettre en évidence certains des mécanismes permettant notre flexibilité et notre adaptation lors de changement de l'environnement. De manière plus générale, ils mettent en évidence un des intérêts possibles de la combinaisons de systèmes d'apprentissages.

3.2 Des modèles sous contraintes biologiques

L'objectif de cette section est de brièvement présenter les modèles de neurones et logiciels utilisés lors de cette thèse, et de présenter les différents choix

d'implémentation qui ont été effectués pour garantir une modélisation proche de la biologie, tout en restant simple, permettant ainsi de mieux mettre en avant la vision fonctionnelle utilisée dans les modèles.

3.2.1 Modèle de neurones utilisé

Les neurones communiquent entre eux (de manière directe, la neuromodulation est évidemment un moyen de communiquer) en générant des impulsions électriques, ou spikes. S'il existe une très grande variété de modèles de neurones, ils peuvent être regroupés en deux approches : ceux utilisant le taux moyen d'émission de spikes pour encoder l'information, ou "rate coding", ou ceux utilisant le timing précis des différents spikes pour encoder l'information entre deux neurones, ou "temporal coding". Dans l'optique de cette thèse, qui veut expliquer par une vision systémique les interactions entre les différents systèmes, utiliser l'un ou l'autre est équivalent, car nous voulons mettre en évidence des phénomènes d'apprentissages qui émergent par la nature des informations codées par les différentes populations, plus que par la façon dont ces informations sont encodées au niveau de chaque neurone.

J'utilise dans les modèles de cette thèse le "rate coding", car présentant usuellement des équations plus simples, et donc à la fois plus efficaces en terme de calcul, et possédant un "pouvoir explicatif" plus fort. Comme la règle d'équation de neurones utilisée est simple, les comportements émergents observés sont plus facilement interprétables en terme d'information encodée. A l'opposé, comme nous cherchons à modéliser la biologie, il est toutefois nécessaire que nos modèles soient cohérents avec les mécanismes biologiques.

Ainsi une unité (ou neurone, mais c'est un abus de langage) des modèles que j'ai réalisés représente le comportement moyen d'une petite population de neurones, ou le taux de décharge moyen U est fonction du potentiel de membrane moyen V , lui-même diminuant spontanément au cours du temps, et augmentant en fonction de l'entrée I fournie par les différents neurones afférents suivant une équation de la forme :

$$\frac{dV}{dt} = (-V + F(I))/\tau$$

$$U = G(V)$$

Avec F et G deux fonctions, souvent sigmoïdes ou fonctions de seuillage, représentant respectivement la résistance en entrée (résistance des synapses à faire passer l'information) et le niveau nécessaire pour déclencher une décharge du neurone. τ est la constante de temps définissant la vitesse d'évolution de V par unité de temps. A cette forme d'équation générale, l'effet de phénomènes

précis, comme l'impact de neuromodulateur ou d'inhibition, pourra être ajoutés, et sera repris dans les parties concernées.

3.2.2 Apprentissage Hebbien

Donald O. Hebb est connu, du moins dans le milieu des modélisateurs de réseaux neuronaux, pour avoir été le premier à exhiber de façon explicite, en 1949, une règle physiologique sur la modification des synapses [Hebb, 1950]. Cette règle par la suite portera son nom, la règle de Hebb, ou apprentissage hebbien.

Elle établit que quand une cellule A, connectée à une cellule B, prend de manière répétée part à l'activation de B, un processus de croissance, ou métabolique, prend place dans l'une ou l'autre de ces cellules, de façon à augmenter l'efficacité de A à activer B. Elle est souvent généralisée de la façon suivante, avec w_{AB} la connexion entre deux neurones A et B d'activités respectives x_A et x_B :

$$\Delta w_{AB} = k * x_A * x_B$$

K est une constante, le taux d'apprentissage, modulant la vitesse d'évolution de la connexion w_{AB} .

Une observation immédiate est que cette règle est instable, quelle que soit la représentation de neurones utilisée. Plus la connexion prend part à l'activation de B, plus elle va être renforcée, et donc plus elle va pouvoir déclencher l'activation de B. Ainsi si le principe de la règle de Hebb est souvent utilisé, son implémentation en pratique est différente. Un des exemples vus précédemment est la règle de Rescorla-Wagner, dans laquelle le poids de la connexion associé à un CS dépend de la valeur de ce CS (activité présynaptique), et de la valeur de récompense prédite (activité postsynaptique). L'introduction d'une erreur de prédiction permet de diminuer la force de la connexion lorsque celle-ci augmente trop, comme le modèle prédira trop l'US, et permet ainsi de garantir la stabilité (pour un taux d'apprentissage suffisamment faible). La règle de Hebb, ainsi que celle de Rescorla-Wagner, serviront de bases à la totalité des règles d'apprentissages synaptiques implémentées dans cette thèse.

3.3 Principales bibliothèques et logiciels.

Je n'ai pas pour but de mentionner ici chaque bibliothèque ou programme que j'ai utilisé durant cette thèse, pour la gestion, le stockage ou la visualisation des données, des expériences ou des résultats. Au contraire, je ne présente ici que les bibliothèques et logiciels sélectionnés ayant un rôle important sur le fonctionnement des modèles ou l'application des modèles, et les raisons de ces choix.

3.3.1 DANA

Les modèles que je présente ici sont implémentés en python, et utilisent la bibliothèque DANA (acronyme pour Distributed, Asynchronous, Numerical and Adaptative computing) [Rougier et Fix, 2012]. Le python a été choisi pour sa simplicité d'utilisation, dans l'optique où les réseaux modélisés peuvent être utilisés par des personnes issues de sciences non directement liées à l'informatique, et ont donc besoin d'un langage de programmation facile d'utilisation (par exemple non compilé). Également, d'autres modèles développés par l'équipe utilisaient déjà python, permettant ainsi une éventuelle communication entre les modèles plus simples.

Le choix de la bibliothèque DANA est motivé par plusieurs raisons :

- DANA, contrairement à des simulateurs de neurones plus répandus tels que Brian ou Neuron, est spécifiquement dédié aux modèles utilisant des neurones “rate-coding”. Il présente ainsi des optimisations spécifiques à ce genre de neurones, permettant plus de performance sur des tâches similaires [Rougier et Fix, 2012].
- DANA impose une modélisation distribuée, à l'échelle du neurone ou de la population, et permet ainsi d'apporter la garantie que certains algorithmes non biologiquement plausibles ne puissent pas être utilisés, comme par exemple les algorithmes utilisant la rétro-propagation du gradient de l'erreur, ou ceux nécessitant un superviseur global, comme c'est le cas de certains algorithmes de type winner-take-all.
- DANA permet d'utiliser les modèles de neurones de son choix, en lui fournissant les équations correspondantes.
- D'autres modèles, représentant différentes régions du cerveau et développés par l'équipe utilisent également DANA, permettant ainsi de faciliter une éventuelle fusion des modèles.

DANA encode chaque population sous forme de groupe (des matrices), contenant l'ensemble des valeurs stockées pour la population (Fig.3.1), et chaque population communique au moyen de connexions à déclarer. Par exemple,

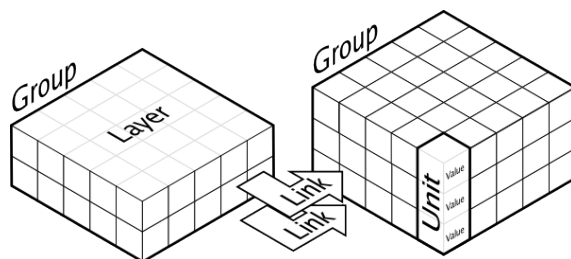


FIGURE 3.1 – Représentation des populations dans DANA. La structure group représente l'ensemble des information stockées pour une population. Une unit représente alors l'ensemble des valeurs stockées pour un élément de la population. Un layer représente un champ particulier pour l'ensemble de la population (illustration de [Rougier et Fix, 2012])

représenter deux populations A et B d'équations quelconques, connectées entre elles via des poids apprenant suivant une équation spécifique, se déclare de la manière suivante :

```
A = zeros(taille_A, equations_A)
B = zeros(taille_B, equations_B)
denseConnection(A,B,poids_initiaux,equations_poids).
```

`zeros()` est une fonction créant une matrice de taille 0, comme par exemple `numpy.zeros()`. `equation` est l'équation de la population ou de la connexion, écrite sous simple forme de chaîne de caractère. DANA permet également un choix de calcul matriciel adapté au type de connexion, comme des connexions peu remplies (éparses) ou au contraire denses. Le modèle peut être simulé sur un certain intervalle de temps avec un pas de temps donné via la commande `run`. La présentation de CS ou l'arrivée d'US, peuvent être déclenchés sous forme d'évènements arrivant à une date précise.

DANA ne possède par contre pas d'outil de visualisation intégrés, la visualisation peut se faire en affichant les activités dans les formats de graphes standards de `numpy`.

3.3.2 Nengo

Si tous les modèles ont été développés, et sont présentés dans cette thèse, via des neurones “rate-coding” en utilisant DANA, certains ont été également modélisés avec des neurones à spikes, comme le modèle d'amygdale décrit dans sa version “rate-coding” dans le chapitre suivant. Pour reproduire le modèle, et l'expérience d'acquisition/extinction/renewal à l'aide de populations de neurones à spikes, nous avons utilisé le simulateur Nengo (<http://www.nengo.ca>). Si cette implémentation n'apporte pas de nouveaux résultats à la thèse, elle permet néanmoins de valider notre vision systémique, en montrant que les résultats obtenus par les modèles ne dépendent moins du type de neurones utilisés, mais davantage des fonctions et interactions des différentes populations.

Nengo utilise le Nef (Neural Engineering Framework), [Eliasmith C., 2003], permettant différentes optimisations lors du calcul de populations. Contrairement à DANA, il a été conçu pour être utilisé avec des neurones à spikes. Il peut être utilisé également avec des neurones “rate-coding”, mais les concepts et optimisations utilisées ne sont pas toutes applicables au neurones de type “rate-coding”. De manière similaire à DANA, Nengo utilise des structures nommées *Ensembles* pour encoder les neurones au niveau de la population, et permet de les connecter entre eux. Si je reprends l'exemple de la partie précédente, pour deux populations de neurones à spikes :

```
A = nengo.Ensemble(taille_A, dimensions=1)
```

```
B = nengo.Ensemble(taille_B, dimensions=1)
nengo.Connection(A, B, synapse=poids_initiaux)
```

`dimensions` indique le nombre de valeurs codées par la population. Relier deux populations entre elles relie ainsi tous les neurones encodant une dimension vers ceux encodant la dimension correspondante dans la seconde population.

Ceci permet par exemple d’encoder un CS de manière distribuée sur toute une population de neurones. L’information ainsi encodée au niveau de la population, peut être décodée, en multipliant la sortie de la population par un ensemble de poids appelé décodeur, à des fins de visualisation, ou de transformation. En particulier, les poids du décodeur peuvent non seulement permettre de retrouver le signal encodé par la population, mais également de lui appliquer une transformation arbitraire. Une des particularités du Nef est ainsi qu’il permet de calculer de manière automatique les poids correspondant à n’importe quelle fonction entre deux populations A et B [Eliasmith C., 2003]. Pour cela, la valeur de A est décodée via un décodeur appliquant la transformation souhaitée, puis de nouveau encodée avant d’être transmise à la population B. Multiplier les poids de l’encodeur et du décodeur permet alors d’obtenir les poids de la connexion entre A et B.

Nengo possède également une interface graphique dédiée, permettant en temps réel de visualiser les activités du réseau, ou de modifier les entrées (https://github.com/nengo/nengo_gui).

3.3.3 Virtual Enaction

Virtual Enaction est une interface avec un jeu présentant un environnement de survie, Minecraft, permettant d’y connecter des modèles comme ceux étudiés dans cette thèse, et auquel j’ai également contribué [Denoyelle *et al.*, 2015]. L’intérêt de Virtual Enaction (VE) est double. D’un point de vue purement démonstratif, il permet à des personnes étrangères aux neurosciences de visualiser à quoi ressemble une expérience, et l’évolution du comportement du modèle lors de l’apprentissage de différents paradigmes. D’un point de vue plus informatique, il permet de proposer un environnement similaire, mais simplifié et avantageux, à ce que pourrait donner l’importation du modèle sur un robot :

- Il permet de donner un corps au modèle, et donc une notion de besoin. L’agent possède ainsi différentes jauges de faim, de soif, de santé. Une variation positive d’une jauge provoquée par l’environnement est ainsi une récompense, et une diminution une punition.
- En donnant un corps au modèle, il lui donne également des commandes motrices précises, telles que “mets toi debout”, ou “tourne de 90 degrés”. Au contraire d’un robot, l’effet de ces commandes est toujours fiable. Par exemple, l’agent ne va pas rater ou être imprécis dans son mouvement, en raison d’une usure des moteurs, ou d’une texture du sol différente, comme cela peut être le cas d’un robot. Ces représentations

d'actions déjà élaborée restent plausible avec la biologie, par exemple en correspondant aux actions élémentaires déjà apprises par la boucle motrice, et mises à disposition des boucles associatives et limbiques.

- Virtual Enaction offre un choix de compromis entre des entrées “artificielles” comme dans les modèles classiques (comme des vecteurs de CS) et des entrées réelles, comme celles d’une caméra. Pour cela Virtual Enaction peut transmettre la vision pixelique du jeu, donc une scène visuelle complète, ou la vision symbolique (présence d’un type de bloc à une certaine distance), plus proches des entrées de modèles classiques. VE peut également transmettre des informations tactiles (la texture des blocs adjacents) ou olfactifs (une valeur correspondant à l’odeur du bloc, pour tous les blocs proches).
- VE met l’environnement en pause durant le temps de calcul du modèle. Ainsi cela n’a pas de conséquence si le modèle n’est pas capable de faire du temps réel, au contraire d’un robot, dont l’environnement peut changer durant la réflexion.

Virtual Enaction est développé en C++, mais peut être utilisé par des modèles en python tels que les miens au moyen de la bibliothèque ctypes. Ctypes permet d’exécuter VE, sous forme de bibliothèque compilée en C, à partir de code python. A chaque pas de temps, VE met en pause le serveur, et appelle une méthode `botDo()` en lui fournissant les différentes visions (pixelique, symbolique, olfactive,...) proposées au modèle. Cette méthode permet à l’utilisateur d’exécuter son modèle sur une entrée, et de communiquer, ou non, une ou plusieurs actions, qui seront envoyées au serveur. `botDo()` permet également à l’utilisateur d’effectuer des traitements entre ceux du modèle et ceux du serveur. Par exemple elle peut permettre à l’utilisateur d’envoyer au serveur des actions non décidées par le modèle, comme envoyer une récompense si le modèle a effectué une certaine action. Dans le cas de modèles strictement pavloviens, le modèle n’effectue par définition pas d’action opérante, et en particulier reste sur place, avec donc peu de variation dans l’environnement. Une solution peut être pour l’utilisateur de rajouter dans `botDo()` un déplacement systématique, ou aléatoire, permettant à un modèle pavlovien de rencontrer quand même différentes situations.

Les modèles informatiques implémentés dans cette thèse et décrits dans la partie suivante utilisent ainsi un formalisme inspiré de la biologie, simulé au moyen de DANA, bibliothèque qui apporte des garanties supplémentaires sur des calculs neuronaux locaux, plausibles biologiquement. L’utilisation de logiciels différents et annexes, tels que Nengo ou VE permet d’appliquer les modèles dans des conditions différentes, et de vérifier ainsi dans un certain point la robustesse des comportements observés. L’usage de DANA permet cependant suffisamment de souplesse pour implémenter l’ensemble des modèles associatifs et motivés, et effets des neuromodulateurs, qui sont l’objet de cette thèse.

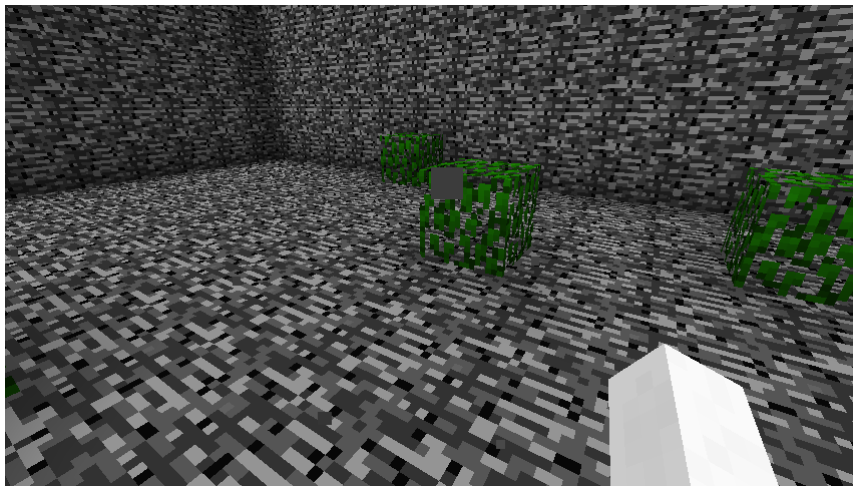


FIGURE 3.2 – Exemple d’expérience pavlovienne réalisée avec Virtual Enaction. La proximité d’un bloc de couleur verte déclenche l’arrivée de récompense. Etant donné que le modèle de décision implémenté est pavlovien, et ne déclenche donc pas d’action opérante comme un déplacement, un déplacement aléatoire est ajouté de manière indépendante du modèle, permettant ainsi que celui-ci puisse déclencher différentes actions. Chaque pièce est dédiée à une phase spécifique d’un paradigme. Pour un paradigme d’extinction, après que le modèle a correctement acquis la réponse pavlovienne dans cette pièce, il est transporté dans une pièce différente (formée par des blocs de couleurs différentes), toujours en présence de blocs vert, et apprend ainsi à inhiber sa prédiction dans un contexte, ici une salle, différent.

Chapitre 4

Le rôle des différents circuits de l'amygdale dans l'acquisition de la réponse pavlovienne

4.1 Les différents circuits de l'amygdale

L'apprentissage associatif, l'apprentissage d'associations pavloviennes, apprend à associer à un CS ou un contexte, une valeur, afin de pouvoir prédire l'arrivée de récompense et de punition, et de fournir cette information à l'apprentissage motivé, de manière à l'utiliser dans la définition de buts et de stratégies. Dans cette partie, nous étudions et reproduisons le fonctionnement de certaines populations de l'amygdale, structure impliquée dans l'apprentissage de telles valeurs. En particulier, nous nous intéresserons aux différents circuits présents dans l'amygdale, et travaillerons sur l'intérêt de ces circuits dans l'apprentissage d'une association adaptée à la situation, notamment dans des cas d'incertitudes, ou de changement de la valeur existante.

L'amygdale joue un rôle crucial dans le conditionnement pavlovien et l'acquisition de la réponse associée. En particulier, elle apprend à associer une valeur aux CS, qu'ils soient de valence positive (boisson, nourriture) ou négative (douleur, décharge électrique...), afin de déclencher la réponse pavlovienne correspondante (orientation, "freezing" ou "gel" de l'animal, qui demeure immobile et crispé, réponse hormonale, salivation). En particulier, trois noyaux de l'amygdale jouent un rôle majeur dans l'acquisition et expression de la réponse de peur : le noyau latéral (LA), le noyau basal (BA), et le noyau central (CeA). LA et BA sont principalement responsables de l'association CS-US ou contexte-US, tandis que CeA est impliqué au niveau "moteur", dans le déclenchement de la réponse pavlovienne [Li et McNally, 2014].

LA est responsable de l'apprentissage de l'association entre le CS sensoriel présent dans le contexte et l'US pour le conditionnement pavlovien [LeDoux, 2007]. LA reçoit des entrées sensorielles du thalamus et du cortex. Les entrées

venant du thalamus sont des informations relativement “brutes” (comme des tonalités pures), tandis que le cortex construit et fournit des représentations plus élaborées (des CS composés de plusieurs traits élémentaires par exemple). Dans ce chapitre, je ne fais pas de différence entre les deux, et parle d'entrée corticale, représentant le gradient des entrées thalamiques “pures” jusqu'aux représentations de plus en plus élaborées du cortex. Les réponses de peur dans LA sont courtes (100ms rapporté par [Burgos-Robles et al. \[2009\]](#)), et des CS présentés avec des délais plus importants avant l'arrivée de l'US nécessitent une activité soutenue dans le prélimbique (PL) [[Pendyam et al., 2013](#)], projetant sur BA, pour pouvoir être associés à un US par l'amygdale. Les expériences que nous reproduisons dans cette partie appartiennent au “delay conditionning” (où l'US arrive avant la fin de la présentation du CS, ou simultanément), et dans cette optique ne nécessitent donc pas PL.

BA est la région de l'amygdale qui s'est le plus étendue lors de l'évolution récente [[Cardinal et al., 2002](#)]. Elle intègre des informations venant de LA, de l'hippocampe, notamment pour le conditionnement au contexte spatial ou temporel [[Goosens et Maren, 2001](#); [Biedenkapp et Rudy, 2009](#); [Eichenbaum et al., 2012](#); [Carretero-Guillén et al., 2013a](#)], de PL et de la partie infralimbique (IL) de mPFC, en particulier pour l'apprentissage de l'extinction [[Sierra-Mercado et al., 2011](#); [Wallis, 2007](#)]. [Herry et al. \[2008\]](#) montre comment deux populations différentes, BAf et BAe, apprennent respectivement l'apprentissage de la peur et de l'extinction, et leur compétition détermine la réponse de l'amygdale par CeA.

CeA est considéré comme le pôle moteur de l'amygdale, déclenchant la réponse pavlovienne de peur. [Cicchi et al. \[2010\]](#) montre que LA et BA projettent sur CeLOn et CeLOff, des populations de la partie latérale de CeA, CeL. Ces projections se font éventuellement par l'intermédiaire de cellules inhibitrices situées entre LA et BA d'une part et CeA d'autre part, les “cellules intercalées”(ITC). CeLOn et CeLOff participent respectivement à l'excitation et l'inhibition de la réponse de peur dans la partie médiale de CeA, CeM. Des projections directes d'ITC sur CeM participent également à la réponse de peur [[Lee et al., 2013](#); [Amano et al., 2010](#)]. Le comportement de “freezing” est déclenché par les projections de CeA sur le periaqueductal grey (PAG), et est généralement considéré comme un marqueur de l'acquisition de la peur [[Sierra-Mercado et al., 2011](#); [Cicchi et al., 2010](#); [LeDoux, 2007](#); [Herry et al., 2008](#)]. Un autre aspect de la réponse pavlovienne est la réponse hormonale, impliquant neuromodulateurs et hormones neuropeptidiques [[Lee et al., 2013](#)], qui influence différents facteurs communs à l'ensemble du cerveau, tels que le niveau d'attention ou le stress liés de la situation présente. En particulier, parmi les différents rôles attribués aux neuromodulateurs, le niveau tonique d'Acetylcholine (ACh), émis par le basal forebrain, a été proposé par [Yu et Dayan \[2005\]](#) comme marqueur du niveau d'incertitude de l'environnement. L'ACh impacte notamment l'hippocampe, en changeant son mode de stockage

4. Le rôle des différents circuits de l'amygdale dans l'acquisition de la réponse pavlovienne

de l'information [Hasselmo, 2006], et par sa capacité à promouvoir une acquisition sensorielle ou contextuelle de la réponse de peur [Calandreau *et al.*, 2006a]. Dans l'amygdale, l'ACh impacte BA, au contraire de LA qui n'en reçoit pas (différence qui est à l'origine de la distinction anatomique entre LA et BA) [Nitecka et Narkiewicz, 1976].

Malgré la présence de différentes populations à l'intérieur de l'amygdale, possédant des rôles différents, ces populations sont peu représentées dans les modèles répliquant l'action de l'amygdale. Les modèles comportementaux vus précédemment, comme ceux de Rescorla-Wagner, de Mackintosh, ou de Le Pelley, s'attachent à reproduire des comportements spécifiques observés dans des expériences sur les animaux, et ce au moyen de règles complexes. En particulier, ils ne reproduisent pas les comportements des différentes populations impliquées dans la tâche. Par exemple pour l'extinction, deux populations différentes encodent respectivement une valeur positive pour le CS et une négative pour le contexte. Or dans beaucoup de modèles, [Schmajuk et DiCarlo, 1992; Kruschke, 2001], le contexte et l'entrée sensorielle sont contenus dans une unique entrée, et traités similairement. Ce type de modèle ne pourrait donc pas par la suite exploiter les différentes dynamiques d'apprentissage pour enrichir d'autres systèmes d'apprentissages, comme par exemple ceux de l'apprentissage motivé. Avec l'avancée des connaissances en neurosciences sur le conditionnement pavlovien, d'autres modèles décrivent les interactions entre l'amygdale et le mPFC, le cortex postérieur et l'hippocampe [Pauli *et al.*, 2011; Meeter *et al.*, 2005; Armony *et al.*, 1997; Moustafa *et al.*, 2013], mais ne différencient pas les différentes populations et associations présentes dans l'amygdale. A l'opposé, d'autres modèles intègrent précisément le comportement de certaines populations de l'amygdale [Pendyam *et al.*, 2013; Li *et al.*, 2009; Kim *et al.*, 2013; Vlachos *et al.*, 2011], mais restent au niveau de ces populations, sans considérer les différents circuits auxquels ces populations appartiennent.

Aucun de ces modèles existants ne prend en compte l'intégralité de l'amygdale en différenciant les rôles des différents circuits, et leur compétition pour sélectionner le circuit le plus adapté à une situation donnée. Je propose dans ce chapitre un modèle implémentant cette vision et intégré dans une vision systémique globale des fonctions des différents afférents et efférents de l'amygdale, dans le but non seulement de valider cette vision par la réplification de différents résultats expérimentaux, mais également de mettre en valeur les différentes associations apprises, et les stratégies mises en place afin d'apprendre l'association adaptée à une situation donnée.

4.2 Description du modèle

Plusieurs décisions doivent être prises au moment de décrire et d'implémenter un modèle. Comme toute représentation de la réalité, un modèle ne

peut être parfait, et ne peut pas décrire absolument tous les phénomènes de la biologie avec un niveau parfait de représentation. Également, cette thèse n'est pas exclusivement dédiée à l'étude du fonctionnement de l'amygdale, mais également à comprendre comment les valeurs des stimulus apprises dans l'amygdale sont utilisées pour prendre des décisions. Par conséquent, tous les flux d'informations, comportements et populations ne seront pas représentés par ce modèle. La discussion de ce chapitre essaiera cependant de donner des éléments de réponse sur d'autres aspects du fonctionnement de l'amygdale, qui n'ont pas pu être inclus dans ce modèle.

Afin de présenter un modèle systémique de l'apprentissage CS-valeur, valeurs utilisables par un apprentissage motivé, nous nous concentrerons donc sur les différentes parties de l'amygdale apprenant ces valeurs et leurs dynamiques d'apprentissage (voir la figure 4.1). Leurs différents afférents inclus dans le modèle, le cortex sensoriel (représentant ici un gradient de représentations sensorielles, des plus primitives partant du thalamus aux plus élaborées dans le cortex perirhinal), l'hippocampe et OFC, seront uniquement représentés par un vecteur d'activités passé en entrée du modèle, représentant leurs différentes activations au cours d'un essai.

Notre modèle d'amygdale comporte 5 populations neuronales différentes, LA, BAf, BAe, CeLON, CeLOff. LA et BAf représentent les "fear neurons", neurones impliqués dans l'apprentissage de la peur, de LA et BA, respectivement. BAe représente les "extinction neurons", neurones impliqués dans l'apprentissage de l'extinction, de BA. CeLON et CeLOff représentent les neurones ON et OFF de CeL, populations observées dans [Herry *et al.*, 2008]. La figure 4.1 résume les différentes populations et connexions modélisées. La valeur et signification de l'ensemble des paramètres utilisés sont décrites dans le tableau 4.2.

Le cortex projette en entrée de LA et fournit un vecteur contenant les différents CS. L'hippocampe et IL projettent respectivement sur BAf, fournissant le contexte, et sur BAe, représentant le contexte en situation d'extinction [Sierra-Mercado *et al.*, 2011]. Plusieurs variables sont également envoyées à l'amygdale. US est un booléen indiquant l'arrivée ou non d'un US, aversif dans les expériences décrites ci-dessous, i.e. une punition. ERR est l'erreur de prédiction, la différence entre l'US rencontré et l'US prédit par le modèle US_{pred} :

$$ERR = US - US_{pred}$$

Deux sources sont possibles comme origine de cette erreur, l'erreur de prédiction calculée dans le PAG et transmise à l'amygdale, ou l'erreur de prédiction transmise par la dopamine [Li et McNally, 2014]. ACh représente le niveau tonique d'Acétylcholine envoyé par le basal forebrain sur l'amygdale, et en particulier à BLA [McGaugh, 2004].

