THÈSE DE DOCTORAT DE
l'UNIVERSITÉ PIERRE ET MARIE CURIE

Préparée à
L'École Normale Supérieure

École doctorale Cerveau, Cognition, Comportement

Présentée par

Muriel EKOVICH

Pour obtenir le grade de
DOCTEUR
de l'UNIVERSITÉ PIERRE ET MARIE CURIE

Contrôle exécutif et mémoire

Inférences probabilistes et récupérations mnésiques dans les lobes frontaux

---

soutenue le 28 octobre 2014

Devant le jury composé de :

Pr. Laurent COHEN	Président du jury
Pr. Etienne KŒCHLIN	Directeur de thèse
Pr. Anne-Lise GIRAUD	Rapporteur
Pr. Simon THORPE	Rapporteur
Pr. Emmanuel PROCYK	Examinateur
Dr. Christopher SUMMERFIELD	Examinateur

"La vie est l'adaptation continue de relations internes à des relations externes"
Herbert Spencer

Remerciements

Nous sommes à l'automne 2008. Il est 8h du matin. J'ai rendez-vous avec Etienne Kœchlin au Laboratoire de Neurosciences Cognitives, fraîchement installé rue d'Ulm. C'est ici je souhaite effectuer mon stage de master 1. Je me dis qu'un rendez-vous à une heure aussi matinale, ce doit être pour tester ma motivation. Quoi qu'il en soit, six ans plus tard, je suis assise au même poste de travail qu'Etienne m'a octroyé non seulement pour mon stage de master 1, mais aussi pour mon stage de master 2 et mon doctorat. Je tiens donc à remercier Etienne pour ces années qui comptent parmi les plus belles. Non seulement pour l'aventure scientifique passionnante dans laquelle il a accepté de me guider, mais aussi pour la liberté qu'il m'a laissée afin de m'épanouir tant intellectuellement qu'humainement, entre les murs du laboratoire et au-delà.

Je tiens également à remercier les rapporteurs et examinateurs de mon jury qui ont accepté d'évaluer ce travail : Anne-Lise Giraud, Simon Thorpe, Laurent Cohen, Emmanuel Procyk et Christopher Summerfield.

Je suis par ailleurs reconnaissante envers la Direction Générale de l'Armement qui a accepté de m'allouer une bourse de thèse pour trois années.

Je souhaite remercier toute l'équipe du Centre de Neuroimagerie de Recherche de l'hôpital de la Pitié-Salpêtrière qui m'a permis d'acquérir mes données dans une atmosphère toujours conviviale.

Des remerciements particuliers vont à Alizée Lopez qui a grandement contribué à l'acquisition des données, ainsi qu'à Jan Drugowitsch qui m'a guidée dans l'étape cruciale d'ajustement des paramètres du modèle computationnel.

Mes remerciements vont bien sûr à l'ensemble des membres, passés et présents, de l'équipe *Frontal Lobe Function Group*, qui ont toujours su se rendre disponible pour répondre à mes questions et créer un environnement de travail à la fois stimulant et décontracté : Anne, Gabriel, Héloïse, Maël, Marion ("bestie"), Philippe, Stefano, Sven, Sylvain, Valentin, Valérian, Vincent.

Je tiens également à remercier l'ensemble des membres du Laboratoire de Neurosciences Cognitives, en particulier Anne-Do, Laura et Marine sans qui nous perdrions la tête.

Je tiens aussi à remercier l'ensemble des participants qui ont accepté de me prêter leur cerveau le temps d'une expérience, en particulier ceux de mon entourage plus ou moins proche : Aurélie, Baptiste, Barbara, Nathalie, Nathaniel, Raphaël, Romain, Stéphanie.

Un grand merci à mon "comité de lecture" dont les membres ont accepté de relire certains passages de cette thèse : Anne, Cécile (ma lectrice la plus assidue !), Elie, Gabriel, Héloïse, Marion, Marwa, Philippe, Stéphane, Sylvain.

Je remercie infiniment mes parents pour leur soutien et leur amour inconditionnels (et aussi pour m'avoir logée, nourrie et blanchie pendant toutes ces années). Je remercie en particulier ma maman qui a accepté de traquer les ir-

réductibles fautes d'orthographe et autre tournures syntaxiques bancales dans une bonne partie du manuscrit. Merci aussi à mon "petit" frère Nathaniel qui a notamment accepté d'être l'un de mes "cobayes".

De tendres "youpi" remerciements vont également à JM pour sa patience, son écoute et son soutien au quotidien.

Enfin, je tiens à remercier tous ceux qui, en dehors du laboratoire, ont contribué à faire de ces quatre dernières années une tranche de vie riche en projets artistiques, en instants scéniques mémorables et en éclats de rire : *Les Évadés de la rue Barrow* (Julien, Manu, Marie, Pauline, Rémi, Sophie, Stéphanie, Tomô, Yann), *Smoking Sofa* (Garo, Guigui, JM, Pipo, Tatiana) et le club de théâtre anglais de Télécom ParisTech (Adrien, Dan, Elie, Fernando, Omar, Tatiana).

Résumé

Les humains doivent sans cesse s'adapter à des environnements incertains, changeants et ouverts. Pour ce faire, il est nécessaire d'explorer, d'ajuster et d'exploiter plusieurs *task-sets*, c'est-à-dire des associations sensorimotrices flexibles entre des stimuli, des actions et des résultats.

Collins et Kœchlin ont proposé un modèle computationnel qui contrôle la création, l'apprentissage, le stockage, la récupération et la sélection de tels *task-sets* pour guider le comportement [54]. Le modèle met à jour les fiabilités d'un ensemble de *task-sets*, c'est-à-dire leur capacité à prédire correctement les conséquences des actions. Cette mise à jour des fiabilités est réalisée par inférence bayésienne en intégrant les résultats des actions et l'information contextuelle. Elle permet d'arbitrer entre l'exploitation d'un *task-set* fiable et l'exploration de nouveaux *task-sets*. En l'absence d'information contextuelle, la fiabilité des *task-sets* appris est maintenue et mise à jour dans le cortex frontopolaire [74].

Ce travail de thèse a pour but d'élucider les mécanismes neuraux permettant le monitorage et la récupération de *task-sets* stockés en mémoire, que ce soit selon les indices contextuels ou les résultats des actions.

Pour ce faire, nous avons élaboré une expérience d'IRM fonctionnelle dans laquelle des sujets sains doivent apprendre, par essais et erreurs, plusieurs *task-sets* associés à divers indices contextuels, puis réutiliser les *task-sets* appris. Les conditions expérimentales varient de façon non prédictible de telle sorte que :

(i) des *task-sets* précédemment appris doivent être réutilisés soit avec le même indice contextuel, soit avec un indice contextuel nouveau ;

(ii) de nouveaux *task-sets* doivent être appris soit avec de nouveaux indices contextuels, soit avec des indices connus.

Les résultats comportementaux montrent que les sujets apprennent, maintiennent, mettent à jour et réutilisent un répertoire grandissant de *task-sets* tel que le prédit le modèle computationnel. Plus précisément :

(i) les indices contextuels connus sont utilisés de façon proactive (avant l'exécution d'une action) pour sélectionner le *task-set* correspondant ;

(ii) lorsque les indices contextuels sont inconnus, les sujets sélectionnent le *task-set* de façon réactive, selon les résultats des actions.

Les résultats d'imagerie révèlent que la fiabilité des *task-sets* est mise à jour dans le cortex frontopolaire et le cortex préfrontal dorso-latéral. De plus, le cortex préfrontal latéral est impliqué dans le processus de sélection proactif et réactif. Cependant, les récupérations proactives et réactives dépendent de réseaux distincts :

(i) d'une part la récupération proactive repose sur un réseau ventral qui inclut le cortex préfrontal ventro-médian, le striatum et l'hippocampe ;

(ii) d'autre part, la récupération réactive repose sur un réseau frontal incluant notamment le cortex frontopolaire.

Abstract

Humans need to adapt to uncertain, changing, and open-ended environments. In such situations, decision-making involves exploring, adjusting and exploiting multiple task-sets – defined as flexible sensorimotor mappings associating stimuli, actions, and expected outcomes.

Collins and Kœchlin proposed a computational model that controls the creation, learning, storage, retrieval, and selection of such task-sets for driving action [54]. The model monitors the reliability of a collection of concurrent task-sets - i.e., the ability of alternative task-sets to correctly predict action outcomes. Task-set reliability is updated in a Bayesian manner according to outcomes and contextual information and arbitrates between exploiting the most reliable task-set or exploring new ones to drive action. It has recently been shown that, without contextual information, the reliability of alternative learned task-sets is monitored in frontopolar cortex [74].

The goal of this study is to investigate the neural mechanisms that subserve the monitoring and retrieval of stored task-sets according to both contextual cues and action outcomes.

We designed an fMRI experiment requiring healthy human subjects to learn by trials and errors and to switch between multiple task-sets associated with various contextual cues. Experimental conditions varied unpredictably such that :

(i) previously learned task-sets re-occurred with either the same or new contextual cues,

(ii) new task-sets that needed to be learned occurred with new cues or previously encountered ones.

Behavioral results and model fits show that subjects learned, monitored and switched across an expanding repertoire of task-sets as predicted by the computational model. More specifically :

(i) Known contextual cues were used proactively (before performing any action) to select the corresponding task-set ;

(ii) When previously learned task-sets re-occurred with unknown contextual cues, subjects selected the stored task-set reactively based on action outcomes.

Model-based fMRI results revealed that task-set reliability is updated in the frontopolar cortex and the lateral prefrontal cortex. Moreover, lateral prefrontal cortex is engaged in the selection process in both cases. However distinct networks are involved depending on whether the retrieval is cue or outcome-based :

(i) on the one hand, proactive retrieval relies on a ventral pathway including ventromedial prefrontal cortex, striatum and bilateral hippocampus ;

(ii) on the other hand, reactive retrieval relies on a frontal network including frontopolar cortex.

Table des matières

Introduction **15**

I Approche théorique des processus d'inférence et de récupération 17

1 Les *task-sets* et le *task-switching* **19**
1.1 Le concept de *task-set* 19
1.2 Le *task-switching* 22
1.2.1 Protocoles de *task-switching* 22
1.2.2 Le coût de *switch* 24

2 Processus d'inférence du système exécutif **27**
2.1 De la logique formelle à l'inférence bayésienne 27
2.1.1 Les humains sont-ils rationnels ? 27
2.1.2 L'approche probabiliste 29
2.1.3 L'inférence bayésienne 30
2.2 Modèles computationnels d'inférence 31
2.2.1 Les modèles de monitorage de l'incertitude 31
2.2.2 Le modèle PROBE 35

3 Processus de récupération en mémoire **45**
3.1 La mémoire de travail 47
3.1.1 Définition 47
3.1.2 Théories psychologiques de la mémoire de travail 47
3.1.3 Capacité de la mémoire de travail 50
3.2 Théorie psychologique de la récupération 50
3.2.1 L'activation en mémoire à long terme 50
3.2.2 La région d'accès direct et le pont 51
3.2.3 Le focus attentionnel et le focus de la réponse 52
3.3 Un modèle computationnel de la récupération en mémoire 54
3.3.1 Module de sélection de la réponse 54
3.3.2 Module de sélection d'un *task-set* 56

II Bases cérébrales des processus d'inférence et de récupération 59

4 Les lobes frontaux et les noyaux sous-corticaux 61
4.1 Les lobes frontaux . . . 62
4.1.1 Un peu d'histoire . . . 62
4.1.2 Anatomie générale du lobe frontal . . . 62
4.1.3 Le cortex préfrontal . . . 64
4.1.4 Les cortex prémoteur et moteur . . . 67
4.2 Les noyaux sous-corticaux . . . 67
4.2.1 Anatomie . . . 67
4.2.2 Rôles fonctionnels . . . 68

5 Bases cérébrales de l'inférence 71
5.1 Cortex préfrontal ventromédian et prédictions . . . 71
5.1.1 Etudes de neuropsychologie . . . 71
5.1.2 Etudes de neuroimagerie . . . 72
5.2 Cortex frontopolaire et monitorage des stratégies alternatives . . . 74
5.2.1 Etudes de neuropsychologie . . . 75
5.2.2 Etudes de neuroimagerie . . . 75

6 Bases cérébrales du *switch* 77
6.1 Cortex cingulaire antérieur . . . 77
6.1.1 La détection d'erreurs . . . 77
6.1.2 La détection du conflit . . . 78
6.1.3 Historique des récompenses et prédictions futures . . . 79
6.1.4 La volatilité . . . 80
6.1.5 Le *switch* . . . 80
6.1.6 Une vision unifiée . . . 82
6.2 Le cortex préfrontal dorso-latéral . . . 83
6.2.1 Etudes de neuropsychologie . . . 83
6.2.2 Etudes de neuroimagerie . . . 83
6.2.3 Etudes d'électrophysiologie . . . 87

7 Bases cérébrales de la récupération 89
7.1 Cortex préfrontal et mémoire de travail . . . 89
7.1.1 Etudes chez le primate non humain . . . 90
7.1.2 Etudes de neuroimagerie . . . 92
7.2 Interactions entre le cortex préfrontal et le lobe temporal médian pour la récupération associative . . . 94
7.2.1 Stockage en mémoire à long terme dans le cortex . . . 94
7.2.2 Rôle du lobe temporal médian . . . 96
7.2.3 Rôle du cortex préfrontal . . . 98
7.2.4 Une vision unifiée . . . 100

III Question de thèse **105**

IV Méthodologie **109**

8 Protocole expérimental **111**
8.1 Contraintes générales . . . 111
8.1.1 Prédictions du modèle PROBE . . . 111
8.1.2 Contraintes techniques . . . 113
8.2 Structure des essais . . . 114
8.2.1 Stimulus . . . 114
8.2.2 Temps de réponse et intervalle inter-stimulus . . . 115
8.2.3 Feedback . . . 115
8.3 Structure des Sessions . . . 117
8.3.1 La Session d'Apprentissage . . . 117
8.3.2 La Session Test . . . 119
8.4 Participants . . . 123
8.4.1 Expérience comportementale pilote . . . 123
8.4.2 Expérience en IRM . . . 123

9 Ajustement des paramètres du modèle sur les données comportementales **125**
9.1 Principe et méthodes d'ajustement de paramètres (ou *fitting*) . . . 125
9.1.1 Principe général . . . 125
9.1.2 Critères d'ajustements . . . 126
9.1.3 Tirage des paramètres . . . 129
9.2 Ajustement des paramètres du modèle PROBE aux données comportementales . 130
9.2.1 Choix des paramètres à ajuster . . . 130
9.2.2 Méthode d'ajustement des paramètres du modèle PROBE . . . 132

10 L'Imagerie par Résonance Magnétique fonctionnelle (IRMf) **133**
10.1 Principe et acquisition des données . . . 133
10.1.1 Un bref historique . . . 133
10.1.2 Le signal *BOLD* (*Blood Oxygen Level Dependant*) . . . 134
10.1.3 Paramètres d'acquisition . . . 134
10.2 Prétraitement du signal (*preprocessing*) . . . 135
10.2.1 La correction du décalage d'acquisition entre coupes (*slice timing*) . . . 135
10.2.2 La correction du mouvement (*realign*) . . . 135
10.2.3 Le recalage (*coregister*) . . . 136
10.2.4 La normalisation spatiale (*normalize*) . . . 136
10.2.5 Le lissage spatial (*smoothing*) . . . 136
10.3 L'approche *model-based* . . . 137
10.3.1 Principe de la méthode *model-based* . . . 137
10.3.2 Etapes de la méthode . . . 137
10.3.3 Un exemple historique : erreur de prédiction et dopamine . . . 138
10.3.4 Limites de la méthode . . . 142
10.4 Modèle linéaire général . . . 142

10.4.1 Spécification du modèle . . . 142
10.4.2 Estimation des paramètres . . . 147
10.4.3 Définition et estimation de contrastes . . . 148
10.4.4 Analyses de groupe . . . 148
10.5 Analyse des régions d'intérêt . . . 149
10.6 Histogrammes péri-stimulus . . . 149
10.6.1 Réponse en impulsions finies (*Finite Impulse Response, FIR*) . . . 149
10.6.2 Résidus . . . 150

V Résultats 151

11 Résultats comportementaux 153
11.1 Classification des essais . . . 153
11.2 Comportement des sujets . . . 154
11.2.1 Session d'apprentissage . . . 154
11.2.2 Session test . . . 157
11.3 Comportement du modèle PROBE . . . 158
11.3.1 Qualité de l'ajustement du modèle sur les données comportementales . . . 158
11.3.2 Points de transition prédits par le modèle . . . 164

12 Résultats d'imagerie 171
12.1 Les signaux de fiabilité . . . 171
12.1.1 Signal de fiabilité du *task-set* acteur et cortex préfrontal ventro-médian . . . 172
12.1.2 Signaux de fiabilité *ex-post*, cortex frontopolaire et préfrontal dorso-latéral 172
12.1.3 Signaux de fiabilité *ex-ante* et cortex préfrontal dorso-latéral . . . 173
12.2 Abandon du *task-set* acteur et pré-SMA . . . 174
12.3 Récupération et cortex préfrontal dorso-latéral . . . 176
12.4 Spécificités des récupérations proactives et réactives . . . 178
12.4.1 Récupération réactive et cortex frontopolaire . . . 178
12.4.2 Récupération proactive : hippocampe, striatum ventral et cortex préfrontal ventro-médian . . . 179
12.5 Consolidation : amygdale et striatum . . . 182

VI Discussion 185

13 Dissociation des processus réactifs et proactifs dans les lobes frontaux 187
13.1 Manifestations comportementales de l'apprentissage et de la récupération de *task-sets* . . . 187
13.2 Réseaux cérébraux des processus d'inférences . . . 188
13.2.1 Réseau médian de l'inférence sur le *task-set* acteur . . . 188
13.2.2 Réseau latéral de l'inférence sur les *task-sets* alternatifs . . . 189
13.3 Réseaux cérébraux des processus de récupération . . . 190
13.3.1 Récupération réactive . . . 190
13.3.2 Récupération proactive . . . 191

14 Limites et perspectives **195**
14.1 Limites du modèle PROBE et du protocole expérimental 195
14.1.1 Schéma de renforcement 195
14.1.2 Modélisation de la mémoire 196
14.1.3 Temporalité des évènements 196
14.2 Neurones et neuromodulateurs 197
14.2.1 Implémentation neuronale 197
14.2.2 Neuromodulateurs 197
14.3 Des processus spécifiques à l'humain ? 198

Conclusion **201**

Annexes **203**

A Les contextes **205**

B Les instructions **207**

C Le débriefing informel **211**

Introduction

Les systèmes biologiques, du plus simple au plus complexe, captivent par leur capacité de s'adapter à l'environnement. N'est-il pas fascinant d'observer les êtres humains évoluer dans un monde aussi complexe que le notre ? Notre espèce semble en effet posséder une incroyable capacité d'adaptation. Afin d'assurer notre survie, l'évolution a façonné notre cerveau pour nous permettre de résoudre des tâches toujours plus complexes. Les fonctions exécutives ont grandement évolué chez le primate pour atteindre un niveau inégalé chez les humains. Ainsi, nous n'agissons pas seulement en réaction aux stimuli extérieurs, mais aussi en fonction de nos objectifs, de nos croyances, du contexte immédiat et de notre passé, avec une grande flexibilité.

Lorsque nous sommes confrontés à une nouvelle situation, nous développons une stratégie comportementale adaptée. Nous avons tout intérêt à mémoriser cette stratégie pour la réutiliser lorsque nous sommes à nouveau confrontés à cette même situation. Cette capacité de stockage et de réutilisation de stratégies précédemment apprises confère un avantage indéniable dans des environnements où des situations sont récurrentes. Cependant, dans des environnements où de nouvelles situations peuvent apparaître, il est nécessaire de choisir entre l'exploitation et l'ajustement des stratégies précédemment apprises, ou l'exploration et l'apprentissage de nouvelles stratégies. Ce dilemme représente un problème complexe mais le cerveau humain semble avoir évolué pour le résoudre.

Choisir entre ces deux options nécessite dans un premier temps d'inférer que les contingences externes ont changé et de quitter la stratégie courante. Par exemple, lorsque les conséquences de nos choix ne correspondent pas à ce que prédit notre stratégie courante, on peut légitimement supposer que les contingences ne sont plus les mêmes. Une telle inférence est dite "réactive" au sens où elle se fonde uniquement sur les résultats de nos actions, donc après avoir agi. Ce processus d'inférence réactive peut toutefois s'avérer inefficace dans le cas de conséquences négatives, voire létales.

Ainsi, un second type d'inférence, fondé sur des indices externes, semble nécessaire. Contrairement aux conséquences de nos actions, les indices externes apparaissent indépendamment de notre comportement mais leur apparition peut informer de changements dans les contingences. Ce type d'inférence est dit "proactif" parce que les indices externes influencent nos choix d'agir.

Ces inférences réactives et proactives permettent uniquement de décider quand il est temps de quitter la stratégie courante. Il est cependant avantageux de savoir également quelle stratégie adopter ensuite. Pour cela il est nécessaire

de maintenir actives plusieurs stratégies alternatives et d'évaluer leur pertinence dans la situation du moment. Idéalement, il faudrait maintenir actif l'ensemble du répertoire de stratégies. Ceci semble toutefois coûteux et peu plausible biologiquement. Il est plus probable que seul un nombre limité de stratégies alternatives reste accessible.

Ainsi, lorsque la stratégie courante ne semble plus adaptée, le système exécutif peut choisir, de façon réactive ou proactive, entre deux options : (i) récupérer l'une des stratégies alternatives si celle-ci semble appropriée, ou (ii) décider d'en créer une nouvelle, pondérée par l'ensemble des stratégies expérimentées par le passé. Les processus d'inférence et de récupération sont donc nécessaires pour résoudre tout problème adaptatif. Ces processus portent sur des représentations abstraites des contingences externes. L'objectif de ce travail de thèse est d'identifier les bases cérébrales de ces processus afin d'en dégager les propriétés communes et les spécificités.

Dans une première partie, nous tâcherons de définir précisément ce que nous entendons par "stratégie comportementale" (ou *task-set*), "inférence" et "récupération". Nous délimiterons ainsi notre cadre théorique sous le prisme de la psychologie cognitive et de la modélisation mathématique. Nous présenterons en particulier le modèle computationnel sur lequel se fonde notre travail.

Dans une seconde partie, nous nous proposons de présenter les connaissances établies concernant les bases cérébrales qui sous-tendent ces types processus. Nous pourrons ainsi établir une série de prédictions quant aux régions cérébrales susceptibles d'émerger dans notre étude.

Nous décrirons ensuite la méthodologie que nous avons employée pour tester ces prédictions et présenterons l'ensemble des résultats que nous avons obtenus.

Enfin, ces résultats seront discutés et nous tenterons d'élargir notre étude à un champ plus vaste que celui de cette thèse.

Première partie

Approche théorique des processus d'inférence et de récupération

Chapitre 1

Les *task-sets* et le *task-switching*

1.1 Le concept de *task-set*

Dans la vie courante, nous sommes régulièrement confrontés à des situations qui nécessitent de sélectionner des actions simples, comme cliquer sur "Répondre" plutôt que "Effacer" lors de la réception d'un email. Mais nous sommes également confrontés à des situations plus complexes, comme le changement de mode opératoire lors du passage d'un Mac à un PC. Dans ces situations, nous faisons appel à des règles, ou *task-sets*, c'est-à-dire des associations adéquates entre stimuli et réponses dans un contexte donné.

Le concept de *task-set* est un concept établi de la psychologie cognitive. Dans leur publication de 1986, Norman et Shallice introduisent la notion comparable de "schéma" de contrôle du comportement [166]. Un schéma s'organise selon des séquences spécifiques d'actions, qu'elles soient "externes" (motrices) ou "internes" (mécanismes cognitifs seuls). Il existe plusieurs schémas en compétition pour guider le comportement et c'est le contrôle cognitif qui supervise la sélection d'un schéma selon les conditions. Norman et Shallice proposent que chaque schéma possède une "valeur d'activation" déterminée par une combinaison de facteurs internes et externes. Un schéma est sélectionné lorsque son niveau d'activation dépasse un seuil. C'est cette "valeur d'activation" qui est nécessaire dans le processus de sélection du schéma. Nous voyons ici l'émergence d'une brique hiérarchique du contrôle cognitif au dessus de la simple association stimulus-réponse.

En 2001, Miller et Cohen définissent plus précisément le *task-set* comme l'association de trois éléments : (i) des règles comportementales, (ii) des indices extérieurs qui déclenchent l'implémentation de ces règles et (iii) les récompenses attendues à la suite de l'exécution de ces règles. Un *task-set* est activé en réponse à des indices externes, et maintenu au cours du temps pour appliquer les règles comportementales correspondantes, aussi longtemps que les récompenses attendues qui lui sont associées sont suffisantes [159].

En 2003, Monsell définit le *task-set* (ou "schéma" procédural) comme la configuration appropriée des ressources mentales pour réaliser une tâche [162]. De même, en 2008, Sakai, considère un *task-set* comme une configuration de processus cognitifs activement maintenue pour accomplir une tâche (voir figure 1.1) [202].

Ainsi, on peut définir un *task-set* comme une configuration des ressources mentales qui, maintenue activement pour une tâche cognitive, permet d'effectuer celle-ci de manière appropriée. C'est donc une fonction permettant de décider de manière non équivoque des réponses à sélectionner dans le cadre d'une tâche cognitive. Le *task-set* représente donc un élément hiérarchique de base du système exécutif au-dessus de la simple association stimulus-réponse. Il correspond à l'objet d'étude de cette thèse.

Le passage d'une tâche à une autre, ou *task-switching*, suppose d'abandonner le *task-set* actif pour en sélectionner un autre. Ceci nécessite un certain contrôle cognitif. Dans la partie suivante, nous allons examiner plus précisément ce processus.

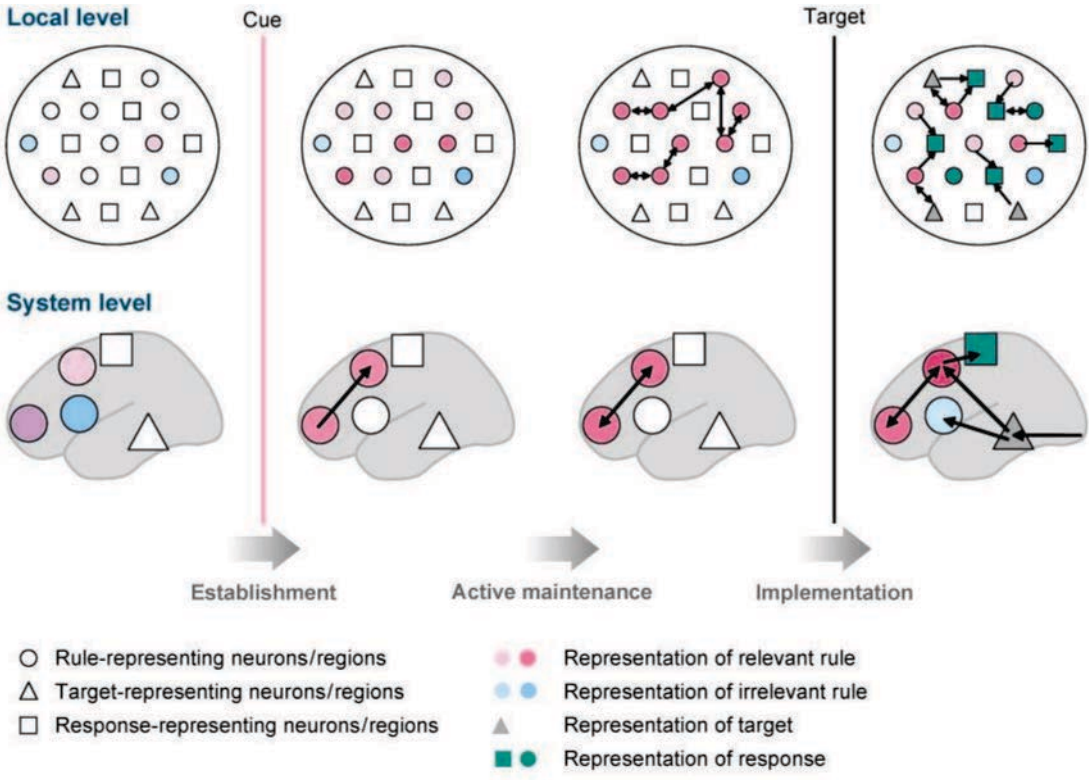


FIGURE 1.1 – **Le concept de *task-set*.** Schéma des mécanismes neuraux associés aux *task-sets*. **Haut** : au niveau local, la création, le maintien et l'implémentation d'un *task-set* s'effectuent via une modification de l'activité de neurones qui représentent la règle de la tâche, le stimulus et la réponse (indiqués respectivement par des cercles, des triangles et des carrés). Après la présentation d'un signal de début de tâche, l'activité et le nombre de neurones qui représentent la règle de la tâche augmentent (rose). L'activité neurale qui représente les autres tâches est toutefois maintenue (cercles bleus). Quand le stimulus est présenté, les neurones représentant la règle interagissent avec les neurones représentant le stimulus (triangles gris) et avec les neurones représentant une réponse donnée (carrés verts). Certains des neurones représentant la règle sont également sélectifs de la réponse (cercles verts). **Bas** : au niveau du système, les processus associés aux *task-sets* reposent sur des interactions entre différentes régions cérébrales. Les régions frontales postérieures représentent de façon sélective une tâche particulière (rose et bleu), et les régions préfrontales antérieures représentent les deux tâches (violet). Lorsque le signal d'exécution de la tâche est présenté, la région préfrontale antérieure interagit avec la région frontale postérieure correspondant à la tâche (rose). Les régions sensorielles de bas niveau impliquées dans le traitement du stimulus sont à leur tour activées (triangle gris). La région frontale postérieure correspondant à la tâche envoie ensuite un signal à la région associée à la réponse (carré vert). Reproduit de [202].

1.2 Le *task-switching*

« Should I stay or should I go now ?
If I go there will be trouble
And if I stay it may be double
So come on and let me know
Should I stay or should I go ? »
The Clash, Combat Rock, 1981

L'environnement est le plus souvent non stationnaire, il change avec le temps et il est nécessaire d'adapter son comportement en conséquence. Ceci suppose d'une part d'inhiber le *task-set* courant et d'autre part de s'engager dans un autre, connu ou nouveau. Expérimentalement, ceci correspond à un changement de tâche (ou *task-switching*). Le sujet effectue une tâche, et à tout moment il peut être amené à en changer. Deux cas peuvent se présenter :

- Un indice extérieur indique au sujet qu'il faut changer (et éventuellement quelle tâche il faut désormais effectuer).
- Les contingences changent sans indices extérieurs et le sujet doit déduire de ses observations qu'il faut changer de tâche.

1.2.1 Protocoles de *task-switching*

Les tâches utilisées dans les paradigmes de *task-switching* sont très variées, s'appuyant souvent sur des processus simples à apprendre. Par exemple : discriminer des nombres pairs et impairs, plus grands ou plus petits que 5, des paires identiques ou différentes, des voyelles ou consonnes, des minuscules ou des majuscules, des mots abstraits ou concrets, etc.

Task-switching fondé sur des indices externes

Dans les protocoles par bloc, une tâche est valide pendant une série d'essais, avant qu'une instruction n'indique le passage à une autre tâche pour une longue durée également. Ce type de protocole permet d'étudier le moment du *switch* mais aussi la période de maintien actif du *task-set* précédant un autre *switch*.

Il existe également des protocoles de type *task-cueing* où un indice (*cue*) est présenté à chaque essai, précédant ou accompagnant le stimulus. Cet indice (visuel en général) instruit le sujet sur la tâche à effectuer, impliquant ainsi un *switch* potentiel à chaque essai.

Task-switching fondé sur les résultats obtenus

Dans certaines études, les participants doivent découvrir par eux-mêmes le moment où une règle n'est plus valide et opter pour une nouvelle règle valide parmi un ensemble d'options connues.

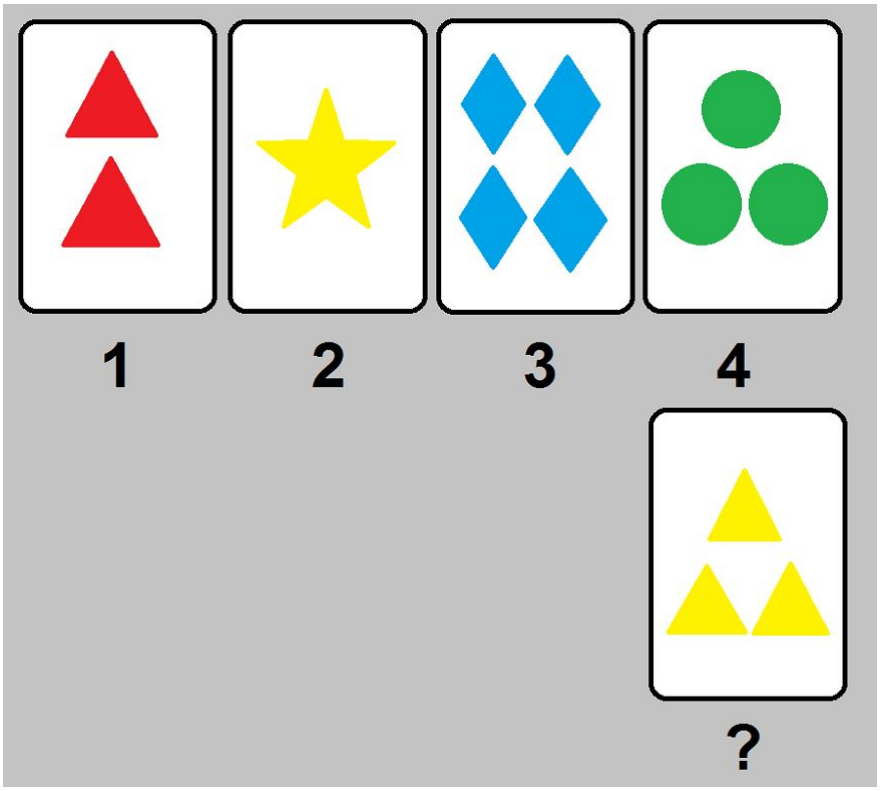


FIGURE 1.2 – **Le *Wisconsin Card Sorting Test*.** Le but de cette tâche est d'associer, à chaque essai, une carte (en bas à droite) avec l'une des cartes présentées (en haut), soit en fonction de la forme (carte 1), soit en fonction de la couleur (carte 2), soit en fonction du nombre d'éléments (carte 4). Le sujet n'est pas informé de la règle à suivre et doit la déduire de ses succès et échecs. La règle peut changer de façon imprévisible.

La tâche du *Wisconsin Card Sorting Test* (WCST) est un exemple classique de ce type de paradigme (voir figure 1.2). Dans cette tâche, des cartes représentant des stimuli variant selon trois dimensions (forme, couleur et nombre) sont présentées au sujet. Celui-ci connaît les trois règles du jeu disponibles : trier les cartes selon l'une des trois dimensions. A chaque essai, il doit sélectionner la carte qui s'accorde avec une carte présentée isolément, mais sans préciser quelle dimension définit cet accord : le sujet qui réalise ce test doit faire des essais, après lesquels l'expérimentateur ou le programme lui indique si sa solution est juste ou fausse, lui permettant ainsi de trouver la dimension pertinente. Cette dernière peut changer en cours d'expérience : après avoir assorti les cartes en fonction du nombre de motifs, il faudra par exemple les assortir en fonction de la couleur, et donc reconfigurer la manière de choisir les cartes.

Ce test permet d'obtenir une mesure de la flexibilité cognitive, une notion assez générale qui couvre à la fois la capacité à maintenir plusieurs tâches simultanément, et l'aptitude à passer d'une tâche à l'autre (mémoire de travail et *task-switching*). Il est très largement utilisé pour la détection de problèmes préfrontaux. En effet, Les patients atteints par exemple de lésions préfrontales ou de maladies neurodégénératives effectuent des erreurs de deux types lors de cette

tâche : persévération (maintien de l'utilisation d'une règle malgré des retours négatifs) et distraction (échec du maintien d'une règle malgré des retours positifs).

Le WCST permet d'appliquer trois *task-sets* différents, et donc d'observer le comportement lors des transitions entre ces *task-sets*. Toutefois, l'espace de *task-sets* étant contraint par la structure de l'expérience, lorsque les trois *task-sets* possibles ont été appris, il n'est pas possible d'en générer d'autres.

1.2.2 Le coût de *switch*

Définition

L'effet principal du *task-switching* est connu sous le nom de *switch-cost*, correspondant au coût comportemental lié au changement de tâche. Cet effet se matérialise par une augmentation du temps de réaction et de la proportion d'erreurs effectuées dans un essais *switch* (changement de tâche) par rapport à un essai *stay*, où la tâche n'a pas changé [157][162].

Cet effet résiste à de nombreuses manipulations :
- Le temps de préparation : si on augmente le temps entre l'indice de changement de tâche et le premier stimulus auquel doit s'appliquer cette tâche, on observe une réduction du *switch-cost* mais pas sa disparition totale [162].
- La quantité d'information : le *switch-cost* persiste même lorsqu'il n'y a pas d'information apportée par l'instruction, c'est-à-dire lorsque l'action sélectionnée pour le stimulus reste identique, quelle que soit la tâche [110][162].

Causes possibles

Plusieurs hypothèses ont été proposées pour expliquer ce phénomène de *switch-cost* :

La reconfiguration du *task-set*. Selon Monsell [162], le *task-switching* nécessite une reconfiguration du *task-set*, ce qui nécessite un certain nombre d'opérations mentales. Il faut par exemple déplacer l'attention vers de nouveaux éléments sensoriels, ou vers des critères plus abstraits conservés en mémoire, récupérer les objectifs internes correspondant à la nouvelle tâche et implémenter l'association stimulus-réponse spécifique au nouveau *task-set*, tout en inhibant l'ancien. Ces différentes opérations demandent un certain temps et certaines ressources, et pourraient expliquer l'effet du *switch-cost*. Cependant, ce processus de reconfiguration peine à expliquer le coût résiduel qui ne disparaît pas, quel que soit le temps de préparation laissé aux sujets. Il est possible qu'une partie de cette reconfiguration nécessite l'apparition du stimulus associé à la nouvelle tâche. Ce processus correspondrait à l'extraction de l'association stimulus-réponse depuis la mémoire de travail, faisant ainsi le lien entre le *task-switching* et la mémoire de travail [193], et justifiant le coût résiduel.

L'inertie du *task-set*. Un deuxième processus pour expliquer le *switch-cost* concerne l'inertie des *task-sets* et le rôle de l'inhibition. Cet effet pourrait s'ajouter à l'effet de reconfiguration précédemment décrit. Ce processus s'appuie sur une observation surprenante : *switcher* d'une tâche moins familière vers une tâche plus familière (par exemple, passer de l'action de nommer la couleur de l'encre à celle de lire un mot désignant une couleur dans une tâche de Stroop [49], ou passer de la langue la moins naturelle à la plus naturelle pour les bilingues) engendre un *switch-cost* plus important [245]. De même, il est plus difficile, après avoir quitté une tâche 1 pour effectuer une tâche 2, de revenir à la tâche 1 plutôt que de passer à une tâche 3 [78]. Cette asymétrie ne peut s'expliquer en termes de reconfiguration de *task-set*, car le *task-set* le mieux maitrisé devrait être le plus facile à configurer. Certains auteurs ont donc proposé l'hypothèse selon laquelle le *task-set* le plus familier doit être fortement inhibé lors de l'utilisation d'un autre *task-set* moins familier. L'effet de *switch-cost* apparait car la levée de cette inhibition est lente et demande des ressources cognitives, ce qui influe sur la performance des sujets.

Si les données expérimentales apportent à la fois des arguments pour et contre chacun de ces deux processus, la plupart des auteurs s'accordent pour dire que le *switch-cost* est le fruit d'une combinaison de plusieurs effets.

La capacité de *task-switching* suggère, de manière implicite, une capacité d'inférence sur les états du monde. Autrement dit, que ce soit en fonction d'indices extérieurs ou de résultats observés, les humains sont capables d'inférer que les contingences de l'environnement ont changé et qu'il est nécessaire de s'adapter (de *switcher*).

Nous détaillons dans le chapitre suivant ce que nous entendons par processus d'inférence ainsi que son implémentation dans le cadre de modèles computationnels de la fonction exécutive humaine.

Chapitre 2

Processus d'inférence du système exécutif

2.1 De la logique formelle à l'inférence bayésienne

Le mot "inférence" évoque souvent une opération de logique formelle. Pourtant, comme nous allons le démontrer dans ce chapitre, l'inférence implémentée dans le système exécutif est tout autre.

2.1.1 Les humains sont-ils rationnels ?

Un peu de logique formelle

Lorsque l'on étudie le processus d'inférence dans le cadre de la logique formelle, on s'intéresse en général à l'inférence conditionnelle. Dans un problème d'inférence conditionnelle, on présente deux éléments :
- un prémisse conditionnel : "si p alors q"
- un des quatre prémisses catégoriels : p, non-p, q ou non-q.

En suivant les principes de logique formelle, on peut effectuer les inférences suivantes :
- Si on a p alors on a q (le *modus ponens*)
- Si on a non-q alors on a non-p (le *modus tollens*).

Cependant, si on a non-p, on ne peut pas en déduire non-q (la "négation de l'antécédent") ; et si on a q, on ne peut pas en déduire p ("l'affirmation du conséquent").

Pour comprendre, illustrons ces principes avec un exemple. Posons p = "être un humain" et q = "être mortel". On a bien "si je suis un humain, alors je suis mortel" et "si je ne suis pas mortel, alors je ne suis pas un humain". En revanche, si je ne suis pas un humain, on ne peut pas en déduire que je ne suis pas mortel. De même, si je suis mortel, on ne peut pas en conclure que je suis un humain.

Par conséquent, dans les tests de logique, les participants devraient approuver le *modus ponens* tout autant que le *modus tollens* mais réfuter la "négation de l'antécédent" et "l'affirmation du conséquent". Or dans les faits, comme nous allons le détailler par la suite, les sujets approuvent le *modus ponens* significativement plus que le *modus tollens* et ne réfutent pas complètement la "négation de l'antécédent" et "l'affirmation du conséquent".

La tâche de sélection de Wason

Dans les tâches classiques de raisonnement, afin de tester la rationalité des gens, on compare la performance des sujets à ce qu'ils auraient dû répondre selon la logique formelle. Une tâche couramment utilisée est la tâche de sélection de Wason [167]. Dans cette dernière, on présente aux sujets une série de cartes, comportant chacune un chiffre ou une lettre : par exemple "A", "K", "2" et "7". Les sujets sont informés que chaque carte comporte une lettre d'un côté et un chiffre de l'autre, et doivent retourner uniquement les cartes qui permettent de prouver ou de réfuter une règle du type : "si une carte porte un A sur une face, alors elle porte un 2 sur l'autre".

Dans le cas présent, il faudrait retourner les cartes "A" et "7". En effet, retourner la carte "A" (p) permet de tester le *modus ponens*, car un "A' sur une face devrait impliquer un "2" (q) sur l'autre. Retourner la carte "7" (non-q) permet de tester le *modus tollens*, car un chiffre différent de "2" sur une face devrait correspondre à une lettre différente de "A" sur l'autre (non-p). La tâche de sélection de Wason teste donc la capacité à appliquer dans des conditions de laboratoire des règles de logique formelle.

Or, dans toutes sortes de tâches de logique de ce type, les sujets réalisent des erreurs nombreuses et systématiques [85]. Pour la tâche de sélection de Wason par exemple, le taux d'erreur peut atteindre 80%. Il a été montré que ce taux d'erreur peut être plus faible lorsque les sujets sont mis dans des situations dites concrètes ou thématiques, tout en gardant la même structure logique et les mêmes exigences de rationalité [121]. Le taux d'erreur s'établit alors autour de 50%, ce qui reste toutefois élevé. Cette observation a conduit certains auteurs à se demander si l'être humain pouvait toujours être considéré comme un être rationnel [224][225].

Selon certaines interprétations, cette déviation systématique de la logique est inévitable et s'expliquerait par les limites de la mémoire de travail qui restreignent la capacité de raisonnement [27][189]. Les humains seraient ainsi rationnels sur le principe mais contraints par des limitations cognitives. Ces approches sont cependant difficiles à concilier avec deux faits : (1) dans la tâche de sélection de Wason, le taux d'erreur peut atteindre 96%, soit bien plus que ce qui pourrait être expliqué par une limite de la mémoire de travail ; (2) Au quotidien, les humains semblent tout à fait rationnels dans leur manière de penser et d'agir. Comment expliquer ce succès des processus d'inférence au quotidien si le système de raisonnement est sujet à tant d'erreurs dans les situations complexes ? La rationalité quotidienne devrait elle aussi être fondée sur des principes de raisonnement.

Des auteurs ont alors suggéré que les erreurs observées sont en réalité dues non pas à des limitations cognitives, mais à une mauvaise norme de comparaison des performances. Ainsi, le système de raisonnement humain ne serait pas formel au sens de la logique formelle, mais fondé sur d'autres principes.

Effectivement, lorsque l'on compare les performances des humains à ce que prédit la théorie probabiliste, les réponses considérées comme des erreurs dans le cadre de la logique formelle deviennent explicables.

2.1.2 L'approche probabiliste

« La théorie des probabilités n'est au fond que le bon sens réduit au calcul »
Laplace

L'approche probabiliste pose l'hypothèse que le raisonnement quotidien est fondé sur les probabilités : très bien adapté à l'environnement naturel, ce mode de raisonnement s'avère moins efficace dans le cadre d'expériences en laboratoire [222]. Ainsi, selon cette approche, les erreurs et les biais observés dans des paradigmes comme la tâche de sélection de Wason se produisent parce que les sujets importent, dans le cadre du laboratoire, leurs stratégies de raisonnement de la vie quotidienne, élaborées dans un monde incertain.

La clé de l'approche probabiliste consiste à prendre en compte le rôle de l'incertitude dans les inférences quotidiennes. Oaksford utilise l'exemple suivant pour illustrer cette idée [167]. Considérons l'affirmation "les oiseaux volent". Si Titi est un oiseau, alors on peut inférer que Titi vole. Toutefois, les implications réalisées au quotidien sont souvent plus complexes. Ajoutons par exemple que Titi est une autruche. La logique formelle nous conduit toujours à déduire que Titi vole. Mais, intuitivement, cette nouvelle information devrait nous amener à reconsidérer notre conclusion : les inférences quotidiennes sont annulables. L'approche probabiliste permet de répondre à ce type de situation en réinterprétant les implications sous forme de probabilités conditionnelles.

Ainsi, la proposition "les oiseaux volent" peut être reformulée sous la forme : la probabilité de voler, sachant qu'il s'agit d'un oiseau, est élevée. Soit par exemple : P (voler/oiseau) = 0.9. Cependant la probabilité de voler en étant un oiseau et une autruche serait plutôt proche de zéro, et la probabilité de voler en étant un oiseau et un perroquet pourrait être par exemple de 0.96.

Cette approche semble ainsi plus adaptée pour comprendre les règles de raisonnement qui se forment au contact d'un environnement incertain, et doivent être apprises à partir d'un historique de cas particuliers.

2.1.3 L'inférence bayésienne

Comme nous l'avons mentionné précédemment, notre cerveau doit faire face au problème de l'incertitude, que ce soit lors de la perception du monde physique qui nous entoure, lors de la prise de décision ou lors de l'exécution d'une action. Cette incertitude se retrouve à différents niveaux : au niveau du système sensoriel, de l'appareil moteur, de nos connaissances et du monde lui-même. La théorie probabiliste, et notamment la théorie d'estimation statistique bayésienne, permet de traiter ces incertitudes.

Le théorème de Bayes [1] est au cœur de cette théorie. Il stipule, dans son interprétation la plus simple, que notre croyance sur le monde devrait être réajustée en fonction du produit de nos croyances a priori et de nos observations.

Le théorème de Bayes se formule ainsi : $P(X/Y) = \frac{P(Y/X)P(X)}{P(Y)}$

Où :
- P(X) correspond à la "probabilité a priori", c'est-à-dire nos croyances a priori, ou encore la probabilité que X soit vraie ;
- P(Y/X) correspond à la "vraisemblance", c'est-à-dire la probabilité que l'hypothèse X puisse générer l'observation Y ;
- P(X/Y) correspond a la "probabilité a posteriori", c'est-à-dire nos croyances réajustées après avoir fait l'observation Y, ou autrement dit la probabilité que X soit vrai après avoir observé Y ;
- P(Y) est la probabilité d'observer Y. Elle sert de normalisateur.

Ainsi, l'inférence bayésienne, en s'appuyant sur le théorème de Bayes, permet de déduire la probabilité d'un événement à partir de celles d'autres événements déjà évaluées, autrement dit en fonction de l'historique des événements passés. Il est à noter que dans la théorie bayésienne, les probabilités ne sont plus interprétées comme les fréquences relatives d'événements mais comme des mesures du degré de connaissance subjective.

O'Reilly et collaborateurs affirment que puisque les fonctions cérébrales nécessitent souvent d'intégrer différentes sources d'informations, et que les processus de traitement impliquent intrinsèquement l'incertitude, l'inférence bayésienne peut s'avérer un bon cadre descriptif d'un certain nombre de sujets d'étude en neurosciences [176]. Par exemple, l'inférence bayésienne peut s'appliquer aux décisions perceptuelles, durant lesquelles le sujet tente de catégoriser un objet ou d'en estimer une mesure à partir d'informations sensorielles incertaines. Les auteurs signalent cependant que le cerveau humain a peu de chances de fonctionner intégralement selon une logique bayésienne, étant donné le coût important en termes de puissance de calcul. En effet, un modèle bayésien doit maintenir en

1. Pasteur de l'Église presbytérienne et mathématicien britannique, le révérend Thomas Bayes (1701-1761) étudie la logique et la théologie à l'université d'Édimbourg. Divers travaux, dont une introduction au calcul différentiel, le font élire à la *Royal Society* le 4 novembre 1742. Ce n'est qu'après la mort de Bayes, en 1761, que son ami Richard Price retrouve, dans ses papiers, un *Essai sur la manière de résoudre un problème dans la doctrine des risques*. Publié par la *Royal Society* en 1763, cet essai applique le principe d'inférence que nous connaissons aujourd'hui sous le nom de "règle de Bayes".

mémoire l'historique des événements nécessaires pour le calcul des probabilités d'intérêt, et au fur et à mesure de l'accroissement de cet historique, le nombre de données nécessaires peut dépasser les limites réalistes de la mémoire humaine. Bien que l'inférence optimale soit trop coûteuse d'un point de vue mnésique, des approximations pertinentes ont été adaptées pour rendre compte du comportement humain [204].