Notre objectif est de représenter les dynamiques et comportements des différentes populations de l'amygdale, en fonction d'une vision systémique de la biologie du cerveau. Par conséquent, les équations et paramètres du modèle

4. Le rôle des différents circuits de l'amygdale dans l'acquisition de la réponse pavlovienne

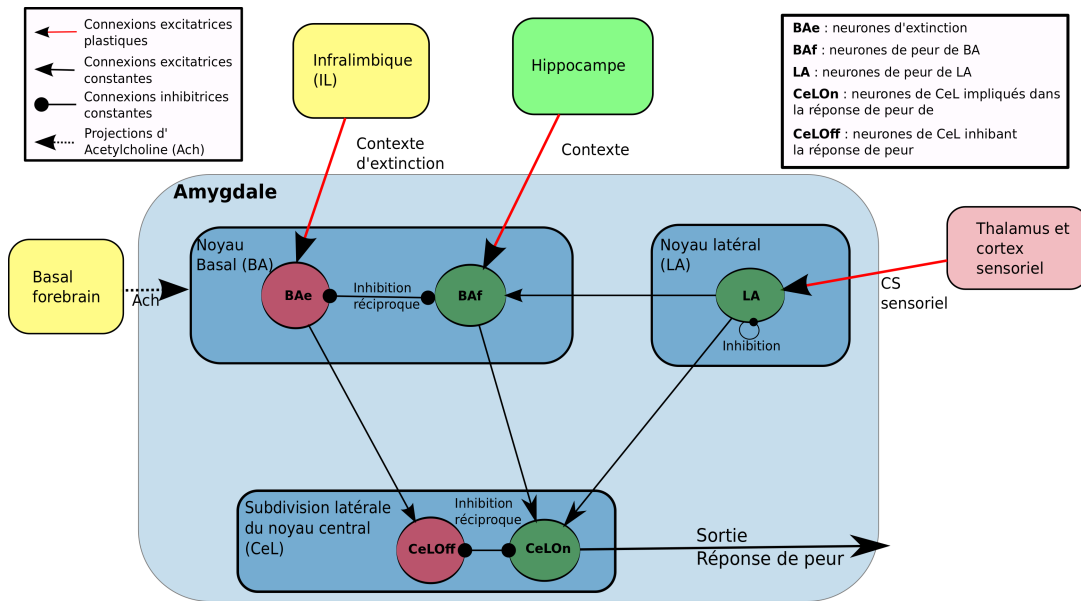


FIGURE 4.1 – Populations et circuits du modèle de l'amygdale. LA et BAF sont deux populations de neurones impliquées dans l'apprentissage de la réponse de peur. LA apprend à prédire la peur à partir d'entrées sensorielles du cortex et du thalamus, et projette sur BAF et CeLOn. BAF reçoit le contexte de l'hippocampe et les prédictions de LA, et prédit ainsi l'arrivée d'US aversif à partir à la fois d'informations contextuelles et d'informations sensorielles. BAe est une population de neurones d'extinction qui reçoit des informations du cortex infralimbique (IL) durant l'extinction. BAe et BAF sont en inhibition réciproque et projettent respectivement sur les populations CeLOn et CeLOff de la partie latérale (CeL) du noyau central de l'amygdale. CeLOn et CeLOff possèdent également des inhibitions réciproques. L'activité de CeLOn est considérée comme le niveau de prédiction de l'US du modèle.

sont simples, et au maximum uniformes d'une population à une autre. De cette façon, les comportements observés, ainsi que les différentes associations apprises par le réseau, résultent essentiellement de l'application de la connectivité biologique du réseau, et non pas de la combinaison de règles d'apprentissage complexes. Les neurones sont représentés au niveau de la population par l'utilisation d'un formalisme classique "mean-field", représentant l'activité moyenne d'une population. La totalité des paramètres utilisés est donnée par la figure fig.5.3. Les connexions et flux d'activité sont décrits dans la figure 4.1. Chaque connexion est totale, i.e. chaque unité de la population d'entrée est systématiquement connectée à toutes les unités de la population de destination. La bibliothèque DANA est utilisée pour la représentation et les dynamiques des neurones [Rougier et Fix, 2012].

Chaque neurone i d'une population de l'amygdale $Amyg$ recevant des entrées excitatrices et inhibitrices E_{exc} et E_{inh} voit son activité décrite par son potentiel de membrane V_i et son taux de décharge U_i :

$$\frac{dV_i}{dt} = (-V_i + F(\sum_{j \in E_{exc}} W_{ij} * U_j)) / \tau$$

$$U_i = noise(sigmoid(V_i)) - \sum_{k \in E_{inh}} W_{ik} * U_k$$

Pour chaque neurone j d'une population d'entrée, W_{ij} désigne le poids de la connexion entre i et j . τ est la constante de temps définissant la dynamique d'évolution du potentiel de membrane, et F une fonction de seuillage non linéaire :

$$F(U) = \max(\min_value, U - \theta)$$

Où \min_value est une constante très faible, dont le rôle est d'éviter les activités inférieures ou égales à 0, qui bloqueraient l'apprentissage hebbien. θ est une constante fixant la valeur du seuil. Une fonction sigmoïde $sigmoid()$ est appliquée à V pour obtenir le taux d'activation U . Un bruit $noise()$ est ajouté pour représenter le bruit interne au neurone. Les expériences ont été effectuées avec un bruit relativement faible (suivant une distribution uniforme, sur un intervalle de 1% de $sigmoid(V)$). Cependant, augmenter le bruit jusqu'à 20% de $sigmoid(V)$ n'influe pas sur la performance du réseau après apprentissage sur les différents paradigmes étudiés. Au delà, les dynamiques sont conservées mais les prédictions du réseau après apprentissage deviennent de moins en moins précises. Les neurones de LA et de BA étant glutamatergiques, leur connexions entrantes et sortantes sont excitatrices. L'inhibition mutuelle entre BAf et BAe, ainsi que celle entre neurones de LA, correspondent aux projections excitatrices des neurones de ces populations vers des neurones intermédiaires inhibiteurs, inhibant la population de destination [Herry *et al.*, 2008]. Au contraire, CeLOn et CeLOff sont gabaergiques et directement en inhibition mutuelle. A l'exception des connexions inhibitrices décrites ci-dessus, toutes les connexions du réseau sont excitatrices 4.1. L'activité de CeLOn, qui projette sur CeA, déclenchant ainsi la réponse de peur, est prise comme prédiction US_{pred} du modèle.

4.2.1 Règles d'apprentissage

Les connexions entre le cortex et LA, entre l'hippocampe et BAf, et entre IL et BAe sont plastiques. Les valeurs des poids des autres connexions du réseau sont fixes. Leurs valeurs initiales sont aléatoirement tirées suivant une distribution uniforme de longueur d'intervalle 0.04 et centrée en 0.03 pour les poids des connexions plastiques, en 0.2 pour les poids reliant LA, BAf et BAe à CeLOn et CeLOff, et en 0.1 pour les poids entre LA et BAf. Les poids des connexions inhibitrices sont centrés en 0.25 entre CeLOn et CeLOff, et en 0.05 entre BAf et BAe. Les poids de ce dernier ensemble de connexion doivent

4. Le rôle des différents circuits de l'amygdale dans l'acquisition de la réponse pavlovienne

Paramètres du modèle		
Paramètre	Description	Valeur
input_size	taille des entrées du cortex, de l'hippocampe et d'IL	10
la_size	nombre d'unités dans LA	10
baf_size	nombre d'unités dans BAf	10
bae_size	nombre d'unités dans BAe	10
CeLOn_size	nombre d'unités de CeLOn	1
CeLOff_size	nombre d'unités de CeLOff	1
Paramètres des équations		
τ	contante de temps	0.05
α	taux d'apprentissage	1
il_tau	constante de temps pour la population générant ACh	5
noise_level	niveau de bruit pour chaque unité	1%
θ	seuil d'entrée d'une unité	0.3
ach_min	niveau minimum d'ACh	1
ach_max	niveau maximum d'ACh	2.5
ach_effect	effet de l'ACh sur BAf et BAe	0.5
uncertainty_effect	effet de l'incertitude sur la génération d'ACh	5
wmin	valeur minimale de l'initialisation des poids non constants	0.01
wmax	valeur maximale de l'initialisation des poids non constants	0.05
Cel_input_W	poids constants reliant LA, BAf et BAe à CeLOn et CeLOff	0.2
LA_to_BAf_W	poids constants reliant LA à BAf	0.1
LA_inhib_W	poids constants de LA sur elle-même	0.1
Cel_inhib_W	poids constants entre CelOn et CelOff	0.25
BA_inhib_W	poids constants entre BAf et BAe	0.05

FIGURE 4.2 – Valeur et description de l'ensemble des paramètres utilisés dans notre modèle. Les paramètres ont au maximum été laissés identiques entre les différentes populations, afin de montrer que les différences de comportement des populations ne viennent pas de leur différences d'équations, mais de leur différences de connectivités.

être fixés à une valeur relativement faible, car une valeur trop importante, permettant à une population de totalement inhiber l'autre, empêcherait ainsi l'apprentissage hebbien de la seconde population de débiter.

Les populations de “fear neurons” de LA et de BAf apprennent à prédire l'arrivée de stimuli aversifs, respectivement à partir du CS sensoriel venant du cortex et du contexte fourni par l'hippocampe. Les poids W_{ij} de la connexion entre un neurone présynaptique i et un neurone postsynaptique j varient suivant une loi hebbienne similaire à celle de Rescorla-Wagner :

$$\frac{dW_{ij}}{dt} = ERR * US * \alpha * U_j^{Pre} * U_i^{Post}$$

Le coefficient d'apprentissage, α , est identique pour les neurones de LA, BAf et BAe. US indique la valeur (magnitude) de l'US si un US est présenté au réseau, et a une valeur nulle sinon. Ceci permet aux “fear neurons” de n'apprendre que lorsqu'un US est présenté au réseau. Par conséquent ceux-ci ne vont pas oublier et diminuer leurs poids en situation d'extinction, et réagissent immédiatement en prédisant l'US lors du renewal.

Au contraire, les neurones d'extinction de BAe n'apprennent pas à prédire l'US, mais à corriger la prédiction lors de l'extinction :

$$\frac{dW_{ij}^{BAe-IL}}{dt} = -ERR * \alpha * U_i^{BAe} * U_j^{IL}$$

Le terme d'erreur opposé à celui d'une règle de Rescorla-Wagner, $-ERR$, leur permet d'augmenter leur réponse si le réseau prédit trop ($US_{pred} > US$), et ainsi d'inhiber davantage les neurones prédisant l'arrivée d'US, conduisant effectivement à une baisse de la prédiction du réseau. Dans les expériences décrites plus bas, l'entrée U_{IL} du réseau n'est active uniquement qu'en situation d'extinction. Détecter une telle situation peut se faire à partir de statistiques sur l'arrivée d'US (par exemple si l'historique récent comporte une suite d'erreurs de prédiction négatives) ce qui est cohérent avec le rôle proposé de IL [Wallis, 2007]. En particulier, cela évite aux neurones d'extinction d'apprendre la prédiction de peur à partir d'un contexte d'extinction lorsque l'arrivée d'US est connue mais stochastique, ou de désapprendre quand ERR est positive.

4.2.2 Le rôle de la neuromodulation

Tel que décrit précédemment, la prédiction du réseau est le résultat d'une compétition entre la prédiction d'US aversifs apprise par LA et BAf, et la prédiction d'extinction apprise par BAe. Cependant, ce modèle met en même temps l'accent sur une autre compétition, celle de l'acquisition de l'US aversif. Dans notre modèle, le réseau reçoit des informations différentes du cortex et de l'hippocampe, représentant les CS et le contexte, qu'il peut associer à l'US

dans LA et BAf. Comment choisir quelle est l'information pertinente à associer à l'US suivant la situation? [Calandrea *et al.*, 2006a] met en évidence le rôle de l'acétylcholine tonique dans cette compétition, rôle que nous reproduisons en modélisant l'effet excitateur de l'ACh tonique sur les neurones de BA. U_{BAf} et U_{BAe} , tels que calculés précédemment, sont multipliés par la valeur représentant le niveau tonique d'acétylcholine, ACh :

$$ACh = ach_effect * (baseline + uncertainty_effect * U^{ACh})$$

avec *baseline* le niveau de base de d'acétylcholine tonique (fixé à 1.0).

uncertainty_effect est une constante modulant la quantité d'ACh émise en fonction de l'activité de la population émettrice U^{ACh} , et *ach_effect* la constante modulant l'effet de la concentration d'ACh sur l'activité des neurones de BLA.

U^{ACh} représente l'activité de la population libérant l'acétylcholine :

$$U^{ACh} = noise(sigmoid(V^{ACh}))$$

$$dV^{ACh}/dt = (-V^{ACh} + F(|ERR|))/\tau_{ACh}$$

Cette population voit son activité augmenter en fonction de la quantité d'erreur commise $|ERR|$, et diminuer en fonction du temps. Ainsi, si l'incertitude est élevée, la quantité d'erreur commise lors des derniers essais est importante, provoquant une libération plus importante d'acétylcholine, conformément au rôle attribué à l'acétylcholine [Yu et Dayan, 2005]. τ_{ACh} est une constante de temps élevée, de façon à ce que l'augmentation d'activité liée à une erreur de prédiction perdure sur plusieurs essais.

Ainsi la présence de stochasticité provoque un niveau tonique d'acétylcholine élevé, qui à son tour excite davantage les neurones de BA, favorisant l'apprentissage de l'US par BA basé sur des informations contextuelles, par rapport à celui de LA, basé sur les indices visuels du cortex.

4.2.3 Protocole expérimental

Les expériences décrites dans cette section ont pour but de reproduire des protocoles expérimentaux effectués chez l'animal, afin d'observer et de comparer les réponses et dynamiques d'apprentissage du réseau à celles rapportées chez l'animal. Nos expériences sont constituées d'une succession d'essais, jusqu'à convergence de l'apprentissage du réseau.

Chaque essai est divisé en trois phases (cf. Fig.4.3). La première phase est la phase de présentation des CS. Les entrées venant du cortex, de l'hippocampe et d'IL sont envoyées au réseau durant 500 cycles (un cycle étant la simulation pour un petit pas de temps dt des équations d'activité et de propagation du

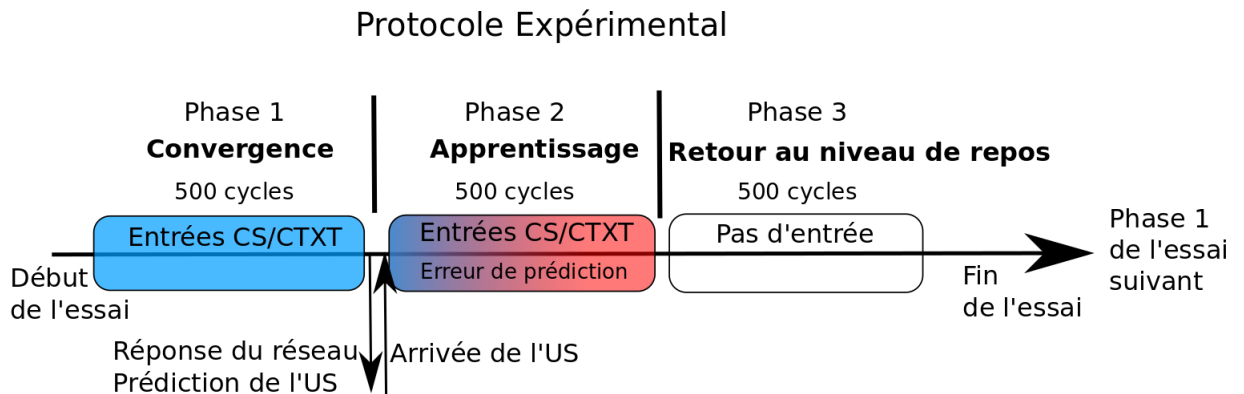


FIGURE 4.3 – Chaque essai est composé de 3 phases de durée identique. Dans la première phase, le CS sensoriel et le contexte sont présentés au réseau. A la fin de cette phase, l’activité de CeLON constitue la prédiction du réseau. Puis l’US est présenté, ce qui permet l’apprentissage du réseau dans la seconde phase grâce à l’erreur de prédiction. Durant la phase 3, aucune entrée n’est présentée au réseau, pour laisser les activités revenir à leur niveau de repos avant le début de l’essai suivant.

réseau). A la fin de cette phase, l’activité des neurones de CeLON donne la prédiction d’US du réseau. La seconde phase est la phase d’apprentissage. L’US est présenté au réseau (avec une valeur nulle en cas d’absence de punition), ainsi que l’erreur de prédiction associée, durant 500 cycles. Dans la troisième phase, durant à nouveau 500 cycles, aucune entrée n’est présentée au réseau, pour laisser le réseau revenir à son état de repos et ainsi préparer le prochain essai.

Les entrées corticales, hippocampiques et infralimbiques sont des vecteurs représentant les caractéristiques sensorielles et contextuelles. Dans ces vecteurs, les informations qui sont présentes sont fortement activées (1 pour les entrées *Hippo* et *IL*, 1.5 pour *cortex*), et les autres présentent un bruit aléatoire, tiré uniformément entre 0 et $n = 1/(\text{taille du vecteur en entrée})$). La forte activation des entrées représente les traits saillants capturés par les processus attentionnels, principalement au niveau du cortex, tandis que les bruits aléatoires représentent les bruits présents dans les structures afférentes, ainsi que les détails de fond, non-significatifs, observés par l’animal.

4.3 Résultats

Dans les expériences rapportées dans ce chapitre, les entrées corticales, hippocampiques et infralimbiques, ainsi que les populations LA, BAF et BAe, comportent chacune 10 unités. Pour CeLON et CeLOff, 1 seule unité est suffisante, comme un seul type d’US (aversif) est considéré dans les expériences ci-dessous. Le premier couple d’expériences, “extinction” et “renewal”, a pour but de reproduire la compétition entre acquisition et extinction de l’association

CS-US par les différentes populations LA, BAf et BAe. Le second, “pairing” et “unpairing”, étudie la compétition entre association sensorielle et contextuelle, et le rôle critique de l'acétylcholine dans cette compétition. Ces expériences ont été choisies car elles mettent en évidence comment les différents circuits de l'amygdale peuvent apprendre une association adaptée à la situation, grâce notamment à la neuromodulation de l'ACh.

4.3.1 Extinction-renewal

Le paradigme étudié ici, l'acquisition de la peur dans un premier contexte, suivi de son extinction dans un second contexte, puis du rétablissement dans le premier contexte, est un paradigme pavlovien classique, pour lequel les réponses des populations de l'amygdale ont en particulier été observées [Herry *et al.*, 2008; Ciocchi *et al.*, 2010].

Dans la phase d'acquisition, un CS sensoriel CS1 et un contexte CTXT1 sont présentés simultanément au réseau, respectivement via l'entrée corticale de LA et l'entrée hippocampique de BAf, et systématiquement suivis de l'US. Cette essai est répété jusqu'à acquisition complète de la réponse de peur et retour de l'ACh à son niveau de base. Dans la phase d'extinction, CS1 est couplé à un contexte différent, CTXT2, et n'est plus suivi d'US, et est répété jusqu'à apprentissage de l'extinction et baisse de l'incertitude. La phase de rétablissement (“renewal”) est constituée d'un seul essai, durant lequel CS1 est à nouveau présenté dans le contexte CTXT1.

La Fig.4.4 montre la réponse des différentes populations de notre modèle aux différentes phases de ce paradigme. Dans la phase de conditionnement, LA et BAf augmentent simultanément leurs connexions respectives au Cortex et à l'Hippocampe, de façon à ce que CeLOn puisse correctement prédire l'arrivée d'US après quelques essais (Fig.4.4.B). BAe et CeLOff voient leur activité décroître durant cette phase en raison de l'inhibition des unités de BAf et de CeLOn, respectivement. La figure 4.4.A montre que la prédiction de peur est distribuée entre LA et BAf. De manière plutôt contre-intuitive, et comme nous le verrons dans le second paradigme, ceci n'implique pas forcément que la prédiction est distribuée à part égale entre le contexte et le CS. En effet, LA projette non seulement sur CeLOn, mais également sur BAf. La réponse de BAf prend donc en compte et le contexte et le CS. La prédiction de CeLOn, basée sur les activités de LA (CS) et BAf (Contexte et CS), prend donc, sous conditions normales, davantage en compte le CS. Comme l'incertitude est élevée dans les premiers essais, la quantité d'acétylcholine tonique augmente, ce qui permet à BAf d'apprendre un peu plus vite, mais les réponses de LA et BAf s'équilibrent par la suite, comme l'incertitude redescend rapidement. La figure 4.4.E illustre l'évolution des poids de LA et BAf, qui augmentent similairement durant l'acquisition.

Durant l'extinction, l'activité de LA reste stable, et celle de BAf augmente

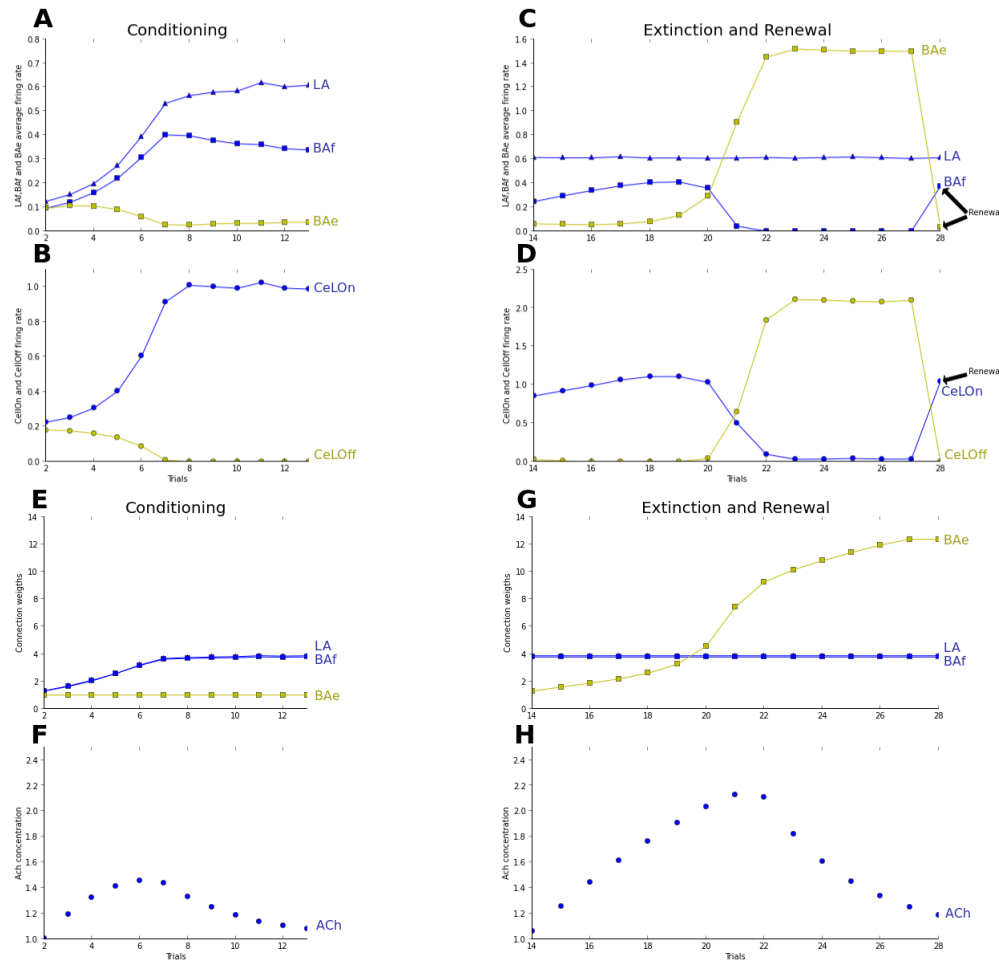


FIGURE 4.4 – Le premier essai, non montré ici, est un essai sans aucun stimulus, pour permettre la stabilisation du réseau aux activités de repos. L’acquisition de la réponse de peur se fait durant les essais 2 à 13, pour laisser l’incertitude et le niveau d’ACh redescendre avant de débuter l’extinction dans un nouveau contexte. Lors du dernier essai, l’animal est de nouveau testé dans son contexte d’acquisition pour le renewal. Les graphes A, B, C et D affichent l’activité moyenne des différentes populations du modèle. E et F montrent l’évolution moyenne des poids des différentes connexions (moyenne des poids normalisée par leur valeur initiale). G et H affichent l’évolution de la concentration d’ACh. (A,B) Les activités de LA et BAf augmentent lors de l’acquisition, déclenchant la réponse de CeLOn. BAf décroît légèrement à la fin de l’acquisition en raison de la diminution de l’incertitude, et du niveau d’ACh. (C,D) L’activité de BAe apprend à inhiber la prédiction de BAf lors de l’extinction. Sa projection sur CeLOff permet d’inhiber la réponse de CeLOn, toujours excité par LA, inhibant ainsi la réponse de peur. Lors du renewal, BAe ne reçoit plus le contexte d’extinction d’IL, ce qui relâche l’inhibition de BAe et de CeLOn, permettant une récupération immédiate de la réponse de peur. (E) LA et BAf apprennent tous deux durant l’acquisition. (F) BAe apprend durant l’extinction. En raison de l’inhibition de LA et BAf lors des premiers essais, l’apprentissage est plus lent au début, et s’accélère lorsque BAe n’est plus totalement inhibé. (G, H) La concentration d’ACh augmente lorsque le réseau fait des erreurs, et que l’incertitude augmente, et diminue lorsque le réseau est régulièrement performant.

légèrement, dû à l'effet de l'acétylcholine, avant de diminuer sous l'inhibition de BAe apprenant progressivement le contexte d'extinction (figure 4.4.F). L'augmentation de l'activité de BAe entraîne l'accroissement de celle de CeLOff, et donc l'inhibition de CeLOn qui restait activé par LA (figure 4.4.C). Lors du renewal, le contexte est à nouveau celui de l'acquisition. BAe qui n'a pas appris ce contexte voit son activité retourner à l'état normal, stoppant l'inhibition de BAF et de CeLOn, et entraînant ainsi le rétablissement immédiat de la réponse de peur. La figure 4.4.G illustre la variation de la quantité d'acétylcholine, correspondant à l'incertitude durant l'extinction, et aidant l'acquisition de l'extinction sur BAe. Les dynamiques de BAe et BAF observées dans ce paradigme reproduisent les résultats rapportés dans [Herry et al. \[2008\]](#), également modélisé dans [Vlachs et al. \[2011\]](#). Le comportement de LA, qui reste active lors de l'extinction, reproduit les résultats de [Repa et al., 2001](#)].

[\[Prado-Alcalá et al., 1994\]](#) rapporte que l'injection d'un antagoniste de l'acétylcholine, la scopolamine, après apprentissage de l'extinction, endommage l'extinction. Nous avons testé et reproduit cet effet en fixant le niveau d'ACh à 0.5 du niveau normal après apprentissage de l'extinction (figure 4.5, dernier essai). L'activité de BAe décroît suite à la diminution d'acétylcholine, entraînant un accroissement léger de BAF et une forte baisse de CeLOff, permettant à CeLOn d'être actif de nouveau. Une diminution plus importante du niveau d'acétylcholine provoque une détérioration plus importante de l'extinction. Nous avons également effectué l'expérience précédente en diminuant la quantité d'acétylcholine avant (et pendant) acquisition et extinction de la réponse de peur. Le niveau d'ACh a été fixé à un niveau constant de 0.5, i.e. 50% du taux normal. L'acquisition de la réponse de peur n'est pas endommagée, même si l'apprentissage de la réponse de peur se fait majoritairement dans LA (figure 4.5.A). En revanche l'extinction est détériorée par la diminution d'ACh. Les neurones de BAe mettent davantage d'essais à apprendre et inhiber BAF, et ne suffisent pas à inhiber CeLOn via CeLOff en raison de la prédiction de peur trop importante dans LA (figure 4.5.C). Notre modèle prédit ainsi qu'une diminution de l'acétylcholine dans l'amygdale avant apprentissage de l'extinction devrait détériorer ce dernier. Cette prédiction est en partie soutenue par [Schroeder et Packard \[2004\]](#), qui montre qu'une augmentation de la concentration d'ACh facilite l'acquisition.

4.3.2 Pairing et unpairing

Dans cette partie, nous étudions la compétition entre le CS et le contexte dans l'acquisition de la peur, ainsi que le rôle de l'acétylcholine dans cette compétition. Pour cela, nous reproduisons les deux paradigmes proposés dans [Calandreau et al. \[2006a\]](#), et regardons leur impact sur LA et BAF. Le premier paradigme est le "pairing". Il consiste en la présentation d'un CS dans un contexte, systématiquement suivi d'un US, similairement à l'acquisition de la

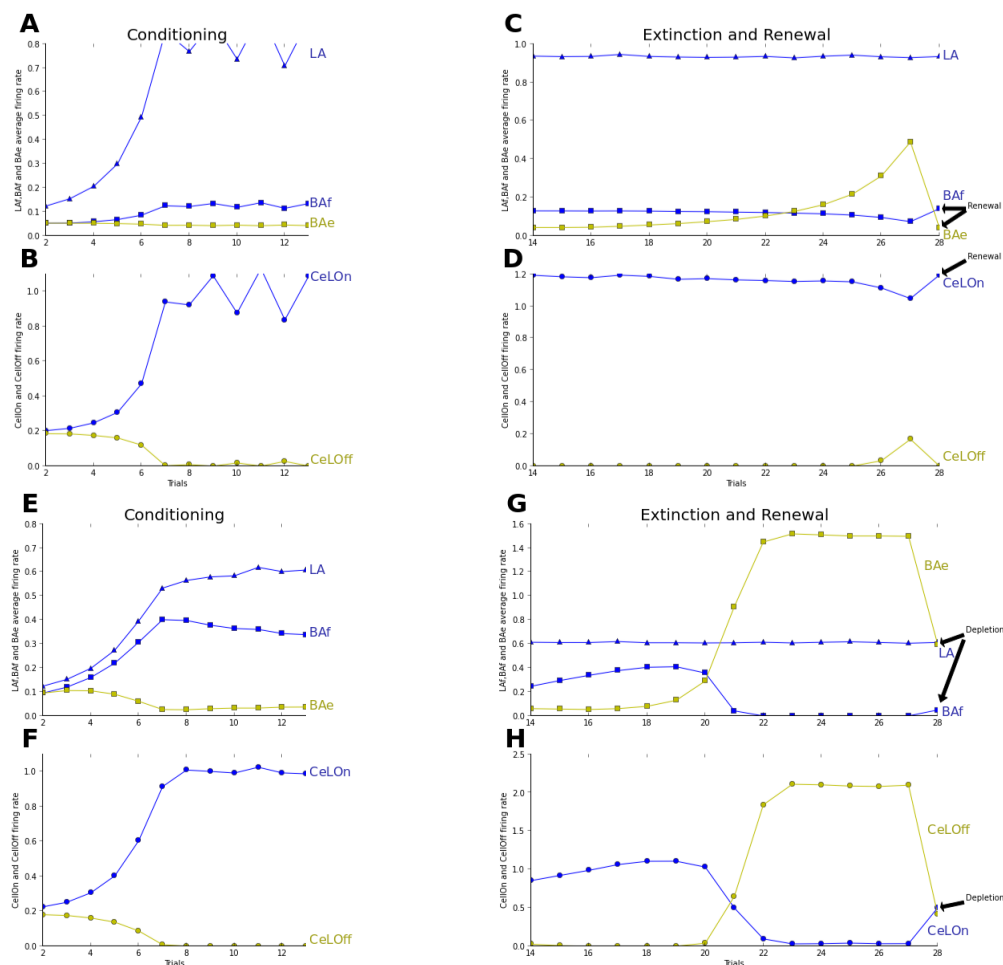


FIGURE 4.5 – Effet de l’acétylcholine sur l’extinction et le renewal. Voir 4.4 pour une description des valeurs représentées par les graphes. (A,B,C,D) montrent l’effet d’une déplétion d’ACh (niveau d’ACh constant fixé à 0.5) durant l’ensemble des essais. (E,F,G,H) montre l’effet d’une déplétion d’acétylcholine suite à un apprentissage normal de l’extinction. Le dernier essai de G et H montre ainsi une déplétion de l’acétylcholine, toujours dans un contexte l’extinction, et non un renewal dans un contexte différent. (A,B) En raison de la déplétion d’ACh, l’apprentissage de l’acquisition CS-US est essentiellement appris par LA, et est suffisant pour déclencher la réponse de CeLON (C,D) L’apprentissage de l’extinction est rendu plus difficile, comme BAe n’a plus l’effet facilitateur de l’ACh, et l’activité de CeLON reste ainsi élevée. Le renewal est automatique, comme l’extinction n’est pas apprise à temps. (E,F) Conditionnement sans déplétion d’ACh, similaire à la figure 4.4. (G,H) Le niveau d’ACh est déplété lors du dernier essai (fixé à 0.5, soit 50% du niveau de base), ce qui provoque une diminution de l’activité de BAe, et une re-acquisition partielle de la réponse de peur.

peur étudiée précédemment. Puis nous testons à quel degré le CS présenté seul et le contexte présenté seul provoquent une réponse de peur. Au contraire, lors de l’ “unpairing”, le CS est aléatoirement décalé dans le temps avant la présentation de l’US, alors que le contexte reste constamment présent. Puis

le CS seul et le contexte seul sont présentés, pour étudier lequel des deux a acquis une valeur négative.