En conclusion, ce que nous appelons "inférence" dans le cadre de cette thèse n'est pas l'inférence au sens de la logique formelle, mais celle au sens probabiliste, qui capture plus fidèlement le comportement humain. On s'intéresse au processus d'inférence bayésienne, appliqué plus précisément à l'échelle des *task-sets* (et non à celle des associations stimulus-réponse). Il s'agit donc d'inférences réalisées à un niveau hiérarchique supérieur.

2.2 Modèles computationnels d'inférence

"I'm not trying to copy Nature. I'm trying to find the principles she's using"
R. Buckminster Fuller

2.2.1 Les modèles de monitorage de l'incertitude

Nous ne détaillerons pas ici le modèle d'apprentissage par renforcement : bien que présent dans l'ensemble des modèles présentés, il ne permet pas à lui seul d'effectuer des inférences au niveau des *task-sets*. En effet, ce modèle est très efficace pour l'apprentissage d'un *task-set* initial et est également capable de s'adapter après un changement de contingences pour apprendre un nouveau *task-set*. Cependant, ce faisant, il efface le précédent *task-set*. A chaque nouvel épisode, le modèle doit donc réapprendre le *task-set* valide. Par ailleurs, le passage d'un *task-set* à l'autre demande de désapprendre ce qui a été appris précédemment et implique donc une adaptation relativement lente (Pour une description succincte du modèles d'apprentissage par renforcement, voir 10.3.3.)

Monitorage de l'incertitude et apprentissage par renforcement simple

Le phénomène de "désapprentissage" d'un modèle d'apprentissage par renforcement simple peut être contourné si l'on ajoute un "détecteur" de changement de contingences, afin de repartir sur un apprentissage nouveau. On peut ainsi combiner un modèle d'apprentissage par renforcement avec un modèle qui estime l'incertitude afin de détecter la nécessité de "switcher".

Plus précisément, ce type de modèle apprend un *task-set* par apprentissage par renforcement et, en parallèle, il apprend également à prédire l'incertitude de ce *task-set*. Au sein d'un *task-set*, un modèle interne de l'incertitude calcule la probabilité d'observer une récompense ($r = 1$) ou non ($r = 0$), pour une action

a face à un stimulus s. Ce modèle interne peut être appris par observation fréquentielle : la probabilité d'obtenir une récompense ($r = 1$) pour le stimulus s et l'action a est estimée par le nombre de fois que la paire (s, a) a conduit à une récompense($r = 1$) divisé par le nombre de fois que la paire a été testée dans le cadre de ce *task-set*.

Yu et Dayan ont implémenté un modèle de ce type [249]. Ils font l'hypothèse que le *task-set* en cours d'utilisation est correct par défaut, jusqu'à ce qu'on observe la nécessité de *switcher*. La confiance dans le *task-set* utilisé est estimée après chaque observation. Cette confiance est *ex-ante* : avant qu'aucune observation ne soit effectuée au temps $t + 1$, on détermine la probabilité que le *task-set* utilisé à t soit également valide à $t + 1$. L'actualisation se fait par inférence bayésienne. Lorsque la probabilité que le *task-set* soit correct devient plus faible que la probabilité qu'il soit incorrect, il est nécessaire de quitter le *task-set* par défaut et d'en utiliser un autre.

Le signal de *switch* permet d'inférer que ce qui est observé n'est plus conforme à ce qui est attendu. Il permet donc d'interrompre l'adaptation et d'envisager le stockage du *task-set* appris avant qu'il n'ait été modifié par apprentissage dans une nouvelle condition. C'est donc un mécanisme qui permet d'apprendre un *task-set* et de déterminer à quel instant il faut interrompre son apprentissage avant qu'il ne soit dégradé.

Le modèle bayésien proposé par Yu et Dayan est donc satisfaisant dans sa capacité à maintenir un *task-set* par défaut en présence d'incertitude et à *switcher* quand nécessaire. Cependant, le mécanisme de *switch* de ce type de modèle ne permet pas de décider quel *task-set* appliquer après le *switch*. En effet, même si les *task-sets* précédemment appris sont stockés, comment décider lequel utiliser autrement qu'en en sélectionnant un au hasard et en le testant pendant un certains nombres d'essais (ce qui est proposé par Yu et Dayan) ? Il est évident que ce processus est sous-optimal. De plus, ce modèle n'est pas satisfaisant car l'espace des *task-sets* à explorer doit être construit avant d'exister. Après chaque *switch* une décision doit être prise entre deux processus : l'apprentissage d'un nouveau *task-set* ou la sélection d'un *task-set* précédemment appris et aucune proposition n'existe pour arbitrer entre ces deux options. Un tel modèle ne permet donc pas de stocker des *task-sets* dans l'espoir de les réutiliser.

Monitorage de l'incertitude et apprentissage par renforcement multiple

Doya et collaborateurs, dans le cadre du contrôle moteur, ont proposé le *multiple model-based reinforcement learning model* ou $MMBRL$ (voir figure 2.1) [77]. Ce modèle repose sur l'existence parallèle de plusieurs modules d'apprentissage par renforcement, chacun pouvant apprendre des valeurs attendues et une stratégie. Chaque module est également muni d'un modèle prédisant les conséquences des commandes motrices qu'il élabore. La comparaison entre la prédiction de ce modèle et le nouvel état observé comme conséquence de la commande motrice effectuée, permet le calcul d'une erreur de prédiction différente de celle liée au modèle d'apprentissage par renforcement.

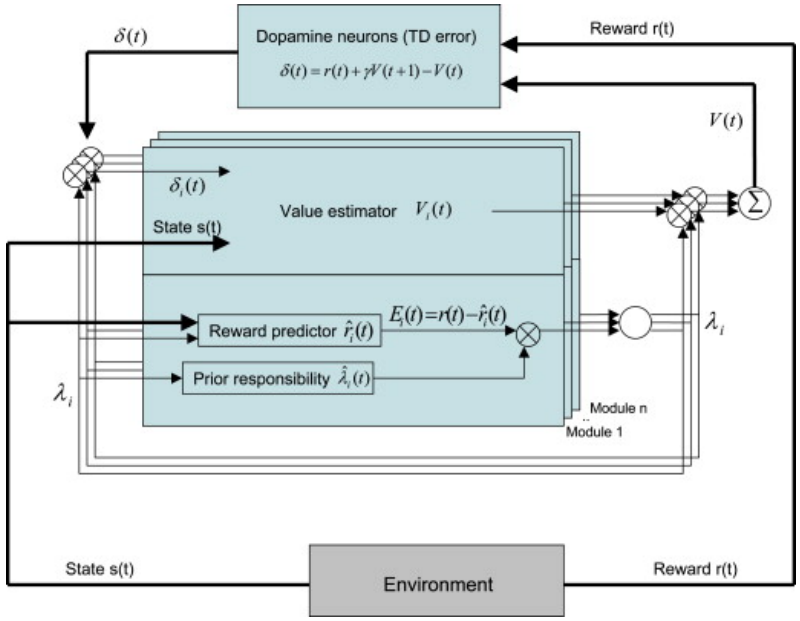


FIGURE 2.1 – **Schéma de principe du modèle de *multiple model-based reinforcement learning***, issu de [20]. Chaque "module" possède sa propre stratégie (représentée par le *Value estimator*) et un modèle interne prédictif (*Reward predictor*) qui permet de calculer un signal de responsabilité (λ_t). Le modèle applique une stratégie globale V(t) qui correspond à la somme des stratégies de chaque module pondérée par les responsabilités. L'apprentissage dans chaque module est proportionnel à l'erreur de prédiction globale δ(t) pondérée par la responsabilité précédente de chaque module (δ_i(t)).

Pour chaque module, on peut ainsi calculer un "signal de responsabilité". Ce signal reflète la validité actuelle des prédictions effectuées par chaque module, et détermine ainsi la responsabilité qu'il devrait avoir dans la sortie motrice. Celle-ci est donc la somme des sorties motrices de chaque module, pondérée par les signaux de responsabilité.

Imamizu et collaborateurs ont étendu ce modèle en prenant en compte l'information contextuelle. Cette information est intégrée a priori sur les différents modules et elle est utilisée, en plus de l'erreur de prédiction du modèle interne, pour calculer le signal de responsabilité. Cela permet au modèle de multiplier les sources d'information pour déterminer quel module il est pertinent d'utiliser, plutôt que de se fonder uniquement sur les erreurs effectuées. On a ainsi des informations a posteriori et a priori [114].

Cependant, dans ce type de modèle, le rôle de chaque module est parfaitement symétrique, il n'existe pas de module par défaut sélectionné tant qu'il est valide. De plus, ce type de modèle exclut la de notion de *switch*, puisque tous les modules prennent part à la décision à chaque essai. Or, contrairement au cadre du contrôle moteur, dans le cadre du contrôle cognitif, adapter son comportement se fait souvent de manière binaire par l'intervention d'un *task-switch* permettant de passer instantanément d'une représentation à l'autre.

Le modèle proposé par Doya et collaborateurs peut être adapté à un environnement discret convenant au contrôle cognitif et ainsi permettre à la fois l'apprentissage et le stockage d'un nombre prédéfini de *task-sets*, l'exploration de l'espace des *task-sets* appris et le passage d'un *task-set* à l'autre [53].

Un tel modèle est muni de N *task-sets*, tous initialisés comme un comportement aléatoire. Chaque *task-set* est muni d'un module d'apprentissage par renforcement et d'un modèle prédictif. Pour chaque *task-set*, on calcule la confiance qu'on a dans ce *task-set* particulier, étant donné l'historique des observations effectuées.

Les signaux de confiance sont utilisés de trois manières par le modèle :

- Ils pondèrent la vitesse d'apprentissage du modèle d'apprentissage par renforcement du *task-set* correspondant. Ainsi un *task-set* n'est appris que lorsque le modèle est confiant quant au fait qu'il se trouve dans l'environnement correspondant à son application.

- Ils pondèrent également l'apprentissage du modèle interne. Ainsi le modèle n'apprend à prédire les conséquences de ses actions que dans l'environnement auquel il est adapté.

- Enfin et essentiellement, ils servent à déterminer quel *task-set* est acteur à un instant donné. Le *task-set* sélectionné applique alors sa stratégie pour déterminer une action.

Ce modèle est capable d'apprendre plusieurs *task-sets* et de savoir quand les utiliser. En effet, initialement, tous les *task-sets* sont vierges et approximativement à égalité et apprennent tous en parallèle. Cependant, la présence de bruit dans l'initialisation assure qu'un *task-set* aura une responsabilité légèrement plus haute que les autres.

A la présentation d'une nouvelle contingence, le même phénomène d'émergence aura lieu entre les *task-sets* restants encore vierges. A la présentation d'une contingence déjà rencontrée précédemment, le modèle peut faire émerger à nouveau le *task-set* correspondant grâce à la mise à jour permanente des signaux de confiance de tous les *task-sets*. En effet, alors, le *task-set* déjà appris est le plus prédictif de ce qui est observé et devient donc le plus responsable.

Ce modèle présente cependant plusieurs limites :

- la nécessité de fixer initialement le nombre N de *task-sets* pouvant être stockés et appris par le modèle. Si le nombre effectif de *task-sets* à apprendre est plus grand que N, le modèle échoue à apprendre et à utiliser une partie des *task-sets*. S'il est plus petit que N, le modèle effectue beaucoup de calculs inutiles.

- la sélection d'un *task-set* à chaque essai ne semble pas correspondre au comportement humain. Il semble en effet que nous tendons à avoir un comportement par défaut et à *switcher* et explorer uniquement lorsque cela est nécessaire.

Pour résumer, ces modèles ne semblent en mesure d'expliquer notre capacité à lier les processus d'inférence et les processus de récupération. Cependant, chacun de ces modèles apporte une pierre à l'édifice en proposant, l'un, la méthode de *switch* après un comportement par défaut pour apprendre un nouveau *task-set*, l'autre, la méthode d'exploration pour la réutilisation des *task-sets*. Le modèle PROBE présenté ci-après combine les différents points forts de ces modèles.

2.2.2 Le modèle PROBE

Le modèle PROBE a été développé par Anne Collins et Etienne Kœchlin [54]. Etant donné qu'il représente notre hypothèse de travail dans cette thèse, nous nous attachons à en faire une description détaillée (voir figure 2.3).

Quelques définitions

Les êtres humains présentent la capacité de s'adapter à des environnements incertains, changeants et ouverts. Un environnement incertain suppose l'existence de processus stochastiques, ou, en d'autres termes, non déterminés ou bruités.

• L'environnement est incertain lorsque les conséquences d'un choix ne peuvent être déterminées avec certitude : il existe plusieurs issues possibles dont les probabilités de réalisation diffèrent (certaines conséquences sont plus probables que d'autres).

• L'environnement est changeant ou volatile lorsque les contingences sont susceptibles de changer de manière imprévue au cours du temps.

• Enfin, un environnement ouvert est un environnement dans lequel le nombre de possibilités différentes est infini ou au moins inconnu.

Pour illustrer ces concepts, prenons l'exemple du choix du moyen de transport pour se rendre le matin sur son lieu de travail. Supposons que le bus de 8h45 soit le moyen le plus rapide. Ce bus est généralement à l'heure, cependant il peut parfois être retardé en raison d'une circulation anormalement dense, ce qui rend son heure d'arrivée incertaine. L'environnement est donc incertain. Admettons maintenant qu'il y ait des travaux de voirie qui dévient la trajectoire du bus pour un certain temps. Il faut alors s'adapter à ce changement de contingences et opter pour un autre moyen de transport, comme le métro. Ainsi l'environnement est volatile. Enfin, il est ouvert car il existe un vaste choix de moyens de locomotion pour effectuer un tel trajet, dont certains n'ont même jamais été empruntés : en bus, en métro, à pied, en vélo, en voiture, en taxi... Dans de telles situations quotidiennes, la prise de décision implique d'explorer, d'ajuster et d'exploiter diverses stratégies comportementales, c'est-à-dire diverses règles d'associations entre stimuli, actions et résultats attendus [126][50]. Cette capacité repose sur la fonction des lobes frontaux qui gère des *task-sets*, soit des représentations mentales actives de stratégies comportementales stockées en mémoire, pour guider le choix [104][162][140][25][202]. Le concept de *task-set* est ici central. Lorsqu'un *task-set* contrôle le comportement, les processus cognitifs sont configurés de manière spécifique pour l'exécution correcte de la tâche en cours. Ceci inclut des processus attentionnels, moteurs, sensoriels, etc. Cette configuration est maintenue activement tout au long de l'exécution de la tâche.

Comment apprenons-nous et gérons-nous une telle collection croissante de stratégies comportementales et comment décidons-nous d'en créer de nouvelles plutôt que de simplement réutiliser et ajuster celles précédemment apprises ?

Plus formellement, comment la fonction exécutive frontale arbitre-t-elle continuellement entre :
- Conserver et ajuster le *task-set* en cours d'utilisation (prendre le bus)
- Sélectionner un autre *task-set* déjà appris (prendre le métro)
- Créer de nouveaux *task-sets* (essayer les vélos en libre service de la municipalité) ?

La sélection et la création de *task-sets* correspondent à des évènements de *switches* comportementaux, lorsque le choix consiste à changer de stratégie plutôt que de persévérer.

De plus, comment ces choix peuvent-ils être effectués d'après les indices contextuels extérieurs ? En effet, en reprenant notre exemple, si davantage de gens que de coutume attendent à l'arrêt de bus, cette information contextuelle permet d'inférer la présence de perturbations. Il est alors sans doute préférable de choisir un autre moyen de transport.

Ce genre de situation soulève un problème computationnel que des modèles d'apprentissage statistiques, fondés sur des processus mathématiques inférentiels tel que les processus de Dirichlet [2], sont en mesure de résoudre de manière optimale [229]. Cependant, ces processus impliquent de réviser à chaque essai tout l'historique des évènements précédents. Autrement dit, ce type de calcul nécessite une mémoire parfaite et infinie et des coûts de calcul croissants. Comment la fonction exécutive frontale pourrait-elle alors implémenter de tels modèles statistiques qui deviennent rapidement intraitables par des systèmes dont la mémoire est limitée ? Se pose ainsi le problème fondamental de comprendre comment, avec des capacités computationnelles et mnésiques limitées, la fonction exécutive humaine parvient à gérer une collection croissante de stratégies comportementales.

Le modèle PROBE, qui peut être vu comme une approximation d'un processus de Dirichlet, tente de répondre à cette question. Il s'agit d'un modèle computationnel de la fonction exécutive frontale qui contrôle la création, l'apprentissage, le stockage, le rappel et la sélection de stratégies comportementales pour guider l'action. Le modèle constitue un algorithme biologiquement plausible en prenant notamment en compte la capacité limitée de la mémoire de travail humaine. Il suppose qu'il existe :
- un nombre fini d'actions possibles,
- un nombre fini de stimuli possibles,
- un nombre fini de contextes possibles,
- une récompense binaire (gagné ou perdu) après chaque action.

Un stimulus est présenté avec son contexte, le sujet effectue une action, il reçoit une récompense de l'environnement et le cycle reprend. Les variables d'entrée du modèle sont donc :
- la séquence de stimuli présentés,
- la séquence de contextes présentés,
- la récompense associée à chaque action possible.

2. Un processus de Dirichlet est un processus complexe d'inférences mathématiques. Il permet de réviser ses croyances sur les états du monde (des lois de probabilité) et en particulier d'en créer de nouveaux ou d'en supprimer.

Structure d'un *task-set*

Le modèle PROBE présuppose que les contingences extérieures sont multiples, distinctes, variables, potentiellement infinies et ne sont pas directement observables. Il construit alors des *task-sets*, c'est-à-dire des représentations mentales des "états cachés externes", ou encore des contingences non observables de l'environnement, pour guider l'action de façon appropriée selon les états externes inférés.
Chaque *task-set* est constitué :
- d'un module sélectif qui encode les associations entre stimuli et actions,
- d'un module prédictif qui encode les associations entre les stimuli, les actions et les résultats espérés,
- d'un module contextuel qui encode les associations entre les indices contextuels externes et la fiabilité des *task-sets*.

Le *task-set* en cours d'utilisation pour guider l'action est dit *task-set acteur*. Il est le seul à pouvoir sélectionner les actions et le seul dont les modules internes sont mis à jour.

Il existe par ailleurs un *task-set* particulier dit "aléatoire". Il est équivalent au hasard et représente donc le comportement aléatoire. La présence de ce *task-set* est important d'un point de vue computationnel, en particulier pour l'initialisation du modèle. En effet, au début, le premier *task-set* acteur est comparé à ce que prédirait un *task-set* agissant au hasard. Lorsque le hasard apparaît meilleur que le *task-set* en cours d'utilisation, le premier *switch* est réalisé (création d'un nouveau *task-set*). Ce *task-set* "aléatoire" ne peut cependant jamais être sélectionné en tant qu'acteur.

Le module sélectif. Après chaque action, le module sélectif s'ajuste selon la valeur de la récompense obtenue via un algorithme d'apprentissage par renforcement classique [227]. L'objectif de ce module est de sélectionner l'action, qui permettra de maximiser la récompense obtenue. Pour se faire, il calcule les valeurs attendues pour chaque action sélectionnée face à un stimulus donné. Ces valeurs sont ajustées selon la règle de Rescorla-Wagner qui tient compte de l'erreur de prédiction après chaque essai (l'écart entre la récompense obtenue et celle attendue) [239]. Le poids de cette erreur de prédiction dans la mise à jour des valeurs dépend de la vitesse d'apprentissage. De plus, on normalise de manière à ce que pour chaque action, la somme des valeurs correspondant aux différents stimuli soit égale à 1. On fait ainsi l'hypothèse que chaque action n'est correcte que pour un seul stimulus. La sélection d'une action s'appuie sur ces valeurs. Afin de permettre un premier niveau d'exploration, ce n'est pas l'action dont la valeur est maximale qui est nécessairement sélectionnée. La maximisation de la récompense attendue est "adoucie" par une procédure de *softmax* qui permet aux autres actions d'être ponctuellement "explorées".

Trois paramètres sont associés à ce module :
- α : le taux d'apprentissage de la règle de Rescorla-Wagner ($0 \leq \alpha \leq 1$)
- β : la température inverse du *softmax* ($0 \leq \beta \leq 100$). Ce paramètre gouverne le degré d'exploration lors du choix d'une action. Lorsqu'il vaut 0, le choix est

entièrement aléatoire (exploration pure). Lorsqu'il tend vers l'infini, le choix de l'action se fait sur la base de l'action dont la valeur est la plus élevée (procédure de maximisation standard, exploitation pure).
- ϵ : le "bruit" général ($0 \leq \epsilon \leq 0.5$)

Le module prédictif. Ce module permet d'estimer la probabilité d'obtenir une récompense sachant le stimulus et l'action effectuée. On suppose qu'une seule action peut être récompensée (la somme des probabilités de récompense des différentes actions est égale à 1). Après obtention de la récompense, le module prédictif met à jour la prédiction des récompenses via une approximation d'un processus de Dirichlet et des inférences bayésiennes.

Le module contextuel. Avec les modules précédents seuls, le contrôle cognitif ne peut s'ajuster qu'à partir d'informations a posteriori : les récompenses. Ceci implique que l'acquisition de *task-sets* ainsi que les *switches* (les changements de *task-set*) sont possibles uniquement en réaction à des succès ou des erreurs, rétroactivement. Or, un aspect important du contrôle cognitif, le contrôle contextuel, permet de *switcher* a priori en fonction du contexte dans lequel on se trouve, de manière proactive. Le module contextuel permet ainsi de calculer la probabilité qu'un *task-set* soit correct compte tenu de l'indice contextuel courant. On peut considérer que ce module agit selon deux effets :

- un effet *top-down* : lorsque le *task-set* associé au contexte courant est présent en mémoire de travail, l'association entre contexte et *task-set* permet la récupération directe de ce *task-set*.

- un effet *bottom-up* : lorsque le *task-set* associé au contexte courant est en mémoire à long terme uniquement, le contexte influence la création d'un nouveau *task-set* en le biaisant vers le *task-set* correspondant au contexte.

Notons que le contexte est totalement indépendant des choix effectués.

Un paramètre est associé à ce module : κ : le poids a priori de l'information contextuelle (≥ 1), c'est à dire l'importance donnée à l'information contextuelle a priori.

Notons qu'un aspect délicat de ce module est celui de la définition du contexte. En effet, si on peut dire que les contextes, tout comme les stimuli, sont des entrées sensorielles, il est difficile de dire ce qui les différencie des stimuli. Typiquement, dans les expériences de contrôle cognitif proposées par Kœchlin et collaborateurs, les entrées sensorielles sont généralement bidimensionnelles : des lettres de couleur. La couleur, qui joue le rôle de contexte, est fusionnée avec lettre, qui joue le rôle de stimulus. Pourtant il est possible de dissocier ces deux types d'information. Le modèle PROBE propose ainsi une définition fonctionnelle de la notion de contexte : il peut être défini comme une partie de l'entrée sensorielle qui apporte de l'information pour la sélection d'un *task-set*.

Taille de la mémoire de travail

Le modèle suppose que la fonction exécutive frontale ne contrôle en parallèle qu'un nombre limité de *task-sets* concurrents. Cette borne reflète la capacité limitée de la mémoire de travail humaine et correspond à un sous-ensemble des stratégies comportementales stockées en mémoire à long terme [59]. Nous discuterons cet aspect un peu plus en détail par la suite.

La taille de cette mémoire de travail (N) est un paramètre du modèle.

Fiabilité des *task-sets* : sélection et création

Comment sélectionner le *task-set* acteur ou décider d'en créer un nouveau ? Autrement dit, comment choisir entre exploiter des stratégies connues ou en explorer de nouvelles ? Ceci est permis grâce à la mesure de la fiabilité des *task-sets*. La fiabilité de chaque *task-set* mesure la probabilité qu'il corresponde aux contingences externes actuelles, donc qu'il soit le mieux adapté pour guider le comportement. Cette mesure s'appuie sur les évènements observables tels que les indices contextuels et l'historique des récompenses obtenues pour chaque action. La fiabilité, en tant que probabilité, est une mesure variant entre 0 et 1. La somme des fiabilités des *task-sets* maintenus en mémoire de travail est égale à 1. Elle est évaluée en ligne, via des inférences bayésiennes, juste après l'obtention de la récompense (fiabilité *ex-post* : μ puis juste avant la sélection de l'action, en intégrant la volatilité perçue de l'environnement ainsi que les indices contextuels extérieurs (fiabilité *ex-ante* : λ). Le module sélectif de l'acteur détermine ensuite la réponse au stimulus. On suppose que la sélection (et l'exploration) a lieu au niveau des *task-sets* d'abord, sur la base des signaux de fiabilité, puis au niveau des actions au sein d'un *task-set*.

La fiabilité est une variable clé du modèle car c'est ce critère qui définit la notion de *task-set* acteur et de *switch*. Lorsqu'un *task-set* acteur existe, le modèle est dans une période d'exploitation. Le *task-set* acteur est celui qui apprend activement alors que tous les autres *task-sets* sont simplement conservés sans apprentissage, bien que leur fiabilité continue d'être estimée. La fin d'une période d'exploitation définit un *switch*. Deux cas de figure peuvent se présenter : si un autre *task-set* émerge immédiatement, on passe directement à une nouvelle période d'exploitation ; sinon, on entre dans une période d'exploration. Ces processus sont détaillés par la suite et dans la figure 2.2.

Deux paramètres sont associés au calcul de la fiabilité :
- τ : la volatilité perçue ($0 \leq \tau \leq 1$), ou autrement dit la probabilité supposée que les contingences changent à chaque essai.
- δ : le biais du contexte en faveur d'un switch ($0 \leq \delta \leq 1$). Un changement de contexte peut être considéré comme une information en soi qui pourrait inciter à changer de comportement. Ainsi, lors d'un changement de contexte, le paramètre δ augmente la probabilité d'effectuer un *switch*.

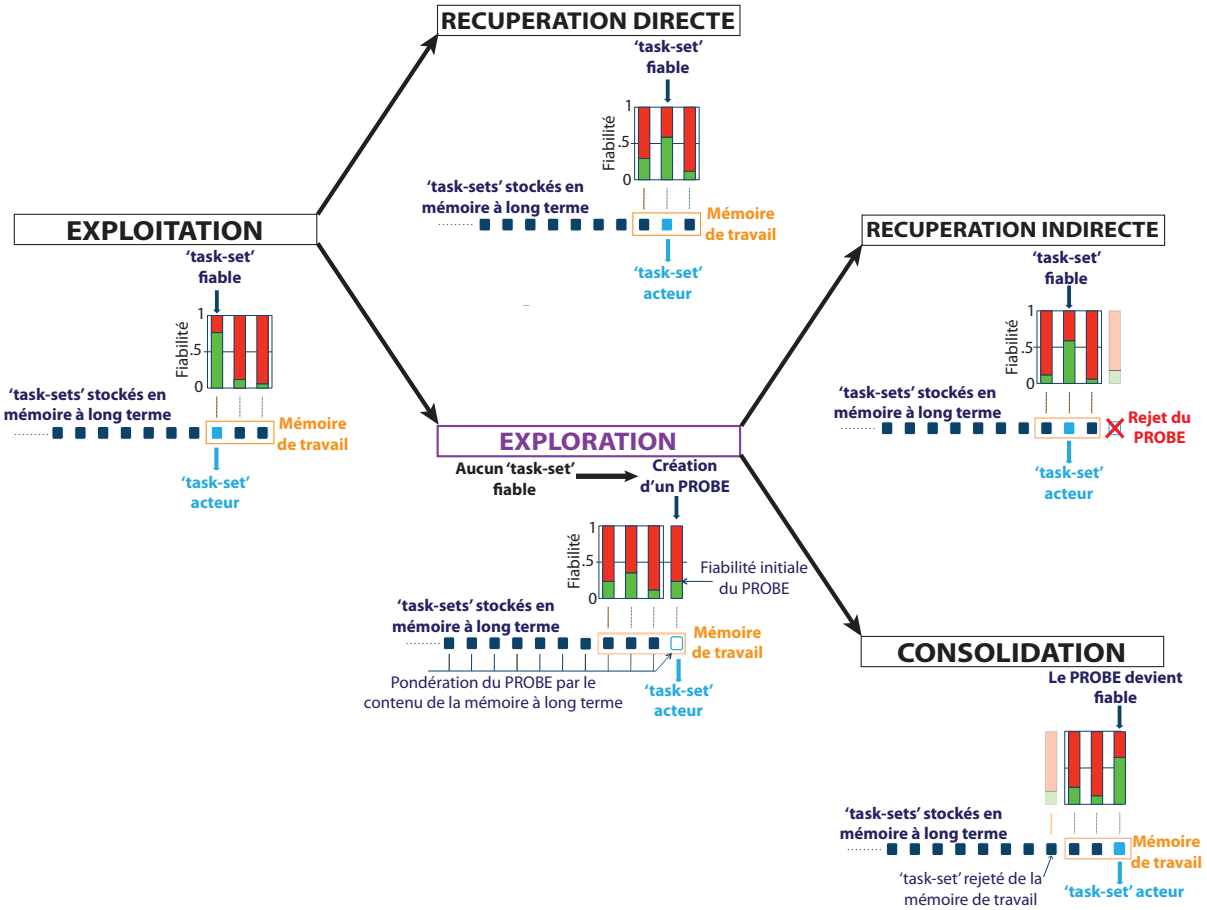


FIGURE 2.2 – **Mécanisme de *switch***

Sélection d'un *task-set* en mémoire de travail. Pour sélectionner un *task-set* dans le répertoire de *task-sets* déjà appris et maintenu en mémoire de travail, le modèle PROBE évalue la fiabilité **absolue** de chaque *task-set*. Ce n'est pas le *task-set* dont la fiabilité est la plus élevée par rapport aux autres qui est sélectionné (fiabilité "relative") mais celui dont la fiabilité dépasse le seuil de 0.5. En effet, lorsque la fiabilité d'un *task-set* est supérieure à 0.5, le *task-set* est plus fiable que non fiable et inversement. Rappelons ici qu'étant donné que la somme des fiabilités de tous les *task-sets* présents en mémoire de travail est égale à un, si la fiabilité d'un *task-set* atteint le seuil de 0.5, il est unique. Par conséquent, lorsqu'un *task-set* est fiable, c'est-à-dire lorsque sa fiabilité dépasse 0.5, il devient l'acteur.

Création d'un nouveau *task-set* Si aucun *task-set* présent en mémoire de travail ne possède une fiabilité suffisamment élevée, c'est-à-dire qu'aucun *task-set* ne semble correspondre aux contingences extérieures, le modèle PROBE créé un nouveau *task-set* qui est testé en tant qu'acteur. Ce *task-set* est appelé *probe task-set*. *Probe* signifie "sonde" en anglais : il s'agit d'un *task-set* nouveau qui, telle une sonde, "explore" l'environnement. Ce *task-set* comporte de nouveaux modules sélectifs et prédictifs qui ne sont pas aléatoires mais formés à partir de

l'ensemble des modules sélectifs et prédictifs stockés en mémoire à long terme [3]. Ce *task-set* est de plus pondéré par les indices contextuels selon les modules contextuels. L'acteur *probe* est initialement non fiable mais sa fiabilité correspond globalement à la probabilité de se retrouver dans une contingence déjà rencontrée. Les modules sélectif et prédictif du *probe* s'ajustent aux contingences extérieures par apprentissage. Deux cas peuvent alors se présenter :
- Rejet du nouveau *task-set* : si un autre *task-set* de la mémoire de travail devient fiable avant l'acteur *probe*, ce dernier est rejeté et le premier devient l'acteur (il est "récupéré").
- Confirmation (ou consolidation) du nouveau *task-set* : lorsqu'il devient fiable, alors que les autres *task-sets* en mémoire de travail restent non fiables, la création du *probe* est "confirmée" ou "consolidée" : le *probe* est maintenu dans la collection de *task-sets* de la mémoire de travail et, éventuellement, le *task-set* utilisé le moins récemment est rejeté de la mémoire de travail vers la mémoire à long terme si la capacité limite est atteinte. L'idée sous-jacente est que des situations plus anciennes sont moins fréquentes et moins susceptibles de réapparaitre à court terme. La mémoire de travail maintient donc les *task-set* qui ont le plus de chance de correspondre à la prochaine contingence.

Dans le modèle PROBE, les *task-sets* non sélectionnés sont supposés non fiables (c'est-à-dire non relié aux états externes courants). Le modèle PROBE suppose donc qu'aucun apprentissage n'a lieu dans les modules sélectifs et prédictifs au sein des *task-sets* non sélectionnés. Ainsi, seuls les modules sélectifs et prédictifs de l'acteur sont ajustés selon les récompenses obtenues. Cette hypothèse est en adéquation avec des résultats empiriques qui montrent que lors de changement de tâches, la sélection d'un *task-set* inhibe les modules internes des *task-sets* non sélectionnés (réf).

Deux paramètres sont associés à la création d'un *probe* :
- θ : le biais de confiance initiale du *task-set probe* ($-2 \leq \theta \leq 3$). Ce paramètre permet de biaiser la fiabilité initiale du *probe* en faveur de la réutilisation d'un *task-set* déjà connu (fiabilité initiale proche de 0, θ tendant vers -2) ou en faveur d'un nouvel apprentissage (fiabilité initiale proche de 0.5, θ tendant vers 3). En effet, plus la valeur de la fiabilité initiale du *probe* est proche de 0.5, plus il risque d'émerger comme acteur (en dépassant le seuil de 0.5). A l'inverse, plus sa fiabilité est proche de 0, plus il risque d'être rejeté en faveur d'un *task-set* présent en mémoire de travail. Le paramètre θ règle ce biais. Il permet en particulier de distinguer une stratégie plutôt exploratoire (biais en faveur de la conservation du

3. L'utilisation de la mémoire à long terme repose sur une logique fréquentielle : le répertoire de *task-sets* en mémoire à long terme s'agrandit à chaque fois qu'un nouvel acteur est créé ; ainsi, les fréquences d'apparition des contingences ont une influence prépondérante sur l'initialisation d'un probe. En effet, à chaque fois qu'une contingence réapparait, si un nouvel acteur est créé, ses modèles sélectif et prédictif apprennent à nouveau les contingences externes, produisant une réplique de ce *task-set* déjà appris lors d'occurrences passées de cette contingence. Donc plus une contingence externe apparaît, plus ses *task-sets* sont reproduits dans la mémoire à long terme. Par conséquent, les *task-sets* appris lors de situations qui réapparaissent fréquemment contribuent davantage à la formation du probe. Il semble en effet judicieux, face à un environnement nouveau, de se fonder sur les comportements les plus fréquemment employés dans le passé.

probe) d'une stratégie plutôt d'exploitation (biais en faveur du rejet du *probe* et donc de la récupération d'un *task-set*).
- η : le biais d'initialisation par la mémoire à long terme du *task-set* probe ($0 \leq \eta \leq 1$). Naturellement on peut envisager que la stratégie initiale du *probe* devrait être uniforme. Cependant, il est probable que les associations déjà utilisées avec succès soient favorisées. Le paramètre η pondère ainsi les valeurs initiales des variables des modèles sélectif et prédictif du probe vers les valeurs de *task-sets* déjà appris.

Le modèle PROBE a déjà été validé dans plusieurs études [54][74].

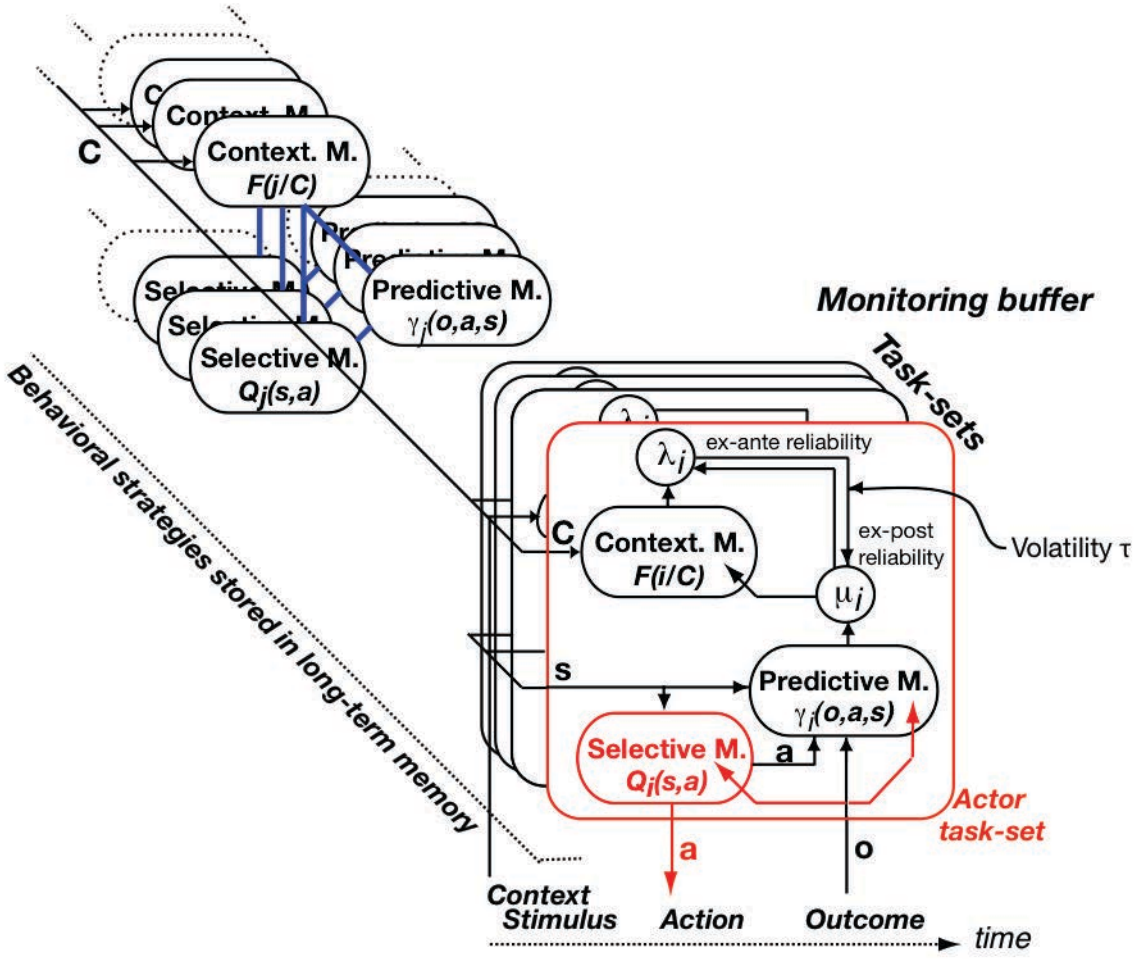


FIGURE 2.3 – **Le modèle PROBE**. La mémoire de travail contient un nombre limité de *task-sets*, contenant chacun des modules sélectif, prédictif et contextuel (M.). La fiabilité de chaque *task-set* est mise à jour à deux moments : juste avant l'exécution de l'action (fiabilité *ex-ante*, λ_i) et juste après la récompense (fiabilité *ex-post*, μ_i). La fiabilité *ex-ante* λ_i est calculée à partir de la fiabilité *ex-post* de l'essai précédent et en tenant compte des indices contextuels C (selon le module contextuel) et de la volatilité perçue τ. La fiabilité *ex-post* μ_i est calculée à partir de la fiabilité *ex-ante* précédant l'action et de la récompense obtenue r (selon le module prédictif). La fiabilité *ex-ante* sert à sélectionner l'acteur pour contrôler le comportement immédiat. Le module sélectif de l'acteur détermine ensuite la réponse au stimulus. Les modules sélectif et prédictif de l'acteur s'ajustent en fonction des récompenses obtenues. Les modules contextuels des *task-sets* s'ajustent selon la fiabilité *ex-post* et apprennent ainsi les indices contextuels qui prédisent la fiabilité du *task-sets*. Reproduit de [54].

Chapitre 3

Processus de récupération en mémoire

"Sans mémoire, tout est nouveau."
Maurice Roche

Inférer que les contingences de l'environnement ont changé et qu'il est nécessaire de quitter le comportement courant représente la première étape du processus adaptatif. Reste à remplacer le *task-set* qui vient d'être quitté par un autre. Une option possible est de récupérer un *task-set* pertinent depuis la mémoire. Dans cette partie, nous tâchons d'expliciter le processus de récupération en tant que tel.

La mémoire est une fonction cognitive indispensable. C'est un outil qui nous permet de garder une trace du passé pour nous aider à faire des choix au présent et à imaginer le futur. C'est la fonction cérébrale qui nous permet de conserver et de restituer des informations. La mémoire et l'apprentissage sont indissociables. Nous ne pouvons pas nous souvenir de choses que nous n'avons pas apprises ou d'événements que nous n'avons pas vécus. L'apprentissage est le mécanisme par lequel la connaissance du monde est acquise alors que la mémoire est le processus par lequel cette connaissance est encodée, consolidée, stockée et plus tard récupérée.

Le champ d'étude de la mémoire est très vaste. Il n'existe d'ailleurs pas *une* mémoire mais *des* mémoires que l'on peut classer selon divers critères :

- Le facteur temps :
  - Mémoire à court terme dont la capacité et la durée de stockage sont limitées
  - Mémoire à long terme dont la capacité et la durée de stockage sont quasiment illimitées
- Le type d'élément à mémoriser :
  - Mémoire déclarative
    - sémantique ("savoir que", faits)
    - épisodique (événements)
  - Mémoire procédurale ("savoir faire")
- Le niveau de conscience
  - Mémoire implicite (inconsciente)
  - Mémoire explicite (consciente)

On distingue également les mémoires spatiale et prospective (tournée vers le futur).

Dans le cadre de cette thèse, nous nous intéressons plus spécifiquement à l'étape de récupération de *task-sets* appris et mémorisés en mémoire à long terme.

Le *task-set* correspond à un niveau hiérarchique au dessus du simple stimulus (objet ou item). Il ne s'agit donc pas ici de mémoriser des éléments discrets (comme mémoriser une liste de mots), mais de mémoriser la règle à appliquer à des items ainsi que l'association entre cette règle et le contexte dans lequel elle s'applique. Le *task-set* s'apparente donc à un objet de mémoire procédurale.

Ainsi que nous l'avons exposé dans la description du modèle PROBE, nous envisageons la récupération des *task-sets* depuis la mémoire de travail vue comme une "fenêtre" de la mémoire à long terme. L'objectif de cette partie est de développer ce concept de mémoire de travail très largement étudié dans la littérature ainsi qu'une théorie psychologique de la récupération de *task-sets* qui soutient l'architecture du modèle PROBE.

3.1 La mémoire de travail

« Everyone knows what attention is. It is the taking possession by the mind in clear and vivid form of one out of what seem several simultaneous objects or train of thought »
James, 1890

3.1.1 Définition

On définit la mémoire de travail comme un système actif de mémoire permettant de stocker temporairement et de manipuler de l'information nécessaire à l'exécution de tâches cognitives complexes [6].

Cependant, dans de nombreuses études, notamment chez les animaux, le terme "mémoire de travail" est souvent synonyme de "mémoire à court terme", qui correspond à la capacité de se souvenir ou de reconnaitre un stimulus après une courte période (de l'ordre de quelques secondes). Or le concept de mémoire de travail inclut également la notion de manipulation d'information.

Il existe plusieurs théories concernant la structure théorique de la mémoire de travail. Parmi elles, trois sont prééminentes. Nous les décrivons ci-dessous.

3.1.2 Théories psychologiques de la mémoire de travail

Le modèle de Baddeley et Hitch

En 1974, Baddeley et Hitch ont proposé un modèle de la mémoire de travail composé de plusieurs sous-systèmes [8]. Les quatre composants de ce modèle sont décrits ci-après.

La boucle phonologique. La boucle phonologique est un système de la mémoire de travail spécialisé dans le stockage et le traitement des informations verbales et symboliques : mots, chiffres, lettres, syllabes, etc. Elle est impliquée dans la lecture, l'écriture, la compréhension orale, et dans le calcul mental. Elle est composée de deux sous-systèmes :
• Un "entrepôt phonologique" qui sert de mémoire verbale à court terme,
• Un système de répétition chargé de répéter mentalement le contenu de l'entrepôt phonologique afin d'en prévenir le déclin. Ce système permet par exemple de maintenir en mémoire un numéro de téléphone tant qu'il est répété.

Le calepin visuo-spatial. Le "calepin visuo-spatial" correspond à la portion de la mémoire de travail chargée de mémoriser des images ou des stimuli visuels. Il peut mémoriser jusqu'à quatre informations de différents types : couleur, forme, vitesse et position d'un objet perçu par le système visuel.

L'administrateur central. Ce sous-système correspond au mécanisme attentionnel de contrôle et de coordination des deux sous-systèmes précédents (les systèmes "esclaves"). Il intègre les informations issues des deux sous-systèmes et les met en relation avec les connaissances conservées en mémoire à long terme. Il dirige l'attention vers l'information pertinente, inhibe l'information non pertinente et les actions inappropriées, et coordonne les processus cognitifs.

Le tampon (ou *buffer*) épisodique. Ce quatrième composant a été ajouté ultérieurement au modèle [7]. C'est au niveau de ce tampon que l'administrateur central peut regrouper les informations issues des sous-systèmes et de la mémoire à long terme.

Le modèle de Cowan

Selon Cowan [58], la mémoire de travail ne représente que la partie activée de la mémoire à long terme. Autrement dit, les représentations en mémoire de travail sont un sous-ensemble des représentations en mémoire à long terme. Il existe donc deux niveaux imbriqués : (i) le premier niveau correspond aux représentations activées en mémoire à long terme. Elles peuvent être très nombreuses car la capacité de la mémoire à long terme n'est pas limitée ; (ii) Le second niveau correspond à la partie la plus activée, le "focus attentionnel", qui, lui, a une capacité limitée : il peut contenir quatre représentations tout au plus [59].

Ce modèle est de type connexionniste et automatiste : il n'existe qu'une seule structure composée d'unités fortement reliées entre elles et couplées à une fonction énergétique, l'activation, qui se localise dans certaines zones du réseau en fonction des besoins. Ce modèle est automatiste puisqu'il ne fait pas appel à des structures de contrôle ou de supervision : les propriétés physiques et mathématiques du réseau, des unités et de la fonction énergétique suffisent à rendre compte de l'ensemble des éléments.

Le modèle d'Oberauer

Oberauer a étendu le modèle de Cowan en ajoutant un troisième composant, un focus attentionnel plus étroit qui ne peut contenir qu'un seul élément à la fois. Par exemple, il est possible de maintenir en mémoire quatre chiffres en même temps dans le "focus attentionnel" tel que défini par Cowan. Cependant, si l'on souhaite effectuer une opération sur chacun de ces chiffres, par exemple leur ajouter deux, il est nécessaire de traiter séparément chaque chiffre. Le composant attentionnel d'Oberauer sélectionne l'un des chiffres pour l'opération puis déplace le "focus attentionnel" au chiffre suivant, jusqu'à ce que l'ensemble des chiffres aient été traités [168].

Oberauer suggère même que la mémoire de travail n'est pas une mémoire à proprement parler, mais plutôt un système attentionnel ; c'est-à-dire un mécanisme ou un processus qui donne la priorité à un sous-ensemble de représentations

par rapport à d'autres, donnant ainsi aux représentations sélectionnées une influence plus importante sur les processus cognitifs à venir. La mémoire de travail serait ainsi de l'attention dirigée vers des représentations mnésiques, qui sélectionne à tout moment un sous-ensemble de la mémoire comme contenu courant de la pensée.

Cette redéfinition de la mémoire de travail la rend indépendante du degré de nouveauté de son contenu. L'information peut tout aussi bien être nouvelle ou déjà bien connue. Les mêmes mécanismes s'appliquent dans les deux cas. Ceci est important pour notre propos. En effet, les tâches utilisées pour étudier la mémoire de travail ont quasiment toujours recours à des ensembles d'items nouveaux (nombre, mots ou objets) qui doivent être maintenus en mémoire de travail à chaque essai. Au contraire, les tâches de contrôle cognitif utilisent presque toujours les mêmes *task-sets* tout au long de l'expérience.

Oberauer est d'ailleurs l'un des premiers à distinguer la mémoire de travail déclarative de la mémoire de travail procédurale. Ou, formulé autrement, la mémoire de travail pour "penser" et celle pour "agir". Selon Oberauer, les mémoires de travail déclarative et procédurale opèrent selon des principes analogues mais ont des capacités limitées distinctes.

La mémoire de travail déclarative. Les représentations déclaratives sont celles qui fournissent de l'information au sujet du monde : l'état actuel de l'environnement (interne et externe), les souvenirs des évènements passés et la connaissance de faits. Cette information est représentée dans un format qui peut être utilisé de façon flexible pour la communiquer (d'où son appellation "déclarative"), et la manipuler.

Le contenu de la mémoire de travail déclarative correspond à des représentations présentes en mémoire déclarative à long terme sur lesquelles l'attention est portée, afin de pouvoir les utiliser lors d'opérations cognitives.

La mémoire de travail procédurale. Les représentations procédurales sont celles qui guident les opérations cognitives et le choix de l'action en spécifiant ce qu'il faut faire selon les circonstances (les circonstances sont des représentations déclaratives). Les représentations procédurales sont des sortes de règles condition-action, où la condition décrit les circonstances pour lesquelles la procédure s'applique, et la composante action définit ce qu'il faut faire.

Même si les représentations procédurales doivent être traduites dans un format déclaratif pour être communiquées (comme une règle verbale lors d'instructions), elles ne sont pas déclaratives en elles-mêmes. L'idée centrale est que la cognition et l'action sont contrôlées par des représentations procédurales. Les représentations déclaratives, telles que des instructions, sont systématiquement traduites en procédures qui génèrent les actions.

Le contenu de la mémoire de travail procédurale correspond à des représentations de la mémoire procédurale à long terme sur lesquelles l'attention est portée, les sélectionnant ainsi pour guider les opérations cognitives effectuées sur les représentations déclaratives.

3.1.3 Capacité de la mémoire de travail

Lorsque l'on mentionne la limite de la mémoire de travail, on pense généralement au nombre d'éléments qui peuvent être mémorisés, et non à la limite des processus de traitement de l'information [61]. Or la capacité de stockage et la capacité de traitement de l'information de la mémoire de travail présentent toutes deux des limites.