Notre modèle ne possède pas de mémoire de travail, et ne peut pas apprendre l'association CS-US si la présentation de l'US est trop éloignée dans le temps de la présentation du CS (ce qui est effectivement le cas de l'amygdale si PL n'est pas présent Wallis [2007]). Nous avons donc choisi de modéliser l' "unpairing" en faisant varier aléatoirement l'intensité du CS plutôt que son arrivée dans le temps. Nous gardons l'idée générale de l' "unpairing", qui est que le CS, par son caractère aléatoire, est moins adapté à prédire l'US que le contexte qui reste invariant. Notre modélisation de l'expérience pourrait correspondre au fait que le CS est, suivant sa date de présentation, plus ou moins saillant dans notre mémoire de travail au moment de l'arrivée de l'US.

Dans notre expérience reproduisant le "pairing", le contexte et le CS sont présentés au réseau et suivis de l'US durant 11 essais. Puis nous testons dans un essai la prédiction du réseau à partir du contexte seul, et dans un deuxième essai la prédiction à partir du CS seul. Le protocole est identique pour l'unpairing, à l'exception du fait que le CS, dans la phase d'activation, est multiplié par une constante tirée aléatoirement suivant une distribution uniforme entre 0 et 1 (correspondant aux cas extrêmes où le CS n'est pas présenté, ou présenté en même temps que l'US), et présente ainsi une forte variabilité dans sa saillance.

La figure 4.6 montre les prédictions du réseau lors du pairing. Les figures 4.6.A et B sont similaires au paradigme précédent. La figure 4.6.C montre les prédictions du réseau, i.e. l'activité de CeLON, suite au conditionnement. La prédiction basée sur le contexte seul (en vert) est peu différente du niveau d'activité au repos, tandis que la prédiction basée sur le CS sensoriel est plus importante. La même expérience a été réalisée avec un niveau d'ACh tonique plus élevé, fixé à 3.0 (soit trois fois le niveau au repos), représentant l'injection d'un agoniste de l'acétylcholine (figures 4.6.D,E,F). Fig.4.6.E montre que cette injection ne perturbe pas les prédictions du réseau. Contrairement au pairing normal où l'apprentissage est distribué entre LA et BAf, Fig.4.6.D montre que BAf est majoritairement responsable de la prédiction dans ce cas. L'apprentissage est alors essentiellement basé sur le contexte, au lieu du CS en condition normale (figure 4.6.F). Ces résultats reproduisent bien ceux obtenus par Calandreau *et al.* [2006a] lors du pairing.

La figure 4.7 affiche les prédictions de notre modèle lors de l'unpairing. La figure 4.7.B montre que le réseau apprend correctement à prédire l'arrivée de l'US, en se basant majoritairement sur l'association acquise dans BAf (figure 4.7.A). De plus, 4.7.C montre que notre unpairing provoque bien une association basée sur le contexte, validant ainsi notre choix de saillance aléatoire du CS. La même expérience effectuée avec un niveau d'ACh réduit (fixé à 50% du niveau de base) provoque l'association inverse, le CS est associé à l'US (figure 4.7.E). La réduction d'ACh ralentit l'apprentissage dans BAf, et permet à LA d'apprendre l'association malgré le fait que le CS soit moins bon prédicteur

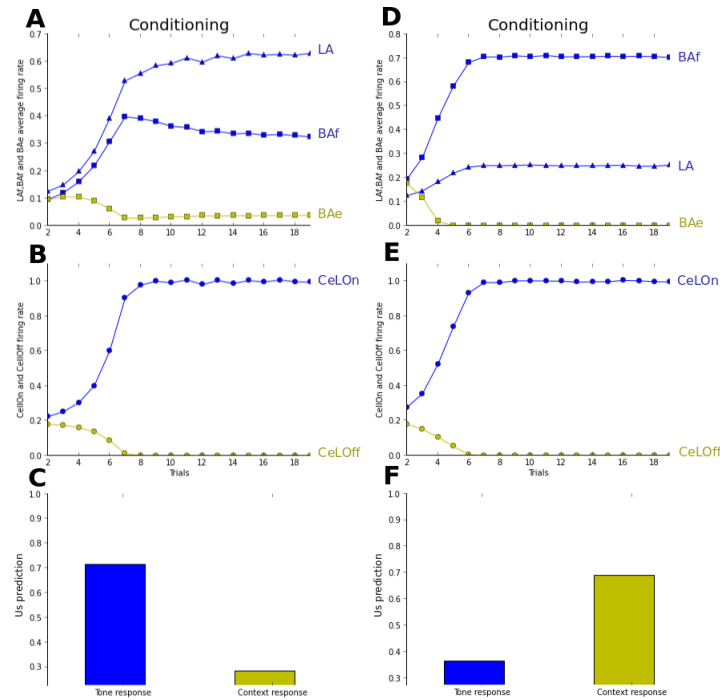


FIGURE 4.6 – (A,B,C) Expérience de pairing (cf. texte pour plus de détails). (D,E,F) Expérience de pairing avec augmentation du niveau de ACh (concentration d’ACh fixée à 3 fois celle du niveau normal au repos) (B,E) L’activité de CeLON augmente, et prédit correctement l’arrivée de l’US après quelques essais dans les deux expériences. (A) Acquisition sous des conditions normales (cf 4.4) (C) Le modèle est testé séparément avec uniquement le CS (bleu) ou uniquement le contexte (vert). Lors du pairing, l’association CS-US est préférentiellement apprise. (D) L’augmentation du niveau d’ACh provoque un apprentissage majoritairement de BAf. (F) Après apprentissage, la réponse est majoritairement basée sur le contexte, contrairement au pairing sous niveau normal d’ACh. L’ACh permet ainsi de moduler la nature de l’association apprise.

(figure 4.7.D). Nos résultats reproduisent donc bien là aussi les résultats obtenus par Calandreau *et al.* [2006a] pour l’unpairing normal et sous ajout d’un antagoniste de l’acétylcholine.

4.4 Discussion

Ce modèle d’association CS-US, basé sur des données de la biologie, est relativement simple. Il ne comporte que 5 populations, avec une dizaine d’unités par population, et une dizaine de connexions entre populations, dont la majorité ne possède pas d’apprentissage. Cependant il distingue différentes associations réalisées par l’amygdale et la compétition entre ces associations, illustrant ainsi certaines des stratégies mises en oeuvre pour un apprentissage pertinent dans une situation donnée. De plus, le fait qu’il soit simple permet d’identifier précisément les dynamiques impliquées et induites par des fonctions

4. Le rôle des différents circuits de l'amygdale dans l'acquisition de la réponse pavlovienne

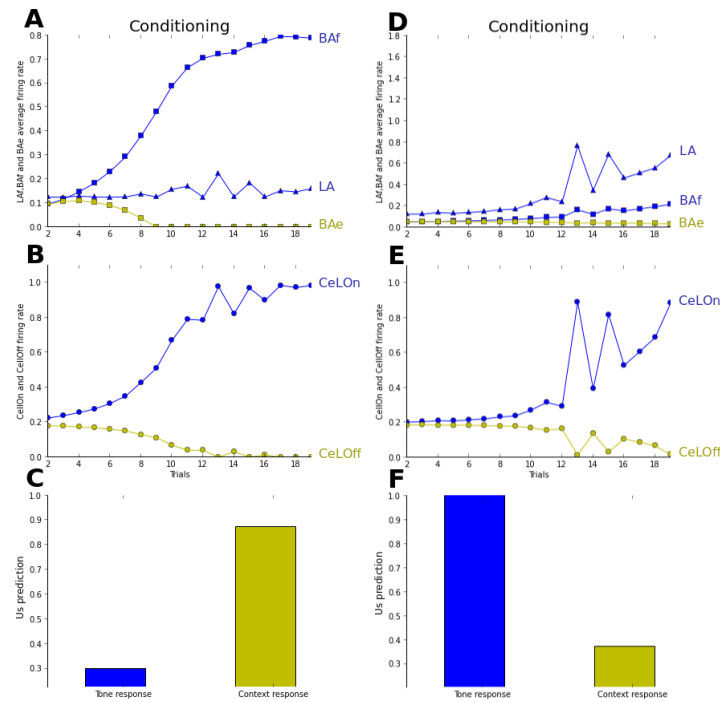


FIGURE 4.7 – (A,B,C) Expérience d'unpairing (cf. texte pour plus de détails) (D,E,F) Expérience d'unpairing sous déplétion du niveau d'ACh (50% du niveau au repos). (B,E) L'activité de CeLOn augmente jusqu'à prédire l'arrivée de l'US dans les deux expériences. (A) Contrairement au pairing et à l'acquisition normale où l'apprentissage est distribué, ici BAf apprend principalement à prédire l'US. (C) L'association apprise par BAf est une association contexte-US. Le CS sensoriel, par son caractère aléatoire, est moins bon prédicteur et n'a pas été associé à l'US. (D) Sous déplétion d'acétylcholine, LA apprend majoritairement. La réponse du réseau peut varier d'un essai sur l'autre en raison du caractère aléatoire du CS dans cette expérience. (F) Après apprentissage sous déplétion d'ACh, la réponse de peur est majoritairement associée au contexte. Le niveau d'ACh permet ainsi de moduler l'association apprise dans l'amygdale.

et flux d'informations similaires à la biologie, plutôt que de règles d'apprentissage complexes. Dans la même idée, nous avons également reproduit ce même modèle avec des neurones à spikes via le simulateur Nengo, montrant que ses propriétés ne sont pas dépendantes du modèle de neurones utilisé. Les expériences reproduites ont à la fois pour but de valider le modèle par rapport à la biologie, et d'analyser et mettre en valeur les différentes stratégies utilisées par l'amygdale pour un apprentissage pertinent.

4.4.1 L'apprentissage complémentaire de LA et BAf

Dans ce modèle, les neurones de LA et BAf reçoivent des entrées externes, respectivement les stimuli via le cortex sensoriel et le contexte via l'hippocampe, et apprennent à anticiper l'arrivée d'un US aversif.

Contrairement au modèle proposé par [Vlachos et al. \[2011\]](#), reproduisant

l'expérience de [Herry et al. \[2008\]](#), notre modèle ne se concentre pas uniquement sur BA et peut ainsi étudier les rôles respectifs de LA et de BAf. En particulier, [Calandreau et al. \[2006a\]](#) montre que lors de l'acquisition de la peur, l'association formée est préférentiellement basée sur le CS sensoriel, appris dans notre modèle par LA. Or [\[Anglada-Figueroa et Quirk, 2005\]](#) montre qu'une lésion de BA après cette même acquisition détériore l'apprentissage de la peur. Notre modèle explique cette apparente contradiction en montrant que la réponse de peur associée au CS est à la fois basée sur les réponses de LA et BAf. En particulier, il existe deux circuits participant à une réponse de peur basée sur l'association CS-US, les circuits LA-CeLON et LA-BAf-CeLON, tandis qu'il n'existe qu'un circuit, Hippo-BAf-CeLON, apprenant en fonction du contexte. Lors du pairing, le réseau est donc biaisé pour une réponse basée sur l'apprentissage CS-US, conformément aux résultats rapportés par [Calandreau et al. \[2006a\]](#). Et une lésion de BA après apprentissage coupe le circuit LA-BAf-CeLON, et nuit donc bien à la réponse de peur acquise basée sur le CS. Dans cette optique, notre modèle prédit également qu'une lésion de BA, après une acquisition de la peur présentant une concentration d'ACH diminuée (par exemple sous l'effet d'un antagoniste de l'acétylcholine comme la scopolamine), devrait moins détériorer la réponse de peur. En effet l'activation des neurones de BA étant moins facilitée lors de l'apprentissage, le circuit LA-BAf-CeLON aura moins d'importance dans la décision de CeLON, et le circuit intact LA-CeLON apprendra davantage.

Les résultats décrits plus haut ont été obtenus avec une connexion non plastique entre LA et BAf. Dans des expériences, non rapportées ici, nous avons obtenu des résultats identiques à ceux décrits dans la section précédente. Ceci suggère que l'apprentissage entre LA et BAf n'est pas nécessaire pour les paradigmes testés ici. Une hypothèse pourrait être que cette plasticité est nécessaire à l'apprentissage de valeurs non utilisées dans nos expériences et apprises par BA [\[Belova et al., 2015\]](#). Une autre hypothèse pourrait être son rôle pour apprendre des discriminations plus précises, comme par exemple moduler la valeur d'un contexte en fonction d'un CS appris dans LA.

4.4.2 L'extinction apprise par BAe

Dans notre modèle, les neurones de BAe apprennent à inhiber dynamiquement la réponse de peur grâce aux projections venant d'IL, sans que cette réponse ne soit oubliée. Nos expériences reproduisent ainsi le fait que les neurones de LA peuvent continuer à réagir durant l'extinction [\[Maren, 2005; Repa et al., 2001\]](#) et que la réponse de peur est immédiatement rétablie en dehors du contexte d'extinction [\[Herry et al., 2008\]](#).

Notre modèle se concentre sur l'amygdale, et n'implémente pas comment une situation d'extinction est détectée dans IL. Une extension possible du modèle serait de proposer une implémentation d'IL permettant une telle dé-

tection, par exemple au moyen d'une population similaire à celle utilisée pour l'ACh, qui stockerait pour un certain contexte les erreurs de prédiction négatives (contrairement à U_{ACh} qui prend en compte les valeurs absolues quel que soit le signe de l'erreur), et enverrait un signal à l'amygdale dès que son activité est suffisante.

Notre modèle propose un apprentissage de l'extinction dans BAe, basé sur un signal envoyé par IL. Une hypothèse alternative serait que l'apprentissage d'extinction soit uniquement dans IL. BAe n'apprendrait donc rien, et aurait uniquement pour rôle de relayer l'activité d'IL pour inhiber la réponse de peur. Cette hypothèse nous paraît cependant moins probable, car comme montré dans [Herry et al. \[2008\]](#), l'inactivation de BA empêche complètement l'extinction. Après réactivation, les animaux font de nouveau preuve d'un niveau de "freezing" élevé, ce qui montre que l'apprentissage qui a lieu dans IL, lorsque BA est inactivé, n'est pas suffisant pour apprendre l'extinction, et semble en faveur d'un apprentissage dans BAe (mais n'exclut pas un apprentissage partiellement dans IL).

L'apprentissage de l'extinction du modèle prend plus de temps que l'acquisition, et les neurones de BAf commencent à décroître leur activité suite à l'augmentation de l'activité de BAe, en cohérence avec [[Herry et al., 2008](#)]. Notre modèle explique le délai dans l'acquisition de l'extinction par le fait que durant les premiers essais de l'extinction, les neurones de BAf ont encore une activité élevée, et inhibent donc les neurones de BAe, rendant l'apprentissage hebbien des connexions d'IL à BAe plus long, comme l'activité postsynaptique, de BAe, est plus faible. Une autre explication, qui n'exclut pas la précédente, est qu'IL doit au préalable avoir appris et envoyé le contexte d'extinction à BAe, avant que BAe ne puisse apprendre, retardant donc de plusieurs essais le début de l'apprentissage dans BAe. Ces deux hypothèses restent en faveur d'un apprentissage de l'extinction par BAe.

D'autres études montrent cependant un apprentissage de l'extinction ne nécessitant pas l'amygdale. En particulier, [Anglada-Figueroa et Quirk \[2005\]](#) rapporte que des rats avec des lésions de BA ne sont pas gênés dans leur acquisition de l'extinction. Cette expérience est de plus reproduite par [Moustafa et al. \[2013\]](#) dans son modèle de l'implication de l'hippocampe dans le conditionnement à la peur. Il prend en compte les mêmes régions que notre modèle, mais l'extinction dépend d'un circuit venant d'IL vers les ITC ("intercalated cells", neurones "intercalés" entre les noyaux latéral et basal et le noyau central de l'amygdale), qui à leur tour inhibent la sortie de l'amygdale et donc la réponse de peur. Ces résultats semblent donc en contradiction avec certains des résultats mentionnés précédemment ([Herry et al. \[2008\]](#), [Vlachos et al. \[2011\]](#)), et reproduits par notre modèle. Cependant, dans une série d'expériences plus extensives, [Laurent et al. \[2008\]](#) montre que des rats avec des lésions de BA peuvent en effet apprendre l'extinction, mais contrairement aux rats intacts ne s'en souviennent pas si testés le jour d'après. Ainsi ces résultats peuvent

être expliqués par la présence de deux circuits parallèles mais impliqués à deux échelles de temps différentes [Pauli *et al.*, 2011]. Le chemin IL-ITC-CeM serait impliqué dans un apprentissage à court terme, de type mémoire de travail, de l’extinction, tandis que le circuit IL-BA-CeL serait impliqué dans un apprentissage à plus long terme. Cette vision est également un argument en faveur d’une extension de notre modèle par une représentation plus réaliste d’IL. De manière plus générale, elle montre que les valeurs calculées dans les boucles frontales de prise de décision, dans un apprentissage dirigé par un but, peuvent aider l’apprentissage répondant, en lui fournissant de la flexibilité lors de situations changeantes comme l’extinction.

4.4.3 Un rôle simplifié de CeLOn.

Dans notre modèle, CeLOn et CeLOff intègrent les prédictions venant de LA et de BA, afin de déterminer la réponse du réseau. Cependant, nous proposons un rôle différent pour ces deux populations. CeLOn représente la sortie du réseau, tandis que CeLOff a uniquement pour rôle de moduler l’activité de CeLOn. Cette dissymétrie se retrouve dans Ciocchi *et al.* [2010], qui montre que CeLOn est responsable des pics d’activité du noyau médial de CeA prédisant l’arrivée d’US, tandis que CeLOff a un effet plus tonique sur CeM. Une extension intéressante à notre modèle serait de prendre en compte l’influence de CeLOff et d’ITC dans la modulation de la réponse de peur [Lee *et al.*, 2013; Amano *et al.*, 2010].

Une autre simplification est l’absence de plasticité dans CeA, plasticité pourtant présente, particulièrement dans CeL [Pare *et al.*, 2004]. Cette plasticité ne devrait pas jouer un rôle majeur dans les paradigmes que nous avons reproduits, mais pourrait participer à l’association de réponses à partir des entrées thalamiques de CeA. Ces entrées thalamiques, comme rapporté par Yu et Dayan [2005], permettraient dans les cas d’ “incertitude imprévue” [Yu et Dayan, 2005] de trouver une nouvelle règle plus rapidement qu’à partir des entrées de haut niveau du cortex, n’ayant pas forcément formé la représentation adéquate. La noradrénaline pourrait de plus permettre de biaiser l’amygdale en faveur des entrées thalamiques, comme décrit dans Johnson *et al.* [2011a], et favoriser ainsi leur association à l’US. Un autre rôle possible pour la plasticité de CeL pourrait être l’apprentissage d’une réponse adaptée à partir des valeurs prédites de LA, BA, mais aussi d’IL via les ITC, lorsque différents US sont présents. Par exemple si deux CS associés à des US différents sont présentés simultanément, une plasticité au niveau d’IL pourrait permettre d’apprendre lequel des deux prédomine, sans remettre en cause leurs valeurs respectives apprises dans LA ou BA, valeurs toujours pertinentes si ces CS sont présentés seuls.

4.4.4 Impact d'ACh sur l'apprentissage

D'un point de vue physiologique, l'ACh tonique module le fonctionnement des neurones en augmentant le ratio signal-sur-bruit (RSB) dans le cortex [Pauli et O'Reilly, 2008], l'hippocampe [Hasselmo, 2006], et BA [Unal *et al.*, 2015]. D'un point de vue fonctionnel, il est considéré que l'ACh tonique représente l'incertitude connue (stochasticité) [Yu et Dayan, 2005].

Nous avons représenté l'effet de l'ACh par une augmentation des réponses des neurones de BA proportionnellement à leurs niveaux d'activation. Ceci augmente donc le RSB. D'autre part, ceci permet de favoriser l'acquisition d'une réponse basée sur le contexte dans BA, plutôt que basée sur le CS sensoriel dans LA. En effet, dans nos expériences reproduisant celles de Calandreau *et al.* [2006a], augmenter la concentration d'acétylcholine permet au réseau d'apprendre à partir du contexte plutôt qu'à partir du CS, et dans le cas de l'unpairing, la stochasticité provoquée par le caractère aléatoire du CS provoque un niveau d'acétylcholine plus haut, et aide à la formation d'une association dans BA plutôt que dans LA. Dans les deux cas, l'acétylcholine aide à la formation d'une règle contextuelle plutôt que sensorielle.

La principale différence entre les expériences de Calandreau *et al.* [2006a] modifiant l'effet de l'ACh et notre modèle est la région sur laquelle les niveaux d'ACh sont modifiés (respectivement l'hippocampe et l'amygdale). Cependant cette modification impacte dans les deux cas le même circuit fonctionnel, hippocampe-BA-CeLON impliqué dans l'apprentissage du contexte et donc son influence dans la compétition entre acquisition sensorielle et contextuelle. Ainsi, le modèle prédit qu'effectuer l'expérience de Calandreau *et al.* [2006a] en modifiant le niveau d'ACh dans BA devrait aboutir aux mêmes résultats. Cette prédiction est en partie supportée par McIntyre *et al.* [2003], qui montre une corrélation positive entre le niveau d'ACh dans l'amygdale et la performance dans une tâche dépendant de l'hippocampe.

Les variations du niveau d'ACh dans BA ont également un impact sur l'autre compétition étudiée ici, la compétition entre acquisition et extinction de la peur. En particulier, Schroeder et Packard [2004] montre un effet facilitateur de l'ajout d'un agoniste de l'ACh lors de l'extinction contextuelle. De plus Prado-Alcalá *et al.* [1994] rapporte, après apprentissage de l'extinction, un rétablissement de la réponse de peur après injection d'un antagoniste de l'ACh. Notre modèle reproduit un effet similaire avec une diminution de l'ACh après extinction (Fig.4.5.G,H).

Ces résultats prédits ou reproduits mettent l'accent sur le rôle central de l'ACh dans la compétition entre les différentes associations effectuées par l'amygdale. L'ACh, en augmentant les activités de BA, permet l'acquisition d'associations alternatives à l'association simple CS-US lorsque les prédictions basées sur cette dernière sont trop incertaines.

Un point qui pourrait être une extension future du modèle est la manière

dont l'acétylcholine est générée. En accord avec le rôle proposé de l'ACh avec l'incertitude connue [Yu et Dayan, 2005], celle-ci augmente en fonction de l'incertitude récemment observée. Cependant, notre génération ne dépend uniquement de l'incertitude récente, générale, et en particulier n'est pas situation-spécifique, ou stimulus spécifique. Si nous supposons une expérience qui alterne une règle dans laquelle la prédiction est incertaine, et une règle dans laquelle notre prédiction est fiable, l'incertitude telle que mesurée dans notre modèle va être élevée dans les deux règles, alors qu'il pourrait être intéressant de mesurer une incertitude propre à chaque règle. De même, si je repasse brutalement dans un contexte que je sais incertain, il peut être intéressant qu'ACh reflète directement le niveau d'incertitude de ce contexte, et n'ait pas à augmenter petit à petit au fur et à mesure des erreurs. En particulier, nous avons vu précédemment que l'amygdale influençait le stockage des contextes dans l'hippocampe, directement au niveau de l'entorhinal, ou indirectement via l'acétylcholine. Ainsi les contextes incertains, pour lesquels l'amygdale fait des erreurs de prédictions, seraient davantage présents dans l'hippocampe, libérant ainsi davantage d'acétylcholine en situation incertaine. Considérer les interactions réciproques entre l'amygdale et l'hippocampe permettrait ainsi d'obtenir une acétylcholine reprenant une incertitude contexte-spécifique et émergente de ces interactions.

4.4.5 Relations avec les paradigmes de conditionnement

Parmi la multitude de paradigmes du conditionnement pavlovien, nous avons choisi de nous concentrer sur l'extinction, le pairing et l'unpairing, comme ils mettaient en avant le rôle des différents circuits de l'amygdale dans la souplesse de l'apprentissage pavlovien.

Notre modèle est capable également d'expliquer le blocking, mais sans apporter d'originalité par rapport à l'explication proposée par la règle de Rescorla-Wagner : comme nos règles d'apprentissage sont dépendent multiplicativement de l'erreur de prédiction, si celle-ci est nulle, nous n'apprenons pas. Dans le cas du blocking, il n'y a donc pas d'apprentissage si un second stimulus est présenté en même temps qu'un stimulus prédisant déjà suffisamment l'US.

L'inhibition conditionnée pourrait être un exemple d'interaction entre apprentissage motivé et opérant, mais les mécanismes précis de celle-ci sont sujets à débat. [Christianson *et al.*, 2012] suggère néanmoins l'implication de structures de l'apprentissage motivé comme OFC, vmPFC et le striatum, et de leurs interactions avec l'amygdale. Notre modèle pourrait éventuellement reproduire l'inhibition conditionnée par l'ajout de populations similaires à BAe, mais recevant des informations relatives au CS à inhiber plutôt que relatives au contexte.

Dans une certaine mesure, notre modèle reproduit l'effet mis en évidence par la règle de Mackintosh, "les meilleurs prédictors apprennent plus". Notre

implémentation basée sur un apprentissage hébbien implique que les neurones avec une activité post-synaptique apprennent davantage. Ainsi, un CS bon prédicteur aura des activités post-synaptiques dans LA et BAf plus élevée, et apprendra donc davantage, qu'un CS pas encore prédicteur. La présence d'interneurones dans LA et BLA pourrait fournir un substrat biologique additionnel : l'activité des CS déjà bons prédicteur inhiberait celles, plus faibles, des autres CS par les interneurones inhibiteur, renforçant ainsi leur différence de vitesse d'apprentissage. [Lee *et al.*, 2013] propose notamment un effet de la dopamine sur ces interneurones, augmentant le ratio signal-sur-bruit, et favorisant ainsi davantage les bons prédicteurs, avec des activités plus élevées. L'erreur de prédiction présente dans l'associabilité de la règle de Mackintosh pourrait ainsi représenter l'action de l'erreur de prédiction de la dopamine sur les interneurones de LA et BA, augmentant l'inhibitions des faibles prédicteur, et favorisant un apprentissage basé sur les meilleurs prédicteurs.

Au contraire, notre modèle ne permet pas d'expliquer la règle de Pearce-Hall [Pearce et Hall, 1980]. Cette règle est cependant reliée à la notion d'incertitude, et pourrait être expliquée par l'effet de l'acétylcholine, et donc de l'incertitude, sur l'entrée de l'amygdale. En favorisant la création de catégories sensorielles dans le cortex, et l'apprentissage dans l'hippocampe, l'ACh provoquerait une représentation de ceux-ci particulièrement importante dans le cortex et l'hippocampe. Ainsi, comparé à un stimulus nouveau, ces CS seraient mieux représentés et plus saillants dans les différentes entrées de l'amygdale. Notre apprentissage hébbien de LA et de BAf, dépendant de l'activité pré-synaptique, apprendrait ainsi davantage à partir de ces stimuli qu'à partir de stimuli nouveaux et moins présents en entrée. Ajouter un modèle plus réaliste de cortex sensoriel et d'hippocampe en entrée pourrait ainsi nous permettre d'expliquer également les phénomènes attentionnels de la règle de Pearce-Hall.

4.4.6 Conclusion

Notre modèle d'amygdale possède trois circuits pouvant apprendre une association, ainsi qu'un mécanisme de compétition entre eux, pour sélectionner l'association pertinente dans une situation donnée. Au niveau de complexité le plus simple, un CS sensoriel peut être associé à un US dans LA, comme déjà proposé par de nombreux modèles basés sur la règle de Rescorla-Wagner [Rescorla et Wagner, 1972]. La valeur de l'US peut ensuite directement être transmise à CeA, pour produire la réponse adaptée. Cependant, dès que l'arrivée d'US est régie par une règle plus complexe, ce circuit simple peut ne pas être adapté à la prédiction de l'US. Le modèle commet alors des erreurs, faisant augmenter la stochasticité apparente de la situation. Ces erreurs provoquent ainsi une augmentation du niveau d'ACh tonique, excitant BA, et promouvant ainsi des circuits correspondant à des stratégies alternatives, comme l'apprentissage de la règle basé sur le contexte, ou l'extinction d'une réponse devenue incor-

recte. L’extinction illustre particulièrement comment les valeurs plus flexibles apprises par l’apprentissage motivé peuvent aider l’apprentissage répondeur. Dans la même idée, un autre circuit, non représenté par notre modèle, est le circuit formé par les projections de PL sur BA [Pendyam *et al.* \[2013\]](#), permettant d’apprendre des associations quand le CS et l’US sont séparés dans le temps, grâce à une activité soutenue dans PL.