Il est souvent difficile de quantifier la capacité de la mémoire de travail notamment parce que son unité de mesure n'est pas toujours clairement définie. Cowan considère que la capacité de la mémoire déclarative se mesure en terme du nombre de "blocs" (*chunks*) d'information et estime que le "nombre magique" est d'environ quatre blocs [59]. Non seulement le nombre de blocs joue un rôle important mais également leur complexité (la longueur des mots, le nombre de caractéristiques d'un objet) [3][46]. Concernant la mémoire de travail procédurale, des études suggèrent que seulement un *task-set* peut être maintenu à chaque instant [237].

Quelles sont les raisons de ces limites ? Plusieurs hypothèses permettent d'expliquer cette capacité restreinte :

• Les représentations maintenues en mémoire de travail pourraient interférer les unes avec les autres. Si davantage d'éléments devaient être maintenus, des interférences pourraient se produire et brouiller la cohérence des représentations [60].

• Les ressources cognitives nécessaires pour maintenir les représentations actives pourraient être limitées [124].

• Il est possible que les représentations mnésiques se dégradent au cours du temps (au bout de quelques secondes) à moins qu'elles ne soient rafraîchies par un processus de répétition. Or la vitesse de répétition peut être limitée [230], d'où une limitation du nombre d'éléments qu'il est possible maintenir.

Le modèle PROBE, qui constitue notre hypothèse de travail, s'inscrit davantage dans le cadre théorique d'Oberauer. Nous développons par la suite cette théorie afin d'explorer plus en détails le processus de récupération.

3.2 Théorie psychologique de la récupération

Oberauer a proposé un cadre théorique pour décrire la récupération de représentations d'objets ou de tâches en mémoire de travail [168]. Cette théorie distingue trois niveaux présentés ci-dessous (voir figure 3.1).

3.2.1 L'activation en mémoire à long terme

La première étape pour permettre la récupération consiste à rendre les représentations pertinentes disponibles et les dissocier des représentations non pertinentes. Anderson et Lebiere ont proposé que les représentations en mémoire à

long terme sont à tout moment activées à un degré qui reflète leur aptitude à devenir bientôt pertinentes [4]. Plus une représentation est activée, plus elle est facile à récupérer. Ainsi, l'activation est la première étape de la sélection d'une représentation. La partie activée de la mémoire à long terme contient des représentations déclaratives et procédurales qui, selon Oberauer, sont reliées entre elles. En effet, des études montrent que les individus acquièrent des associations entre des représentations déclaratives (les stimuli) et les représentations procédurales (les *task-sets*) qui leur sont appliquées [242].

Ainsi, dans un premier temps, au sein de l'ensemble des représentations stockées en mémoire à long terme, un sous-ensemble est activé au dessus d'un certain seuil. Les représentations de ce sous-ensemble sont ainsi prédisposées à la récupération. Dans le cadre du modèle PROBE, ce niveau correspondrait à l'ensemble des *task-sets* présents dans ce que nous appelons la "mémoire de travail". Leur niveau d'activation correspondrait à la valeur de leur fiabilité. Plus leur fiabilité est élevée (plus ils sont activés), plus ils sont susceptibles d'être récupérés.

3.2.2 La région d'accès direct et le pont

Au second niveau, Oberauer dissocie une région dédiée à la mémoire déclarative (la région d'accès direct) d'une autre dédiée à la mémoire procédurale (le pont). La région d'accès direct et le pont sélectionnent un sous-ensemble de représentations dans l'ensemble activé en mémoire à long terme. Leur capacité est limitée.

La région d'accès direct

La région d'accès direct correspond à un sous-ensemble des représentations activées en mémoire à long terme. Seules les représentations maintenues dans la région d'accès direct peuvent être sélectionnées pour un traitement cognitif. Ainsi, les représentations ne sont pas récupérées directement depuis la mémoire à long terme. L'ensemble des représentations pertinentes est d'abord "chargé" dans la région d'accès direct.

Des représentations qui ne sont pas nécessaires peuvent rapidement être rejetées de la région d'accès direct vers la mémoire à long terme activée puis récupérées à nouveau dans la région d'accès direct quand elles s'avèrent à nouveau nécessaires.

Les limites de la capacité de la mémoire de travail déclarative reflètent le nombre limité d'éléments indépendants en mesure d'être maintenus en même temps dans la région d'accès direct [59][170].

Le pont

Les *task-sets* représentent des règles d'associations stimulus-réponse. Par exemple, considérons l'instruction "Appuyer sur le bouton gauche à l'apparition

d'un triangle ou un cercle et appuyer sur le bouton droit à l'apparition d'un carré ou d'une croix". Cette instruction (déclarative) peut être traduite en deux *task-sets* (représentations procédurales) : l'association "triangle" ou "cercle" / "bouton gauche" et l'association "carré" ou "croix" / "bouton droit". Le *task-set* utilisé à tout moment est celui qui occupe le pont.

Les représentations des réponses au sein du *task-set* maintenu dans le pont sont directement disponibles via leurs liens avec le stimulus correspondant. Par exemple, si le premier *task-set* de notre exemple est maintenu dans le pont, toute représentation d'un triangle qui entre en mémoire de travail déclarative enclenche directement un mouvement vers le bouton réponse gauche. Ainsi le *task-set* actuellement maintenu dans le pont agit comme un "réflexe préparé" [109] : une fois mise en place, la représentation d'un stimulus déclenche automatiquement la sélection de la réponse qui lui correspond.

Ainsi, lors du *task-switching*, seul le *task-set* actuellement pertinent est maintenu dans le pont, tout en maintenant les autres *task-sets* actifs en mémoire à long terme afin qu'ils puissent être rapidement récupérés dans le pont si nécessaire [153][192].

Le pont ne peut maintenir qu'une seule représentation à la fois. Dans le cadre du modèle PROBE, ce niveau correspondrait au *task-set* acteur qui guide le comportement en cours.

3.2.3 Le focus attentionnel et le focus de la réponse

A ce troisième niveau, Oberauer dissocie également un processus lié à la mémoire déclarative d'un processus lié à la mémoire procédurale : le focus attentionnel et le focus de la réponse.

Le focus attentionnel

Le focus attentionnel sélectionne à tout moment un élément parmi ceux actuellement maintenus dans la région d'accès direct. L'objet est sélectionné par le focus attentionnel pour être traité et reste dans ce dernier jusqu'à ce qu'il soit remplacé.

Oberauer appelle ce niveau le focus attentionnel par analogie au focus de l'attention visuelle [83] [231]. A ce niveau, une unique représentation est sélectionnée pour réaliser une opération cognitive. La capacité du focus de l'attention est limitée, mais l'information qui y est traitée est à l'abri de toute interférence avec d'autres représentations.

C'est le contrôle exécutif qui gère la sélection d'une représentation parmi les alternatives en compétition [153]. La représentation sélectionnée est alors maintenue dans un état privilégié qui la différencie des alternatives non sélectionnées.

Le focus de la réponse

La sélection d'une action s'effectue dans un focus de la réponse qui maintient la représentation d'une réponse à tout moment, lui donnant ainsi un statut privilégié qui indique qu'il s'agit de la réponse à exécuter. La sélection d'une réponse parmi les alternatives est nécessaire seulement jusqu'à l'exécution de la réponse. Cependant, par analogie avec le focus attentionnel en mémoire déclarative qui semble maintenir le dernier objet sélectionné même une fois le traitement terminé, il est possible que le focus de réponse maintienne la dernière réponse sélectionnée même après son exécution.

Dans le cadre du modèle PROBE, le focus de la réponse correspondrait à la sélection de l'action pour l'essai en cours.

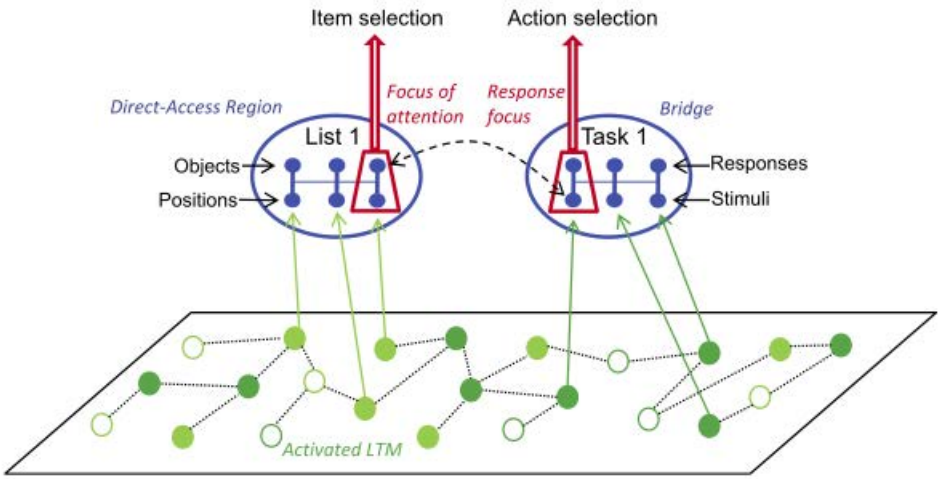


FIGURE 3.1 – **Schéma du modèle de récupération en mémoire de travail déclarative et procédurale.** Le réseau de mémoire à long terme est représenté en vert (vert clair : représentations déclaratives ; vert foncé : représentations procédurales). Les représentations dans la région d'accès direct ou dans le pont, ainsi que leurs liaisons temporaires, sont représentés en bleu (bien qu'apparaissant comme des groupes séparés sur la figure pour la clarté de l'illustration, ces représentations ne sont pas des copies distinctes de celles en mémoire à long terme ; elles sont identiques aux représentations activées dans la mémoire à long terme avec lesquelles elles sont reliées par des flèches). Le focus attentionnel et le focus de la réponse (en rouge) sélectionnent un objet et une action, respectivement. La position de l'objet sert d'indice pour sa sélection ; le stimulus auquel est liée l'action sert également d'indice de sélection.

3.3 Un modèle computationnel de la récupération en mémoire

Oberauer a proposé un modèle connexionniste pour tenter de modéliser sa théorie de la récupération en mémoire déclarative et procédurale [169]. Nous le décrivons ici en modifiant quelque peu la terminologie afin de relier ce modèle avec notre propre cadre théorique. Nous décrivons le modèle dans son aspect procédural sachant que les mécanismes de mémoire déclarative sont similaires.

Ce modèle est schématisé dans la figure 14.1. Il comporte deux modules : un module de sélection de la réponse et un module de sélection des *task-sets*.

3.3.1 Module de sélection de la réponse

Architecture

Dans la figure 3.2, ce module est représenté par le réseau de droite. Il est constitué de deux couches d'unités fonctionnelles interconnectées : une couche d'entrée (*input layer*) et une couche de sortie (*output layer*). Ce module contient également une couche de "candidats" (*candidate layer*) qui comprend autant d'unités que la couche de sortie : à une unité de la couche de sortie correspond une unité de la couche de candidats.

Le pattern d'activité de la couche d'entrée encode le stimulus. Celui de la couche de sortie encode la réponse associée au stimulus. La couche de candidats sert à représenter la force mnésique de chaque réponse indépendamment de leur lien avec le stimulus. Par défaut, toutes les réponses possibles sont activées par la couche de candidats, ce qui permet de les dissocier d'autres réponses non pertinentes pour la tâche en cours. Par exemple, si la tâche nécessite de répondre uniquement avec la main droite, toutes les actions de la mains droite sont activées dans la couche de candidats alors que les actions de la main gauche ne le sont pas.

Encodage d'un *task-set*

Les associations entre les stimuli et les réponses sont modélisées par les connexions entre les couches d'entrée et de sortie. L'activation dans la couche d'entrée est transmise à la couche de sortie via ces connexions. Ainsi, l'activation résultante dans la couche de sortie correspond à la prédiction du système : elle reflète l'action qui est actuellement associée au stimulus.

Le pattern d'activation des couches d'entrée et de sortie modélise ainsi, dans le cadre théorique d'Oberauer, le focus de la réponse. Le poids des connexions entre les deux couches modélise quant à lui le pont. Dans notre cadre théorique, on peut considérer que ce module de sélection de la réponse correspond au *task-set* acteur.

Les connexions entre les couches d'entrée et de sortie sont doubles. Ce qui

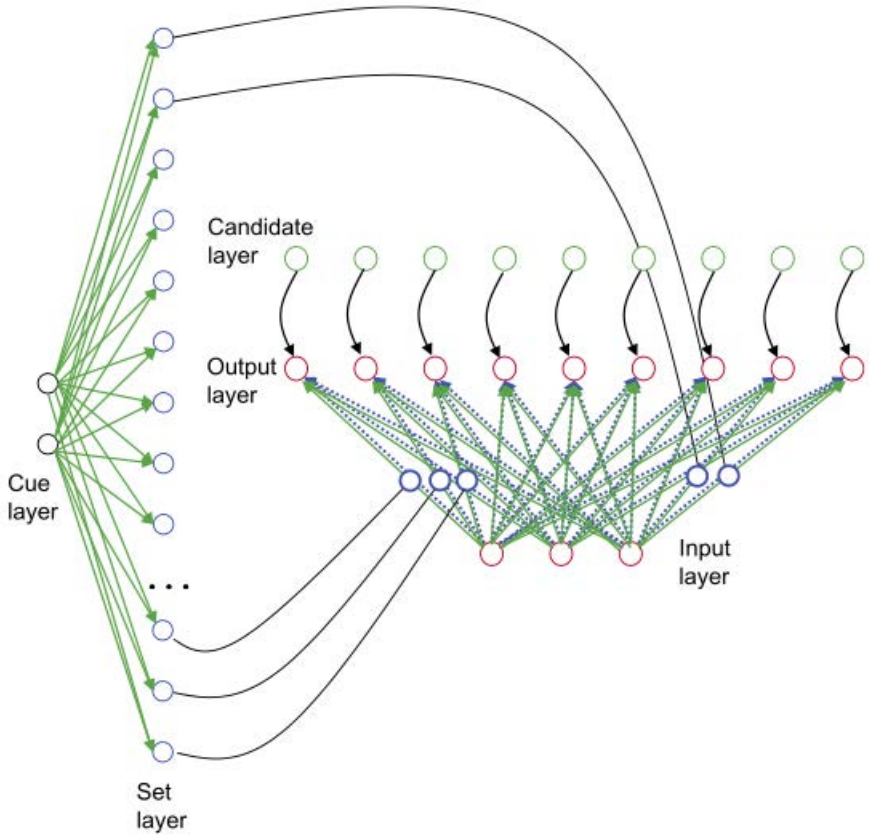


FIGURE 3.2 – **Architecture du modèle computationnel de récupération.** Le modèle comporte un module de sélection de la réponse (structure horizontale à trois couches sur la droite) et un module de sélection des *task-sets* (structure verticale à deux couches sur la gauche). Le module de sélection de la réponse comprend une couche d'entrée qui représente le stimulus, une couche de sortie qui représente la réponse et une couche de "candidats". Les couches d'entrée et de sortie sont connectées par des associations qui changent lentement (lignes continues vertes) et des liens qui changent rapidement (lignes pointillées bleues). Chaque unité de la couche de sortie est liée par une connexion fixe à une unité de la couche de candidats. Le module de sélection des *task-sets* comprend une couche des indices et une couche de *task-sets*. Chaque unité de la couche de *task-sets* est connectée à une association stimulus-réponse du module de sélection de la réponse [169].

les différencie est leur taux d'apprentissage, rapide pour les unes et lent pour les autres. Le taux d'apprentissage rapide (lignes bleues pointillées) modélise les liens temporaires qui peuvent rapidement être mis à jour. Le taux d'apprentissage lent (lignes vertes) modélise les associations établies en mémoire à long terme. La force des associations dans le réseau "lent" de mémoire à long terme correspond à la fréquence cumulée d'utilisation des associations stimulus-réponse. Dans le cadre du modèle PROBE, ces liens traduiraient le biais d'initialisation d'un *probe*.

Sélection d'une réponse

Une fois qu'un *task-set* est ainsi encodé, les réponses individuelles sont accessibles lors de la présentation du stimulus. Un stimulus active sa représentation dans la couche d'entrée. Cette activation est transmise à la couche de sortie. De plus, chaque unité de la couche de sortie reçoit de l'activation de l'unité correspondante dans la couche de candidats. L'activation résultante dans la couche de sortie reflète ainsi la somme de deux sources d'activation, une de la couche d'entrée, l'autre de la couche de candidats. L'activation de la couche d'entrée reflète la force du lien entre stimulus et action. L'activation de la couche de candidats reflète la force mnésique de chaque représentation indépendamment du stimulus. Autrement dit, l'activation parvenant à une unité de sortie depuis la couche de candidats reflète la probabilité que cette représentation soit la réponse correcte, et l'activation provenant de la couche d'entrée reflète la probabilité conditionnelle que cette réponse soit correcte connaissant le stimulus actuellement représenté dans la couche d'entrée.

La durée de sélection de l'action est également modélisée : chaque unité de la couche de sortie est dotée d'un accumulateur linéaire, qui démarre à 0 (avec plus ou moins de bruit) et augmente proportionnellement à l'activation totale reçue par l'unité. Les accumulateurs sont en compétition pour atteindre un seuil fixe. Lorsque le premier accumulateur dépasse ce seuil, la réponse représentée par l'accumulateur gagnant est sélectionnée. Tous les autres accumulateurs sont par ailleurs réinitialisés à 0 afin d'obtenir une représentation non ambiguë de la réponse sélectionnée dans la couche de sortie. L'association entre le stimulus et l'action sélectionnée est ensuite renforcée.

3.3.2 Module de sélection d'un *task-set*

Architecture

La fonction de ce module (réseau de gauche sur la figure 14.1) est de sélectionner le *task-set* pertinent et de l'établir dans le module de sélection de l'action. Le module de sélection du *task-set* opère sur les mêmes principes que le module de sélection de l'action. Il reçoit en entrée une représentation d'un indice qui indique quel *task-set* est pertinent pour l'essai courant. Cette entrée est représentée dans la couche d'indices (*cue layer*) qui est connectée à la couche de *task-sets* (*set layer*). Le pattern d'activité de la couche de *task-set* représente le *task-set* sélectionné. Chaque unité de la couche de *task-sets* est reliée à une connexion à apprentissage rapide du module de sélection de l'action. Ceci permet ainsi la modulation des associations stimulus-réponse.

Ce module de sélection permet de maintenir des représentations de *task-sets* actuellement non sélectionnés. Ainsi, il peut être utilisé pour externaliser l'information actuellement inutile dans le pont mais qui peut être récupérée ultérieurement. Dans le cadre du modèle PROBE, ce module serait l'équivalent de la mémoire de travail dans laquelle sont maintenus les *task-sets* alternatifs.

Mémorisation du *task-set*

Lors de l'apprentissage d'un nouveau *task-set*, comme décrit précédemment, le module de sélection de l'action l'encode en ajustant le poids des connexions entre les couches d'entrée et de sortie. Ainsi, les liens stimulus-réponse du *task-set* sont représentés par ces connexions. Une fois qu'un *task-set* a été entièrement encodé, les activations des unités de modulation (cercles bleus du réseau de droite) sont transférées aux unités correspondantes de la couche de *task-sets*. Le module de sélection du *task-set* acquiert quant à lui une association entre chaque *task-set* et l'indice avec lequel il a été encodé. Les associations entre les indices et les *task-sets* se reflètent dans le poids des connexions avec le module de sélection du *task-set*. Ainsi, ce processus modélise l'acquisition et la mémorisation à long terme d'un *task-set*.

Récupération du *task-set*

Le *task-set* appris et mémorisé peut maintenant être réutilisé. La présentation d'un indice est encodée dans la couche d'indices. Cette activation est transmise via les connexions du réseau pour générer un pattern d'activation dans la couche de *task-sets*. La durée de récupération d'un *task-set* est modélisée de la même manière que la durée de sélection d'une réponse. Une fois que le *task-set* a été récupéré, le pattern d'activation résultant dans la couche de *task-sets* permet de mettre à jour les connexions dans le module de sélection de l'action.

Pour résumer, ce modèle propose un mécanisme explicite pour la formation, la mémorisation et la récupération de *task-sets*. Il est l'un des premiers à offrir une description computationnelle de la relation entre la mémoire de travail et la mémoire à long terme dans un cadre connexionniste. Les simulations de ce modèle permettent par ailleurs de reproduire des effets classiques de *task-switching* tels que le coût de *switch*, les effets de congruence et de répétition.

Notons que les trois composants de la mémoire de travail (mémoire à long terme, pont et focus de la réponse) ne correspondent pas à trois modules distincts dans le modèle. La mémoire à long terme est modélisée de trois façons : par l'apprentissage à long terme des associations entre stimulus et réponse, par l'activation des "candidats", et par le mécanisme d'apprentissage et de récupération des *task-sets*. Le pont correspond aux connexions entre les couches d'entrée et de sortie et à l'activation correspondante dans la couche de *task-sets*. Enfin, le focus de la réponse correspond aux représentations activées dans les couches d'entrée et de sortie.

Ce modèle présente plusieurs similitudes avec le modèle PROBE. Il semble par ailleurs lui être complémentaire sur différents aspects.

Premièrement, il s'agit ici d'un modèle de mémoire et non d'un modèle de la fonction exécutive. Il permet ainsi d'analyser des processus similaires (apprentissage et récupération de *task-set*) sous un autre angle.

Deuxièmement, ce modèle est de type connexionniste. Il permet ainsi de décrire des processus à une échelle qui s'apparente davantage au réseau de neurones plutôt qu'au système, comme c'est le cas dans le modèle PROBE.

Enfin, il intègre une dimension temporelle (durée de récupération des *task-sets* et de sélection de l'action) qui est absente dans le modèle PROBE.

Deuxième partie

Bases cérébrales des processus d'inférence et de récupération

Chapitre 4

Les lobes frontaux et les noyaux sous-corticaux

Le cerveau humain est divisé en deux hémisphères symétriques interconnectés par des commissures inter hémisphériques. Le cortex cérébral, c'est-à-dire la couche supérieure de 2 à 4 mm d'épaisseur, que l'on nomme communément la "matière grise", est constitué d'une assemblée dense de corps cellulaire de neurones, constituée d'environ vingt milliards de neurones. L'autre partie, la "matière blanche", est constituée de cellules gliales et d'axones myélinisés. Ils permettent la communication entre les neurones d'une région à l'autre. Le cortex cérébral est divisé en quatre lobes : le lobe frontal (dont le cortex préfrontal), séparé du lobe pariétal par la fissure de Rolando (ou sillon central). Le lobe temporal, séparé des lobes pariétal et frontal par la fissure de Sylvius (ou sillon latéral), et enfin le lobe occipital, impliqué principalement dans les fonctions de vision. De plus, le cerveau possède des structures internes additionnelles telles que les noyaux sous-corticaux.

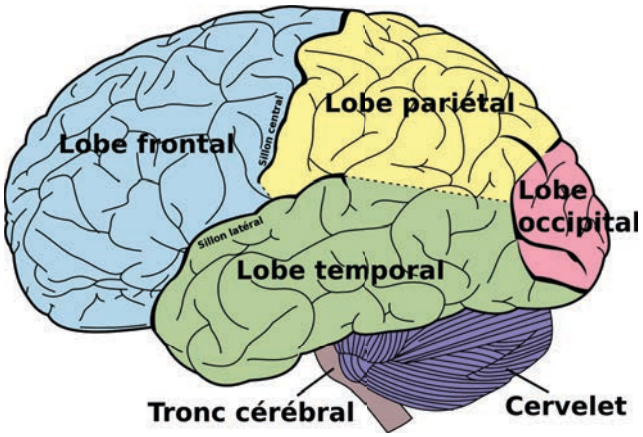


FIGURE 4.1 – **Principaux lobes du cerveau.** Vue latérale gauche. Figue 728 du Gray's Anatomy.

4.1 Les lobes frontaux

4.1.1 Un peu d'histoire

On ne peut pas parler des lobes frontaux sans mentionner le cas historique de Phineas Gage qui amorça le questionnement sur le rôle fonctionnel de cette région. En 1848, Phineas Gage, alors âgé de 25 ans, est contremaître dans la Compagnie Rutland & Burlington qui pose les voies d'une nouvelle ligne de chemin de fer traversant le Vermont, aux Etats-Unis. Le 13 septembre, à 16 heures, alors qu'il bourre une charge d'explosif dans la fissure d'un rocher, une étincelle met feu aux explosifs. Sa barre à mine de 3 cm de diamètre et 90 cm de long est projetée à 30 mètres de là non sans frapper violemment sa pommette gauche. Elle traverse son crâne et provoque une lésion importante de son lobe frontal gauche. La barre ressort par le sommet du crâne en brisant l'os pariétal droit. Malgré la gravité apparente de la blessure, Phineas survit. Secouru par les ouvriers, il demande à être assis sur le chariot du chantier et ramené en ville. A 18 heures, il est pris en charge par le médecin de Cavendish, le Docteur Harlow, qui va non seulement le soigner mais également noter minutieusement l'évolution de son patient. Le 18 novembre, après deux mois de soins, il est considéré comme guéri et rentre à son domicile.

Avant son accident, Phineas jouissait d'une excellente réputation auprès de ses collègues et de ses employeurs. Il était considéré comme sérieux, attentionné, sociable, fiable et de bon jugement. Après son accident, s'il perd l'usage de son œil gauche, son état physique semble quant à lui inchangé. Pourtant son comportement n'est plus même. Il jure, se bagarre et erre sans but en ville. Son humeur changeante, son tempérament devenu grossier et capricieux lui font changer souvent de travail. Il passe même comme attraction au Cirque Barnum. Après plusieurs années d'errance, il rentre dans la ferme familiale où il décède le 21 mai 1861 suite à une crise d'épilepsie. Son crâne est exhumé quelques années plus tard, mais son examen ne permet pas de tirer des conclusions sur son cas. Pour autant, le cortex frontal a depuis rencontré un intérêt clinique et scientifique croissant. Antonio et Hanna Damasio ont même réalisé des reconstitutions sur ordinateur du crâne de Phineas Gage afin de comprendre les causes de ses changements de comportement (voir figure 4.2)[65].

Les modifications comportementales décrites par le Dr Harlow correspondent à ce qui est diagnostiqué de nos jours, au niveau clinique, lors d'une lésion frontale : difficultés à agir lorsque l'action est planifiée, absence de volonté propre, indifférence aux émotions et aux désirs, hyperactivité globale, impulsivité, trouble du comportement social.

4.1.2 Anatomie générale du lobe frontal

Dans le cerveau humain, le sillon central (ou scissure de Rolando) sépare le lobe frontal du lobe pariétal tout au long de la surface du cortex cérébral. Le sillon latéral (ou scissure de Sylvius) sépare le gyrus (ou circonvolution) frontal

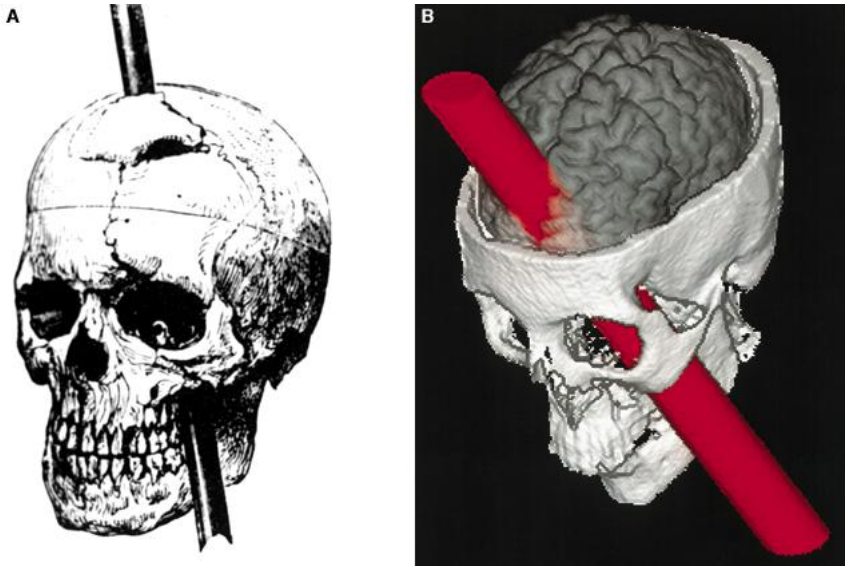


FIGURE 4.2 – **Phineas Gage. A.** Dessin réalisé par le docteur Harlow. **B.** Reproduction en 3D par l'équipe de A. Damasio [65].

inférieur du lobe temporal.

On subdivise le lobe frontal en trois grandes structures depuis le sillon central jusqu'au pôle antérieur :

- Le cortex moteur situé juste en avant du sillon central ;
- Le cortex prémoteur en avant du précédent ;
- Le cortex préfrontal qui regroupe toutes les autres régions antérieures du lobe frontal.

En comparaison avec les primates non humains, les lobes frontaux du cerveau humain sont bien plus développés. Le cortex frontal représente en effet 30% de la surface corticale.

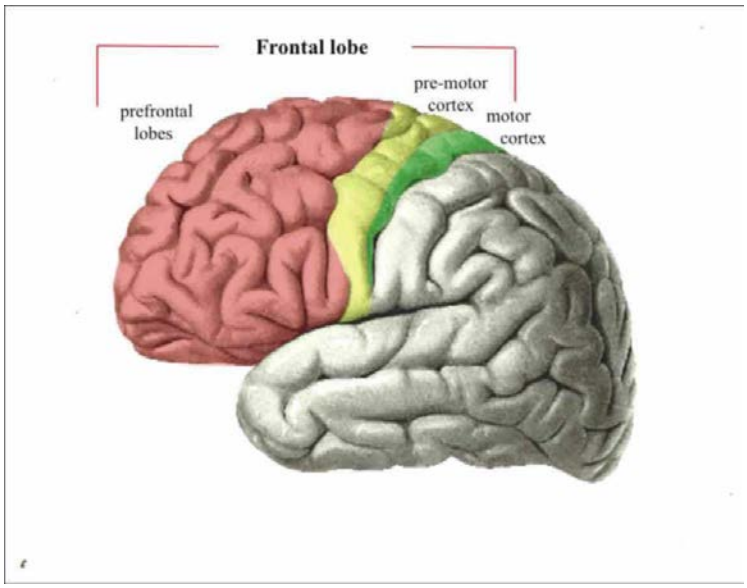


FIGURE 4.3 – **Le lobe Frontal.**

4.1.3 Le cortex préfrontal

Anatomie

Le cortex préfrontal constitue la partie antérieure des lobes frontaux, en position rostrale par rapport aux cortex moteur et prémoteur. Dans la classification de Brodmann [1], il regroupe les aires 8, 9, 10, 11, 44, 45, 46, 47 sur la face latérale, et 12, 32 et 24 sur la face médiale (voir figure 4.4)[30]. Il représente chez l'humain environ un tiers de la surface totale du cortex, cette proportion cortex préfrontal/cortex étant la plus importante de l'ensemble des mammifères, à égalité avec les grands singes [211][210][212]. C'est probablement la région qui a évolué le plus tardivement au cours de l'évolution [119]. C'est aussi une des parties du cerveau qui continue à se développer le plus tard, jusqu'à vingt ans environ [219].

Le cortex préfrontal est la région cérébrale qui entretient le plus de liens réciproques avec la plupart des autres structures [64][99]. C'est un lieu de convergence des informations sensorielles provenant de l'ensemble des aires cérébrales associatives sur lesquelles il projette également. Il est aussi interconnecté avec les structures hippocampiques informant des expériences passées du sujet, et avec les structures limbiques et hypothalamiques permettant une représentation de l'état interne du sujet et renseignant sur son état affectif et motivationnel [190]. Il projette également sur les régions motrices et prémotrices et sur les noyaux sous-corticaux, impliqués dans la préparation et l'exécution des mouvements [22][183]. Le cortex préfrontal envoie des fibres sur pratiquement toutes les structures desquelles il en reçoit. Cette structure occupe donc une position clé pour orchestrer les différentes étapes de l'adaptation comportementale.

Neuropsychologie des patients préfrontaux

L'illustration la plus frappante d'inadaptation comportementale des patients préfrontaux est sans doute l'observation des comportements pathologiques tels que la tendance à saisir involontairement les objets qui sont à leur portée (comportement de préhension), à imiter spontanément les gestes de l'examinateur (comportement d'imitation) ou à utiliser les objets à leur portée alors qu'ils ont pour consigne de ne pas le faire (comportement d'utilisation). Ces comportements de dépendance à l'environnement décrits par Lhermitte et collaborateurs ne sont jamais observés chez les sujets sains en l'absence d'instruction [145][144]. En revanche, ils sont exprimés de façon spontanée et parfois irrépressible chez les patients frontaux, reflétant l'activation de réponses comportementales automatiques, dictées par l'environnement et non par la volonté du sujet. Ces comportements peuvent être observés à des degrés variés [238]. Ceci suggère que le cortex

1. Les aires de Brodmann sont des délimitations du cortex du cerveau humain définies par le neurologue et neurophysiologiste Korbinian Brodmann sur une base cytoarchitectonique. Cela signifie que les aires correspondent à l'organisation structurale apparente du cortex (nombre de couches, épaisseurs des couches, arborisation dendritique etc.). Ainsi, chaque région du cortex ayant la même organisation cellulaire a un numéro allant de 1 à 52. Brodmann a également relié chacune de ces 52 aires à une fonction propre.

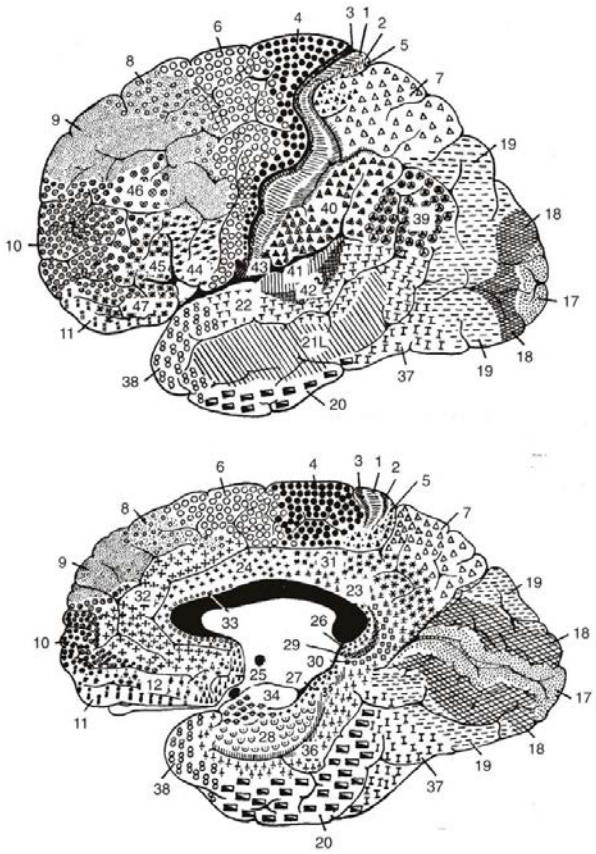


FIGURE 4.4 – **Carte cytoarchitectonique de Brodmann.** Les aires correspondent à l'organisation structurale apparente du cortex (nombre de couches, épaisseurs des couches, arborisation dendritique etc.). Ainsi, chaque région du cortex ayant la même organisation cellulaire porte un numéro allant de 1 à 52.

préfrontal est nécessaire à l'élaboration de comportements nouveaux et volontaires, dirigés vers un but, qui nécessite de rompre la boucle réflexe perception-action.

Rôles fonctionnels

Les recherches ont confirmé le rôle du cortex préfrontal pour différentes fonctions cognitives dites "supérieures", telles que le langage, la mémoire de travail, le raisonnement et plus généralement les fonctions exécutives qui incluent notamment la capacité du cerveau humain à prendre des décisions et à planifier des actions en accord avec des buts internes et des évènements extérieurs [159]. Le cortex préfrontal est donc au cœur de la coordination entre la pensée et l'action.

Plus précisément, Miller et Cohen proposent que le cortex préfrontal occupe une fonction spécifique dans le contrôle cognitif : il permet le maintien actif de patterns d'activité qui représentent les buts et les moyens de les atteindre. Cette activité envoie des signaux à l'ensemble du cerveau, qui affecte les processus visuels mais aussi les autres modalités sensorielles, ainsi que les systèmes

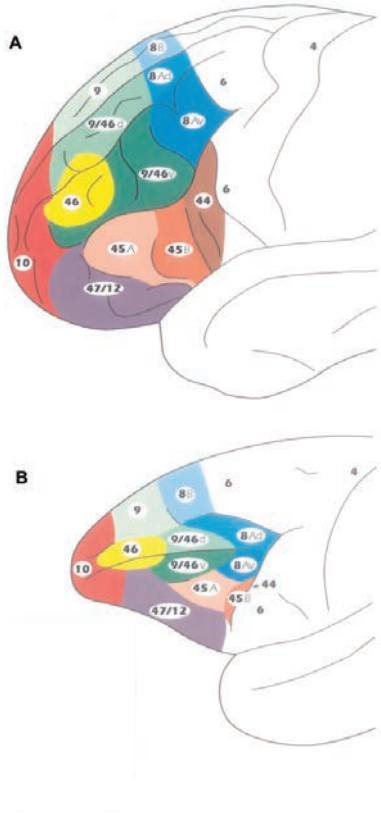


FIGURE 4.5 – **Cartes architectoniques du cortex préfrontal humain et de macaque. (A)** Suface latérale du cortex préfrontal humain. **(B)** Surface latérale du cortex préfrontal de macaque. Extrait de [183].

responsables de l'exécution de la réponse, la récupération mnésique, l'évaluation émotionnelle, etc. Ces signaux sont importants lorsque les stimuli sont ambigus ou lorsque plusieurs réponses sont possibles et que la réponse appropriée à la tâche est en compétition avec des alternatives.

Kœchlin et collaborateurs, ont proposé une organisation fonctionnelle hiérarchique du cortex préfrontal sur plusieurs niveaux de contrôle [132] :
- Le contrôle sensoriel, c'est-à-dire la sélection d'une action en fonction d'un stimulus ;
- Le contrôle contextuel : la sélection d'un ensemble d'associations stimulus-action (soit d'un *task-set*) en fonction d'un contexte présent ;
- Le contrôle épisodique : la sélection d'un ensemble d'associations contexte/*task-sets* en fonction d'un événement (instruction) passé.

Les auteurs observent une activation caudale pour le contrôle sensoriel. Ils montrent que cette activation s'étend de plus en plus rostralement avec les deux autres niveaux de contrôle. Par ailleurs, les analyses de connectivité fonctionnelle montrent que cette organisation fonctionnelle s'effectue en cascade, avec une influence *top-down* des régions rostrales vers les régions caudales.

Kœchlin et collaborateurs montrent également qu'un dernier niveau de hiérarchie est encodé dans le cortex frontopolaire dans le cadre du *branching*, soit la mise en attente d'une tâche pendant l'exécution d'une autre [136][140][110].

L'importance de cette région frontopolaire dans la mise en attente d'une tâche pendant qu'une autre est effectuée est confirmée par Boorman et collaborateurs, qui montrent que cette région encode la valeur d'une action non sélectionnée pendant que le cortex préfrontal ventromédian encode la valeur de l'action sélectionnée [23].

4.1.4 Les cortex prémoteur et moteur

Le cortex prémoteur est la partie du lobe frontal du cerveau située en avant des régions motrices. Le cortex prémoteur participe aux opérations cognitives qui précèdent la réponse motrice, telles que la planification et la préparation de la réponse motrice [177][184]. Le cortex moteur quant à lui implémente l'exécution précise de la commande motrice. Le cortex moteur est en interaction constante avec d'autres structures nerveuses impliquées dans le mouvement comme le système des noyaux sous-corticaux et le cervelet

4.2 Les noyaux sous-corticaux

4.2.1 Anatomie

Contrairement au cortex préfrontal, les noyaux sous-corticaux constituent un ensemble de structures cérébrales évolutivement anciennes. Ils incluent :
- Le striatum composé du noyau caudé et du putamen
- Le pallidum composé des globus pallidus interne et externe
- Le noyau sous-thalamique
- La substance noire compacte et réticulée (voir figure 4.6)

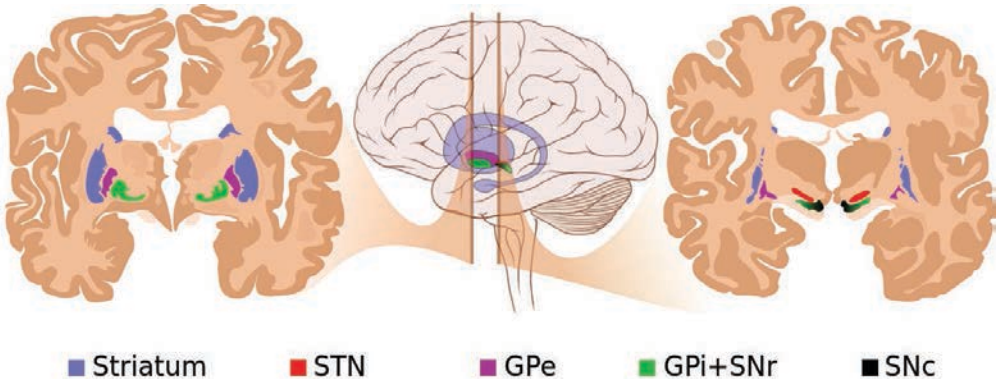


FIGURE 4.6 – **Anatomie des noyaux sous-corticaux.** STN = noyau sous-thalamique, GPe = globus pallidus externe ; GPi = globus pallidus interne ; SNr = substance noire réticulée ; SNc = substance noire compacte.

4.2.2 Rôles fonctionnels

Les noyaux sous-corticaux sont essentiels à de très nombreuses fonctions cérébrales, comme la sélection de l'action, l'apprentissage, la motivation et l'émotion, et leur dysfonctionnement peut provoquer ou participer à des maladies graves. La structure qui va particulièrement nous intéresser est le striatum. Il est au moins impliqué dans :
• Le mouvement volontaire ou automatique ;
• Les comportements appétitifs ou aversifs, en particulier la motivation alimentaire ou sexuelle (striatum ventral) ;
• La nociception ou la gestion de la douleur, via le système dopaminergique et/ou les endorphines, et dans l'effet placebo ;
• La cicatrisation et parfois la régénérescence de certains tissus cérébraux (neurogenèse secondaire) après une blessure neuronale ou un stress ayant induit la mort de nombreux neurones (ischémie / accident vasculaire cérébral) ;
• Certaines maladies :

- La maladie de Huntington est une maladie héréditaire due une dégénérescence du striatum, avec comme conséquence une motricité fortement perturbée (chorée).

- La maladie de Parkinson est due à une perte de l'innervation dopaminergique du striatum provenant de la substance noire (laquelle dégénère au cours de la maladie).

- Les phénomènes d'addiction dépendent de la plasticité des synapses du striatum.

La connectivité dense qui relie le striatum au cortex préfrontal implique un lien important entre ces deux régions, via des boucles fronto-basales, dans un grand nombre de processus cognitifs : des théories récentes présentent les noyaux sous-corticaux comme le nœud cérébral primitivement responsable du tri de l'information, et de son acheminement sélectif aux régions adéquates [226]. Cet ensemble de structures formerait donc une "boîte à outils" centrale, conservée par l'évolution, permettant l'organisation basique de l'information cérébrale et son traitement par les systèmes moteurs, limbiques et cognitifs.

L'intégrité des noyaux sous-corticaux est nécessaire pour observer des comportements de *task-switching*. Lorsque l'on compare les performances de deux groupes de patients, un premier groupe avec des lésions préfrontales, et un second avec des lésions basales, les patients du deuxième groupe montrent de grandes difficultés à changer de tâche [246]. Ce déficit est également observé chez des patients atteints de la maladie de Parkinson [41]. Cependant, le rôle des noyaux sous-corticaux ne se limite pas à l'implémentation des commandes motrices nécessaires au changement de tâche : de nombreux travaux prouvent en effet que ces structures sont essentielles à l'apprentissage fondé sur la valeur, tel que l'apprentissage par renforcement.

L'apprentissage par renforcement correspond à l'apprentissage d'informations au sujet de stimuli ou d'actions uniquement sur la base des récompenses et des punitions (les renforcements) qui leur sont associées. L'apprentissage par renforcement est un apprentissage peu supervisé dans la mesure où aucune in-

formation explicite n'est fournie pour déterminer quelle action doit être effectuée dans une situation donnée. Seuls les renforcements sont informatifs pour l'apprentissage d'associations stimulus-réponse.

Plusieurs études s'appuient sur la méthode *model-based* (voir 10.3) pour illustrer la façon dont les humains apprennent à prendre des décisions adaptatives dans un milieu incertain afin de maximiser leurs récompenses, notamment à l'aide de modèles du type apprentissage par renforcement [66][102][105] [171] . La prise de décision peut alors être dissociée en deux composantes distinctes selon le modèle dit "acteur/critique" [227] (voir figure 4.7) :
- Le "critique" prédit les récompenses futures associées à un ensemble particulier de stimuli ou d'actions au cours de l'apprentissage. Il s'appuie sur la notion centrale d'erreur de prédiction qui correspond à la différence entre la récompense effectivement obtenue et celle prédite.
- L'"acteur" sélectionne l'action selon une stratégie de choix qui permet de maximiser les gains.

Wolfram Schultz et ses collaborateurs ont été les premiers à avoir mis en évidence l'existence de signaux d'erreur de prédiction dans le cerveau en enregistrant l'activité de neurones dopaminergiques chez le primate non humain éveillé pendant une tâche de conditionnement pavlovien classique [207]. O'Doherty et ses collaborateurs ont mis en évidence le même type de signal d'erreur de prédiction dans le cerveau humain, en particulier le striatum ventral et le cortex orbitofrontal, deux régions cibles des neurones dopaminergiques [171]. Par la suite, Montague puis O'Doherty et leurs collaborateurs ont proposé que le striatum ventral et le striatum dorsal implémentent respectivement le "critique" et l'"acteur" dans des tâches de choix d'actions.

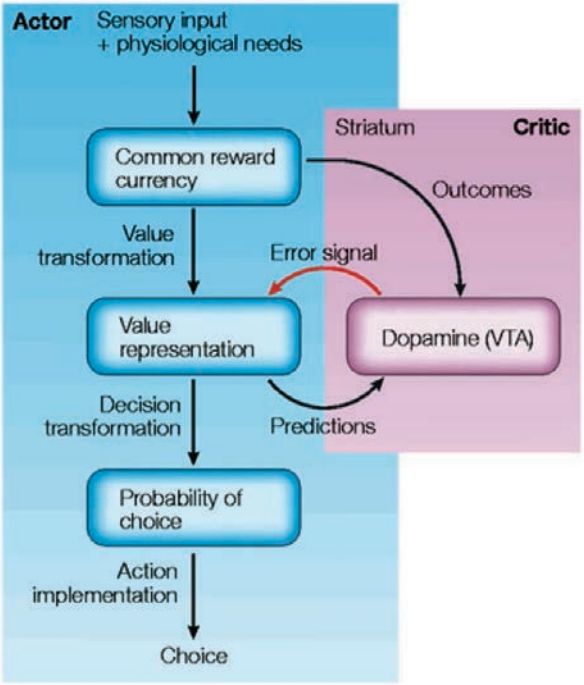


FIGURE 4.7 – **Le modèle acteur/critique.**

Chapitre 5

Bases cérébrales de l'inférence

5.1 Cortex préfrontal ventromédian et prédictions

La distinction entre le cortex orbito-frontal et le cortex préfrontal ventro-médian n'est pas toujours claire étant donné la grande proximité anatomique des deux structures et l'usage variable que différents auteurs font de cette terminologie. La séparation entre ces deux aires se fait parfois plutôt sur des critères de connectivité que sur des critères anatomiques [14].

Le cortex préfrontal ventro-médian possède des connexions majeures avec des structures sous corticales telles que l'amygdale et le striatum ventral, ainsi que d'autres aires préfrontales (principalement le cortex préfrontal dorso-latéral), le lobe temporal supérieur antérieur et le cortex cingulaire postérieur [44]. De par sa connectivité avec de nombreuses structures impliquées dans l'apprentissage, la mémoire, l'encodage de la récompense et les indices de récompense, le cortex préfrontal ventro-médian est bien placé pour intégrer des informations et produire un comportement adapté [163].

5.1.1 Etudes de neuropsychologie

Suite à une lésion du cortex préfrontal ventral, la vie quotidienne des patients est bouleversée : les activités dans lesquelles ils s'engagent aboutissent à des pertes dans différents domaines (financier, relationnel, social, etc.). Les choix qu'ils font sont désormais non avantageux pour eux, et surtout remarquablement différents de ceux qu'ils auraient fait avant leur lésion. Tout ce passe comme si ces patients, sensibles uniquement aux conséquences immédiates de leurs actions, étaient incapables d'en envisager les conséquences différées, positives ou négatives [15][42]. De façon surprenante, ils ne sont plus capables d'apprendre de leurs erreurs passées, comme en témoigne des engagements répétés dans des choix aux conséquences négatives. En revanche, paradoxalement à ce trouble de prise de décision dans la vie courante, leurs capacités de résolution de problème, de jugement et leurs fonctions intellectuelles sont normales en situation de test

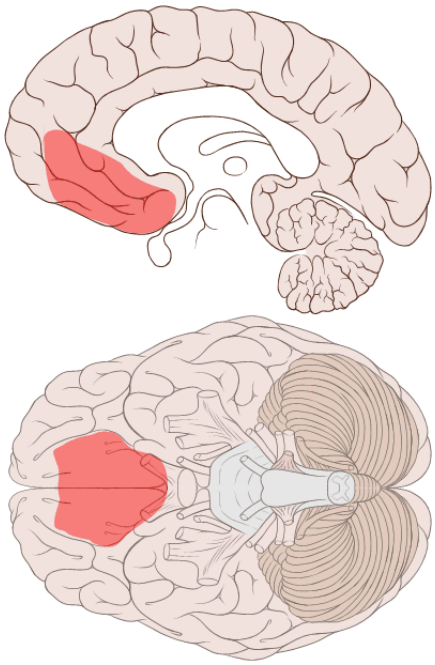


FIGURE 5.1 – **Le cortex préfrontal ventro-médian.** Vue médiane (haut) et du dessous (bas).

de laboratoire[65].

Le cortex préfrontal ventro-médian semble donc nécessaire pour adopter un comportement adapté en fonction des conséquences futures prédictibles.