Considérée seule, l’amygdale est une structure simple, avec des régions relativement primitives [[Cardinal *et al.*, 2002](#)], apprenant les valeurs des différents CS, pour préparer leur arrivée par des réponses pavloviennes immédiates, telles que l’action d’orientation, le “freezing”, la réponse hormonale, la salivation. Considérée dans une vue systémique, les valeurs et associations apprises par l’amygdale sont également transmises aux systèmes responsables de l’apprentissage motivé, comme par exemple la nature de l’US [[Holland et Gallagher, 2004](#)], leur donnant ainsi des informations sur la valeur des différents buts, pour pouvoir déterminer le plus intéressant en fonction du besoin et de la situation actuelle. Un autre apport de l’apprentissage de l’amygdale pour l’apprentissage motivé est la possibilité de faire des statistiques sur les prédictions, illustré ici par l’acétylcholine, permettant d’apporter une information cruciale, la stochasticité de la situation. De manière similaire à notre modèle, un apprentissage motivé pourrait prendre en compte des informations statistiques similaires pour choisir des stratégies alternatives, comme sélectionner les buts et actions sur lesquels il a la plus grande confiance, et qui ont donc de bonnes chances d’obtenir des résultats malgré l’incertitude de la situation.

Les stratégies, modélisées ici dans le but d’expliquer la flexibilité, l’adaptation dont nous font preuve nos différents systèmes d’apprentissage, pourraient être exploitées pour améliorer des algorithmes de machine learning, comme par exemple le deep learning. Les premières couches du deep learning apprennent des régularités de manière non supervisée, codant ainsi des valeurs similaires à celles du cortex sensoriel. La dernière couche, de sortie, apprend en fonction des couches précédentes la réponse adaptée, et pourrait être vue comme un modèle plus complexe de notre apprentissage de LA-CeL. Ainsi, nous pourrions ajouter en entrée directe de la dernière couche un système apprenant certains cas particuliers, représentant l’hippocampe, et privilégier ce système pour les entrées à fortes stochasticité, similaire à notre ACh. Les entrées sur lesquelles la partie originelle de l’algorithme fait des erreurs, détectées par une mesure de la stochasticité, seraient ainsi apprises en tant que cas particuliers par la couche de sortie, ce qui devrait permettre une amélioration du taux de succès de l’algorithme.

Chapitre 5

L'apprentissage pavlovien de la boucle cortico-striatale limbique guide l'apprentissage opérant.

5.1 Introduction

Dans le chapitre précédent, nous regardions du point de vue de l'apprentissage pavlovien, comment apprendre différentes associations et valeurs pertinentes en fonction de la situation, notamment en fonction de valeurs fournies par l'apprentissage motivé. Dans ce chapitre, nous étudions au contraire l'apprentissage opérant, et comment l'apprentissage pavlovien peut aider celui-ci. Nous nous intéresserons à comment les valeurs associées au CS peuvent être utilisées dans la sélection de l'action, permettant de sélectionner l'action la plus adaptée à la situation.

Pour cela, nous reprenons et améliorons ici un modèle des boucles cortico-striatales existant, le modèle de [Guthrie *et al.* \[2013\]](#), repris par [\[Topalidou *et al.*, 2015\]](#). En effet, si les boucles impliquant le striatum dorsal et dorso-médial sont davantage associées à l'apprentissage opérant, apprentissage d'actions dirigées vers un but, ou apprentissage d'habitudes [Balleine et O'Doherty \[2010\]](#), le shell de Nacc correspondrait à un apprentissage pavlovien, [Penner et Mizumori \[2012\]](#), apprenant à associer les CS et "états" présents dans OFC à leurs valeurs estimées. Ainsi nous regardons l'apprentissage de la boucle limbique, en cohérence avec les similitudes observées par [\[Cassell *et al.*, 1999\]](#) entre l'amygdale et NAcc, dans un rôle proche du conditionnement répondant de l'amygdale.

Le modèle considéré possède un apprentissage de valeur pour les CS, reprenant l'apprentissage attribué au shell, et étudie comment cet apprentissage associatif permet de sélectionner l'action adaptée dans des paradigmes de conditionnement opérant tel que [Pasquereau *et al.* \[2007\]](#). Il modélise également la formation d'habitudes dans les parties plus dorsales du striatum, et combine

ces deux types d'apprentissage, apprentissages dirigés ou non vers le but, pour l'opérant. Nous nous servons de ce modèle pour étudier comment l'apprentissage associatif peut aider l'apprentissage opérant, et plus particulièrement, comment l'association de valeurs aux différents états sensoriels présents dans OFC interagit avec l'apprentissage d'actions directement à partir des CS.

5.2 Modèle préexistant

Le modèle de [Guthrie et al. \[2013\]](#) est un modèle reprenant plusieurs boucles des ganglions de la base, nommées cognitive, associative et motrice, et montrant leur rôle dans le conditionnement opérant (Figure (figure 5.1)). Pour respecter les notations du modèle que j'ai repris, et comme la description du modèle que j'utilise repose sur celle de [Guthrie et al. \[2013\]](#), je parle dans ce chapitre et le suivant de boucle cognitive pour désigner la boucle limbique. La boucle cognitive (OFC-striatum ventral-gpi cognitif-thalamus cognitif) apprend ainsi à donner une valeur aux CS présentés au réseau par renforcement des connexions entre l'OFC et le shell. Ces CS peuvent représenter à la fois des CS simples ou plus intégrés, ou états sensoriels complexes de OFC.

Ces valeurs sont transmises par les connexions entre OFC et dmS à la boucle associative (dlPFC-dmS) qui associe les CS aux actions à effectuer pour les atteindre (et dépendant donc de la position des CS). La boucle motrice reçoit les différentes actions possible et leur saillance respectives grâce aux connexions entre le striatum associatif et le GPi moteur, et permet d'accentuer leurs différences de saillance. Quand une action est devenue suffisamment saillante, elle est sélectionnée comme réponse du réseau.

L'apprentissage associatif CS-US est appris dans les poids w entre l'OFC et le shell en fonction de l'erreur de prédiction de manière similaire à la règle de Rescorla-Wagner :

$$dw = err * V_{ctx} * V_{str} * \alpha$$

où α est le taux d'apprentissage, V_{ctx} est l'activité corticale présynaptique, V_{str} l'activité striatale postsynaptique, et err l'erreur de prédiction envoyée par une partie extérieure au modèle, et non implémentée par des neurones. Cette partie fonctionne également de manière similaire à Rescorla-Wagner, effectuant des prédictions, et calculant l'erreur entre l'US reçu et celui attendu.

Ce modèle est ensuite testé sur la tâche de [\[Pasquereau et al., 2007\]](#), dans laquelle 4 CS avec différentes probabilités de récompense peuvent être soumis à l'animal. A chaque essai, 2 des CS sont choisis aléatoirement et présentés dans deux positions aléatoires parmi quatre positions possibles. Le modèle a un temps limité pour prendre une décision, et recevoir la récompense associée au CS sélectionné. Si aucune décision n'est prise, le modèle n'est pas récompensé. Un essai est considéré réussi quand le CS le plus récompensé des deux présentés est sélectionné. La figure 5.2 tirée de [Guthrie et al. \[2013\]](#) montre que le modèle

5. L'apprentissage pavlovien de la boucle cortico-striatale limbique guide l'apprentissage opérant.

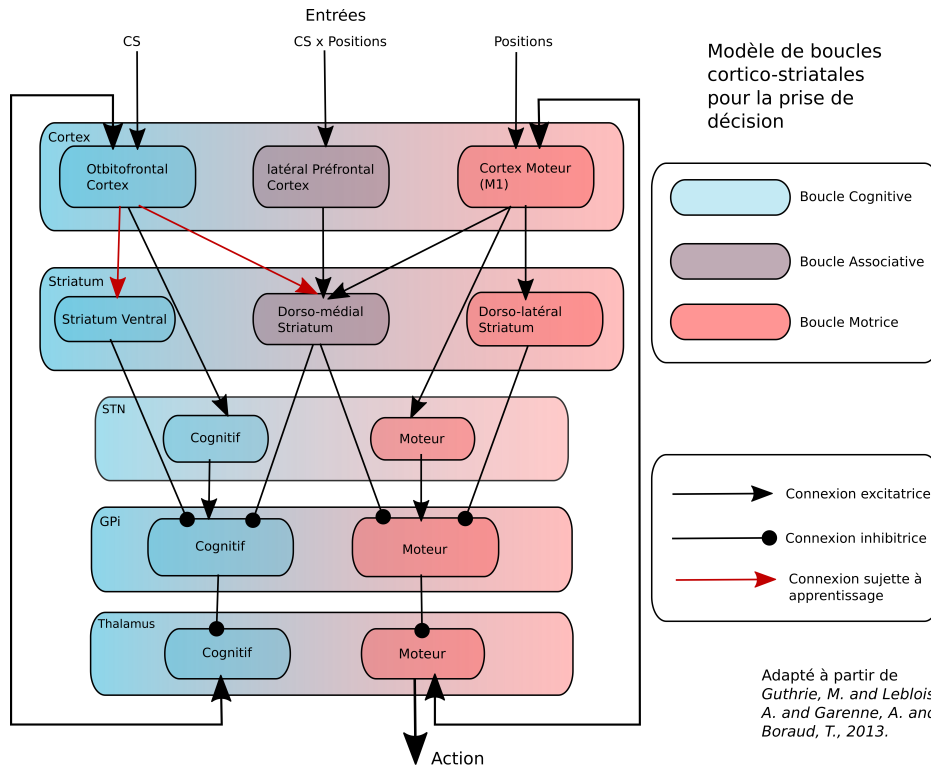


FIGURE 5.1 – Modèle des boucles cortico-striatales de Guthrie *et al.* [2013]. La boucle cognitive (OFC-striatum ventral-gpi-thalamus dans ce schéma) apprend par renforcement des connexions entre l’OFC et le shell à donner une valeur aux CS présentés au réseau. Ces valeurs sont reçues par la boucle associative (dIPFC-dmS) via les connexions OFC-dmS, qui associent aux CS les actions pour les atteindre (dépendant donc de leur positions). La boucle motrice reçoit les différentes actions possibles et leur saillance respectives via les connexions entre le striatum associatif et le GPi moteur, et fait émerger l’une. Quand une action est suffisamment saillante, elle est sélectionnée comme réponse du réseau. Un apprentissage hebbien entre OFC et dmS a été ajouté par [Topalidou *et al.*, 2015], et permet de renforcer les actions habituellement sélectionnées à partir d’un CS, et permet la formation d’habitude. Ainsi, ce modèle combine un apprentissage pavlovien de la boucle cognitive, qui apprend les valeurs des différentes situations présentes, et un apprentissage d’actions habituelles, apprenant des associations directes entre CS et actions.

parvient correctement à apprendre le CS le plus récompensé, ce quel que soit la combinaison (6 possibles) de CS présentés. De plus, il montre que la présence de bruit est suffisante pour permettre une exploration dans un modèle “naïf”, i.e. avec des poids initialement uniformes, en l’absence de règle apprise (car dans le cas contraire le modèle serait resté bloqué sur son choix initial).

Nous utiliserons par la suite l’implémentation de Guthrie *et al.* [2013] reprise par [Topalidou *et al.*, 2015]. Cette implémentation rajoute un apprentissage hebbien entre OFC et dmS (figure 5.1, et explique grâce à cet apprentissage la formation d’habitudes : si un CS est couramment associé à une action, la connexion correspondante va être renforcée entre OFC et dmS, et va per-

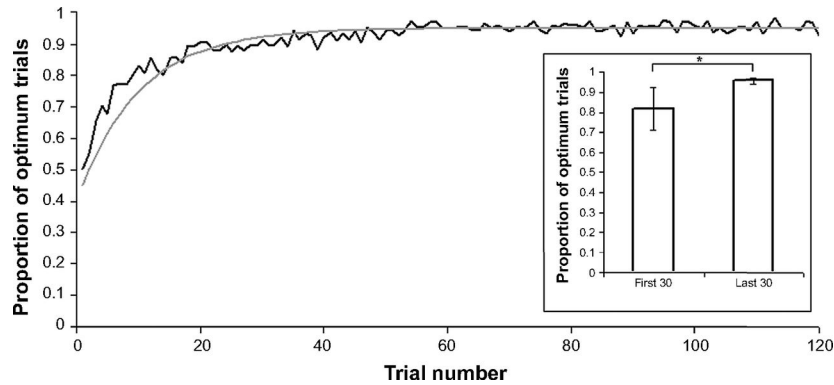


FIGURE 5.2 – Figure tirée de Guthrie *et al.* [2013]. Taux de réussite du modèle dans la tâche de [Pasquereau *et al.*, 2007]. Un essai est considéré réussi si le CS avec la plus grande probabilité de récompense est sélectionné. Ce modèle parvient donc bien à apprendre la tâche de [Pasquereau *et al.*, 2007], et correctement distinguer le meilleur choix dans chacune des 6 combinaisons de 2 CS parmi les 4 CS possibles.

mettre au CS de déclencher directement l’action, sans avoir besoin de la valeur apprise par l’apprentissage par renforcement entre l’OFC et le shell.

$$dw = V_{ofc} * V_{str-asso} * \alpha$$

Où w est le poids des connexions entre l’OFC et le dmS, V_{ofc} l’activité du neurone présynaptique, $V_{str-asso}$ l’activité du neurone postsynaptique, et α le taux d’apprentissage. Ce modèle combine donc deux types différents d’apprentissages. L’apprentissage opérant, modélisé par les boucles associatives et motrices, apprend sur le long terme, de manière indépendante de la récompense. Cet apprentissage permet la formation d’habitudes, qui ont pour avantage d’être robustes, résistantes au bruit, et de permettre une action rapide, en contrepartie de quoi elles sont lentes à s’adapter au changement de règle. En revanche, l’apprentissage pavlovien associant états et valeurs entre l’OFC et le shell apprend plus vite, et peut plus facilement s’adapter à un changement de la règle actuelle.

Ce modèle illustre la synergie possible entre ces deux différents type d’apprentissages, tout deux coexistant dans le cerveau [Packard et Knowlton, 2002]. La synergie est obtenue en permettant au système pavlovien d’influencer la saillance des entrées du système opérant. Ainsi, ce dernier, plus lent, est guidé par l’apprentissage plus rapide du premier, qui lui met en valeur les meilleurs éléments, ou états intéressants. Au contraire, lorsque les habitudes sont formées, l’arrivée du CS permet de déclencher directement l’action, évitant le calcul plus coûteux d’évaluer les différents CS pour choisir le meilleur. Les interactions permettent ainsi au système global de bénéficier à la fois de l’avantage de temps de convergence du système modèle based associatif, et de la rapidité de décision du système habituel. Ces interactions sont donc particulièrement intéressantes pour le reversal, en permettant un changement guidé

par la boucle cognitive.

5.3 Le calcul de l'erreur de prédiction, via la neuromodulation

Un des points du modèle de Guthrie *et al.* [2013] qui peut être amélioré dans une approche centrée sur les mécanismes de la biologie est son “critique”, son système calculant l'erreur de prédiction pour guider l'apprentissage du shell. Dans le modèle d'origine, ce “critique” est modélisé de manière non bio-inspirée. Il apprend à prédire l'arrivée de récompense pour le CS ciblé par l'action d'une manière similaire à Rescorla-Wagner, en se corrigeant via la différence entre la récompense qu'il prédit et celle arrivée. Cette récompense est également envoyée aux connexions entre OFC et le shell. Deux remarques peuvent être émises face à ce système de critique : il n'est que peu plausible biologiquement, et est “dé-corrélé” de l'apprentissage du modèle. En effet si ce “critique” prédit correctement la récompense de l'action sélectionnée, il n'a plus d'erreur de prédiction, ce qui arrête l'apprentissage entre l'OFC et le striatum. Le fait que le critique ait appris, par exploration, qu'un autre CS soit plus intéressant n'influence pas le choix de la boucle sensorielle. Par exemple, si le critique apprend trop rapidement lors d'un reversal, il apprendra à prédire que le CS précédemment récompensé n'est plus suivi de récompense avant que le réseau ne puisse changer sa décision. Ainsi, le modèle peut rester bloqué sur un choix non-optimal, sans pouvoir apprendre une meilleure action, alors que le système de critique a pourtant appris le nouveau choix optimal. Au contraire, s'il apprend trop lentement, ses erreurs continueront à entraîner le réseau longtemps après que celui-ci a appris le bon CS, et ce CS sera donc renforcé au niveau des connexions OFC-shell, entraînant un reversal très long.

D'un point de vue fonctionnel, nous voyons donc la nécessité de corréliser ces deux types d'apprentissage. D'un point de vue biologique, cette erreur transmise par le critique peut être attribuée aux projections dopaminergiques de VTA sur les connexions entre OFC et le Shell Joel *et al.* [2002]. VTA, qui reçoit l'US, confronterait cet US à la prédiction effectuée par le shell pour émettre l'erreur de prédiction, comme proposé par des règles d'apprentissage basées sur la biologie O'Reilly *et al.* [2007a]. Ainsi nous prenons l'erreur de prédiction, $US - US_{\text{prédit}}$, comme la différence, effectuée dans VTA, entre l'arrivée excitatrice de l'US et son inhibition par les prédictions d'US effectuées par le striatum ventral. Notre changement consiste donc à modifier l' $US_{\text{prédit}}$ calculé précédemment, et à remplacer son apprentissage précédent, une variable par CS stockant l'historique de récompense associé à ce CS, par la valeur donnée par l'activité du striatum ventral. De plus le striatum ventral apprend dans notre modèle en fonction de l'erreur de prédiction, comme la valeur de $US_{\text{prédit}}$ calculée à part dans le modèle de Guthrie *et al.* [2013].

5.3. Le calcul de l'erreur de prédiction, via la neuromodulation

Paramètres du modèle		
Paramètre	Description	Valeur
$init_critic$	Valeur initiale du critique (ancien critique)	0.25
α_critic	Taux d'apprentissage du critique (ancien critique)	0.2
α_str	Constante modélisant le poids des connexions de STR au système dopaminergique	0.04
α_LTP	Taux d'apprentissage lors de LTP	0.00008
α_LTD	Taux d'apprentissage lors de LTD	0.00005
g_cog_cog	Gain des connexions entre cortex cognitif et striatum ventral	1.2
g_cog_ass	Gain des connexions entre cortex cognitif et striatum dorsal	0.3

FIGURE 5.3 – Description et valeur des paramètres ajoutés ou modifiés par rapport à [Topalidou et al. \[2015\]](#).

Une autre solution envisagée, et testée dans un premier temps, est de conserver le calcul de l'erreur de prédiction de [Guthrie et al. \[2013\]](#), mais d'ajouter une influence excitatrice des valeurs prédites pour les CS par le critique sur l'aire OFC. Ainsi, si le critique apprenait le choix optimal plus rapidement que le modèle, les valeurs apprises influençaient le modèle, et le guidait vers la bonne décision, évitant ainsi que le réseau ne converge vers un stimulus non récompensé. A l'opposée, si le critique était long à apprendre, ces faibles prédictions n'influençaient que peu le réseau, et ralentissaient donc son apprentissage hebbien, évitant un l'over-training trop important des connexions. Si cette solution était fonctionnelle, elle n'avait pas les bases biologiques de la précédente, et possédait une redondance dans l'apprentissage de $US_{predict}$, appris à la fois par le critique et dans le shell.

Ainsi, de manière identique à [Guthrie et al. \[2013\]](#), les valeurs des CS sont apprises dans les poids w entre l'OFC et le shell du striatum ventral en fonction de l'erreur de prédiction :

$$dw = err * V_{ctx} * V_{str} * \alpha_LT$$

où α_LT est le taux d'apprentissage constant, mais différent suivant qu'il s'agisse d'augmentation des poids sur le long terme (LTP, pour "Long Term Potentiation") ou de LTD (pour "Long Term Depletion", diminution sur le long terme). V_{ctx} est l'activité corticale présynaptique, V_{str} l'activité striatale postsynaptique, et err l'erreur de prédiction, calculée dans notre modèle par :

$$ERR = reward - \alpha_str * V_{str}$$

où α_str est une constante de normalisation, permettant de ramener les valeurs de V_{str} entre 0 et 1. Ainsi, cette erreur de prédiction permet à la fois d'empêcher le réseau d'apprendre avant d'avoir acquis la valeur pertinente dans le shell, et est cohérent avec les données biologiques, reposant sur des projections inhibitrices du striatum à VTA, comme exploitées par d'autres modèles des circuits dopaminergiques ([O'Reilly et al. \[2007a\]](#)).

5.4 Expérience de reversal

Pour montrer l'apport de notre erreur de prédiction, en plus de sa cohérence avec la biologie, nous exposons ici son comportement lors d'une tâche d'acquisition et de reversal, et comparons sa convergence à celle d'un réseau paramétré de manière identique, mais possédant l'erreur de prédiction de [Guthrie et al. \[2013\]](#).

Notre protocole est adapté par rapport à celui de [Guthrie et al. \[2013\]](#) pour correspondre à une tâche de reversal. En particulier, le déroulement d'un essai suit le protocole de [Guthrie et al. \[2013\]](#), mais la règle de contingence testée va correspondre à un reversal. A chaque essai, deux CS sensoriels sont présentés simultanément au réseau, à deux positions aléatoires, pendant 2500ms, et le réseau doit effectuer une action en direction de l'un des CS. Dès que le modèle effectue une action (avant la fin des 2500ms), la récompense est distribuée en fonction de la probabilité de récompense associée au CS choisi. Si le modèle n'a pas effectué d'action au bout des 2500ms de présentation, il ne reçoit pas de récompense. Les activités neuronales sont ensuite laissées revenir à leur niveau de base, puis l'on procède à l'essai suivant. La phase d'acquisition dure jusqu'à acquisition de la réponse et formation d'habitudes par l'apprentissage opérant. Parmi les deux CS présentés, un des deux est récompensé (probabilité de 100%), tandis que l'autre ne l'est pas (probabilité nulle). Lors du reversal, les taux de récompense associés aux CS sont inversés par rapport à ceux de l'acquisition. Le réseau doit alors apprendre le changement de contingence, et changer sa réponse apprise, afin de sélectionner le CS qui n'était pas récompensé lors de la phase d'acquisition. Cette phase dure jusqu'à acquisition du reversal.

La figure 5.5 montre la moyenne des résultats obtenus lors de 100 expériences effectuées par des modèles "naïfs", i.e avec des poids initialisés aléatoirement, avec le calcul de l'erreur de prédiction tel que décrit dans le modèle de [Guthrie et al. \[2013\]](#) (courbe rouge) et avec l'erreur de prédiction calculée en fonction de l'activité du shell (courbe bleue). Dans les deux cas, l'acquisition et le reversal sont correctement acquis (figure 5.5.A), c'est à dire que le réseau apprend correctement à faire une action vers le CS le plus récompensé, et le temps de décision évolue de manière similaire quelle que soit le critique (figure 5.5.B). Ces résultats valident que notre implémentation bio-inspiré du calcul de l'erreur de prédiction n'endommage pas la performance lors d'une tâche de reversal. Dans la même optique, notre modèle observe bien des dynamiques d'évolution de boucles similaires à celles de [Guthrie et al. \[2013\]](#) (figure 5.4). De plus, la figure 5.5.C montre que la valeur de l'US (de 1 dans les expériences) utilisée pour l'erreur de prédiction est exacte en utilisant le critique basé sur la prédiction striatale, et montrent une évolution directe de la valeur dès les premiers essais du reversal. Au contraire, l'ancien critique ne possède qu'une valeur approximative de l'US, et met davantage de temps avant de changer

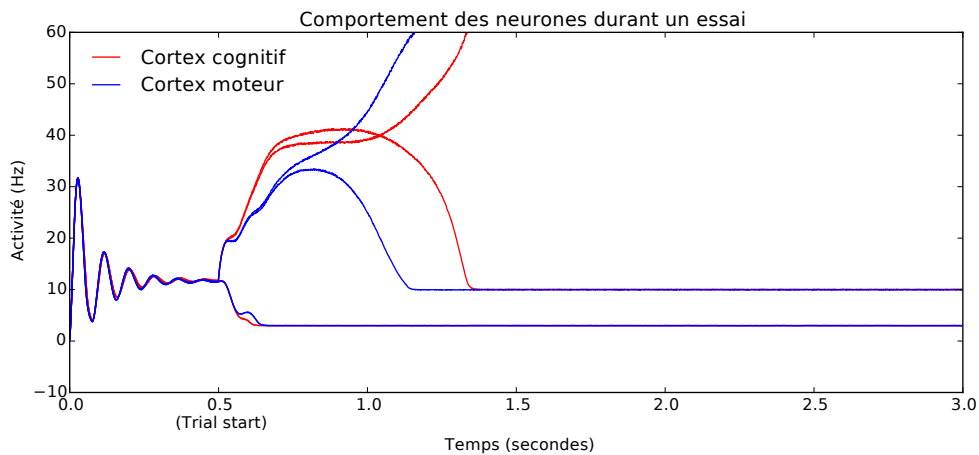


FIGURE 5.4 – Dynamiques d'évolution des activités des boucles cognitives et motrices au cours d'un essai (en absence d'apprentissage). Les dynamiques sont similaires à celles observées dans [Guthrie et al. \[2013\]](#). Dans cet essai, la boucle motrice guide la réponse du réseau, et la convergence de la boucle cognitive. Les choix non-présentés au réseau diminuent immédiatement après le début de l'essai, tandis que la différence entre les choix présents met plus de temps à émerger.

la valeur stockée. Ceci explique pourquoi dans la figure 5.5.A, l'apprentissage du modèle possédant le nouveau critique commence légèrement plus tôt que son concurrent. Cette apprentissage prend néanmoins un peu plus d'essais, en raison du fait que les connexions entre CTX et STR ont appris pendant davantage de temps lors de l'acquisition (comme le critique met plus de temps à atteindre la valeur exacte de l'US), sont donc plus renforcées et plus lente à être remplacées par un nouveau choix.

5.5 Conclusion

Nous voyons dans ce chapitre comment deux systèmes d'apprentissage, apprentissage associatif pavlovien et apprentissage d'habitudes opérant, sont combinés dans les ganglions de la base. En particulier, nous avons montré que l'apprentissage d'associations CS-US dans le shell permet un apprentissage rapide du système global, compensant la lenteur d'apprentissage du système habituel. Réciproquement, l'apprentissage d'habitudes permet d'accélérer le temps de décision dans une situation connue, et éviter des coûts de calculs par des passages répétés dans la boucle OFC-shell, pour évaluer les valeurs des différents CS ou états sensoriels.

Dans ce système, le rôle de l'OFC est crucial, en tant que plateforme d'échange, entre l'apprentissage associatif de la boucle cognitive et l'apprentissage opérant. L'OFC reçoit également d'importantes projections de l'amygdale, en particulier de BLA, qui apprend différents aspects relatifs à l'US [[Schoen-](#)

5. L'apprentissage pavlovien de la boucle cortico-striatale limbique guide l'apprentissage opérant.

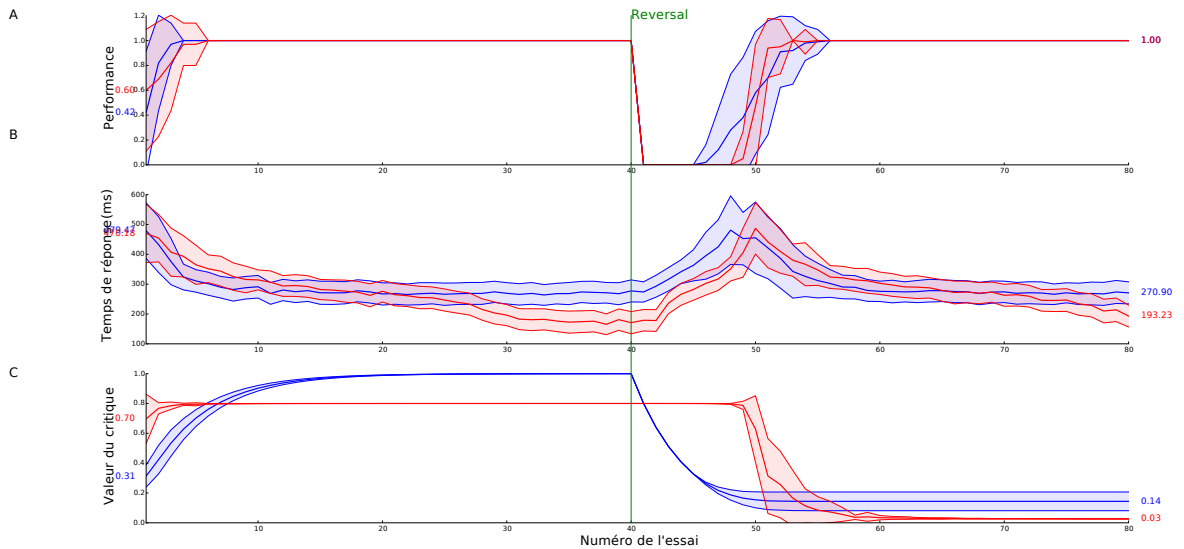


FIGURE 5.5 – Expérience de Reversal effectuée par le modèle de Guthrie *et al.* [2013] (rouge) et par la version utilisant le striatum pour le calcul de l’erreur de prédiction (bleue). Tous les paramètres communs entre les deux versions sont identiques, seule la façon dont est calculé le critique, et les paramètres spécifiques, changent. Chaque courbe est la moyenne de 100 expériences effectuées à partir d’un modèle “naïf”(sans connaissance préalable). Les aires colorées indiquent la déviation standard de chaque courbe. (A) Performance moyenne en fonction des essais. Les deux modèles apprennent correctement l’acquisition des valeurs des CS et le reversal, dans des constantes de temps similaires. (B) Temps de réponse moyen. Les deux versions du modèles présentent des temps de réponse similaires. (C) Valeur encodée par le critique pour le CS prédicteur d’US lors de la phase d’acquisition, et utilisée pour l’erreur de prédiction. Même sur un paradigme simple, l’ancien critique peut ne pas apprendre la valeur exacte, sans pour autant gêner la performance du modèle. Le nouveau système basé sur les prédictions striatale arrive à calculer correctement la valeur de l’US prédit, et à déclencher la dopamine correspondante. Il est notamment plus rapide à adapter sa valeur lors de changement de contingence.

baum *et al.*, 2003]. Dans cette optique, le rôle de l’amygdale serait donc similaire à celui que nous avons montré ici pour le shell, c’est-à-dire apprendre des valeurs servant à moduler les états passés en entrée de l’apprentissage opérant, influençant celui-ci en mettant rapidement en évidence les buts, CS à forte valeur associée, pour la prise de décision. L’amygdale permettrait donc bien d’apporter les valeurs des US associés aux CS dans l’OFC, pour ensuite être utilisé par l’apprentissage opérant, comme illustré par notre modèle. Schoenbaum *et al.* [2003]; Holland et Gallagher [2004]; Saddoris *et al.* [2005] observent deux associations dans l’OFC, une association CS-outcome, autrement dit des neurones réactifs à l’arrivée d’outcome qui se mettent à réagir au moment du CS, et des associations “Action-outcome”, des neurones qui, suite à l’action, se mettent à prédire l’arrivée de récompense. L’augmentation de performance

de l'animal dans la tâche opérante est observée quand les neurones Action-outcome se mettent à devenir sensibles au CS, réagissant au moment de l'arrivée du CS, encodant ainsi l'association CS(ou état sensoriel)-Action-Outcome. En particulier, [Schoenbaum et al. \[2003\]](#) montre que, si déconnecter l'amygdale de l'OFC n'altère pas la performance dans la tâche, cela diminue tout de même le nombre de neurones prédisant la récompense qui répondront au CS auditif dans l'OFC. Les neurones d'OFC réagissant au CS sont alors davantage influencés par les propriétés sensorielles du CS que celles de l'US. Ainsi, un apprentissage dirigé vers le but, apprentissage Action-Outcome a lieu et est reflété dans les activités de l'OFC, et l'amygdale permet d'apporter la valeur associée à un CS précis, permettant à l'apprentissage opérant de distinguer quel CS est pertinent pour le choix de l'action.