5.1.2 Etudes de neuroimagerie

Les études de neuroimagerie impliquent le cortex préfrontal ventro-médian dans de très nombreuses fonctions dont la régulation endocrinienne [62], les émotions et leur régulation [84], le conditionnement par la peur et l'extinction [206], la mémoire épisodique et sémantique, [21], la prospection [35] ou encore l'estimation économique [187]. Nous ne détaillons ici que certaines de ces fonctions.

Cortex préfrontal ventro-médian et prédiction de l'avenir

Prédiction de la récompense. Plusieurs études d'imagerie cérébrale chez l'humain montrent l'implication du cortex préfrontal ventro-médian dans la prédiction des résultats d'un choix. Son activité corrèle notamment avec la valeur espérée de l'option choisie, indépendamment de la nature du choix [244][94]. Autrement dit, les représentations de valeur dans cette région peuvent correspondre à des associations action-résultat ou stimulus-résultat.

Prédiction de scénarios. Le cortex préfrontal ventro-médian est directement connecté à des régions de la mémoire, comme l'hippocampe, qui peuvent favo-

riser à la fois la mémoire et la projection dans le futur [35]. Dans leur étude de 2009, Burton et collaborateurs montrent d'ailleurs que l'activité d'anticipation des neurones du cortex préfrontal ventro-médian est abolie chez les rats dont l'hippocampe est lésé, ce qui confirme la connexion forte entre ces deux structures [38].

Le cortex préfrontal ventro-médian, convertisseur de devise

Dans les expériences réalisées chez l'humain, la récompense est en général monétaire. L'activation observée dans le cortex préfrontal ventro-médian est-elle une représentation spécifique de l'argent ? Plusieurs études montrent que ce n'est pas le cas. En effet, dans des expériences qui utilisent d'autres types de récompense (de la nourriture, des consommables non alimentaire, une odeur plaisante, l'attractivité de visages, etc.), l'activité du cortex préfrontal ventro-médian corrèle toujours avec la prédiction de la récompense [47]. Plus spécifiquement, une région précise du cortex préfrontal ventro-médian semble impliquée dans la représentation de la valeur des différentes catégories de biens. Ce résultat suggère que le cerveau encode les valeurs de différents biens à l'aide d'une "devise" cérébrale commune, qui permet une évaluation partagée des différentes catégories de biens. Cette intégration en un système de valorisation commun semble d'ailleurs la manière la plus simple, pour le cerveau, de réaliser des choix économiques qui impliquent des biens de nature différente.

Cortex préfrontal ventro-médian et niveau de confiance

De Martino et collaborateurs démontrent que l'activité du cortex préfrontal ventro-médian corrèle non seulement avec la valeur de l'option choisie, mais également avec la confiance attribuée à ce choix [67].

Donoso, Collins et Kœchlin montrent de plus une corrélation du cortex préfrontal ventro-médian avec la confiance attribuée au *task-set* acteur (donc celui choisi), dans leur étude d'IRM fonctionnelle *model-based* qui s'appuie sur le modèle PROBE sans l'information contextuelle. Cette mesure permet d'inférer un changement dans les contingences extérieures lorsque le *task-set* courant n'apparaît plus fiable [103].

Une vision unifiée

Dans leur revue de 2012, Roy et collaborateurs effectuent une méta-analyse des études portant sur le cortex préfrontal ventro-médian et en proposent une vision unifiée [191]. Selon ces auteurs, cette région représente un carrefour entre divers systèmes impliqués dans la mémoire épisodique, les émotions, la cognition sociale, les signaux sensoriels et les réponses comportementales et physiologiques. Le cortex préfrontal ventro-médian permettrait ainsi, en intégrant l'ensemble des

informations internes et externes, de former une représentation abstraite pertinente de l'environnement afin de déterminer les réponses comportementales et physiologiques adaptées. Autrement dit, cette région représenterait le *task-set* acteur. Le rôle du cortex préfrontal ventro-médian regroupe ainsi plusieurs processus :

i) L'élaboration d'une représentation de l'environnement d'après les signaux perçus ;

ii) L'utilisation de l'expérience passée pour établir une prédiction des conséquences des choix à venir ;

iii) L'évaluation des conséquences (positives ou négatives) ;

iv) Le déclenchement des réponses physiologiques et émotionnelles appropriées.

Ainsi, le rôle du cortex préfrontal ventro-médian ne se réduit par à l'évaluation d'une "valeur" unique d'un stimulus ou d'une action, mais permet de lier une situation particulière à un comportement spécifique. Autrement dit, le rôle du cortex préfrontal ventro-médian n'est pas de déterminer si un objet est menaçant ou appâtant, mais plutôt de déterminer la réponse comportementale (fuir, affronter, consommer, protéger, nourrir, se reposer, etc.) la plus adaptée à la situation donnée.

Le processus clé correspond à la combinaison d'unités élémentaires d'informations provenant des systèmes sensoriels ou de la mémoire à long terme en une représentation de la manière dont l'organisme se situe dans son environnement, ce qui permet ensuite d'établir des prédictions au sujet des évènements futurs.

Le cortex préfrontal ventro-médian semble donc représenter un bon candidat pour représenter le *task-set* acteur pour guider le comportement et inférer lorsque celui-ci n'est plus pertinent.

5.2 Cortex frontopolaire et monitorage des stratégies alternatives

Le cortex frontopolaire est la partie la plus antérieure du cortex préfrontal latéral. Il est défini d'après ses caractéristiques cytoarchitectonique (aire de Brodmann 10). Il se développe tardivement, à la fois d'un point de vue phylogénétique et ontogénétique [210] et est particulièrement développé chez l'humain, en comparaison des autres primates. Le cortex frontopolaire possède des connections réciproques importantes avec des régions préfrontales voisines : le cortex préfrontal ventro-médian et le cortex latéral [183].

Il est resté longtemps un mystère. On en savait peu sur cette région, principalement parce que sa localisation, sous un sinus (présence d'air), rend les enregistrements électrophysiologiques assez difficiles. Ceci a changé avec l'avènement des techniques modernes d'imagerie.

Il a été montré que le cortex frontopolaire est impliqué dans la cognition humaine de haut niveau, dont le raisonnement, la résolution de problème et le *multitasking*. Il se situe au sommet d'une hiérarchie de processus de contrôle

cognitif qui sous-tendent la coordination d'actions et de pensées en fonctions de buts internes [10][139].

5.2.1 Etudes de neuropsychologie

Les patients lésés au niveau du cortex frontopolaire ne souffrent pas de handicaps majeurs dans des tests de perception, de langage et d'intelligence, mais montrent de grandes difficultés à prendre des décisions dans des environnements incertains et ouverts, typiques de la vie quotidienne. En particulier, ils ne parviennent plus à réaliser des tâches de *multitasking* [37].

5.2.2 Etudes de neuroimagerie

Des études de neuroimagerie chez des sujets sains ont montré que le *multitasking* implique de façon robuste le cortex frontopolaire [28][137][138]. Il est plus précisément impliqué dans le maintien de tâches alternatives comme nous allons le décrire plus en détails.

Cortex frontopolaire et *branching*

Kœchlin et collaborateurs ont montré l'implication du cortex frontopolaire dans le processus de *branching* (d'embranchement), c'est-à-dire la mise en attente d'une tâche secondaire pendant l'exécution d'une tâche primaire puis le retour à la tâche secondaire lorsque la tâche primaire est achevée. Le cortex frontopolaire constitue ainsi un outil de stockage temporaire nécessaire pour le maintient d'une tâche en attente d'exécution future et permet d'éviter des interférences avec la tâche en cours [137][139].

Cette mise en attente de règle permet de contourner l'une des limitations majeurs du système exécutif qui est sont caractère sériel : une seule règle est exécutée à la fois.

Par la suite, Charron et Kœchlin ont construit un paradigme dans lequel des sujets doivent exécuter des tâches sur une série de lettres. Une tâche primaire peut être soit mise en attente, soit terminée, afin de permettre l'exécution d'une tâche secondaire [45]. La mise en attente correspond à la condition de *branching*, et la terminaison à une condition de *switch*. De plus, le niveau de récompense de chaque tâche est modulé. Les données IRM montrent que le cortex cingulaire antérieur dorsal gauche et droit, qui sont conjointement responsables de la motivation dans le cas d'une tâche isolée, se divisent en présence de deux tâches dans la condition de *branching*, chaque côté encodant alors la récompense associée à l'une des tâches. Par ailleurs, des activations bilatérales apparaissent dans le cortex frontopolaire dans la condition de *branching*. Cette étude explore ainsi la capacité du cerveau, non seulement à maintenir des règles en attente, mais également à conserver le niveau de motivation associé à une règle alternative.

Cortex frontopolaire et monitorage de strategies alternatives

Dans les expériences précédentes, le passage d'une règle à une autre est indiqué par un indice contextuel, en l'occurrence un changement de casse (passage de lettres majuscules à minuscules). Cependant, la flexibilité comportementale nécessite de pouvoir s'adapter même en l'absence de signes explicites de l'environnement. Boorman et collaborateurs étudient cette question à l'aide d'un paradigme où les sujets doivent choisir de presser un bouton, à gauche ou à droite, pour obtenir une récompense monétaire affichée à l'écran et variant d'un essai à l'autre [23]. Chaque bouton permet d'obtenir la récompense selon une certaine probabilité mais aucun indice n'indique quelle est la meilleure option (celle qui a la plus forte probabilité d'obtention de la récompense). Les sujets doivent donc déduire cette probabilité par essais et erreurs. De plus, la "meilleure" option peut changer. Les auteurs utilisent un modèle bayésien pour estimer les probabilités de récompense associée à chaque option, et les données IRM montrent une corrélation entre l'activité du cortex frontopolaire et la valeur de l'option non choisie, relativement à l'option choisie.

Par ailleurs, ce modèle bayésien permet de prédire les *switches* comportementaux, c'est-à-dire les points à partir desquels les valeurs des options s'inversent, conduisant les sujets à changer d'option. Les données montrent une activation du cortex frontopolaire dans les essais précédant immédiatement le *switch*, et un changement de connectivité fonctionnelle avec le cortex pariétal et les régions prémotrices. Le cortex préfrontal ventro-médian apparaît quant à lui coder la valeur relative de l'option choisie, un résultat qui rejoint les études déjà citées. Les auteurs proposent ainsi que le cortex frontopolaire et le cortex préfrontal ventro-médian jouent deux rôles complémentaires dans la prise de décision : tandis que le cortex frontopolaire collecte sur le long terme les preuves en faveur d'un éventuel *switch* destiné à adapter le comportement, le cortex préfrontal ventro-médian reflète sur le court terme la valeur de l'option courante par rapport aux options alternatives.

Ce résultat est confirmé par l'étude de Donoso, Collins et Kœchlin qui observent une corrélation du cortex frontopolaire avec la confiance attribuée aux *task-sets* alternatifs, une mesure qui permet d'inférer si l'une de ces stratégies alternatives est adaptée aux contingences de l'environnement [74].

Chapitre 6

Bases cérébrales du *switch*

Le cortex cingulaire antérieur et le cortex préfrontal dorsolatéral sont souvent activés ensemble lors de tâches cognitives telles que le *task-switching* (voir figure 6.1). Ils représentent donc de bons candidats de substrats cérébraux du *switch*. Les fonctions de ces régions sont cependant dissociables.

Le *task-switching* implique en effet au moins deux mécanismes neuraux distincts, l'un de surveillance et de détection du conflit (preuve d'un changement de contingences et de la nécessité d'effectuer un *switch*) et l'autre d'organisation des priorités entre les *task-sets* en compétition [63][122]. Ces deux fonctions, comme le suggère une théorie influente, seraient prises en charge par le cortex cingulaire antérieur et le cortex préfrontal dorsolatéral, respectivement [129][24].

Cependant, on trouve dans la littérature un grand nombre de fonctions attribuées à ces deux régions. Nous en faisons ici une brève synthèse.

6.1 Cortex cingulaire antérieur

Le cortex cingulaire antérieur est l'une des régions cérébrale dont les fonctions sont le plus débattues dans la littérature. Nous répertorions ici certains des différents rôles attribués à cette région.

6.1.1 La détection d'erreurs

Il arrive que même dans des tâches simples de discrimination de cible, les sujets fassent des erreurs. C'est dans ce type de situations que l'on peut faire l'expérience subjective du 'Oups ! Je me suis trompé'. Si l'on enregistre l'activité cérébrale par électroencéphalographie lors de ces évènements, on peut déterminer les substrats cérébraux correspondant à cette expérience [86][93][205]. Les potentiels évoqués enregistrés quand la réponse est correcte sont soustraits des potentiels enregistrés lors d'une erreur pour déterminer le signal correspondant à l'erreur. La reconstruction de source montre que le cortex cingulaire antérieur est lié à ce signal (*error-related negativity*)[107]. De même, en IRM fonctionnelle, on

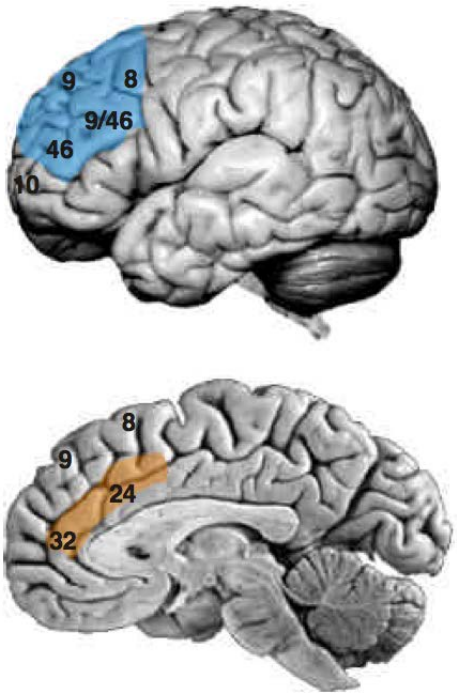


FIGURE 6.1 – **Le cortex préfrontal dorso-latéral (haut) et le cortex cingulaire antérieur (bas).**

observe une activation significative dans le cortex préfrontal médian qui corrèle avec les essais où le sujet a fait une erreur [130]. Ce type de résultats suggère donc que le cortex cingulaire antérieur, en particulier dans sa partie dorsale, réagirait aux erreurs et aux renforcements négatifs. Ces observations sont d'ailleurs nombreuses et reproductibles [108][165][236][188][199].

6.1.2 La détection du conflit

Un ensemble de travaux suggèrent quant à eux que le cortex préfrontal médian, dont le cortex cingulaire antérieur mais aussi l'aire motrice supplémentaire, joue un rôle dans la détection du conflit entre plusieurs réponses possibles, notamment lorsque des réponses automatiques sont inappropriées et doivent être inhibées [147][26][228][129][188][197][31][13]. Le cortex cingulaire antérieur (ou plus généralement des régions médiales du cortex préfrontal) signalerait ainsi au cortex préfrontal dorsolatéral la nécessité d'augmenter le niveau de contrôle *top-down* qu'il exerce, afin de faciliter le prochain changement de *task-set*.

Le cortex cingulaire antérieur pourrait ainsi être considéré comme une région cérébrale qui signale le besoin d'ajuster le comportement par la détection de conflit, d'erreurs ou de renforcements négatifs. Toutefois, d'autres études montrent que le cortex cingulaire antérieur réagit également aux renforcements positifs, en particulier lorsque ceux-ci apportent des informations qui permettent de guider le comportement futur [241][125][151][203]. Par exemple, le cortex cingulaire antérieur s'active au moment d'une récompense positive inattendue. Ainsi,

il est possible que la littérature qui prône la détection des renforcements négatifs comme fonction du cortex cingulaire antérieur, émerge en partie de biais expérimentaux. En effet, dans la plupart des protocoles, les récompenses positives sont en général plus nombreuses que les récompenses négatives. Les récompenses négatives sont donc à la fois plus incertaines mais aussi plus informatives. La fonction du cortex cingulaire antérieur ne peut donc se limiter à la détection du conflit et de l'erreur mais doit intégrer une notion d'incertitude et de prédiction.

6.1.3 Historique des récompenses et prédictions futures

Sélectionner la meilleure action possible parmi plusieurs alternatives nécessite souvent de tenir compte de l'expérience passée. Les choix peuvent être guidés par des indices ou des instructions, mais dans beaucoup de situations le comportement volontaire nécessite la sélection d'actions sur la base de la valeur intrinsèque de chaque option disponible [156]. Selon des théories de l'apprentissage, cette valeur intrinsèque est calculée d'après l'historique récent des résultats obtenus suite à ce choix, que les résultats soient positifs ou négatifs [227]. Certains auteurs suggèrent que le cortex cingulaire antérieur joue un rôle essentiel dans ce processus.

Ainsi, dans une étude réalisée chez le singe, Kennerley et collaborateurs montrent que le rôle du cortex cingulaire antérieur n'est pas la détection ou la correction des erreurs en tant que telles, mais consiste de manière plus générale à guider le choix en se fondant sur l'historique des actions réalisées et des renforcements obtenus [128]. Dans un premier temps, les auteurs observent que les singes lésés au niveau du cortex cingulaire antérieur ne sont pas moins bons que les singes contrôles dans les essais suivant une erreur. Cependant, les singes lésés sont moins enclins à répéter une réponse précédemment récompensée. Ainsi, plutôt qu'un déficit de détection d'erreur, ce type d'observation suggère que le cortex cingulaire antérieur est essentiel pour la sélection d'actions récompensées par le passé. Autrement dit, le cortex cingulaire antérieur n'utiliserait pas uniquement les erreurs pour guider l'action, mais intégrerait l'historique des renforcements au cours du temps pour représenter la valeur de chaque option. Et effectivement, dans un second temps, les auteurs observent que l'influence de l'historique des récompenses est significativement moins importantes chez les singes lésés que chez les singes contrôles. Les singes lésés semblent incapables d'utiliser l'historique des récompenses pour estimer la valeur de chaque action et agir de façon optimale. Le cortex cingulaire antérieur permettrait donc de construire, modifier et/ou utiliser les contingences actions-résultats pour estimer les valeurs de chaque option disponible. Ces valeurs peuvent ensuite être utilisées pour guider le comportement de façon optimale. Ce processus est indépendant du fait que l'historique des renforcements contient des résultats positifs (récompenses) ou négatifs (erreurs). Ainsi, selon ces auteurs, même en l'absence de cortex cingulaire antérieur, les singes conservent des représentations de la valence des récompenses, certains aspects du coût de l'action ainsi que la mesure combinée de ces deux facteurs, mais sont incapables d'utiliser la probabilité, associée à chaque option, d'obtenir

une récompense.

Dans la même ligne de pensée, Rushworth et Behrens suggèrent que l'activité du cortex cingulaire antérieur, dans la période précédant une décision, encode la propension d'une action à faciliter les décisions futures, que ce soit en terme de coûts et de gains immédiats, ou en terme de fiabilité de l'information [200]. Lors de l'observation du résultat d'une action, l'activité du cortex cingulaire antérieur encoderait une mesure de la pertinence de l'information obtenue pour les décisions futures.

6.1.4 La volatilité

Afin de prendre des décisions appropriées, il est nécessaire de tenir compte de l'historique des actions et des renforcements. Cependant, le poids de cet historique pour une décision dépend de la volatilité (ou la variabilité) de l'environnement. En effet, lorsqu'un environnement change fréquemment, il faut donner peu de poids aux évènements anciens ; tandis que dans un environnement stable, il faut limiter l'influence des évènements immédiats.

Behrens et collaborateurs démontrent que le cortex cingulaire antérieur contribue à évaluer la volatilité et à régler en conséquence la vitesse d'apprentissage [18]. Dans cette étude, les sujets effectuent une tâche de "suivi de probabilité" (*probability tracking*) dans laquelle un des deux stimuli présentés à l'écran est récompensé avec la probabilité p, l'autre avec la probabilité 1-p. La valeur de cette probabilité p change plusieurs fois au cours de l'expérience. De façon cruciale, la durée de stabilité de p change au cours de l'expérience, représentant un environnement incertain changeant plus ou moins vite (la volatilité est plus ou moins grande). Les auteurs de cette étude modélisent, à l'aide d'un modèle bayésien, l'apprentissage non seulement des probabilités p, mais aussi de la volatilité de l'environnement à tout instant. Cette volatilité guide la vitesse d'apprentissage de p. L'étude montre une corrélation de l'activité du cortex cingulaire antérieur avec la volatilité perçue par les sujets, comme inférée par le modèle. L'étude conclut donc sur le rôle du cortex cingulaire antérieur pour le guidage de l'apprentissage dans des situations complexes, en prenant notamment en compte l'incertitude intrinsèque de l'environnement.

6.1.5 Le *switch*

Plusieurs études chez le singe et l'humain ont montré une corrélation entre l'activité du cortex cingulaire antérieur et un changement abrupt de comportement (*switch*).

Quilodran et collaborateurs ont par exemple étudié le rôle du cortex cingulaire antérieur à l'aide d'une expérience réalisée chez le singe [185]. Dans ce paradigme, les animaux doivent choisir entre quatre stimuli différents, dont un seul est récompensé. Chaque épisode, ou suite d'essais, se compose d'une période

d'exploration dévolue à la recherche du stimulus récompensant, et d'une période d'exploitation durant laquelle la bonne réponse est répétée plusieurs fois. Les données électrophysiologiques montrent un rôle du cortex cingulaire antérieur lors du changement d'état entre exploration et exploitation.

Chez l'humain, Kolling et collaborateurs ont élaboré un protocole d'IRM fonctionnelle dans lequel les participants sont amenés à effectuer des choix d'exploration ou d'exploitation [141]. Les données montrent que l'activité du cortex cingulaire antérieur corrèle, lors du choix entre exploration et exploitation, avec la valeur estimée de l'exploration, et lors du choix entre deux stimuli, avec la différence de valeur entre l'option choisie et l'option non choisie. Ainsi le cortex cingulaire antérieur coderait à la fois le bénéfice d'une entrée en exploration, et le bénéfice relatif d'un choix par rapport à un autre. Les auteurs suggèrent donc que le cortex cingulaire antérieur estime la valeur du *switch*, c'est-à-dire l'intérêt à changer de comportement en faveur d'une action alternative à l'action choisie, ou d'un choix alternatif au choix par défaut.

Hyafil et Kœchlin utilisent un paradigme de *task-switching* combiné à un test de Stroop spatial pour explorer la flexibilité du système de contrôle exécutif [110]. Dans la tâche proposée aux sujets, des stimuli composés d'un triangle et d'un mot sont présentés à l'écran. Ces deux éléments peuvent être congruents (par exemple un triangle pointant vers le haut en même temps que le mot "haut"), ou incongruents (par exemple un triangle pointant vers la gauche en même temps que le mot "bas"). Les sujets doivent répondre en déplaçant un joystick, soit en fonction de l'orientation du triangle, soit en fonction de la signification du mot. Ces deux tâches, spatiale et verbale, sont indicées par la couleur du triangle. D'un essai à l'autre, si la couleur est inchangée, les participants doivent poursuivre la même tâche, si la couleur change, les participants doivent changer de tâche. Lors des essais de *switch*, les temps de réactions sont plus longs que pour les autres essais (résultat classique de *task-switching*). Les temps de réactions sont d'autant plus longs lorsque les sujets passent d'un essai incongruent à un autre essai incongruent. Les données d'imagerie montrent une activité du cortex cingulaire antérieur au moment des *switches*, indépendamment du type d'essai : cet effet ne différencie donc pas les essais congruents ou incongruents. Les auteurs concluent que le cortex cingulaire antérieur met à jour la priorité associée au nouveau *task-set*, quelque soit le niveau de contrôle nécessaire. Ceci est cohérent avec l'idée que le cortex cingulaire antérieur contribue à la prise de décision en associant des renforcements aux actions, permettant ainsi aux ressources d'être allouées à une tâche donnée sur la base de sa signification motivationnelle [198][240]. De plus, le cortex cingulaire antérieur semble jouer un rôle lors de la configuration d'un *task-set* pour un épisode futur : par exemple en répondant à des indices d'instruction [75], particulièrement lorsque ces indices portent une information motivationnelle claire [31].

Cette fonction de reconfiguration des associations liées à un nouveau *task-set* est également confirmée avec l'étude menée par Karlsson et collaborateurs chez le rat [127]. Les auteurs constatent en effet que des augmentations transitoires de l'activité dans le cortex préfrontal médian accompagnent les périodes où les croyances de l'animal sont modifiées après un changement de l'environnement.

L'activité dans la majorité des neurones de cette région subit des changements marqués, abrupts et coordonnés lorsque la croyance antérieure est abandonnée en faveur de l'exploration de stratégies alternatives. Cette dynamique reflète un *switch* du réseau vers un état d'instabilité, qui diminue au cours de la période d'exploration au fur et à mesure que de nouvelles représentations stables sont formées.