La question se pose alors de la différence entre les valeurs apprises par le shell et par l'amygdale, malgré les similitudes observées dans nos deux modèles, et malgré une règle d'apprentissage similaire en fonction de leurs erreurs de prédictions. Une première réponse pourrait être liée aux différentes réponses pavloviennes déclenchées par le shell et l'amygdale. Le shell apprend la réponse préparatoire (orientation, avoidance) [Balleine et Killcross \[2006\]](#), réponses dépendant de la valence du CS (aversif ou appétitif) mais pas de sa nature (il se boit, se mange) tandis que l'amygdale est plus basiquement associée aux réponses hormonales, salivation, freezing, et stocke les informations relatives à la nature du CS dans BLA [Belova et al. \[2015\]](#). L'information transmise par l'amygdale à l'OFC contiendrait donc des informations sur l'US prédit (sucré, gras, ...) tandis que celle apprise par le shell serait davantage liée à la valence, à quel point ce CS est appétitif ou aversif. Une seconde réponse pourrait être liée aux différences d'entrée. L'amygdale forme des associations à partir de CS simples transmis par le cortex sensoriel, tandis que le shell reçoit en entrées des représentations plus élaborées d'OFC, des états sensoriels [\[Rolls, 2004a\]](#). Pour résumer cette hypothèse, l'amygdale aiderait la prise de décision et l'apprentissage motivé en répondant à la question "quel est l'US qui va arriver" tandis que le shell répondrait à la question "quelle est la valeur de la situation?". Cette vision est cohérente avec notre modèle d'amygdale, qui reçoit d'IL (correspondant à l'OFC chez les rats) la "situation d'extinction", situation dont la valence n'est plus positive, et qui pourrait ainsi être apprise par la boucle OFC-shell, pour aider l'amygdale. Il n'est donc pas exclu que l'interaction entre systèmes associatifs soit également intéressante pour aider l'apprentissage. Ce modèle répond donc par deux aspects à notre problématique. D'une part l'apprentissage répondant, que ce soit par l'amygdale ou par la boucle cognitive, permet de fournir des buts sur lesquels l'apprentissage opérant peut effectuer des actions. Lors de changement de situation, comme le reversal, la valeur des buts peut ainsi changer, et être apportée à l'apprentissage opérant, sans pour autant que celui-ci n'ait à modifier les actions déjà apprises, restant ainsi à disposition des autres comportements.

5. L'apprentissage pavlovien de la boucle cortico-striatale limbique guide l'apprentissage opérant.

D'autre part, combiner apprentissage motivé et associatif permet de s'opposer à l'apprentissage d'habitude, et d'apporter ainsi un compromis entre des décisions souples, flexibles, mais coûteuses en temps et en réflexion, et décision robustes et rapides.

Chapitre 6

Le rôle de la noradrénaline dans le compromis exploration-exploitation.

6.1 Introduction

Dans la partie précédente, nous avons repris et amélioré la plausibilité biologique du modèle développé par [Guthrie *et al.*, 2013]. Ceci nous a permis d’observer comment deux modes d’apprentissages différents présents dans ce modèles, un pavlovien et un opérant, parviennent à compenser leurs défauts respectifs lors de l’apprentissage ou du reversal d’une règle. En particulier, la souplesse et la vitesse d’apprentissage de la boucle sensorielle permet de s’adapter plus rapidement à une règle, tandis que l’apprentissage lent et statistique de la boucle associative permet l’acquisition robuste d’une règle efficace sur le long terme. Lors d’un changement de règle, la boucle sensorielle apprend la modification de la valeur des CS, augmente la saillance du nouveau CS pertinent transmise à la boucle motrice, jusqu’à permettre la sélection de cette action à la place de l’action habituelle. Puis au fur à mesure que cette nouvelle action est effectuée, une nouvelle habitude robuste se forme dans la partie motrice.

Cependant, si ce système aborde le compromis entre exploration et exploitation, il peut dans certains cas rester long à s’adapter au changement de situation, en raison de son absence d’exploration. En effet il est nécessaire d’explorer au moins une fois le nouveau choix optimal, afin que la boucle sensorielle puisse commencer à apprendre sa valeur et permettre à l’action correspondante de surpasser l’action habituelle. Dans le modèle actuel (ainsi que dans [Guthrie *et al.*, 2013]), la boucle sensorielle apprend d’abord, lors du reversal, que le CS habituellement sélectionné n’est plus suivi de récompense, et désapprend donc sa valeur précédente, et ce jusqu’à ce qu’un autre CS ne devienne plus important, et soit sélectionné comme nouvelle cible de l’action. Si cette stratégie fonctionne s’il n’y a que deux options à disposition, il est possible de rester blo-

quer sur un choix sous-optimal, comme ce système ne garanti pas suffisamment d'exploration des différentes possibilités (par exemple si deux choix après le reversal apportent de la récompense à 50% et 100% de chance respectivement, et que le modèle commence à regarder le premier). Ainsi, pour combiner les deux apprentissages présents dans le modèle de façon efficace, et en particulier améliorer l'adaptation lors d'un changement de contingence, il est nécessaire d'inclure un mécanisme d'exploration à la façon dont apprentissage opérant et motivés sont combinés.

Lors de notre étude de l'apprentissage de l'amygdale, nous avons vu que la neuromodulation pouvait, en influençant le comportement d'une partie spécifique dans un circuit, amener à une prise en charge adaptée à l'incertitude de la situation. En particulier, le neuromodulateur Noradrénaline (NE, pour norepinephrine), attribué à la détection de l'incertitude inconnue, ou imprévue, causée par un changement de la règle existante [Yu et Dayan, 2005], est particulièrement adapté à la détection de changement de contingence. Dans son modèle, Aston-Jones et Cohen [2005] propose que le niveau tonique de NE, dépendant de l'incertitude inconnue, contrôle le caractère aléatoire de la sélection de l'action, et donc l'équilibre entre exploitation et exploration. Son modèle est composé de deux unités sensorielles représentant deux choix alternatifs, dont le gain est modulé par NE. NE est générée en fonction de l'incertitude imprévue par deux accumulateurs en inhibition mutuelle, permettant d'optimiser la décision lors de changement de règle [Brown *et al.*, 2005]. Cependant, il est limité à des décisions binaires, et n'explique donc pas comment effectuer de l'exploration si celle-ci possède différentes cibles potentielles [Pouget *et al.*, 2013]. De plus, il se concentre sur les effets de l'exploration dans le cortex sensoriel, alors que cette exploration a lieu dans un réseau de décision complexe, et incluant en particulier le cortex préfrontal [Daw *et al.*, 2006].

Dans cette partie, nous étudions la génération et l'effet de NE sur le modèle de [Guthrie *et al.*, 2013], afin non seulement d'améliorer la performance de ce modèle lors du reversal, mais également de proposer, par l'explication des mécanismes liés à NE, comment un système combinant apprentissage motivé et associatif peut être performant dans le compromis entre exploration et exploitation. En particulier, notre modèle systémique nous permettra d'étudier non seulement l'effet excitateur de NE au niveau cortical, comme proposé des modèles précédents [Aston-Jones et Cohen, 2005; McClure *et al.*, 2006], mais également de prendre en compte de manière originale son effet inhibiteur rapporté dans le striatum [Nicola et Malenka, 1998], et de discuter l'éventuelle complémentarité entre ces deux effets.

6.2 Le système noradrenergique

La source principale de noradrénaline est le Locus Coeruleus (LC), un noyau du tronc cérébral [Berridge et Waterhouse, 2003]. Un des rôles reconnu de NE est la modification des activités du Thalamus et du Cortex lors du traitement de l'information sensorielle [Sara et Bouret, 2012], en fonction du niveau d'attention requis par la situation supposé encodé par le niveau tonique de NE [Berridge et Waterhouse, 2003]. En particulier, les niveaux les plus importants sont observés lors de changements brusques dans l'environnement, lorsque par exemple les contingences d'une règle connue changent, qui demandent une exploration des différents stimuli présents pour apprendre les nouvelles contingences, et ainsi arriver à trouver la stratégie la plus optimale [Aston-Jones et Cohen, 2005].

Aston-Jones *et al.* [1997] montre l'implication de NE lors du reversal, lorsque le critère sensoriel prédisant l'arrivée d'US devient invalide et demande de trouver le nouveau critère pertinent. Yu et Dayan [2005] font ainsi la distinction entre la stochasticité, incertitude connue, encodée par l'ACh, et l'incertitude imprévue, encodée par NE. ACh, comme que modélisée précédemment, et l'incertitude connue sont liées à l'exploitation : si j'observe une arrivée de récompense incertaine, mais que je connais cette incertitude, je sais prédire les chances d'avoir de la récompense, je n'ai pas donc pas forcément intérêt à chercher une nouvelle règle, mais peut-être à préciser la règle existante. En revanche, NE est liée à l'exploration : si mes prédictions fiables deviennent soudainement incertaines, c'est que quelque-chose dans mon environnement a changé, et j'ai donc tout intérêt à explorer pour découvrir ce qui a changé.

LC reçoit plusieurs types différents d'informations permettant la génération de NE. Il reçoit des informations de bas niveau telles que le niveau d' "arousal" (niveau d'éveil ou d'excitation) du système nerveux sympathique, et des informations relatives aux stimulus sensoriels saillants du système oculomoteur [Berridge et Waterhouse, 2003]. Des éléments plus élaborés sont envoyés par CeA et mPFC (et en particulier ACC) [Sara et Bouret, 2012], comme la nouveauté [Krebs *et al.*, 2013], différentes prédictions de récompense ou des conflits de décisions. Le rôle généralement proposé de ces connexions est de permettre l'évaluation de l'incertitude de la situation [Aston-Jones et Cohen, 2005]. Également, d'autres centres impliqués dans la neuromodulation sont en influence réciproque avec LC [Sara et Bouret, 2012].

LC projette sur la plupart des régions du cerveau, et en particulier sur les structures attentionnelles comme le cortex sensoriel pariétal, dont NE facilite l'activation, mécanisme souvent lié à l'exploration [Aston-Jones *et al.*, 1999]. Il est intéressant de noter que les ganglions de la base sont la seule structure cérébrale qui ne reçoit pas de projections de LC, à l'exception notable du Shell du nucléus accumbens, où NE est rapporté avoir un effet inhibiteur [Nicola et Malenka, 1998]. Ceci, lié à la spécificité de ces projections sur le shell

[Berridge *et al.*, 1997], ainsi que le rôle important du Shell dans notre modèle de décision, nous a amenés à modéliser également cet effet, et à en proposer un rôle fonctionnel différent de celui de NE sur le cortex.

6.3 Modéliser l'exploration

Dans cette section, nous reprenons donc le modèle inspiré de [Guthrie *et al.*, 2013] vu précédemment, auquel nous ajoutons un module représentant la génération de NE, et ainsi que l'effet de NE sur le cortex et striatum. Nous étudions la réaction de ce réseau dans une tâche comprenant un changement soudain de règle, le reversal. Détecter ce changement et promouvoir un comportement d'exploration permet de découvrir et d'apprendre les nouvelles contingences plus rapidement, s'adaptant ainsi plus vite aux nouvelles situations. Je ne décrirai dans cette partie que les modifications effectuées par rapport au chapitre précédent, illustrées dans la figure 6.1. Les paramètres modifiés par rapport au chapitre précédent sont tous mentionnés dans 6.3. L'implémentation utilise, comme précédemment, la bibliothèque DANA pour la modélisation et l'évaluation des neurones.

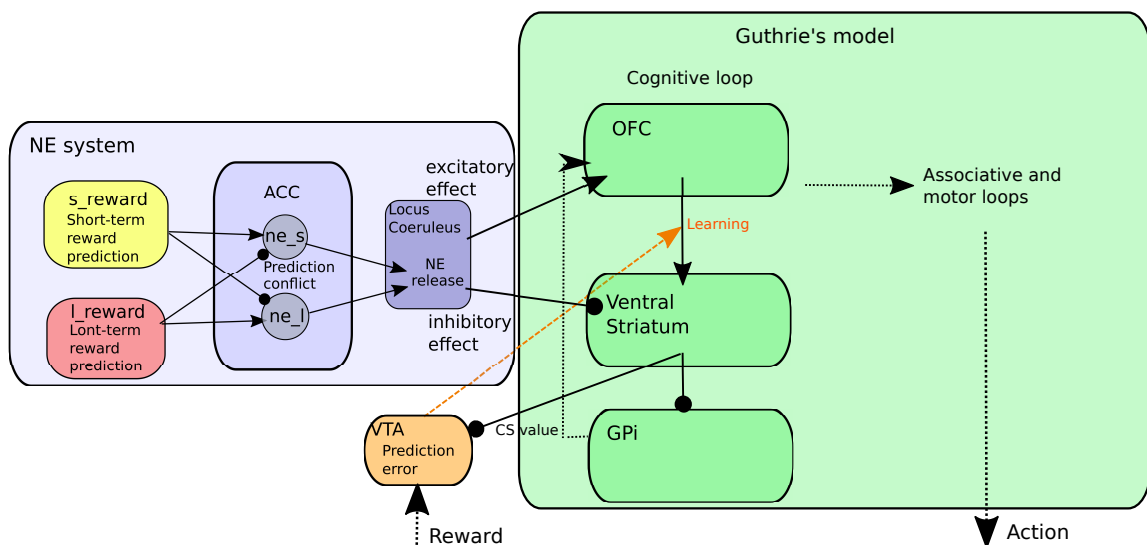


FIGURE 6.1 – Ajout d'un système modélisant la génération de NE, et son effet sur le modèle de boucles cortico-striatales. CT-PRED et LT-PRED sont respectivement des populations prédisant l'arrivée de récompense à partir respectivement de l'historique récent et de l'historique à plus long terme. Chaque population inhibe l'entrée excitatrice de l'autre sur ACC, résultant en une activité de ACC et une libération de NE uniquement en cas de différence entre les deux prédictions. Plus l'historique récent et l'historique plus ancien sont en contradiction, plus le changement de règle et l'incertitude "imprévue" sont marqués, et plus le niveau tonique de NE sera élevé. NE est ensuite utilisé pour déclencher de l'exploration dans le modèle de décision, à la fois par facilitation de l'excitation corticale et par l'inhibition des entrées striatales.

6. Le rôle de la noradrénaline dans le compromis exploration-exploitation.

Le niveau tonique de NE est calculé comme la différence (ou conflit) entre deux systèmes prédisant l'arrivée de récompense basés respectivement sur un historique récent $ct_prediction$ et sur un historique à plus long terme $lt_prediction$. Ces prédictions peuvent correspondre aux prédictions de récompense respectives de l'hippocampe, via ACC, et de l'amygdale, transmises ensuite à LC. Ces prédictions sont calculées comme la récompense moyenne respectivement sur les n_ct_essais et n_lt_essais , et envoyées en entrée au système noradrenergique.

$$lt_prediction = (\sum_{k \in n_lt_essais} reward_k) / n_lt_essais$$

$$ct_prediction = (\sum_{k \in n_st_essais} reward_k) / n_st_essais$$

avec $reward_k$ la récompense reçue à l'essai k .

Ces prédictions sont respectivement passées en entrées respectives de deux populations de l'ACC calculant le conflit entre deux prédictions, ne_c et ne_l , comme illustré par la figure 6.1. De plus, chaque prédiction inhibe la population recevant l'autre prédiction en entrée :

$$\frac{dV_{ne_c}}{dt} = \tau * (-V_{ne_c} + ct_prediction - lt_prediction)$$

$$U_{ne_c} = clamp(V_{ne_c} * (1 + noise))$$

$Noise()$ est le niveau de bruit tiré suivant une distribution uniforme, et $clamp()$ une fonction limitant la valeur de V_{ne_c} à un intervalle positif.

Symétriquement :

$$\frac{dV_{ne_l}}{dt} = \tau * (-V_{ne_l} + lt_prediction - ct_prediction)$$

$$U_{ne_l} = clamp(V_{ne_l} * (1 + noise))$$

Le niveau de NE, ne , est pris comme la somme des activités U_{ne_c} et U_{ne_l} . Par conséquent, si $lt_prediction$ et $ct_prediction$ ne prédisent toutes deux aucune récompense, ne_s et ne_l sont toutes deux faibles, et le niveau tonique de NE est lui aussi faible. De même, si $lt_prediction$ et $ct_prediction$ prédisent toutes deux l'arrivée de récompense, ne_s et ne_l sont excitées par une des prédictions mais inhibées par l'autre, et la concentration de NE reste faible. Ainsi, tant que l'historique récent et l'historique sur plus long terme présentent le même taux d'arrivée de récompense, le niveau de NE est faible. Au contraire, si la règle change brusquement, dans le cas d'incertitude imprévue, les historiques récent et ancien ne vont pas présenter les mêmes taux d'arrivée de récompense lors des premiers essais suivant le changement de règle. L'un prédira donc plus de récompense que l'autre, et sa population correspondante, ne_c ou ne_l , sera plus excitée qu'elle ne sera inhibée, conduisant donc ainsi à une augmentation

de la concentration de NE. Notre système noradrenargique, en représentant le calcul d’une fonction “ou exclusif” entre une prédiction de récompense basée sur un historique récent et sur un historique à plus long terme, libère donc bien NE lors de changements de contingence brusque, lorsque l’incertitude imprévue est élevée.

De manière similaire aux modèles classiques modélisant l’effet de NE sur le cortex, nous avons implémenté un effet facilitateur de NE sur l’excitation des neurones corticaux par un gain multiplicatif : U_{ctx} :

$$\frac{dU_{ctx}}{dt} = f(V_{ctx} * (1 + ne) * (1 + noise))$$

où V_{ctx} et U_{ctx} sont respectivement le potentiel de membrane et le taux de décharge des neurones corticaux, f et $noise$ sont les fonctions respectivement sigmoïde et de bruit interne au neurone, utilisées par [Guthrie *et al.*, 2013]. L’effet inhibiteur de NE, nouveau dans notre modèle, impacte le gain de sortie des projections du cortex sur le Shell :

$$gain = g_{ctx_str} * ne_mod$$

avec g_{ctx_str} le gain constant entre le cortex et le striatum dans la boucle sensorielle, et ne_mod l’effet modulateur de NE :

$$ne_mod = max(0.5, 1 - ne_ef * ne)$$

ne_ef est une constante modulant l’effet de NE sur les connexions, fixée de façon à ce qu’uniquement les valeurs maximum de ne provoquent une valeur minimale de ne_mod , et donc une inhibition maximale du Shell. Ces constantes sont de plus réglées pour ce que l’inhibition maximale diminue de moitié les projections excitatrices du cortex au Shell, reprenant l’effet rapporté par [Nicola et Malenka, 1998].

6.4 Expériences et résultats

Pour mettre en évidence l’avantage de l’exploration dans l’aide à la décision lors d’un changement de tâche, nous avons testé notre modèle sur l’apprentissage du reversal. A chaque essai, deux CS sensoriels sont présentés simultanément au réseau, à deux positions aléatoires, pendant 2500ms, et le réseau doit effectuer une action en direction d’un des CS. Dès que le modèle effectue une action (ce qui peut arriver avant la fin des 2500ms), la récompense est distribuée en fonction de la probabilité de récompense associée au CS choisi. Si le modèle n’a pas effectué d’action au bout des 2500ms de présentation, il ne reçoit pas de récompense. Les activités neuronales sont ensuite laissées revenir à leur niveau de base, puis l’on procède à l’essai suivant. La phase

6. Le rôle de la noradrénaline dans le compromis exploration-exploitation.

Paramètres du modèle		
Paramètre	Description	Valeur
α_{str}	Constante modélisant le poids des connexions de STR au système dopaminergique	0.04
α_{LTP}	Taux d'apprentissage lors de LTP	0.00008
α_{LTD}	Taux d'apprentissage lors de LTD	0.00005
g_{cog_cog}	Gain des connexions entre cortex cognitif et striatum ventral	1.2
g_{cog_ass}	Gain des connexions entre cortex cognitif et striatum dorsal	0.3
g_{ne_exc}	Gain des connexions excitatrices des populations de NE	1.0
g_{ne_inh}	Gain des connexions inhibitrices des populations de NE	-1.0
n_{lt_essais}	Nombre d'essais pris en compte par l'historique à long terme	30
n_{ct_essais}	Nombre d'essais pris en compte par l'historique récent	3
ne_ef	constante modulant l'effet des projections de NE sur les connexions	0.8

FIGURE 6.2 – Description et valeur des paramètres modifiés par rapport à [Topalidou *et al.*, 2015].

d'acquisition est composée d'un nombre important d'essais pour effectuer de l'over-training et permettre la formation d'habitudes, ainsi que pour permettre à NE de revenir à son niveau de repos. Parmi les deux CS présentés, un est systématiquement récompensé (récompensé à 100% de probabilité), tandis que l'autre ne l'est pas. Lors du reversal, les taux de récompense associés aux CS sont inversés par rapport à ceux de l'acquisition. Le réseau doit alors apprendre le changement de contingence, et changer sa réponse apprise, afin de sélectionner le CS qui n'était pas récompensé lors de l'acquisition.

La figure 6.3 montre la moyenne des résultats obtenus lors de 100 expériences effectuée par des modèles "naïfs" (ie. modèle sans connaissance initiale, avec des connexions initialisées aléatoirement comme au chapitre précédent) avec (courbe bleue) et sans (courbe rouge) notre module représentant NE. Dans les deux cas, l'acquisition et le reversal sont correctement appris, c'est à dire que le réseau apprend correctement à effectuer une action vers le CS le plus récompensé. Cependant, pour les modèles avec NE, l'augmentation du niveau de NE survient rapidement suite au changement de contingence et permet au réseau d'effectuer directement de l'exploration, résultant en une convergence moyenne plus rapide que les réseaux sans NE, qui doivent oublier la règle précédente avant de commencer à apprendre la nouvelle (figure 6.3.A). Le niveau de noradrénaline libéré est particulièrement important au début de l'acquisition et du reversal, et correspond donc bien à l'incertitude imprévue (figure 6.3.C). Son action sur l'exploration permet effectivement une exploration et un apprentissage plus rapide du changement de contingence. Concernant le temps de réponse avant décision, Mason et Iversen [1977] montre que des animaux avec une diminution de NE répondent plus rapidement aux tâches demandées (moins de réflexion avant de déclencher une action) mais persévèrent davantage dans leur choix. Nous reproduisons bien cet effet de NE figure 6.3.B) : le temps de prise de décision des modèles sans NE est en moyenne plus important que ceux avec NE, tout en persévérant davantage avant de s'adapter au

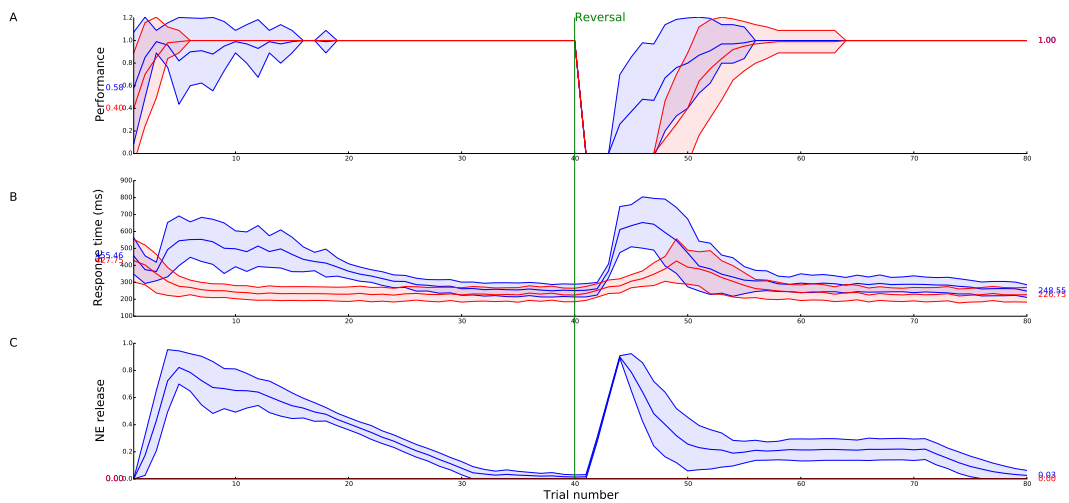


FIGURE 6.3 – Expérience de Reversal effectuée par le modèle de [Guthrie *et al.*, 2013]. La courbe rouge affiche la performance du modèle de sans NE, tandis que la bleue montre sa performance avec ajout de l’effet de NE. Chaque courbe est la moyenne de 100 expériences effectuée à partir d’un modèle “naïf”. Les aires colorées indiquent la déviation standard de chaque courbe. (A) Performance moyenne en fonction des essais. Les deux modèles apprennent correctement l’acquisition des valeurs des CS et le reversal. Cependant, la noradrénaline permet de détecter plus rapidement le changement de contingence de récompense, et donc d’apprendre le reversal plus rapidement que le modèle sans NE. (B) Temps de convergence moyen. Durant les premiers essais du reversal, l’exploration par inhibition du striatum induit une réponse plus importante de NE. (C) Concentration moyenne de NE. Le niveau de NE libéré est particulièrement important au début de l’acquisition, ainsi qu’au début du reversal, et correspond donc bien à l’incertitude imprévue.

changement de contingence.

6.5 Conclusion

Nous avons discuté dans le chapitre précédent de l’intérêt de combiner différents systèmes d’apprentissage, pavlovien et habituel, pour posséder à la fois un apprentissage flexible et un apprentissage robuste permettant des décisions rapides. Dans ce chapitre, nous avons amélioré la façon dont le modèle de la section précédent s’adapte au changement de situation. Pour cela, nous avons ajouté un système noradrénergique reproduisant l’incertitude imprévue, et son rôle dans l’équilibre entre exploration et incertitude, rôle couramment proposé pour NE [McClure *et al.*, 2006; Aston-Jones et Cohen, 2005]. Le système noradrénergique détecte l’incertitude imprévue par la différence entre deux estimations de la récompense, basées sur un historique des récompenses à court et à long terme. Si l’arrivée de récompense change brusquement, les

deux historique sont en contradiction, provoquant ainsi une différence et une libération de NE.

Comme montré lors du reversal, l'exploration induite par NE améliorer la vitesse d'adaptation d'un modèle de décision à une situation imprévue liée à un changement de règle, avantage adaptatif crucial pour les animaux et systèmes autonomes. En plus de l'aspect excitateur de NE sur les neurones corticaux, effet déjà implémenté dans de précédents modèles [McClure *et al.*, 2006], notre modèle a pour originalité de proposer un second mécanisme dans le déclenchement de l'exploration, le rôle inhibiteur de NE dans le Shell du striatum ventral, conformément à l'effet rapporté par [Nicola et Malenka, 1998]. De plus, nous proposons que ces deux effets ne soient pas contradictoires, mais en revanche complémentaires. Comme proposé par [Aston-Jones et Cohen, 2005], la Noradrénaline corticale faciliterait l'exploration de nouvelles règles par la création de nouveaux états dans le cortex, plus adaptés à la situation actuelle. Ainsi, celle-ci serait en particulier essentielle lorsque le changement de règle est extra-dimensionnel, c'est à dire lorsque que la nouvelle règle dépend d'une dimension auparavant non pertinente [Tait *et al.*, 2007]. En revanche, l'effet de NE que nous proposons dans le striatum, en inhibant les valeurs des états déjà appris, permet l'exploration de nouvelles valeurs basées sur les états existants dans le cortex. Suivant cette hypothèse, cet effet de NE aurait peu d'importance dans les cas où la nouvelle règle est basée sur une dimension différente, mais serait particulièrement important dans le cadre du reversal où ce sont les mêmes stimuli qui voient leur valeurs modifiées.

De manière intéressante, un rôle similaire pour la dopamine tonique, qui possède également un effet inhibiteur sur le striatum [Nicola et Malenka, 1998], a été proposé par Humphries *et al.* [2012], lorsque les changements de contingence concernent l'aspect moteur. Par exemple, il est possible que la règle sensorielle ne change pas, mais que l'action à faire pour obtenir la récompense change. Dans ce cas, l'exploration "sensorielle" effectuée par NE, que ce soit par l'exploration de nouvelles valeurs ou de nouveaux états et dimensions, n'est pas suffisante pour trouver la nouvelle règle, nécessitant une exploration à un niveau différent, une "exploration motrice" de la dopamine. Cet ajout, testé dans une tâche différenciant "exploration sensorielle" et "exploration motrice" serait une extension intéressante au modèle.

Mis en relation avec notre problématique, ce chapitre montre non seulement l'intérêt combiner les apprentissages associatifs et motivés, mais également l'intérêt de faire évoluer la façon dont ceux ci sont combinés pour notre prise de décision. NE affecte ici les valeurs calculées par l'apprentissage répondant, et utilisées par l'apprentissage opérant. En perturbant ces valeurs, ie. en "déconnectant" le système répondant du système opérant lors de situation incertaines, le système opérant n'a plus de valeur pour les différents buts, et les teste ainsi de manière aléatoire. D'une certaine façon, nous montrons que ne pas faire interagir les systèmes opérant et répondant de manière temporaire est permet

également d'obtenir des apprentissages répondants et opérants plus flexibles, adaptés aux situations d'exploration.

Conclusion

Dans cette thèse, nous avons étudié les apprentissages associatifs, dans le sens apprentissage répondant ou pavlovien d'associations CS-US, et opérants ou motivés, dans le sens apprentissage de l'action motivé par la présence d'un but. Si leurs interactions sont multiples, et peuvent avoir des intérêts divers, nous nous sommes concentrés sur la façon dont ces interactions peuvent permettre un apprentissage adaptatif indépendant, permettant de réagir dans différentes situations, ou changements de situation, comme dans les paradigmes d'extinction ou de reversal. A l'aide de données de la biologie, et de la modélisation informatique, j'ai reproduit et mis en évidence plusieurs mécanismes expliquant le rôle des interactions entre apprentissage répondant et opérant dans la capacité d'adaptation de notre apprentissage.

L'étude et la modélisation de l'amygdale permet de se placer du point de vue de l'apprentissage répondant, et de se poser la question "qu'est ce que peut m'apporter l'apprentissage opérant ?".

Fonctionnellement, l'apprentissage opérant est plus complexe que l'apprentissage répondant (nécessité de choisir le but le plus intéressant par rapport aux besoins, aux coûts, nécessité d'élaborer des stratégies et de dérouler ces stratégies jusqu'à obtention du but,...). Il nécessite ainsi des données plus complexes telles que des mémoires de travail dans PL, ou de faire des statistiques pour obtenir les informations nécessaires à la décision, comme présentes dans OFC, IL ou Acc. Notre modèle d'amygdale montre la nécessité de fournir ces valeurs à l'apprentissage répondant dans le cas de changement de situation comme l'extinction. De plus, notre modèle montre que la capacité à former un apprentissage adapté à la situation est obtenue par la présence et la compétition de différents circuits. Ces différents circuits possèdent des fonctions différentes de par leurs connectivités, recevant des entrées différentes (CS, contexte, extinction), et projetant sur différentes populations de sortie (CeLOn et CeLOff). La compétition entre ces différents circuits est de plus loin d'être faite au hasard, mais permet dans la façon dont elle est implémentée dans le cerveau, de former des associations adaptés à la situation. Nous illustrons ce fait en montrant l'effet de ACh sur cette compétition, qui permet de mettre en avant un apprentissage basé sur les informations plus complexes de l'hippocampe ou de l'apprentissage opérant quand l'incertitude est élevée, et des associations plus simples et robustes, CS-US dans le latéral, lorsque l'incertitude est faible. ACh

permet ainsi de faciliter les interactions entre apprentissage répondant et motivé, pour permettre un apprentissage répondant à partir de valeurs complexes dans des situations incertaines.

L'étude et la modélisation des boucles cortico-striatales nous permet au contraire de nous placer du point de vue de l'opérant, et de nous poser la question "qu'est ce que nous apporte l'apprentissage répondant ?".

Une des réponses mises en avant par notre modèle est l'évaluation des buts, condition nécessaire pour pouvoir effectuer la meilleure action, celle qui va permettre l'obtention du but le plus optimal possible dans une situation. Les interactions entre apprentissages motivés et associatifs sont ainsi dissymétriques : si l'apprentissage répondant peut, comme dans le cas de l'association CS-US ou contexte-US, fonctionner sans l'apprentissage motivé, l'apprentissage motivé nécessite lui la présence de motivation et de but, et ne peut se faire sans apprentissage répondant. Dans notre modèle, le lieu précis de ce flux d'information est le cortex, en particulier l'OFC, qui contient les valeurs calculées par la boucle limbique, et projette notamment sur le striatum dorsal pour fournir ces valeurs à la boucle associative. Ce modèle met également en évidence l'intérêt de combiner apprentissage répondant et apprentissage opérant, ici CS-action ou habitudes, en permettant à la fois la capacité d'apprendre des valeurs flexibles, pouvant être rapidement modifiée dans le cas de changement, et la capacité d'apprendre des réponses rapides et robustes, peu coûteuses en réflexion, par la formation d'habitude.

L'étude et la modélisation du système noradrénergique et de son effet sur la boucle limbique nous permet d'étudier le rapport entre le compromis entre exploitation et exploration et la combinaison des apprentissages associatifs et motivés. Pour cela nous montrons que générer NE par la détection de conflit entre systèmes de prédiction permet de détecter l'incertitude imprévue déclenchée lors de changements de situation. Puis, en modélisant le rôle excitateur de NE sur le cortex frontal, et inhibiteur sur le striatum ventral, nous montrons que NE parvient à provoquer l'exploration. Cette exploration s'explique par le fait que l'action de NE perturbe les valeurs des buts calculées par la boucle limbique, en inhibant la valeur des buts déjà appris, au niveau du striatum ventral, et en excitant les représentations du cortex, augmentant ainsi la valeur de buts alternatifs. Ce modèle répond ainsi de deux manières à la problématique. Il montre que le calcul de statistiques obtenues à partir des systèmes associatifs et motivés peut permettre de s'adapter lors de changement soudain de situation, en permettant l'exploration. Également, il montre que la neuro-modulation peut supprimer les interactions entre répondant et opérant, et que cette suppression dans les cas d'incertitude imprévue va permettre l'exploration de nouveaux buts, en l'absence de l'accès à la valeur des buts précédents.

Pour résumer ces trois approches en quelques mots (figure 6.4), l'opérant aide la flexibilité de l'apprentissage répondant en lui fournissant des évaluations complexes de la situation, déclenchant différentes associations au moyen de cir-

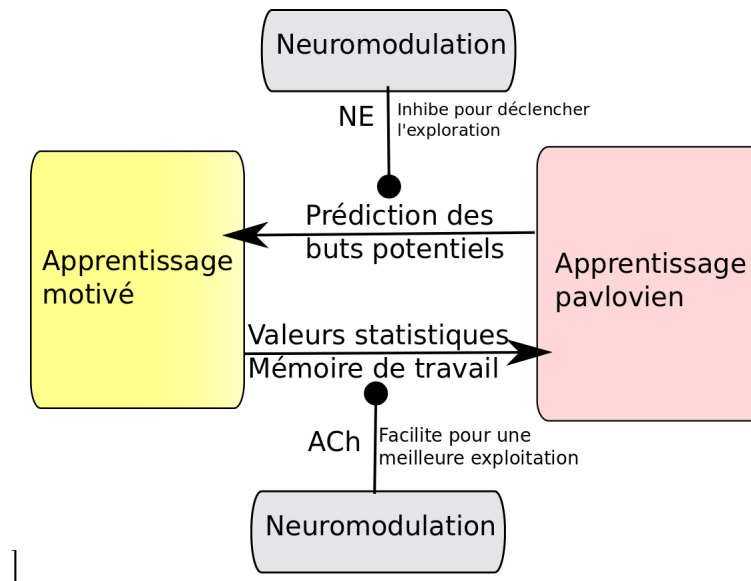


FIGURE 6.4 – Résumé des interactions étudiées et modélisées dans cette thèse. L'apprentissage répondant permet de fournir la prédiction d'informations relatives aux différents USs pour définir les buts du système opérant. Les valeurs utilisées par le système opérant pour définir ses stratégies (statistique sur les situations, informations stockées en mémoire de travail) sont fournies à l'apprentissage répondant pour lui permettre d'acquérir des associations différentes, à base d'entrées plus complexes que de simples CS. La neuromodulation permet moduler ces interactions en fonction de statistiques collectées sur les différents apprentissages. Elle favorise ainsi l'acquisition de ces associations dans des situations stochastiques par l'effet de l'ACh sur l'amygdale. NE permet d'inhiber les valeurs de la boucle limbique transmises aux boucles associatives, et ainsi de provoquer l'exploration de nouveaux buts en coupant la transmission de la valeur des buts présents à la boucle associative.

cuits en compétition, tandis que l'apprentissage répondant aide l'apprentissage opérant en lui fournissant la valeur des buts, par la prédiction d'informations relatives aux US associés aux stimuli. La neuromodulation permet de modifier la façon dont les systèmes opérant et répondant sont combinés, en favorisant ou en supprimant leurs interactions, de manière adaptée à la situation par différentes mesures de l'incertitude de celle-ci.

Du point de vue des neurosciences, cette thèse contribue en proposant certains modèles nouveaux, qui mettent en avant des mécanismes pas encore modélisés dans la littérature. C'est le cas de la modélisation de la compétition entre les différents circuits de l'amygdale, mais également du rôle de l'acétylcholine dans cette compétition, et du rôle de la noradrénaline dans le striatum ventral. Ce n'est pas le cas du modèle étudié dans le chapitre 5, mais transformer le modèle existant en un modèle reprenant explicitement un apprentissage pavlovien et l'utilisant pour guider l'apprentissage motivé était une étape indispensable afin de pouvoir aller plus loin dans l'étude des interactions entre apprentissages répondant et motivés, et nous permet ici d'étudier l'effet de la noradrénaline sur celles-ci. Ces modèles permettent ainsi de reproduire ou

d'expliquer des résultats expérimentaux observés, mais pas encore modélisés jusqu'à présent, comme l'effet des différentes lésions de LA ou de BA pré ou post-training, ou l'effet de déplétion de NE corticale gênant l'exploration lors de reversal extra-dimensionnel, mais pas intra-dimensionnel. De plus, ils nous ont permis d'émettre différentes prédictions. C'est le cas de l'effet de l'ACh dans l'amygdale lors du pairing et de l'impairing, que nous prédisons similaire à celui observés pour l'ACh dans l'hippocampe, car affectant le même circuit fonctionnel. C'est également le cas de l'effet de déplétion d'ACh lors de l'extinction, qui devrait détériorer celle-ci. Également, nous prédisons qu'une déplétion de NE dans le striatum devrait gêner l'exploration lors du reversal intra-dimensionnel. Ces prédictions découlent directement de notre vision fonctionnelle des interactions entre apprentissage pavlovien et opérant.

D'un point de vue informatique, nous montrons comment combiner différents systèmes d'apprentissage pour aller vers plus d'autonomie ou de flexibilité. L'apprentissage répondant permet de prédire la valeur des différents buts, et peut ainsi, éventuellement mis en relation avec un système évaluant les besoins de l'agent, permettre à un agent de choisir de manière autonome ses propres objectifs à partir de ceux prédits, ainsi potentiellement accessibles, dans la situation. Un autre intérêt de l'apprentissage répondant que nous illustrons est la possibilité caractériser la situation présente à partir de statistiques sur la performance de sa prédiction. Notre modèle d'ACh permet par exemple de définir si une situation est "fiable" ou au contraire "incertaine, stochastique". Notre modèle de NE permet de définir si une situation a "brusquement changée", ou au contraire est "constante". Ces statistiques permettent ainsi d'évaluer la situation, et de déclencher des actions ou stratégies d'apprentissages adaptés, comme par exemple s'engager dans de l'exploration si la situation a changée, ou promouvoir des stratégies d'apprentissage plus complexes si celles par défaut mènent à une situation incertaine. Nous proposons ainsi des mécanismes non seulement pour combiner les systèmes d'apprentissages, mais également des stratégies pour adapter la façon dont un agent combine ses différents systèmes d'apprentissages à la situation.

Comme les interactions entre apprentissages associatifs et motivés sont multiples et complexes, cette thèse ne couvre qu'une fraction des phénomènes impliqués, et d'autres mécanismes auraient pu être sélectionnés et étudiés. En restant sur la même problématique, l'étude de plusieurs systèmes pourrait apporter des explications complémentaires :

- l'effet de la dopamine tonique sur l'exploration. Comme discuté précédemment, ce rôle pourrait être similaire à celui de NE dans le shell [Nicola et Malenka, 1998], ou consister en une exploration plus motrice, comme modélisée par Humphries *et al.* [2012] sur la sortie des ganglions de la base. Ainsi la dopamine pourrait permettre une exploration plus motrice, en explorant différentes actions, tandis que NE permettrait d'explorer différents buts. La spécificité de NE à l'exploration liée aux

but est ce qui nous a fait privilégier NE par rapport à DA.

- Le rôle singulier du shell, et ses interactions, similarités et différences avec l'amygdale, aurait pu être davantage exploré. En particulier, les afférents et efférents très spécifiques du shell au sein de la boucle limbique, et du système dopaminergique Penner et Mizumori [2012]; Cardinal *et al.* [2002], suggèrent un rôle fonctionnel distinct de celui-ci par rapport aux autres structures du striatum ventral. Les projections OFC-shell, OFC-amygdale, amygdale-OFC montrent deux circuits différents, OFC-amygdale-shell, et amygdale-OFC-shell, dont les rôles distincts restent à expliquer.
- Étudier le rôle d'Acc et de la motivation dans le choix du but pourrait permettre d'inclure les besoins, ainsi que la détection de conflit et de nouveauté dans la définition du but, et de proposer un modèle de NE tonique dépendant du conflit et de la nouveauté [Jodo *et al.*, 1998; Krebs *et al.*, 2013]. Ceci permettrait par exemple l'exploration de stimuli nouveaux, même en absence d'incertitude sur l'arrivée de récompense. La prise en compte des besoins et de la motivation pourrait permettre une intégration plus réaliste dans Virtual Enaction, utilisant la valeur des jauges pour définir les buts.

Également, différentes problématiques auraient pu être proposées dans le cadre de l'étude des apprentissages associatifs et motivés.

- Un sujet intéressant est l'étude de la dopamine phasique et l'erreur de prédiction, comme moyen de communication entre apprentissage pavlovien et apprentissage opérant. En particulier, ce sujet pourrait aborder les interactions réciproques entre l'amygdale, le striatum et le système dopaminergique, et sur les rôles respectifs des deux premiers dans le calcul de l'erreur de prédiction.
- Un autre axe aurait pu consister en l'étude spécifique de différents neuromodulateurs ou l'étude des interactions réciproques entre différents neuromodulateurs, comme étudier les rôles respectifs de DA et 5HT sur les apprentissages opérants et répondants.
- Un dernier sujet qui me semble intéressant est d'étudier comment les apprentissages associatifs et motivés gèrent les différents US positifs, et aversifs. Les modèles que j'utilise dans cette thèse ne regarde en effet qu'un seul type d'US, et n'apporte pas de réponse sur les mécanismes impliqués dans la gestion de buts de types différents.

Chacun des problèmes ci-dessus est intéressant dans l'étude de la combinaison des systèmes d'apprentissage associatifs et motivés, et pourrait probablement faire l'objet d'une thèse en soi. Également, dans une étude de systèmes d'apprentissage différents, des travaux futurs pourraient concerner les interactions entre cortex, amygdale et hippocampe, et au rôle de l'acétylcholine dans ces interactions, dans l'optique de reproduire l'incertitude connue de l'acétylcholine en tant que phénomène émergent de l'interaction amygdale-hippocampe,

ou de reproduire la règle de Pearce-Hall comme émergent des interactions cholinergiques amygdale-cortex.

Pour reprendre Aristote, notre tout est plus que la somme des parties d'apprentissages car nos apprentissages intéragissent, et s'aident les un les autres en se fournissant des valeurs adaptées. De plus, nous sommes capables de modifier de façon dont nous sommons les apprentissages, comme par exemple au moyen de la neuromodulation.

Liste des publications

- Article dans une revue :
 - Maxime Carrere, Frédéric Alexandre. A pavlovian model of the amygdala and its influence within the medial temporal lobe. *Frontiers in Systems Neuroscience*, Frontiers, 2015, pp.14.
- Communications dans un congrès :
 - Maxime Carrere, Frédéric Alexandre. A model of amygdalar modulations and interactions in pavlovian conditioning. *Neural Computation and Psychology Workshop*, Aug 2014, Lancaster, United Kingdom.
 - Frédéric Alexandre, Maxime Carrere, Randa Kassab. Feature, Configuration, History : a bio-inspired framework for information representation in neural networks. *International Conference on Neural Computation Theory and Applications*, Oct 2014, Rome, Italie.
 - Maxime Carrere, Frédéric Alexandre. Modeling pavlovian conditioning with multiple neuronal populations. *IEEE International Joint Conference on Neural Networks*, Jul 2015, Killarney, Irlande.
 - Maxime Carrere, Frédéric Alexandre. Modeling the Sensory Roles of Noradrenaline in Action Selection. *6th International Conference on Development and Learning and on Epigenetic Robotics*, Sept 2016, Paris, France
 - Maxime Carrere, Frédéric Alexandre. A System-level Model of Noradrenergic function. *25th International Conference on Artificial Neural Networks*, Sept 2016, Barcelone, Espagne.
- Posters :
 - Frédéric Alexandre, Maxime Carrere. Modeling Neuromodulation as a Framework to Integrate Uncertainty in General Cognitive Architectures *The Ninth Conference on Artificial General Intelligence*, Jul 2016, New-York, United States. 2016
 - Maxime Carrere, Frédéric Alexandre. Using the Amygdala in decision making *Fifth International Symposium on Biology of Decision Making (SBDM)*, May 2015, Paris, France
- Participation à un chapitre d'ouvrage :
 - Nicolas Denoyelle, Maxime Carrere, Florian Pouget, Thierry Viéville, Frédéric Alexandre. From biological to numerical experiments

in systemic neuroscience : a simulation platform. A.R. Londral, P. Encarnação. Advances in Neurotechnology, Electronics and Informatics, 12, Springer, 2015, Biosystems & Biorobotics.

Bibliographie

- ADAMS, S., KESNER, R. P. et RAGOZZINO, M. E., 2001. Role of the medial and lateral caudate-putamen in mediating an auditory conditional response association. *Neurobiology of learning and memory*, 76(1) :106–116.
URL <http://view.ncbi.nlm.nih.gov/pubmed/11525249>
- AHISSAR, Merav et HOCHSTEIN, Shaul, 1993. Attentional control of early perceptual learning. *Proc. Natl. Acad. Sci. USA*, 90 :5718–5722.
- ALEXANDER, G. E. et CRUTCHER, M. D., 1990. Functional architecture of basal ganglia circuits : neural substrates of parallel processing. *Trends in neurosciences*, 13(7) :266–271.
URL <http://view.ncbi.nlm.nih.gov/pubmed/1695401>
- AMANO, Taiju, UNAL, Cagri T. et PARÉ, Denis, 2010. Synaptic correlates of fear extinction in the amygdala. *Nature neuroscience*, 13(4) :489–494.
- AMARAL, D.G. et WITTER, M.P., 1995. Hippocampal formation. Dans G. Paxinos, rédacteur, *The Rat Nervous System*, Biosystems & Biorobotics. Springer.
- ANDERSEN P., Morris R., Amaral D. Bliss T. O’Keefe J., 2007. The hippocampus Book. Oxford University Press.
- ANGLADA-FIGUEROA, David et QUIRK, Gregory J., 2005. Lesions of the basal amygdala block expression of conditioned fear but not extinction. *The Journal of Neuroscience*, 25(42) :9680–9685. doi :10.1523/jneurosci.2600-05.2005.
- ARMONY, Jorge L., SERVAN-SCHREIBER, David, COHEN, Jonathan D. et LE-DOUX, Joseph E., 1997. Computational modeling of emotion : explorations through the anatomy and physiology of fear conditioning. *Trends Cogn Sci*, 1(1) :28–34. doi :10.1016/s1364-6613(97)01007-3.
- ASHBY, Gregory F., ALFONSO-REESE, Leola A., TURKEN, U. et WALDRON, Elliott M., 1998. A neuropsychological theory of multiple systems in category learning. *Psychological Review*, 105 :442–481.

- ASTON-JONES, G., RAJKOWSKI, J. et KUBIAK, P., 1997. Conditioned responses of monkey locus coeruleus neurons anticipate acquisition of discriminative behavior in a vigilance task. *Neuroscience*, 80(3) :697–715.
- ASTON-JONES, Gary et COHEN, Jonathan D., 2005. An integrative theory of Locus Coeruleus-Norepinephrine function : Adaptive Gain and Optimal Performance. *Annual Review of Neuroscience*, 28(1) :403–450.
- ASTON-JONES, Gary, RAJKOWSKI, Janusz et COHEN, Jonathan, 1999. Role of locus coeruleus in attention and behavioral flexibility. *Biological Psychiatry*, 46(9) :1309–1320.
- BACHEVALIER, J., MEUNIER, M., LU, M. X. et UNGERLEIDER, L. G., 1997. Thalamic and temporal cortex input to medial prefrontal cortex in rhesus monkeys. *Experimental brain research*, 115(3) :430–444.
URL <http://view.ncbi.nlm.nih.gov/pubmed/9262198>
- BADRE, David, 2008. Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends in Cognitive Sciences*, 12(5) :193–200.
doi :10.1016/j.tics.2008.02.004.
- BAKER, Chris I., BEHRMANN, Marlene et OLSON, Carl R., 2002. Impact of learning on representation of parts and wholes in monkey inferotemporal cortex. *Nature neuroscience*, 5(11) :1210–1216. doi :10.1038/nn960.
URL <http://dx.doi.org/10.1038/nn960>
- BALASUBRAMANI, Pragathi P., CHAKRAVARTHY, V. Srinivasa, RAVINDRAN, Balaraman et MOUSTAFA, Ahmed A., 2014. An extended reinforcement learning model of basal ganglia to understand the contributions of serotonin and dopamine in risk-based decision making, reward prediction, and punishment learning. *Frontiers in computational neuroscience*, 8. doi : 10.3389/fncom.2014.00047.
URL <http://dx.doi.org/10.3389/fncom.2014.00047>
- BALLEINE, B. W., LILJEHOLM, M. et OSTLUND, S. B., 2009. The integrative function of the basal ganglia in instrumental conditioning. *Behav Brain Res*, 199(1) :43–52+.
- BALLEINE, Bernard W. et KILLCROSS, Simon, 2006. Parallel incentive processing : an integrated view of amygdala function. *Trends Neurosci*, 29(5) :272–279. doi :10.1016/j.tins.2006.03.002.
- BALLEINE, Bernard W. et O'DOHERTY, John P., 2010. Human and rodent homologies in action control : corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology*, 35(1) :48–69. doi :10.1038/npp.2009.131.
URL <http://dx.doi.org/10.1038/npp.2009.131>

BIBLIOGRAPHIE

- BARNES, Terra D., KUBOTA, Yasuo, HU, Dan, JIN, Dezhe Z. et GRAYBIEL, Ann M., 2005. Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature*, 437(7062) :1158–1161. doi :10.1038/nature04053.
URL <http://dx.doi.org/10.1038/nature04053>
- BASKERVILLE, K. A., SCHWEITZER, J. B. et HERRON, P., 1997. Effects of cholinergic depletion on experience-dependent plasticity in the cortex of the rat. *Neuroscience*, 80(4) :1159–1169.
URL <http://view.ncbi.nlm.nih.gov/pubmed/9284068>
- BEAR, M. F. et SINGER, W., 1986. Modulation of visual cortical plasticity by acetylcholine and noradrenaline. *Nature*, 320(6058) :172–176. doi :10.1038/320172a0.
URL <http://dx.doi.org/10.1038/320172a0>
- BECHARA, Antoine, DAMASIO, Hanna, DAMASIO, Antonio R. et LEE, Gregory P., 1999. Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *The Journal of Neuroscience*, 19(13) :5473–5481.
URL <http://www.jneurosci.org/content/19/13/5473.abstract>
- BECHARA, Antoine et NAQVI, Nasir, 2004. Listening to your heart : interoceptive awareness as a gateway to feeling. *Nature Neuroscience*, 7(2) :102–103. doi :10.1038/nn0204-102.
URL <http://dx.doi.org/10.1038/nn0204-102>
- BECKER, S, 1992. An information-theoretic, unsupervised learning algorithm for neural networks. *Doctoral dissertation. Department of Computer Science. University of Toronto.*
- BELOVA, Marina A., PATON, Joseph J., MORRISON, Sara E. et SALZMAN, C. Daniel, 2015. Expectation modulates neural responses to pleasant and aversive stimuli in primate amygdala. *Neuron*, 55(6) :970–984. doi :10.1016/j.neuron.2007.08.004.
URL <http://dx.doi.org/10.1016/j.neuron.2007.08.004>
- BERNS, Gregory S., MCCLURE, Samuel M., PAGNONI, Giuseppe et MONTAGUE, P. Read, 2001. Predictability modulates human brain response to reward. *The Journal of Neuroscience*, 21(8) :2793–2798.
URL <http://www.jneurosci.org/content/21/8/2793.abstract>
- BERRIDGE, C. W., STRATFORD, T. L., FOOTE, S. L. et KELLEY, A. E., 1997. Distribution of dopamine beta-hydroxylase-like immunoreactive fibers within the shell subregion of the nucleus accumbens. *Synapse (New York,*

- N. Y.*), 27(3) :230–241.
URL <http://view.ncbi.nlm.nih.gov/pubmed/9329158>
- BERRIDGE, Craig W. et WATERHOUSE, Barry D., 2003. The locus coeruleus-noradrenergic system : modulation of behavioral state and state-dependent cognitive processes. *Brain Research Reviews*, 42(1) :33–84.
- BERRIDGE, Kent C. et ROBINSON, Terry E., 1998. What is the role of dopamine in reward : hedonic impact, reward learning, or incentive salience? *Brain Research Reviews*, 28(3) :309–369. doi :10.1016/s0165-0173(98)00019-8.
URL [http://dx.doi.org/10.1016/s0165-0173\(98\)00019-8](http://dx.doi.org/10.1016/s0165-0173(98)00019-8)
- BIEDENKAPP, Joseph C. et RUDY, Jerry W., 2009. Hippocampal and extrahippocampal systems compete for control of contextual fear : Role of ventral subiculum and amygdala. *Learning & Memory*, 16(1) :38–45. doi : 10.1101/lm.1099109.
- BISSIERE, Stephanie, HUMEAU, Yann et LUTHI, Andreas, 2003. Dopamine gates ltp induction in lateral amygdala by suppressing feedforward inhibition. *Nature Neuroscience*, 6(6) :587–592. doi :10.1038/nn1058.
URL <http://dx.doi.org/10.1038/nn1058>
- BLISS, T. V. et COLLINGRIDGE, G. L., 1993. A synaptic model of memory : long-term potentiation in the hippocampus. *Nature*, 361(6407) :31–39. doi : 10.1038/361031a0.
URL <http://dx.doi.org/10.1038/361031a0>
- BORNSTEIN, Aaron M. et DAW, Nathaniel D., 2011. Multiplicity of control in the basal ganglia : computational roles of striatal subregions. *Current opinion in neurobiology*, 21(3) :374–380.
URL <http://dx.doi.org/10.1016/j.conb.2011.02.009>
- BOTVINICK, Matthew M., COHEN, Jonathan D. et CARTER, Cameron S., 2004. Conflict monitoring and anterior cingulate cortex : an update. *Trends in cognitive sciences*, 8(12) :539–546.
URL <http://dx.doi.org/10.1016/j.tics.2004.10.003>
- BOURET, Sebastien, DUVEL, Adam, ONAT, Selim et SARA, Susan J., 2003. Phasic activation of locus ceruleus neurons by the central nucleus of the amygdala. *The Journal of Neuroscience*, 23(8) :3491–3497.
URL <http://www.jneurosci.org/content/23/8/3491.abstract>
- BOURET, Sebastien et SARA, Susan J., 2005. Network reset : a simplified overarching theory of locus coeruleus noradrenaline function. *Trends in Neurosciences*, 28(11) :574–582.

BIBLIOGRAPHIE

- BROWN, Eric, GAO, Juan, HOLMES, Philip, BOGACZ, Rafal, GILZENRAT, Mark et COHEN, Jonathan D., 2005. Simple Neural Networks that optimize decisions. *Int. J. Bifurcation Chaos*, 15(03) :803–826.
URL <http://dx.doi.org/10.1142/s0218127405012478>
- BURGOS-ROBLES, Anthony, VIDAL-GONZALEZ, Ivan et QUIRK, Gregory J., 2009. Sustained Conditioned Responses in Prelimbic Prefrontal Neurons Are Correlated with Fear Expression and Extinction Failure. *J. Neurosci.*, 29(26) :8474–8482. doi :10.1523/jneurosci.0378-09.2009.
- BURWELL, R. D., WITTER, M. P. et AMARAL, D. G., 1995. Perirhinal and postrhinal cortices of the rat : a review of the neuroanatomical literature and comparison with findings from the monkey brain. *Hippocampus*, 5(5) :390–408.
URL <http://view.ncbi.nlm.nih.gov/pubmed/8773253>
- BUSH, Robert R. et MOSTELLER, Frederick, 1951. A mathematical model for simple learning. *Psychological Review*, 58(5) :313–323. doi :10.1037/h0054388.
URL <http://dx.doi.org/10.1037/h0054388>
- BUSSEY, Timothy J. et SAKSIDA, Lisa M., 2002. The organization of visual object representations : a connectionist model of effects of lesions in perirhinal cortex. *The European journal of neuroscience*, 15(2) :355–364.
URL <http://view.ncbi.nlm.nih.gov/pubmed/11849301>
- BUSTI, Daniela, GERACITANO, Raffaella, WHITTLE, Nigel, DALEZIOS, Yannis, MAŃKO, Mirosława, KAUFMANN, Walter, SÄTZLER, Kurt, SINGEWALD, Nicolas, CAPOGNA, Marco et FERRAGUTI, Francesco, 2011. Different fear states engage distinct networks within the intercalated cell clusters of the amygdala. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 31(13) :5131–5144. doi :10.1523/jneurosci.6100-10.2011.
URL <http://dx.doi.org/10.1523/jneurosci.6100-10.2011>
- CALANDREAU, Ludovic, TRIFILIEFF, Pierre, MONS, Nicole, COSTES, Laurence, MARIEN, Marc, MARIGHETTO, Aline, MICHEAU, Jacques, JAFFARD, Robert et DESMEDT, Aline, 2006a. Extracellular hippocampal acetylcholine level controls amygdala function and promotes adaptive conditioned emotional response. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 26(52) :13556–13566.
- CALANDREAU, Ludovic, TRIFILIEFF, Pierre, MONS, Nicole, COSTES, Laurence, MARIEN, Marc, MARIGHETTO, Aline, MICHEAU, Jacques, JAFFARD, Robert et DESMEDT, Aline, 2006b. Extracellular hippocampal acetylcholine level controls amygdala function and promotes adaptive conditioned emotional response. *The Journal of neuroscience : the official journal of the*

Society for Neuroscience, 26(52) :13556–13566.

URL <http://view.ncbi.nlm.nih.gov/pubmed/17192439>

CARDINAL, Rudolf N., PARKINSON, John A., HALL, Jeremy et EVERITT, Barry J., 2002. Emotion and motivation : the role of the amygdala, ventral striatum, and prefrontal cortex. *Neuroscience & Biobehavioral Reviews*, 26(3) :321–352. doi :10.1016/s0149-7634(02)00007-6.

URL [http://dx.doi.org/10.1016/s0149-7634\(02\)00007-6](http://dx.doi.org/10.1016/s0149-7634(02)00007-6)

CARLSEN, Jørn, 1988. Immunocytochemical localization of glutamate decarboxylase in the rat basolateral amygdaloid nucleus, with special reference to gabaergic innervation of amygdalostriatal projection neurons. *J. Comp. Neurol.*, 273(4) :513–526. doi :10.1002/cne.902730407.

URL <http://dx.doi.org/10.1002/cne.902730407>

CARRERE, Maxime et ALEXANDRE, Frédéric, 2013. Émergence de catégories par interaction entre systèmes d'apprentissage. Dans Philippe Preux et Marc Tommasi, rédacteurs, *Conférence Francophone sur l'Apprentissage Automatique (CAP)*. Lille, France.

URL <https://hal.inria.fr/hal-00826554>

CARRERE, Maxime et ALEXANDRE, Frédéric, 2015. A pavlovian model of the amygdala and its influence within the medial temporal lobe. *Frontiers in systems neuroscience*, 9. doi :10.3389/fnsys.2015.00041.

URL <http://dx.doi.org/10.3389/fnsys.2015.00041>

CARRETERO-GUILLÉN, Alejandro, PACHECO-CALDERÓN, Renny, DELGADO-GARCÍA, José M. et GRUART, Agnès, 2013a. Involvement of Hippocampal Inputs and Intrinsic Circuit in the Acquisition of Context and Cues During Classical Conditioning in Behaving Rabbits. *Cerebral cortex (New York, N.Y. : 1991)*.

CARRETERO-GUILLÉN, Alejandro, PACHECO-CALDERÓN, Renny, DELGADO-GARCÍA, José M. et GRUART, Agnès, 2013b. Involvement of hippocampal inputs and intrinsic circuit in the acquisition of context and cues during classical conditioning in behaving rabbits. *Cerebral cortex (New York, N.Y. : 1991)*. doi :10.1093/cercor/bht321.

URL <http://dx.doi.org/10.1093/cercor/bht321>

CARTER, C. S., BRAVER, T. S., BARCH, D. M., BOTVINICK, M. M., NOLL, D. et COHEN, J. D., 1998. Anterior cingulate cortex, error detection, and the on-line monitoring of performance. *Science (New York, N.Y.)*, 280(5364) :747–749. doi :10.1126/science.280.5364.747.

URL <http://dx.doi.org/10.1126/science.280.5364.747>

BIBLIOGRAPHIE

- CASSELL, M. D., FREEDMAN, L. J. et SHI, C., 1999. The intrinsic organization of the central extended amygdala. *Annals of the New York Academy of Sciences*, 877 :217–241.
URL <http://view.ncbi.nlm.nih.gov/pubmed/10415652>
- CHAO, L. L., HAXBY, J. V. et MARTIN, A., 1999. Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature neuroscience*, 2(10) :913–919.
URL <http://view.ncbi.nlm.nih.gov/pubmed/10491613>
- CHRISTIAN, Kimberly M. et THOMPSON, Richard F., 2003a. Neural substrates of eyeblink conditioning : Acquisition and retention. *Learning & Memory*, 10(6) :427–455. doi :10.1101/lm.59603.
URL <http://dx.doi.org/10.1101/lm.59603>
- CHRISTIAN, Kimberly M. et THOMPSON, Richard F., 2003b. Neural substrates of eyeblink conditioning : Acquisition and retention. *Learning & Memory*, 10(6) :427–455. doi :10.1101/lm.59603.
URL <http://dx.doi.org/10.1101/lm.59603>
- CHRISTIANSON, John P., FERNANDO, Anushka B., KAZAMA, Andy M., JOVANOVIC, Tanja, OSTROFF, Linnaea E. et SANGHA, Susan, 2012. Inhibition of fear by learned safety signals : a mini-symposium review. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 32(41) :14118–14124. doi :10.1523/jneurosci.3340-12.2012.
URL <http://dx.doi.org/10.1523/jneurosci.3340-12.2012>
- CIOCCHI, Stephane, HERRY, Cyril, GRENIER, Francois, WOLFF, Steffen B. E., LETZKUS, Johannes J., VLACHOS, Ioannis, EHRLICH, Ingrid, SPRENGEL, Rolf, DEISSEROTH, Karl, STADLER, Michael B., MULLER, Christian et LUTHI, Andreas, 2010. Encoding of conditioned fear in central amygdala inhibitory circuits. *Nature*, 468(7321) :277–282. doi :10.1038/nature09559.
URL <http://dx.doi.org/10.1038/nature09559>
- CISEK, Paul, 2006. Integrated neural processes for defining potential actions and deciding between them : a computational model. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 26(38) :9761–9770. doi :10.1523/jneurosci.5605-05.2006.
URL <http://dx.doi.org/10.1523/jneurosci.5605-05.2006>
- COHEN, Jonathan D., MCCLURE, Samuel M. et YU, Angela J., 2007. Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 362(1481) :933–942.
URL <http://dx.doi.org/10.1098/rstb.2007.2098>

- COOLS, Roshan, NAKAMURA, Kae et DAW, Nathaniel D., 2010. Serotonin and dopamine : Unifying affective, activational, and decision functions. *Neuropsychopharmacology*, 36(1) :98–113. doi :10.1038/npp.2010.121.
URL <http://dx.doi.org/10.1038/npp.2010.121>
- CORBIT, Laura H., MUIR, Janice L. et BALLEINE, Bernard W., 2003. Lesions of mediodorsal thalamus and anterior thalamic nuclei produce dissociable effects on instrumental conditioning in rats. *The European journal of neuroscience*, 18(5) :1286–1294.
URL <http://view.ncbi.nlm.nih.gov/pubmed/12956727>
- CORCORAN, Kevin A. et QUIRK, Gregory J., 2007. Activity in prelimbic cortex is necessary for the expression of learned, but not innate, fears. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 27(4) :840–844. doi :10.1523/jneurosci.5327-06.2007.
URL <http://dx.doi.org/10.1523/jneurosci.5327-06.2007>
- DAW, Nathaniel D., KAKADE, Sham et DAYAN, Peter, 2002. Opponent interactions between serotonin and dopamine. *Neural networks : the official journal of the International Neural Network Society*, 15(4-6) :603–616.
URL <http://view.ncbi.nlm.nih.gov/pubmed/12371515>
- DAW, Nathaniel D., NIV, Yael et DAYAN, Peter, 2005. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature neuroscience*, 8(12) :1704–1711. doi :10.1038/nn1560.
URL <http://dx.doi.org/10.1038/nn1560>
- DAW, Nathaniel D., O'DOHERTY, John P., DAYAN, Peter, SEYMOUR, Ben et DOLAN, Raymond J., 2006. Cortical substrates for exploratory decisions in humans. *Nature*, 441(7095) :876–879.
URL <http://dx.doi.org/10.1038/nature04766>
- DAYAN, Peter, 2012. Twenty-five lessons from computational neuromodulation. *Neuron*, 76(1) :240–256.
URL <http://view.ncbi.nlm.nih.gov/pubmed/23040818>
- DAYAN, Peter et DAW, Nathaniel D., 2008. Decision theory, reinforcement learning, and the brain. *Cognitive, affective & behavioral neuroscience*, 8(4) :429–453.
URL <http://dx.doi.org/10.3758/cabn.8.4.429>
- DENOYELLE, Nicolas, CARRERE, Maxime, POUGET, Florian, VIÉVILLE, Thierry et ALEXANDRE, Frédéric, 2015. From biological to numerical experiments in systemic neuroscience : a simulation platform. Dans P. Encarnaçã

BIBLIOGRAPHIE

- A.R. Londral, rédacteur, *Advances in Neurotechnology, Electronics and Informatics*, tome 12 de *Biosystems & Biorobotics*. Springer.
URL <https://hal.inria.fr/hal-01227968>
- DESMEDT, Aline, MARIGHETTO, Aline, RICHTER-LEVIN, Gal et CALANDREAU, Ludovic, 2015. Adaptive emotional memory : the key hippocampal-amygdalar interaction. *Stress*, 18(3) :297–308. doi :10.3109/10253890.2015.1067676.
URL <http://dx.doi.org/10.3109/10253890.2015.1067676>
- DEZFOULI, Amir et BALLEINE, Bernard W., 2012. Habits, action sequences and reinforcement learning. *The European journal of neuroscience*, 35(7) :1036–1051. doi :10.1111/j.1460-9568.2012.08050.x.
URL <http://dx.doi.org/10.1111/j.1460-9568.2012.08050.x>
- DI MARTINO, A., SCHERES, A., MARGULIES, D. S., KELLY, A. M., UDDIN, L. Q., SHEHZAD, Z., BISWAL, B., WALTERS, J. R., CASTELLANOS, F. X. et MILHAM, M. P., 2008. Functional connectivity of human striatum : a resting state fmri study. *Cerebral cortex (New York, N.Y. : 1991)*, 18(12) :2735–2747. doi :10.1093/cercor/bhn041.
URL <http://dx.doi.org/10.1093/cercor/bhn041>
- DO-MONTE, Fabricio H., QUIÑONES LARACUENTE, Kelvin et QUIRK, Gregory J., 2015. A temporal shift in the circuits mediating retrieval of fear memory. *Nature*, 519(7544) :460–463. doi :10.1038/nature14030.
URL <http://dx.doi.org/10.1038/nature14030>
- DOYA, Kenji, 2002. Metalearning and neuromodulation. *Neural Networks*, 15(4-6) :495–506.
URL [http://dx.doi.org/10.1016/s0893-6080\(02\)00044-8](http://dx.doi.org/10.1016/s0893-6080(02)00044-8)
- DRANIAS, Mark R., GROSSBERG, Stephen et BULLOCK, Daniel, 2008. Dopaminergic and non-dopaminergic value systems in conditioning and outcome-specific revaluation. *Brain Research*, 1238 :239–287. doi :10.1016/j.brainres.2008.07.013.
URL <http://dx.doi.org/10.1016/j.brainres.2008.07.013>
- EICHENBAUM, Howard, SAUVAGE, Magdalena, FORTIN, Norbert, KOMOROWSKI, Robert et LIPTON, Paul, 2012. Towards a functional organization of episodic memory in the medial temporal lobe. *Neurosci Biobehav Rev.*, 36(7) :1597–1608. doi :10.1016/j.neubiorev.2011.07.006.
- EL-AMAMY, Heather et HOLLAND, Peter C., 2007. Dissociable effects of disconnecting amygdala central nucleus from the ventral tegmental area or substantia nigra on learned orienting and incentive motivation. *The European journal of neuroscience*, 25(5) :1557–1567. doi :10.1111/j.1460-9568.

2007.05402.x.

URL <http://dx.doi.org/10.1111/j.1460-9568.2007.05402.x>

ELIASMITH C., Anderson C. H., 2003. Neural Engineering : Computation, Representation, and Dynamics in Neurobiological Systems. MIT Press.

EVERITT, B. J. et ROBBINS, T. W., 1997a. Central cholinergic systems and cognition. *Annual review of psychology*, 48 :649–684. doi :10.1146/annurev.psych.48.1.649.

URL <http://dx.doi.org/10.1146/annurev.psych.48.1.649>

EVERITT, B. J. et ROBBINS, T. W., 1997b. Central cholinergic systems and cognition. *Annual review of psychology*, 48 :649–684. doi :10.1146/annurev.psych.48.1.649.

URL <http://dx.doi.org/10.1146/annurev.psych.48.1.649>

FANSELOW, M., 2000. Contextual fear, gestalt memories, and the hippocampus. *Behavioural Brain Research*, 110(1-2) :73–81. doi :10.1016/S0166-4328(99)00186-2.

FEATHERSTONE, R. E. et McDONALD, R. J., 2005. Lesions of the dorsolateral or dorsomedial striatum impair performance of a previously acquired simple discrimination task. *Neurobiology of learning and memory*, 84(3) :159–167.

URL <http://view.ncbi.nlm.nih.gov/pubmed/16165379>

FIX, Jérémy, 2008. *Mécanisme numériques et distribués de l'anticipation motrice*. Thèse de doctorat, Laboratoire Lorrain de Recherche en Informatique et ses Applications.

FRISTON, Karl, 2002. Functional integration and inference in the brain. *Progress in neurobiology*, 68(2) :113–143.

URL <http://view.ncbi.nlm.nih.gov/pubmed/12450490>

FUSTER, Joaquin M., 2001. The Prefrontal Cortex—An Update : Time Is of the Essence. *Neuron*, 30(2) :319–333. doi :10.1016/S0896-6273(01)00285-9.

FUXE, Kjell, 1965. Evidence for the existence of monoamine neurons in the central nervous system. 65(4) :573–596. doi :10.1007/bf00337069.

URL <http://dx.doi.org/10.1007/bf00337069>

GILZENRAT, Mark S., HOLMES, Benjamin D., RAJKOWSKI, Janusz, ASTON-JONES, Gary et COHEN, Jonathan D., 2002. Simplified dynamics in a model of noradrenergic modulation of cognitive performance. *Neural networks : the official journal of the International Neural Network Society*, 15(4-6) :647–663.

URL <http://view.ncbi.nlm.nih.gov/pubmed/12371518>

BIBLIOGRAPHIE

- GOLMAYO, L., NUÑEZ, A. et ZABORSZKY, L., 2003. Electrophysiological evidence for the existence of a posterior cortical-prefrontal-basal forebrain circuitry in modulating sensory responses in visual and somatosensory rat cortical areas. *Neuroscience*, 119(2) :597–609.
URL <http://view.ncbi.nlm.nih.gov/pubmed/12770572>
- GOOSENS, Ki A. et MAREN, Stephen, 2001. Contextual and Auditory Fear Conditioning are Mediated by the Lateral, Basal, and Central Amygdaloid Nuclei in Rats. *Learning & Memory*, 8(3) :148–155. doi :10.1101/lm.37601.
- GREBA, Q., GIFKINS, A. et KOKKINIDIS, L., 2001. Inhibition of amygdaloid dopamine d2 receptors impairs emotional learning measured with fear-potentiated startle. *Brain research*, 899(1-2) :218–226.
URL <http://view.ncbi.nlm.nih.gov/pubmed/11311883>
- GROSSBERG, S. et RAIZADA, R. D., 2000. Contrast-sensitive perceptual grouping and object-based attention in the laminar circuits of primary visual cortex. *Vision research*, 40(10-12) :1413–1432.
URL <http://view.ncbi.nlm.nih.gov/pubmed/10788649>
- GUTHRIE, M., LEBLOIS, A., GARENNE, A. et BORAUD, T., 2013. Interaction between cognitive and motor cortico-basal ganglia loops during decision making : a computational study. *Journal of Neurophysiology*, 109(12) :3025–3040.
- HABER, Suzanne, 2011. *Neuroanatomy of Reward*, tome 20114980, pages 235–261. CRC Press. ISBN 978-1-4200-6726-2. doi :10.1201/b10776-15.
URL <http://dx.doi.org/10.1201/b10776-15>
- HABER, Suzanne N. et KNUTSON, Brian, 2010. The reward circuit : Linking primate anatomy and human imaging. *Neuropsychopharmacology REVIEWS*, 35 :4–26.
- HASSELMO, Michael E., 2006. The role of acetylcholine in learning and memory. *Curr Opin Neurobiol*, 16(6) :710–715. doi :10.1016/j.conb.2006.09.002.
- HAXBY, J. V., GOBBINI, M. I., FUREY, M. L., ISHAI, A., SCHOUTEN, J. L. et PIETRINI, P., 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science (New York, N.Y.)*, 293(5539) :2425–2430. doi :10.1126/science.1063736.
URL <http://dx.doi.org/10.1126/science.1063736>
- HEALY, Michael J., CAUDELL, Thomas P. et GOLDSMITH, Timothy E., 2008. A model of human categorization and similarity based upon category theory. Rapport technique, University of New Mexico, Albuquerque, New Mexico 87131, USA.

- HEBB, D. O., 1950. Organization of behavior. *J. Clin. Psychol.*, 6(3) :307.
URL [http://dx.doi.org/10.1002/1097-4679\(195007\)6:3%3C307::aid-jclp2270060338%3E3.0.co;2-k](http://dx.doi.org/10.1002/1097-4679(195007)6:3%3C307::aid-jclp2270060338%3E3.0.co;2-k)
- HERRY, Cyril, CIOCCHI, Stephane, SENN, Verena, DEMMOU, Lynda, MULLER, Christian et LUTHI, Andreas, 2008. Switching on and off fear by distinct neuronal circuits. *Nature*, 454(7204) :600–606. doi :10.1038/nature07166.
URL <http://dx.doi.org/10.1038/nature07166>
- HERRY, Cyril et JOHANSEN, Joshua P., 2014. Encoding of fear learning and memory in distributed neuronal circuits. *Nature neuroscience*, 17(12) :1644–1654.
URL <http://view.ncbi.nlm.nih.gov/pubmed/25413091>
- HOLLAND, P. C. et GALLAGHER, M., 1999. Amygdala circuitry in attentional and representational processes. *Trends in cognitive sciences*, 3(2) :65–73.
URL <http://view.ncbi.nlm.nih.gov/pubmed/10234229>
- HOLLAND, Peter C. et GALLAGHER, Michela, 2004. Amygdala-frontal interactions and reward expectancy. *Current Opinion in Neurobiology*, 14(2). doi :10.1016/j.conb.2004.03.007.
- HUFF, Nicole C. et RUDY, Jerry W., 2004. The amygdala modulates hippocampus-dependent context memory formation and stores cue-shock associations. *Behavioral neuroscience*, 118(1) :53–62.
URL <http://view.ncbi.nlm.nih.gov/pubmed/14979782>
- HUMPHRIES, Mark D., KHAMASSI, Mehdi et GURNEY, Kevin, 2012. Dopaminergic control of the exploration-exploitation trade-off via the basal ganglia. *Frontiers in neuroscience*, 6.
URL <http://view.ncbi.nlm.nih.gov/pubmed/22347155>
- JAHFARI, Sara, WALDORP, Lourens, VAN DEN WILDENBERG, Wery P., SCHOLTE, H. Steven, RIDDERINKHOF, K. Richard et FORSTMANN, Birte U., 2011. Effective connectivity reveals important roles for both the hyperdirect (fronto-subthalamic) and the indirect (fronto-striatal-pallidal) fronto-basal ganglia pathways during response inhibition. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 31(18) :6891–6899.
URL <http://view.ncbi.nlm.nih.gov/pubmed/21543619>
- JODO, E., CHIANG, C. et ASTON-JONES, G., 1998. Potent excitatory influence of prefrontal cortex activity on noradrenergic locus coeruleus neurons. *Neuroscience*, 83(1) :63–79.
URL <http://view.ncbi.nlm.nih.gov/pubmed/9466399>

BIBLIOGRAPHIE

- JOEL, Daphna, NIV, Yael et RUPPIN, Eytan, 2002. Actor-critic models of the basal ganglia : new anatomical and computational perspectives. *Neural Networks*, 15(4-6) :535–547.
URL [http://dx.doi.org/10.1016/s0893-6080\(02\)00047-3](http://dx.doi.org/10.1016/s0893-6080(02)00047-3)
- JOHNSON, Luke R., HOU, Mian, PRAGER, Eric M. et LEDOUX, Joseph E., 2011a. Regulation of the Fear Network by Mediators of Stress : Norepinephrine Alters the Balance between Cortical and Subcortical Afferent Excitation of the Lateral Amygdala. *Frontiers in Behavioral Neuroscience*, 5. doi :10.3389/fnbeh.2011.00023.
URL <http://dx.doi.org/10.3389/fnbeh.2011.00023>
- JOHNSON, Luke R., HOU, Mian, PRAGER, Eric M. et LEDOUX, Joseph E., 2011b. Regulation of the fear network by mediators of stress : Norepinephrine alters the balance between cortical and subcortical afferent excitation of the lateral amygdala. *Frontiers in behavioral neuroscience*, 5.
URL <http://view.ncbi.nlm.nih.gov/pubmed/21647395>
- KASSAB, Randa et ALEXANDRE, Frédéric, 2015. Integration of exteroceptive and interoceptive information within the hippocampus : a computational study. *Frontiers in systems neuroscience*, 9. doi :10.3389/fnsys.2015.00087.
URL <http://dx.doi.org/10.3389/fnsys.2015.00087>
- KAZAMA, Andy M., HEUER, Eric, DAVIS, Michael et BACHEVALIER, Jocelyne, 2012. Effects of neonatal amygdala lesions on fear learning, conditioned inhibition, and extinction in adult macaques. *Behavioral neuroscience*, 126(3) :392–403.
URL <http://view.ncbi.nlm.nih.gov/pubmed/22642884>
- KILLCROSS, Simon et COUTUREAU, Etienne, 2003. Coordination of actions and habits in the medial prefrontal cortex of rats. *Cerebral cortex*, 13(4) :400–408.
URL <http://dx.doi.org/10.1093/cercor/13.4.400>
- KIM, Dongbeom, PARÉ, Denis et NAIR, Satish S., 2013. Mechanisms contributing to the induction and storage of Pavlovian fear memories in the lateral amygdala. *Learning & memory (Cold Spring Harbor, N.Y.)*, 20(8) :421–430.
- KLAVIR, Oded, GENUD-GABAI, Rotem et PAZ, Rony, 2013. Functional connectivity between amygdala and cingulate cortex for adaptive aversive learning. *Neuron*, 80(5) :1290–1300.
URL <http://view.ncbi.nlm.nih.gov/pubmed/24314732>
- KREBS, Ruth M., FIAS, Wim, ACHTEN, Eric et BOEHLER, Carsten N., 2013. Picture novelty attenuates semantic interference and modulates concomitant neural activity in the anterior cingulate cortex and the locus coeruleus.

- NeuroImage*, 74 :179–187.
URL <http://view.ncbi.nlm.nih.gov/pubmed/23454569>
- KRINGELBACH, Morten L., 2005. The human orbitofrontal cortex : linking reward to hedonic experience. *Nat Rev Neurosci*, 6(9) :691–702.
- KRINGELBACH, Morten L. et ROLLS, Edmund T., 2004. The functional neuroanatomy of the human orbitofrontal cortex : evidence from neuroimaging and neuropsychology. *Progress in neurobiology*, 72(5) :341–372. doi : 10.1016/j.pneurobio.2004.03.006.
URL <http://dx.doi.org/10.1016/j.pneurobio.2004.03.006>
- KRUSCHKE, John K., 2001. Toward a unified model of attention in associative learning. *Journal of Mathematical Psychology*, 45 :812–863.
- LAMME, V. A., 2001. Blindsight : the role of feedforward and feedback corticocortical connections. *Acta psychologica*, 107(1-3) :209–228.
URL <http://view.ncbi.nlm.nih.gov/pubmed/11388136>
- LAMME, V. A. et ROELFSEMA, P. R., 2000. The distinct modes of vision offered by feedforward and recurrent processing. *Trends in neurosciences*, 23(11) :571–579.
URL <http://view.ncbi.nlm.nih.gov/pubmed/11074267>
- LAMME, Victor A. F., 2003. Why visual attention and awareness are different. *Trends Cogn Sci*, 7(1) :12–18. doi :au%20%20-%20lamme,%20victor%20a.f.
URL <http://dx.doi.org/au%20%20-%20lamme,%20victor%20a.f.>
- LAMONT, E. W. et KOKKINIDIS, L., 1998. Infusion of the dopamine d1 receptor antagonist sch 23390 into the amygdala blocks fear expression in a potentiated startle paradigm. *Brain research*, 795(1-2) :128–136.
URL <http://view.ncbi.nlm.nih.gov/pubmed/9622611>
- LAURENT, Vincent, MARCHAND, Alain R. et WESTBROOK, R. Frederick, 2008. The basolateral amygdala is necessary for learning but not relearning extinction of context conditioned fear. *Learning & memory (Cold Spring Harbor, N.Y.)*, 15(5) :304–314.
- LE PELLEY, M. E., 2004a. The role of associative history in models of associative learning : a selective review and a hybrid model. *The Quarterly journal of experimental psychology. B, Comparative and physiological psychology*, 57(3) :193–243.
URL <http://view.ncbi.nlm.nih.gov/pubmed/15204108>
- LE PELLEY, M. E., 2004b. The role of associative history in models of associative learning : a selective review and a hybrid model. *The Quarterly Journal of Experimental Psychology*, 57(3) :193–243. doi :10.1080/

BIBLIOGRAPHIE

02724990344000141.

URL <http://dx.doi.org/10.1080/02724990344000141>

LEDOUX, Joseph, 2003. The emotional brain, fear, and the amygdala. *Cellular and molecular neurobiology*, 23(4-5) :727–738.

URL <http://view.ncbi.nlm.nih.gov/pubmed/14514027>

LEDOUX, Joseph, 2007. The amygdala. *Current Biology*, 17(20) :R868–R874.

LEDOUX, Joseph E., 2000. Emotion circuits in the brain. *Annual Review of Neuroscience*, 23(1) :155–184. doi :10.1146/annurev.neuro.23.1.155.

URL <http://dx.doi.org/10.1146/annurev.neuro.23.1.155>

LEE, Seungho, KIM, Su-Jeong, KWON, Oh-Bin, LEE, Joo H. et KIM, Joung-Hun, 2013. Inhibitory networks of the amygdala for emotional memory. *Frontiers in Neural Circuits*, 7(129). doi :10.3389/fncir.2013.00129.

URL http://www.frontiersin.org/neural_circuits/10.3389/fncir.2013.00129/abstract

LEVY, W. B., 1996. A sequence predicting ca3 is a flexible associator that learns and uses context to solve hippocampal-like tasks. *Hippocampus*, 6(6) :579–590. doi :10.1002/(sici)1098-1063(1996)6:6<579::aid-hipo3>3.0.co;2-c.

URL [http://dx.doi.org/10.1002/\(sici\)1098-1063\(1996\)6:6<579::aid-hipo3>3.0.co;2-c](http://dx.doi.org/10.1002/(sici)1098-1063(1996)6:6<579::aid-hipo3>3.0.co;2-c)

LEVY, William B., 1989. *A Computational Approach to Hippocampal Function*, tome 23, pages 243–305. Elsevier. ISBN 9780125433235. doi :10.1016/s0079-7421(08)60113-9.

URL [http://dx.doi.org/10.1016/s0079-7421\(08\)60113-9](http://dx.doi.org/10.1016/s0079-7421(08)60113-9)

LI, Guoshi, NAIR, Satish S. et QUIRK, Gregory J., 2009. A Biologically Realistic Network Model of Acquisition and Extinction of Conditioned Fear Associations in Lateral Amygdala Neurons. *Journal of Neurophysiology*, 101(3) :1629–1646. doi :10.1152/jn.90765.2008.

LI, Susan S. et MCNALLY, Gavan P., 2014. The conditions that promote fear learning : Prediction error and Pavlovian fear conditioning. *Neurobiology of Learning and Memory*, 108(0) :14–21. doi :<http://dx.doi.org/10.1016/j.nlm.2013.05.002>.

LI, Z., 1996. A theory of the visual motion coding in the primary visual cortex. *Neural computation*, 8(4) :705–730.

URL <http://view.ncbi.nlm.nih.gov/pubmed/8624959>

- LISMAN, J. E., 1999. Relating hippocampal circuitry to function : recall of memory sequences by reciprocal dentate-ca3 interactions. *Neuron*, 22(2) :233–242.
URL <http://view.ncbi.nlm.nih.gov/pubmed/10069330>
- LONG, Arwen B., KUHN, Cynthia M. et PLATT, Michael L., 2009. Serotonin shapes risky decision making in monkeys. *Social cognitive and affective neuroscience*, 4(4) :346–356. doi :10.1093/scan/nsp020.
URL <http://dx.doi.org/10.1093/scan/nsp020>
- LYND-BALTA, E. et HABER, S. N., 1994. Primate striatonigral projections : a comparison of the sensorimotor-related striatum and the ventral striatum. *The Journal of comparative neurology*, 345(4) :562–578.
URL <http://view.ncbi.nlm.nih.gov/pubmed/7962700>
- MACKINTOSH, N. J., 1975. A theory of attention : Variations in the associability of stimuli with reinforcement. *Psychological Review*, 82(4) :276–298.
- MAIA, Tiago V., COONEY, Rebecca E. et PETERSON, Bradley S., 2008. The neural bases of obsessive-compulsive disorder in children and adults. *Development and psychopathology*, 20(4) :1251–1283.
URL <http://view.ncbi.nlm.nih.gov/pubmed/18838041>
- MANNELLA, Francesco, GURNEY, Kevin et BALDASSARRE, Gianluca, 2013. The nucleus accumbens as a nexus between values and goals in goal-directed behavior : a review and a new hypothesis. *Frontiers in behavioral neuroscience*, 7.
- MAREN, Stephen, 2005. Building and Burying Fear Memories in the Brain. *The Neuroscientist*, 11(1) :89–99. doi :10.1177/1073858404269232.
- MARR, D., 1971. Simple memory : A theory for archicortex. *Philosophical Transactions of the Royal Society B : Biological Sciences*, 262(841) :23–81. doi :10.1098/rstb.1971.0078.
URL <http://dx.doi.org/10.1098/rstb.1971.0078>
- MARSHALL, Lisa et BORN, Jan, 2007. The contribution of sleep to hippocampus-dependent memory consolidation. *Trends in cognitive sciences*, 11(10) :442–450. doi :10.1016/j.tics.2007.09.001.
URL <http://dx.doi.org/10.1016/j.tics.2007.09.001>
- MASON, Stephen T. et IVERSEN, Susan D., 1977. An investigation of the role of cortical and cerebellar noradrenaline in associative motor learning in the rat. *Brain Research*, 134(3) :513–527. doi :10.1016/0006-8993(77)90826-5.
URL [http://dx.doi.org/10.1016/0006-8993\(77\)90826-5](http://dx.doi.org/10.1016/0006-8993(77)90826-5)

- MCCLELLAND, J. L., MCNAUGHTON, B. L. et O'REILLY, R. C., 1995. Why there are complementary learning systems in the hippocampus and neocortex : insights from the successes and failures of connectionist models of learning and memory. *Psychological review*, 102(3) :419–457.
- MCCLURE, Samuel, GILZENRAT, Mark et COHEN, Jonathan, 2006. An exploration-exploitation model based on norepinephrine and dopamine activity. Dans Y. Weiss, B. Schölkopf et J. Platt, rédacteurs, *Advances in Neural Information Processing Systems 18*, pages 867–874. MIT Press.
URL http://www.csmb.princeton.edu/~smcclure/pdf/MGC_NIPS.pdf
- MCGAUGH, James L., 2004. The amygdala modulates the consolidation of memories of emotionally arousing experiences. *Annual review of neuroscience*, 27(1) :1–28. doi :10.1146/annurev.neuro.27.070203.144157.
- MCINTYRE, C. K., MARRIOTT, L. K. et GOLD, P. E., 2003. Cooperation Between Memory Systems : Acetylcholine Release in the Amygdala Correlates Positively With Performance on a Hippocampus-Dependent Task. *Behavioral Neuroscience*, 117 :320–326.
- MCNAUGHTON, B. et NADEL, L., 1990. Hebb-Marr networks and the neurobiological representation of action in space.
- MCNAUGHTON, B. L. et MORRIS, R. G. M., 1987. Hippocampal synaptic enhancement and information storage within a distributed memory system. *Trends in Neurosciences*, 10(10) :408–415. doi :10.1016/0166-2236(87)90011-7.
URL [http://dx.doi.org/10.1016/0166-2236\(87\)90011-7](http://dx.doi.org/10.1016/0166-2236(87)90011-7)
- MEETER, M., MYERS, C. E. et GLUCK, M. A., 2005. Integrating Incremental Learning and Episodic Memory Models of the Hippocampal Region. *Psychological Review*, 112(3) :560–585. doi :10.1037/0033-295x.112.3.560.
- MIKKULAINEN, Risto, BEDNAR, James A., CHOE, Yoonsuck et SIROSH, Joseph, 1997. Self-organization, plasticity, and low-level visual phenomena in a laterally connected map model of the primary visual cortex. Dans R. L. Goldstone, P. G. Schyns et D. L. Medin, rédacteurs, *Perceptual Learning, Psychology of Learning and Motivation*, pages 257–308.
URL <http://z.cs.utexas.edu/users/ai-lab/pub-view.php?PubID=97094>
- MILLER, M. W. et VOGT, B. A., 1984. Direct connections of rat visual cortex with sensory, motor, and association cortices. *The Journal of comparative neurology*, 226(2) :184–202.
URL <http://view.ncbi.nlm.nih.gov/pubmed/6736299>

- MINCES, V. H., ALEXANDER, A. S., DATLOW, M., ALFONSO, S. I. et CHIBA, A. A., 2013. The role of visual cortex acetylcholine in learning to discriminate temporally modulated visual stimuli. *Frontiers in behavioral neuroscience*, 7. doi :10.3389/fnbeh.2013.00016.
URL <http://dx.doi.org/10.3389/fnbeh.2013.00016>
- MINK, J. W., 1996. The basal ganglia : focused selection and inhibition of competing motor programs. *Progress in neurobiology*, 50(4) :381–425.
URL <http://view.ncbi.nlm.nih.gov/pubmed/9004351>
- MORÉN, Jan, 2002. *Emotion and Learning A Computational Model of the Amygdala*. Thèse de doctorat, Lund University Cognitive Studies 93.
- MOUSTAFA, Ahmed A., GILBERTSON, Mark W., ORR, Scott P., HERZALLAH, Mohammad M., SERVATIUS, Richard J. et MYERS, Catherine E., 2013. A model of amygdala–hippocampal–prefrontal interaction in fear conditioning and extinction in animals. *Brain and Cognition*, 81(1) :29–43.
- NICOLA, S. M. et MALENKA, R. C., 1998. Modulation of synaptic transmission by dopamine and norepinephrine in ventral but not dorsal striatum. *Journal of neurophysiology*, 79(4) :1768–1776.
URL <http://view.ncbi.nlm.nih.gov/pubmed/9535946>
- NITECKA, L. et NARKIEWICZ, O., 1976. Localization of acetylcholinesterase activity in the amygdaloid body of man. *Acta neurobiologiae experimentalis*, 36(3) :333–351.
URL <http://view.ncbi.nlm.nih.gov/pubmed/970243>
- NIV, Yael, 2009. Reinforcement learning in the brain. *Journal of Mathematical Psychology*, 53(3) :139–154.
URL <http://linkinghub.elsevier.com/retrieve/pii/S0022249608001181>
- NIV, Yael, DAW, Nathaniel D., JOEL, Daphna et DAYAN, Peter, 2007. Tonic dopamine : opportunity costs and the control of response vigor. *Psychopharmacology*, 191(3) :507–520. doi :10.1007/s00213-006-0502-4.
URL <http://dx.doi.org/10.1007/s00213-006-0502-4>
- OLSHAUSEN, Bruno A. et FIELD, David J., 1996. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(6583) :607–609. doi :10.1038/381607a0.
URL <http://dx.doi.org/10.1038/381607a0>
- O'REILLY, Randall C. et FRANK, Michael J., 2006. Making Working Memory Work : A Computational Model of Learning in the Prefrontal Cortex and Basal Ganglia. *Neural Comp.*, 18(2) :283–328.

BIBLIOGRAPHIE

- O'REILLY, Randall C., FRANK, Michael J., HAZY, Thomas E. et WATZ, Brandon, 2007a. PVLV : The primary value and learned value Pavlovian learning algorithm. *Behavioral neuroscience*, 121(1) :31–49.
URL <http://dx.doi.org/10.1037/0735-7044.121.1.31>
- O'REILLY, Randall C., FRANK, Michael J., HAZY, Thomas E. et WATZ, Brandon, 2007b. Pvlv : The primary value and learned value pavlovian learning algorithm. *Behavioral Neuroscience*, 121 :31–49.
- O'REILLY, Randall C. et RUDY, Jerry W., 2001. Conjunctive Representations in Learning and Memory : Principles of Cortical and Hippocampal Function. *Psychological Review*, 108(2) :311–345.
- O'REILLY, R.C. et MUNAKATA, Y., 2000. *Computational Explorations in Cognitive Neuroscience : Understanding the Mind by Simulating the Brain*. MIT Press, Cambridge, MA, USA.
- PACKARD, Mark G. et KNOWLTON, Barbara J., 2002. Learning and memory functions of the basal ganglia. *Annual Review of Neuroscience*, 25(1) :563–593. doi :10.1146/annurev.neuro.25.112701.142937.
URL <http://dx.doi.org/10.1146/annurev.neuro.25.112701.142937>
- PARE, D., QUIRK, G. J. et LEDOUX, J. E., 2004. New vistas on amygdala networks in conditioned fear. *J Neurophysiol*, 92(1) :1–9.
- PASQUEREAU, Benjamin, NADJAR, Agnes, ARKADIR, David, BEZARD, Erwan, GOILLANDEAU, Michel, BIOULAC, Bernard, GROSS, Christian Eric E. et BORAUD, Thomas, 2007. Shaping of motor responses by incentive values through the basal ganglia. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 27(5) :1176–1183. doi :10.1523/jneurosci.3745-06.2007.
URL <http://dx.doi.org/10.1523/jneurosci.3745-06.2007>
- PAULI, Wolfgang M., HAZY, Thomas E. et O'REILLY, Randall C., 2011. Expectancy, Ambiguity, and Behavioral Flexibility : Separable and Complementary Roles of the Orbital Frontal Cortex and Amygdala in Processing Reward Expectancies. *Journal of Cognitive Neuroscience*, 24(2) :351–366.
URL http://dx.doi.org/10.1162/jocn_a_00155
- PAULI, Wolfgang M., HAZY, Thomas E. et O'REILLY, Randall C., 2012. Expectancy, ambiguity, and behavioral flexibility : separable and complementary roles of the orbital frontal cortex and amygdala in processing reward expectancies. *Journal of cognitive neuroscience*, 24(2) :351–366.
URL <http://view.ncbi.nlm.nih.gov/pubmed/22004047>

- PAULI, Wolfgang M. et O'REILLY, Randall C., 2008. Attentional control of associative learning—a possible role of the central cholinergic system. *Brain Research*, 1202 :43–53.
- PAZ, Rony et PARÉ, Denis, 2013. Physiological basis for emotional modulation of memory circuits by the amygdala. *Current Opinion in Neurobiology*, 23(3) :381–386. doi :<http://dx.doi.org/10.1016/j.conb.2013.01.008>.
URL <http://www.sciencedirect.com/science/article/pii/S095943881300024X>
- PEARCE, J. M., 1987. A model for stimulus generalization in pavlovian conditioning. *Psychological review*, 94(1) :61–73.
URL <http://view.ncbi.nlm.nih.gov/pubmed/3823305>
- PEARCE, J. M. et HALL, G., 1980. A model for pavlovian learning : variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological review*, 87(6) :532–552.
URL <http://view.ncbi.nlm.nih.gov/pubmed/7443916>
- PENDYAM, Sandeep, BRAVO-RIVERA, Christian, BURGOS-ROBLES, Anthony, SOTRES-BAYON, Francisco, QUIRK, Gregory J. et NAIR, Satish S., 2013. Fear signaling in the prelimbic-amygdala circuit : a computational modeling and recording study. *Journal of Neurophysiology*, 110(4) :844–861. doi : 10.1152/jn.00961.2012.
- PENNER, Marsha R. et MIZUMORI, Sheri J. Y., 2012. Neural systems analysis of decision making during goal-directed navigation. *Progress in neurobiology*, 96(1) :96–135.
- PESSOA, Luiz, 2010. Emotion and cognition and the amygdala : from "what is it?" to "what's to be done?". *Neuropsychologia*, 48(12) :3416–3429. doi : 10.1016/j.neuropsychologia.2010.06.038.
URL <http://dx.doi.org/10.1016/j.neuropsychologia.2010.06.038>
- PHELPS, Elizabeth A., 2004. Human emotion and memory : interactions of the amygdala and hippocampal complex. *Current opinion in neurobiology*, 14(2) :198–202. doi :10.1016/j.conb.2004.03.015.
URL <http://dx.doi.org/10.1016/j.conb.2004.03.015>
- POORT, Jasper, RAUDIES, Florian, WANNIG, Aurel, LAMME, Victor A., NEUMANN, Heiko et ROELFSEMA, Pieter R., 2012. The role of attention in figure-ground segregation in areas v1 and v4 of the visual cortex. *Neuron*, 75(1) :143–156. doi :10.1016/j.neuron.2012.04.032.
URL <http://dx.doi.org/10.1016/j.neuron.2012.04.032>

BIBLIOGRAPHIE

- POUGET, Alexandre, BECK, Jeffrey M., MA, Wei Ji J. et LATHAM, Peter E., 2013. Probabilistic brains : knowns and unknowns. *Nature neuroscience*, 16(9) :1170–1178.
URL <http://dx.doi.org/10.1038/nn.3495>
- PRADO-ALCALÁ, R. A., HAIEK, M., RIVAS, S., ROLDAN-ROLDAN, G. et QUIRARTE, G. L., 1994. Reversal of extinction by scopolamine. *Physiology & Behavior*, 56(1) :27–30.
- PRICE, Joseph L., 2003. Comparative Aspects of Amygdala Connectivity. *Annals of the New York Academy of Sciences*, 985(1) :50–58. doi :10.1111/j.1749-6632.2003.tb07070.x.
URL <http://dx.doi.org/10.1111/j.1749-6632.2003.tb07070.x>
- PUCE, A., ALLISON, T., ASGARI, M., GORE, J. C. et MCCARTHY, G., 1996. Differential sensitivity of human visual cortex to faces, letterstrings, and textures : a functional magnetic resonance imaging study. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 16(16) :5205–5215.
URL <http://www.jneurosci.org/cgi/content/abstract/16/16/5205>
- RAGOZZINO, Michael E., 2007. The contribution of the medial prefrontal cortex, orbitofrontal cortex, and dorsomedial striatum to behavioral flexibility. *Annals of the New York Academy of Sciences*, 1121 :355–375. doi :10.1196/annals.1401.013.
URL <http://dx.doi.org/10.1196/annals.1401.013>
- RAJKOWSKI, J., KUBIAK, P. et ASTON-JONES, G., 1994. Locus coeruleus activity in monkey : Phasic and tonic changes are associated with altered vigilance. *Brain Research Bulletin*, 35(5-6) :607–616. doi :10.1016/0361-9230(94)90175-9.
URL [http://dx.doi.org/10.1016/0361-9230\(94\)90175-9](http://dx.doi.org/10.1016/0361-9230(94)90175-9)
- RENDEIRO, Diogo, SACRAMENTO, João et WICHERT, Andreas, 2014. Taxonomical associative memory. 6(1) :45–65. doi :10.1007/s12559-012-9198-4.
URL <http://dx.doi.org/10.1007/s12559-012-9198-4>
- REPA, J. Christopher, MULLER, Jeff, APERGIS, John, DESROCHERS, Theresa M., ZHOU, Yu et LEDOUX, Joseph E., 2001. Two different lateral amygdala cell populations contribute to the initiation and storage of memory. *Nat Neurosci*, 4(7) :724–731. doi :10.1038/89512.
- RESCORLA, R.A. et WAGNER, A.R., 1972. A theory of pavlovian conditioning : Variations in the effectiveness of reinforcement and nonreinforcement. Dans *Classical Conditioning II : Current Research and Theory*, pages 64–99. Appleton Century Crofts.

- RESCORLA, Robert A., 1988. Pavlovian conditioning. it's not what you think it is. Dans *American Psychologist*, pages 151–160.
URL <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.294.1082>
- RICHTER-LEVIN, Gal, 2004. The amygdala, the hippocampus, and emotional modulation of memory. *The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry*, 10(1) :31–39. doi :10.1177/1073858403259955.
URL <http://dx.doi.org/10.1177/1073858403259955>
- ROELFSEMA, Pieter R. et VAN OOYEN, Arjen, 2005. Attention-gated reinforcement learning of internal representations for classification. *Neural Computation*, 17 :2176–2214.
- ROELFSEMA, Pieter R., VAN OOYEN, Arjen et WATANABE, Takeo, 2010a. Perceptual learning rules based on reinforcers and attention. *Trends in Cognitive Sciences*, 14(2) :64–71. doi :10.1016/j.tics.2009.11.005.
URL <http://dx.doi.org/10.1016/j.tics.2009.11.005>
- ROELFSEMA, Pieter R., VAN OOYEN, Arjen et WATANABE, Takeo, 2010b. Perceptual learning rules based on reinforcers and attention. *Trends in Cognitive Sciences*, 14(2) :64–71. doi :10.1016/j.tics.2009.11.005.
URL <http://dx.doi.org/10.1016/j.tics.2009.11.005>
- ROESCH, Matthew R. et OLSON, Carl R., 2004. Neuronal activity related to reward value and motivation in primate frontal cortex. *Science (New York, N. Y.)*, 304(5668) :307–310. doi :10.1126/science.1093223.
URL <http://dx.doi.org/10.1126/science.1093223>
- ROLLS, Edmund T., 2004a. Convergence of sensory systems in the orbitofrontal cortex in primates and brain design for emotion. *The anatomical record. Part A, Discoveries in molecular, cellular, and evolutionary biology*, 281(1) :1212–1225.
URL <http://view.ncbi.nlm.nih.gov/pubmed/15470678>
- ROLLS, Edmund T., 2004b. The functions of the orbitofrontal cortex. *Brain and cognition*, 55(1) :11–29. doi :10.1016/s0278-2626(03)00277-x.
URL [http://dx.doi.org/10.1016/s0278-2626\(03\)00277-x](http://dx.doi.org/10.1016/s0278-2626(03)00277-x)
- ROUGIER, Nicolas P. et FIX, Jérémy, 2012. DANA : Distributed (asynchronous) Numerical and Adaptive modelling framework. *Network : Computation in Neural Systems*, 23(4) :237–253. doi :10.3109/0954898X.2012.721573.
- RUMELHART, David E., HINTON, Geoffrey E. et WILLIAMS, Ronald J., 1986. Learning representations by back-propagating errors. *Nature*,

BIBLIOGRAPHIE

323(6088) :533–536. doi :10.1038/323533a0.

URL <http://dx.doi.org/10.1038/323533a0>

SADACCA, Brian F., WIKENHEISER, Andrew M. et SCHOENBAUM, Geoffrey, 2016. Toward a theoretical role for tonic norepinephrine in the orbitofrontal cortex in facilitating flexible learning. *Neuroscience*.

SADDORIS, Michael P., GALLAGHER, Michela et SCHOENBAUM, Geoffrey, 2005. Rapid associative encoding in basolateral amygdala depends on connections with orbitofrontal cortex. *Neuron*, 46(2) :321–331.

URL <http://view.ncbi.nlm.nih.gov/pubmed/15848809>

SARA, Susan J. et BOURET, Sebastien, 2012. Orienting and Reorienting : The Locus Coeruleus Mediates Cognition through Arousal. *Neuron*, 76(1) :130–141.

SCHMAJUK, N. A. et DICARLO, J. J., 1992. Stimulus configuration, classical conditioning, and hippocampal function. *Psychological review*, 99(2) :268–305.

SCHMAJUK, Nestor A. et DICARLO, James J., 1992. Stimulus configuration, classical conditioning, and hippocampal function. Dans *Psychological Review*, pages 268–305.

URL <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.124.5397>

SCHMAJUK, NestorA et MOORE, JohnW, 1985. Real-time attentional models for classical conditioning and the hippocampus. 13(4) :278–290. doi :10.3758/bf03326534.

URL <http://dx.doi.org/10.3758/bf03326534>

SCHOENBAUM, G., CHIBA, A. A. et GALLAGHER, M., 1999. Neural encoding in orbitofrontal cortex and basolateral amygdala during olfactory discrimination learning. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 19(5) :1876–1884.

URL <http://view.ncbi.nlm.nih.gov/pubmed/10024371>

SCHOENBAUM, Geoffrey, SETLOW, Barry, SADDORIS, Michael P. et GALLAGHER, Michela, 2003. Encoding predicted outcome and acquired value in orbitofrontal cortex during cue sampling depends upon input from basolateral amygdala. *Neuron*, 39(5) :855–867.

URL <http://view.ncbi.nlm.nih.gov/pubmed/12948451>

SCHOUPS, A., VOGELS, R., QIAN, N. et ORBAN, G., 2001. Practising orientation identification improves orientation coding in v1 neurons. *Nature*, 412(6846) :549–553. doi :10.1038/35087601.

URL <http://dx.doi.org/10.1038/35087601>

- SCHROEDER, Jason P. et PACKARD, Mark G., 2004. Facilitation of memory for extinction of drug-induced conditioned reward : role of amygdala and acetylcholine. *Learning & memory (Cold Spring Harbor, N.Y.)*, 11(5) :641–647.
- SCHROLL, Henning, VITAY, Julien et HAMKER, Fred H., 2012. Working memory and response selection : a computational account of interactions among cortico-basalganglio-thalamic loops. *Neural networks : the official journal of the International Neural Network Society*, 26 :59–74. doi : 10.1016/j.neunet.2011.10.008.
URL <http://dx.doi.org/10.1016/j.neunet.2011.10.008>
- SCHULTZ, W., DAYAN, P. et MONTAGUE, P. R., 1997a. A neural substrate of prediction and reward. *Science (New York, N.Y.)*, 275(5306) :1593–1599.
URL <http://view.ncbi.nlm.nih.gov/pubmed/9054347>
- SCHULTZ, W., TREMBLAY, L. et HOLLERMAN, J. R., 2000. Reward processing in primate orbitofrontal cortex and basal ganglia. *Cerebral cortex (New York, N.Y. : 1991)*, 10(3) :272–284. doi :10.1093/cercor/10.3.272.
URL <http://dx.doi.org/10.1093/cercor/10.3.272>
- SCHULTZ, Wolfram, DAYAN, Peter et MONTAGUE, P. Read, 1997b. A neural substrate of prediction and reward. *Science*, 275(5306) :1593–1599. doi : 10.1126/science.275.5306.1593.
URL <http://dx.doi.org/10.1126/science.275.5306.1593>
- SETLOW, Barry, SCHOENBAUM, Geoffrey et GALLAGHER, Michela, 2003. Neural encoding in ventral striatum during olfactory discrimination learning. *Neuron*, 38(4) :625–636.
URL <http://view.ncbi.nlm.nih.gov/pubmed/12765613>
- SHAPIRO, M. L. et EICHENBAUM, H., 1999. Hippocampus as a memory map : synaptic plasticity and memory encoding by hippocampal neurons. *Hippocampus*, 9(4) :365–384.
URL <http://view.ncbi.nlm.nih.gov/pubmed/10495019>
- SIERRA-MERCADO, Demetrio, PADILLA-COREANO, Nancy et QUIRK, Gregory J., 2011. Dissociable roles of prelimbic and infralimbic cortices, ventral hippocampus, and basolateral amygdala in the expression and extinction of conditioned fear. *Neuropsychopharmacology*, 36(2) :529–538. doi : 10.1038/npp.2010.184.
- SIGALA, Natasha et LOGOTHETIS, Nikos K., 2002. Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature*, 415(6869) :318–320.
URL <http://view.ncbi.nlm.nih.gov/pubmed/11797008>

BIBLIOGRAPHIE

- SILVER, David, HUANG, Aja, MADDISON, Chris J., GUEZ, Arthur, SIFRE, Laurent, VAN DEN DRIESSCHE, George, SCHRITTWIESER, Julian, ANTONOGLOU, Ioannis, PANNEERSHELVAM, Veda, LANCTOT, Marc, DIELEMAN, Sander, GREWE, Dominik, NHAM, John, KALCHBRENNER, Nal, SUTSKEVER, Ilya, LILLICRAP, Timothy, LEACH, Madeleine, KAVUKCUOGLU, Korumay, GRAEPEL, Thore et HASSABIS, Demis, 2016. Mastering the game of go with deep neural networks and tree search. *Nature*, 529(7587) :484–489. doi :10.1038/nature16961.
URL <http://dx.doi.org/10.1038/nature16961>
- SOTRES-BAYON, Francisco, SIERRA-MERCADO, Demetrio, PARDILLA-DELGADO, Enmanuelle et QUIRK, Gregory J., 2012. Gating of fear in pre- limbic cortex by hippocampal and amygdala inputs. *Neuron*, 76(4) :804–812. URL <http://view.ncbi.nlm.nih.gov/pubmed/23177964>
- STEINMETZ, Joseph E., 2000. Brain substrates of classical eyeblink conditioning : a highly localized but also distributed system. *Behavioural Brain Research*, 110(1-2) :13–24.
- STEVENS, Francis L., HURLEY, Robin A. et TABER, Katherine H., 2011. Anterior cingulate cortex : unique role in cognition and emotion. *The Journal of neuropsychiatry and clinical neurosciences*, 23(2) :121–125. URL <http://view.ncbi.nlm.nih.gov/pubmed/21677237>
- SUTTON, R. S. et BARTO, A. G., 1981. Toward a modern theory of adaptive networks : expectation and prediction. *Psychological review*, 88(2) :135–170. URL <http://view.ncbi.nlm.nih.gov/pubmed/7291377>
- SUTTON, R. S. et BARTO, A. G., 1998. *Reinforcement Learning : An introduction*. MIT Press.
- SUZUKI, Wendy A., 1996. The anatomy, physiology and functions of the perirhinal cortex. *Current Opinion in Neurobiology*, 6(2) :179–186. doi : 10.1016/s0959-4388(96)80071-7. URL [http://dx.doi.org/10.1016/s0959-4388\(96\)80071-7](http://dx.doi.org/10.1016/s0959-4388(96)80071-7)
- SWINDALE, N. V., 1996. The development of topography in the visual cortex : a review of models. *Network (Bristol, England)*, 7(2) :161–247. doi :10.1088/0954-898x/7/2/002. URL <http://dx.doi.org/10.1088/0954-898x/7/2/002>
- TAIT, David S., BROWN, Verity J., FAROVIK, Anja, THEOBALD, David E., DALLEY, Jeffrey W. et ROBBINS, Trevor W., 2007. Lesions of the dorsal noradrenergic bundle impair attentional set-shifting in the rat. *The European journal of neuroscience*, 25(12) :3719–3724. URL <http://view.ncbi.nlm.nih.gov/pubmed/17610591>

- TANAKA, Saori C., DOYA, Kenji, OKADA, Go, UEDA, Kazutaka, OKAMOTO, Yasumasa et YAMAWAKI, Shigeto, 2004. Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nature Neuroscience*, 7(8) :887–893. doi :10.1038/nn1279.
URL <http://dx.doi.org/10.1038/nn1279>
- TANAKA, Saori C., SCHWEIGHOFER, Nicolas, ASAH, Shuji, SHISHIDA, Kazuhiro, OKAMOTO, Yasumasa, YAMAWAKI, Shigeto et DOYA, Kenji, 2007. Serotonin differentially regulates short- and long-term prediction of rewards in the ventral and dorsal striatum. *PloS one*, 2(12).
URL <http://view.ncbi.nlm.nih.gov/pubmed/18091999>
- THORPE, S. J. et FABRE-THORPE, M., 2001. Neuroscience. seeking categories in the brain. *Science (New York, N.Y.)*, 291(5502) :260–263. doi :10.1126/science.1058249.
URL <http://dx.doi.org/10.1126/science.1058249>
- TOPALIDOU, Meropi, LEBLOIS, Arthur, BORAUD, Thomas et ROUGIER, Nicolas P., 2015. A long journey into reproducible computational neuroscience. *Frontiers in computational neuroscience*, 9.
URL <http://view.ncbi.nlm.nih.gov/pubmed/25798104>
- TREVES, A. et ROLLS, E. T., 1994. Computational analysis of the role of the hippocampus in memory. *Hippocampus*, 4(3) :374–391. doi :10.1002/hipo.450040319.
URL <http://dx.doi.org/10.1002/hipo.450040319>
- TRICOMI, Elizabeth, BALLEINE, Bernard W. et O'DOHERTY, John P., 2009. A specific role for posterior dorsolateral striatum in human habit learning. *The European journal of neuroscience*, 29(11) :2225–2232. doi :10.1111/j.1460-9568.2009.06796.x.
URL <http://dx.doi.org/10.1111/j.1460-9568.2009.06796.x>
- TULVING, Endel, 2002. Episodic memory : From mind to brain. *Annual Review of Psychology*, 53(1) :1–25. doi :10.1146/annurev.psych.53.100901.135114.
URL <http://dx.doi.org/10.1146/annurev.psych.53.100901.135114>
- UNAL, Cagri T., PARE, Denis et ZABORSZKY, Laszlo, 2015. Impact of Basal Forebrain Cholinergic Inputs on Basolateral Amygdala Neurons. *The Journal of Neuroscience*, 35(2) :853–863.
- VANRULLEN, Rufin et KOCH, Christof, 2005. *Visual attention and visual awareness.*, tome 1, chapitre 3.
- VITAY, Julien, 2006. *Emergence de fonctions sensorimotrices sur un substrat neuronal numérique distribué.* Thèse de doctorat, Université Henri Poincaré - Nancy 1.

BIBLIOGRAPHIE

- VIZI, E. S., 1980. Modulation of cortical release of acetylcholine by noradrenaline released from nerves arising from the rat locus coeruleus. *Neuroscience*, 5(12) :2139–2144. doi :10.1016/0306-4522(80)90129-3.
URL [http://dx.doi.org/10.1016/0306-4522\(80\)90129-3](http://dx.doi.org/10.1016/0306-4522(80)90129-3)
- VLACHOS, Ioannis, HERRY, Cyril, LÜTHI, Andreas, AERTSEN, Ad et KUMAR, Arvind, 2011. Context-dependent encoding of fear and extinction memories in a large-scale network model of the basal amygdala. *PLoS computational biology*, 7(3) :e1001104+. doi :10.1371/journal.pcbi.1001104.
- WALLENSTEIN, G. V. et HASSELMO, M. E., 1997. Gabaergic modulation of hippocampal population activity : sequence learning, place field development, and the phase precession effect. *Journal of neurophysiology*, 78(1) :393–408.
URL <http://jn.physiology.org/cgi/content/abstract/78/1/393>
- WALLIS, Jonathan D., 2007. Orbitofrontal cortex and its contribution to decision-making. *Annual review of neuroscience*, 30(1) :31–56. doi : 10.1146/annurev.neuro.30.051606.094334.
- WANG, Qian, JIN, Jingji et MAREN, Stephen, 2016. Renewal of extinguished fear activates ventral hippocampal neurons projecting to the prelimbic and infralimbic cortices in rats. *Neurobiology of learning and memory*.
URL <http://view.ncbi.nlm.nih.gov/pubmed/27060752>
- WHITE, Norman M., 2009. Some highlights of research on the effects of caudate nucleus lesions over the past 200 years. *Behavioural brain research*, 199(1) :3–23.
URL <http://view.ncbi.nlm.nih.gov/pubmed/19111791>
- WISE, Roy A., 2004. Dopamine, learning and motivation. *Nature Reviews Neuroscience*, 5(6) :483–494. doi :10.1038/nrn1406.
URL <http://dx.doi.org/10.1038/nrn1406>
- YIN, Henry H., MULCARE, Shweta Prasad P., HILÁRIO, Monica R., CLOUSE, Emily, HOLLOWAY, Terrell, DAVIS, Margaret I., HANSSON, Anita C., LOVINGER, David M. et COSTA, Rui M., 2009. Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. *Nature neuroscience*, 12(3) :333–341.
URL <http://view.ncbi.nlm.nih.gov/pubmed/19198605>
- YU, A. J. et DAYAN, P., 2005. Uncertainty, Neuromodulation, and Attention. *Neuron*, 46.
- ZIPSER, D. et RUMELHART, D. E., 1990. Neurobiological significance of new learning models. Dans E. Schwartz, rédacteur, *Computational neuroscience*. Cambridge, MA : MIT Press.