6.1.6 Une vision unifiée

Alexander et Brown tentent de faire une synthèse des différentes théories concernant la fonction du cortex cingulaire antérieur. Ils proposent que cette région a pour rôle de prédire les résultats des actions, qu'il s'agisse ou non d'une récompense [2]. Cette théorie permettrait de réinterpréter les signaux d'erreurs, de conflit et de volatilité. Le cortex cingulaire antérieur serait par ailleurs en mesure de détecter des résultats surprenants, lorsque la conséquence de l'action n'a pas été correctement prédite. De manière plus précise, le modèle suggère que des neurones individuels génèrent des signaux qui reflètent une prédiction de la probabilité et du moment d'obtention de différents résultats possibles d'une action. Ces signaux sont ensuite inhibés lorsque le résultat prédit arrive effectivement. Par conséquent, l'activité de ce réseau de neurones est maximale lorsqu'un résultat attendu ne se réalise pas. Cela revient à dire que le cortex cingulaire antérieur signale, au moins en partie, la non réalisation inattendue d'un résultat prédit, ce qui est compatible avec un certain nombre d'observations expérimentales. Le modèle proposé se distingue par différents aspects. En particulier, il n'apprend pas lui-même les associations stimulus-réponse, mais intègre des *task-sets* existants. Il prédit ainsi les résultats attendus pour des actions effectuées selon ces *task-sets*, pour un stimulus donné. Par ailleurs, il permet le monitorage simultané de différentes prédictions et donc de différents résultats, chacun associé à une certaine probabilité d'occurrence. Enfin, le modèle est capable de détecter lorsqu'un événement, qu'il soit récompensant ou non, ne se produit pas alors qu'il était attendu. Ce modèle s'est montré pertinent pour expliquer une grande variété de données issues d'expériences chez l'humain et le singe, et les auteurs suggèrent que la fonction du cortex cingulaire antérieur ainsi proposée en fait une région clé pour l'évaluation des actions avant leur exécution, pondérant les risques et les bénéfices anticipés, particulièrement dans des situations nouvelles.

6.2 Le cortex préfrontal dorso-latéral

Le cortex préfrontal dorso-latéral inclut les aires de Brodmann 8A, 46, 9/46 et 9, bien que souvent les données sur cette région concernent plus spécifiquement les aires 46 et 9/46 (voir figure 6.1). Plusieurs théories ont été proposées concernant ses fonctions [99][91][159]. Dans l'ensemble, les auteurs s'accordent sur le fait que le cortex préfrontal dorso-latéral est la région cérébrale responsable de l'implémentation du contrôle cognitif, c'est-à-dire de la coordination des pensées et actions en fonctions d'objectifs internes et de l'environnement. Cette fonction repose sur des interactions avec des régions cérébrales prémotrices et associatives [159].

6.2.1 Etudes de neuropsychologie

Les patients atteints d'une lésion de la portion latérale du cortex préfrontal ont en commun des difficultés dans la réalisation de tâches complexes qui requièrent d'inhiber les informations ou les actions non pertinentes, d'utiliser une représentation mentale pour guider l'action, de manipuler mentalement ces représentations, d'organiser une série d'action pour parvenir à un but, de répartir son attention entre plusieurs tâches et de mener de front plusieurs activités simultanément, mais aussi de conduire un raisonnement abstrait, de former des concepts et de catégoriser, de récupérer les informations stockées en mémoire à long terme [97][90][37][96]. Ces troubles sont habituellement regroupés sous le label "syndrome dysexécutif".

Par exemple, les patients atteints de lésions du cortex préfrontal dorso-latéral présentent des déficits dans la tâche du *Wisconsin Card Sorting Test*. Ils continuent à appliquer la première règle de tri des cartes même si celle-ci a changé et qu'ils en sont informés. Autrement dit ils ne parviennent pas à inhiber une réponse précédemment correcte.

6.2.2 Etudes de neuroimagerie

Le modèle en cascade

Différentes études suggèrent une organisation hiérarchique du cortex préfrontal dorso-latéral [9][10]. En particulier, Kœchlin et collaborateurs proposent que le cortex préfrontal dorso-latéral se structure selon une hiérarchie de représentation allant du cortex prémoteur aux régions les plus antérieures du cortex préfrontal [132].

Plus précisément, Kœchlin et collaborateurs font l'hypothèse que le contrôle cognitif fait intervenir au moins trois régions fonctionnelles. La première région, située au niveau du cortex prémoteur, est responsable du contrôle sensoriel : elle sélectionne les actions motrices en réponse aux stimuli observés. La deuxième

région, située au niveau du cortex préfrontal dorso-latéral caudal (typiquement les aires de Brodmann 9, 44 et 45), est responsable du contrôle contextuel : elle sélectionne les associations stimulus-réponse sur la base des signaux contextuels qui accompagnent les stimuli (la règle d'association est connue). Finalement, la troisième région, située au niveau de cortex préfrontal dorso-latéral rostral (typiquement l'aire de Brodmann 46), est responsable du contrôle épisodique : elle sélectionne les *task-sets* adéquats en fonction de l'épisode en cours, donc en fonction d'un indice épisodique passé ou d'un objectif interne. Le modèle proposé prédit que les contrôles épisodique, contextuel et sensoriel contribuent de manière cumulative au contrôle cognitif, l'information disponible pour orienter le comportement s'accumulant progressivement des régions rostrales aux régions caudales. Ce modèle est nommé "modèle en cascade" pour refléter l'idée d'une cascade informationnelle allant du cortex préfrontal dorso-latéral rostral au cortex prémoteur.

Cette hiérarchie fonctionnelle a par ailleurs été validée par des études de lésions. Badre et collaborateurs ont soumis douze patients atteints de lésions focales du lobe frontal à des tâches dont les demandes en contrôle cognitif varient [11]. Plus précisément, les patients étaient soumis à des choix dans des conditions toujours plus abstraites, chaque condition présentant un niveau de contrôle différent. Les données comportementales montrent des déficits de performance d'autant plus importants que les tâches à accomplir sont abstraites. Les données IRM, montrent par ailleurs que le déficit dépend de façon sélective de la région cérébrale lésée, dans la logique du modèle en cascade : le site de la lésion brise la continuité de l'information dans l'axe rostro-caudal, aboutissant à une chute de la performance dans la forme de contrôle associée à ce site, et dans les formes de contrôle d'abstraction supérieure, tout en laissant intactes celles de moindre abstraction.

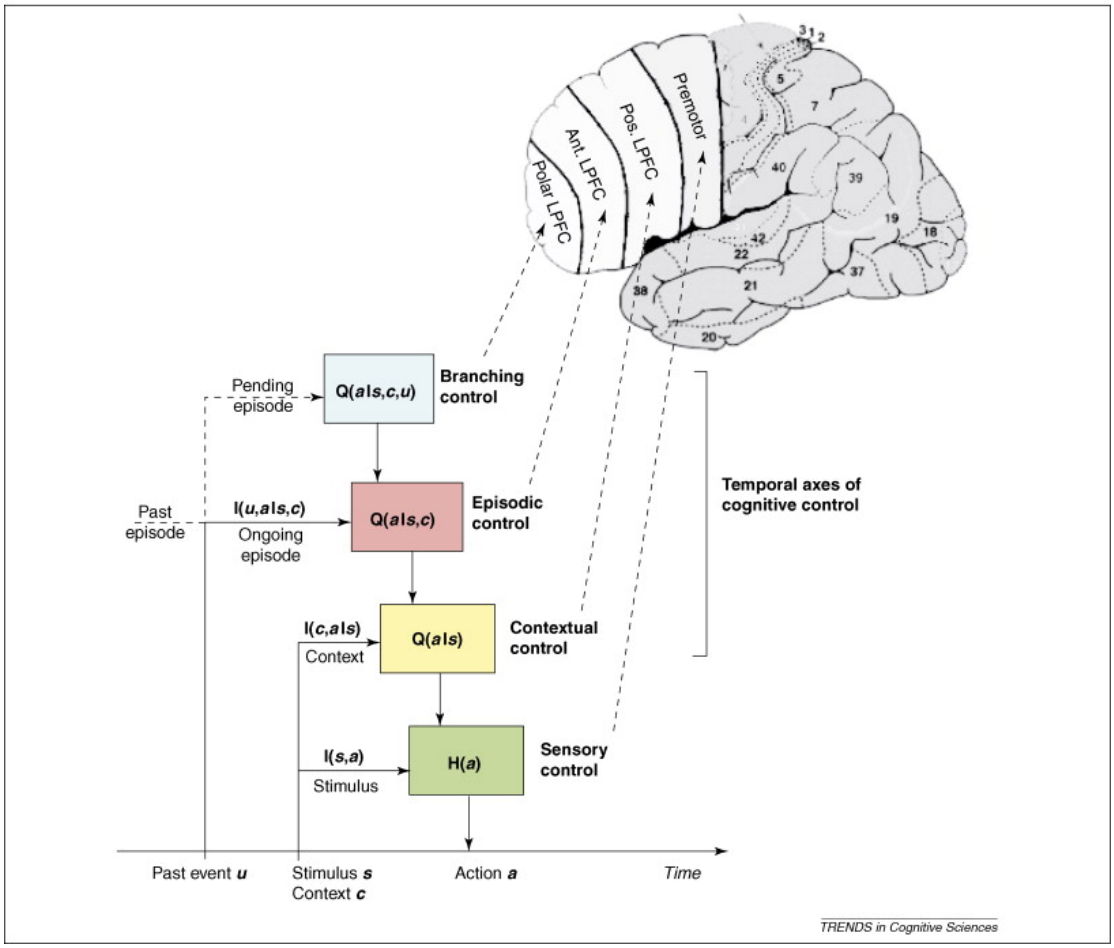


FIGURE 6.2 – **Le modèle en cascade.** Le contrôle cognitif opère selon trois niveaux imbriqués de processus de contrôle (contextuel, épisodique et de *branching*), implémentés des régions préfrontales postérieures aux régions préfrontales polaires. H(a) mesure la quantité totale d'information nécessaire pour sélectionner l'action A. Cette quantité est traitée dans le cortex prémoteur. H(a) est la somme de deux termes : l'information *bottom-up* provenant stimulus s (I(s,a), contrôle sensori-moteur) et l'information *top-down* Q(a|s) traitée dans le cortex préfrontal latéral postérieur (Pos. LPFC), mesure du contrôle cognitif. Le contrôle cognitif, quant à lui, est la somme de deux termes : l'information *bottom-up* provenant du contexte c dans lequel le stimulus s apparait (I(c,a|s), contrôle contextuel) ; et l'information *top-down* Q(a|s,c) traitée dans le cortex préfrontal latéral antérieur (Ant. LPFC). Enfin, ce dernier terme de contrôle est la somme de l'information *bottom-up* provenant d'un évènement passé u (I(u,a)|s,c), contrôle épisodique) et de l'information *top-down* traitée dans le cortex préfrontal latéral polaire (contrôle de *branching*). Le contrôle de *branching* correspond à l'information provenant d'évènements qui précèdent u et sont maintenus dans un état d'attente jusqu'à l'achèvement de l'épisode en cours.

Acquisition d'associations abstraites

Dans le protocole expérimental du modèle en cascade, les associations entre contexte et tâche sont fournies au sujet (pour chaque contexte, le sujet sait quelle tâche il doit réaliser). On peut cependant se demander si, lorsque ces associations sont présentes mais que le sujet n'en est pas informé, elles sont tout de même utilisées par le système exécutif. Badre et collaborateurs ont tenté de répondre à cette question [12]. Dans leur protocole, ils présentent aux sujets des stimuli qui peuvent varier par leur forme et leur orientation. Dans une condition contrôle, chaque objet, défini par sa forme et son orientation, est associé à une réponse qui doit être apprise spécifiquement. Dans une condition hiérarchique, la couleur du cadre qui entoure l'objet fournit une information sur la dimension qui doit être prise en compte : par exemple, si le cadre est rouge, les objets doivent être classés selon leur forme, et s'il est bleu, selon leur orientation. La couleur du cadre fournit ainsi un indice contextuel, qui permet de hiérarchiser l'information en indiçant deux ensembles d'associations, c'est-à-dire deux *task-sets*. Mais cette information n'est pas donnée d'emblée aux sujets : ceux-ci doivent donc apprendre la règle et la réutiliser pour améliorer leur performance. Les données comportementales montrent un brusque gain de performance lorsque les sujets apprennent les *task-sets* et leurs indices, dans la condition hiérarchique. Les données IRM permettent, quant à elles, de visualiser le réseau cérébral qui s'active davantage dans la condition hiérarchique que dans la condition contrôle : ce réseau englobe différentes régions du cortex préfrontal dorso-latéral, en particulier les aires de Brodmann 44 et 46. Les auteurs suggèrent ainsi que face à un nouvel environnement, le cerveau recherche des associations entre contexte et action à plusieurs niveaux d'abstraction simultanément, une capacité qui sous-tend l'adaptation rapide et la généralisation de connaissances à de nouveaux problèmes.

Connectivité

Le rôle central du cortex préfrontal dorso-latéral dans le contrôle est attesté par des études de connectivité. Cole et collaborateurs relèvent ainsi qu'en mesurant la connectivité globale, le cortex préfrontal dorso-latéral apparaît comme une nœud essentiel dans une tâche de mémoire de travail à forte demande en contrôle [52]. Les auteurs démontrent que le cortex préfrontal dorso-latéral s'apparente à une "plaque tournante", influente à l'échelle du cerveau entier, qui facilite la mise en œuvre de processus de contrôle. Par ailleurs, une méta-analyse menée par Gläscher et collaborateurs sur plus de trois cents personnes atteintes de lésions cérébrales focales, dont la moitié impliquant le cortex préfrontal dorso-latéral, montre l'existence de deux réseaux préfrontaux fonctionnels et anatomiques distincts [95]. Le premier réseau, impliquant le cortex préfrontal dorso-latéral et le cortex cingulaire antérieur, est associé au contrôle cognitif, au monitorage des conflits et au *switch*. Le deuxième réseau, impliquant le cortex orbito-frontal, le cortex préfrontal ventro-médian et le cortex frontopolaire, est quant à lui associé à la prise de décision fondée sur la valeur. Cette étude confirme ainsi, par

la grande quantité de données traitées, la robustesse des fonctions associées au cortex préfrontal dorso-latéral.

6.2.3 Etudes d'électrophysiologie

Des travaux d'électrophysiologie permettent de confirmer le rôle du cortex préfrontal dorso-latéral dans le contrôle cognitif et l'exécution de règles comportementales. Buschman et collaborateurs présentent à des singes une série de stimuli dont la couleur et l'orientation sont variables et auxquels les singes doivent répondre en utilisant l'une de ces deux dimensions [39]. La règle à utiliser, répondre en fonction de la couleur ou de l'orientation, reste constante pendant un certain nombre d'essais puis change : les changements sont indicés par la couleur d'une bordure entourant le stimulus. Les animaux montrent l'effet classique de *switch-cost* en répondant plus lentement après un *switch*. Les données électrophysiologiques montrent par ailleurs des phénomènes de synchronisation entre des ensembles de neurones qui représentent des règles. Les auteurs suggèrent que différents types de synchronies sont utilisées pour sélectionner l'ensemble correspondant à la règle pertinente, et pour désélectionner un ensemble inapproprié. Ainsi des phénomènes de synchronisation pourraient façonner dynamiquement des ensembles de neurones pertinents pour la tâche à accomplir, à partir d'un réseau initial composé de grands circuits qui se chevauchent.

Chapitre 7

Bases cérébrales de la récupération

Ainsi que nous l'avons décrit dans la première partie, la récupération correspond au rappel de représentations en mémoire de travail. Par conséquent, dans cette partie, nous passons en revue, d'une part, les études sur la mémoire de travail ; et, d'autre part, les études sur le processus de récupération en tant que tel.

7.1 Cortex préfrontal et mémoire de travail

Lorsque l'on écoute quelqu'un parler, on utilise notre mémoire de travail pour identifier les segments de phrases, associer le sujet au verbe puis aux objets pour en comprendre le sens. Lorsque l'on résout un problème d'arithmétique mentalement, que l'on mémorise un numéro de téléphone, que l'on planifie un mouvement d'échec ou que l'on suit une instruction verbale ou écrite, nous utilisons notre mémoire de travail. Il est en fait difficile de trouver une fonction cognitive qui n'engage pas la mémoire de travail.

De fait, la mémoire de travail est un concept central des sciences cognitives et l'un des sujets les plus étudiés en psychologie cognitive et en imagerie humaine. L'histoire et la richesse de ce concept ont été élaborées et synthétisées de façon exhaustive dans une littérature expérimentale volumineuse [6][124].

La mémoire de travail se définit comme un ensemble d'opérations mentales permettant le maintien, la manipulation et l'utilisation de représentations mentales en prévision d'une action [99]. Elle regroupe au moins deux ensembles de processus cognitifs distincts mais interdépendants :
1) La capacité de maintenir à court terme une représentation mentale,
2) La gestion exécutive, c'est-à-dire l'utilisation des informations maintenues en représentation interne.

Cette dissociation recoupe en grande partie le modèle psychologique de Baddeley [6][7] postulant l'existence d'une part de systèmes "tampons" permettant le

stockage sous une forme "active" de l'information et d'autre part, d'un "administrateur central", gestionnaire allouant les ressources attentionnelles et manipulant l'information maintenue.

Ces deux processus seraient d'ailleurs traités dans des aires cérébrales distinctes : les aires corticales pariétales postérieures pour le stockage et les aires préfrontales pour le traitement selon certains auteurs [217][216] ; ou au sein même du cortex préfrontal, dans différentes aires architectoniques, pour d'autres [180][174].

Quoi qu'il en soit, le cortex préfrontal latéral a été identifié comme un élément critique du système de mémoire de travail qui implique également d'autres aires corticales. Parmi elles le cortex pariétal postérieur, le cortex infero-temporal, le gyrus cingulaire et la formation hippocampique, qui ont tous des connexions avec de cortex préfrontal latéral [14][209][223].

7.1.1 Etudes chez le primate non humain

Des tâches de réponse différée ont permis de montrer l'importance critique du cortex préfrontal pour le maintien d'information en mémoire de travail chez le primate. Dans ces expériences, l'animal regarde de la nourriture placée sous un pot (parmi deux). Un écran opaque est ensuite abaissé et maintenu pendant une durée de plusieurs secondes à plusieurs minutes pendant lequel l'animal doit maintenir en mémoire l'information pertinente afin de choisir le pot correct lorsque l'écran est levé. Les singes avec des lésions du cortex préfrontal dorsolatéral ne parviennent plus à réaliser cette tâche [98]. Dans une variante de cette tâche, la *delayed matching to sample test*, on montre au singe un premier stimulus, s'ensuit un délai variable, puis le singe est récompensé s'il répond à un stimulus "test" identique au premier. La composante cruciale de ce protocole est la période de délai (pouvant aller jusqu'à 20 secondes), pendant laquelle le singe doit maintenir une représentation mentale d'un stimulus pertinent non visible. Des neurones préfrontaux latéraux émettent une réponse soutenue pendant les périodes de délai comme s'ils prolongeaient l'impact du stimulus ou qu'ils anticipaient son apparition [69].

Il existe plusieurs théories sur la spécialisation fonctionnelle du cortex frontal dans la mémoire de travail. Deux en particulier sont en compétition : la théorie "domaine spécifique" et la théorie "processus spécifique". Ces théories se concentrent notamment sur la dissociation entre les régions ventrales et dorsales du cortex préfrontal latéral.

La théorie de spécialisation selon de le domaine

Selon la théorie "domaine spécifique", le cortex préfrontal est le site des processus de mémoire de travail et différentes régions au sein du cortex préfrontal maintiennent temporairement différents types d'information [99][209]. Plus spécifiquement, le cortex préfrontal ventro-latéral serait responsable du maintien

de l'information concernant l'objet, telle que sa forme ; alors que le cortex préfrontal dorso-latéral serait responsable du maintien de sa localisation (position dans l'espace). Cette théorie se fonde en grande partie sur des études d'enregistrements électrophysiologiques chez le primate non humain et représente une extension rostrale de l'organisation "quoi/où" du système visuel que l'on trouve dans les régions postérieures [234]. Ainsi, Wilson et collègues observent que des cellules du cortex préfrontal ventral codent pour l'information de l'objet pendant un délai, alors que des cellules frontales dorsales codent pour l'information spatiale pendant ce même délai. Plus généralement, Goldman-Rakic et collaborateurs suggèrent que la distinction entre les caractéristiques de l'objet et sa localisation spatiale qui correspondrait à une dissociation ventro-dorsale au sein du cortex préfrontal inclut tous les composants de la mémoire de travail : les mécanismes de contrôle attentionnel, mnésiques et de réponse [100]. C'est-à-dire qu'il n'y aurait pas de spécialisation des différents processus de la mémoire de travail dans le cortex préfrontal, mais seulement une spécialisation des domaines sur lesquels ces processus opèrent.

La théorie de spécialisation selon les processus

Une théorie alternative est celle "processus spécifique" qui propose que la différence entre le cortex préfrontal ventro- et dorso-latéral repose non pas sur le type d'information maintenue mais sur le type de processus qui traite cette information [180]. Selon cette théorie, le cortex préfrontal ventro-latéral sous-tend des processus qui transfèrent, maintiennent et associent l'information en mémoire de travail. Cette information peut avoir été perçue récemment ou récupérée depuis la mémoire à long terme. Le cortex préfrontal dorso-latéral, quant à lui, sous-tend des processus plus complexes qui traitent l'information maintenue en mémoire de travail, tels que le monitorage et la planification de haut niveau.

Cette distinction anatomo-fonctionnelle peut être mise en évidence en comparant deux types de tâche de mémoire de travail : les *delayed matching tasks* et les *self-ordered tasks*. Les *delayed matching tasks* nécessitent simplement que le sujet décide si un stimulus test appartient à un ensemble de stimuli maintenu en mémoire. Cette tâche nécessite seulement du maintien. Dans les *self-ordered tasks*, le sujet doit également sélectionner un stimulus appartenant à un ensemble de stimuli présenté précédemment, mais, au cours des essais, chaque stimulus ne doit être sélectionné qu'une fois (pas répétition possible). Ceci nécessite que le sujet non seulement sélectionne des stimuli dans un ensemble maintenu en mémoire mais mette également à jour l'ensemble des réponses précédentes. Selon Petrides et collaborateurs, une tâche de *delayed matching* devrait impliquer le cortex préfrontal ventro-latéral, alors qu'une tache de *self-ordering* devrait impliquer le cortex préfrontal dorso-latéral. En accord avec cette vision, des lésions du cortex préfrontal dorso-latéral chez les primates produisent des déficits dans les tâches de *self-ordering* mais pas spécialement dans les tâches de *delayed-matching*[182][175]. De plus, une revue de la littérature de D'Esposito et Postle montrent que des patients avec des lésions du cortex préfrontal dorso-latéral maintiennent de bonnes performances dans des tâches verbales ou spatiales simples

qui nécessitent uniquement le maintien d'un stimulus en mémoire sans aucune manipulation [80].

Bien que souvent placées en opposition, les théories "domaine spécifique" et "processus spécifique" ne sont pas nécessairement incompatibles. Le cortex préfrontal pourrait en effet être fonctionnellement dissociable selon le type d'information et le type de processus.

7.1.2 Etudes de neuroimagerie

Tâches de maintien d'information

Le modèle théorique de la mémoire de travail de Baddeley et Hitch [8] a beaucoup influencé le cadre des études de neuroimagerie fonctionnelle qui se sont avant tout intéressées au substrats cérébraux du maintien d'information.

En 1973, grâce à l'imagerie fonctionnelle, Risberg et Ingvar ont confirmé l'importance du cortex préfrontal humain pour la mémoire de travail. Ils ont en effet observé des activations hémodynamiques dans les lobes frontaux latéraux lors de tâches de mémorisation de nombres. Depuis, de nombreuses études d'imagerie ont rapporté des activations préfrontales latérales pendant des tâches de mémoire de travail pour des stimuli verbaux, perceptuels et spatiaux [181][70][142].

Maintien d'information verbale. Un test classique de maintien d'information en mémoire de travail est celui de Steinberg. On présente au sujet un ensemble de trois ou neuf stimuli, qui est ensuite retiré pendant plusieurs secondes puis un seul stimulus test est à nouveau présenté. Le sujet doit décider si ce stimulus test faisait partie de l'ensemble initial. En utilisant cette tâche, plusieurs études [178] observent des activations significatives dans plusieurs régions de l'hémisphère gauche : cortex préfrontal ventro-latéral, aires motrices et pariétales.

Maintien d'information spatiale. En utilisant le même type de tâche que précédemment mais où cette fois-ci l'information à mémoriser est l'information spatiale des stimuli, des auteurs observent des activations dans plusieurs régions de l'hémisphère droit, largement homologues à celles observées dans les tâches de maintien verbal [123][216][217] : cortex pariétal droit, cortex prémoteur dorsal droit et cortex préfrontal ventro-latéral droit.

Ainsi, contrairement à ce que suggèrent des études chez les primates, les études d'imagerie ne font pas de distinction ventro-dorsale. L'activation du cortex préfrontal associée au maintien d'information sur l'objet semble davantage latéralisée à gauche par rapport au maintien d'information spatiale, davantage latéralisée à droite.

Tâches de manipulation de l'information

Les tâches de *N-back* sont des tâches très classiques pour étudier la mémoire de travail, dans ses aspects à la fois de maintien et de manipulation d'information.

Dans la tâche du *2-back* par exemple, on montre au sujet une chaine de lettres, une lettre à la fois, et on lui demande de presser un bouton si la lettre est identique à celle apparue avant la précédente (voir figure 7.1). C'est une tâche exigeante qui nécessite le maintien en mémoire d'au moins deux informations à chaque pas de temps. Lors de l'exécution de cette tâche on observe une activation robuste dans le cortex préfrontal latéral [142].

Braver et collaborateurs ont fait varier la charge verbale de la mémoire de travail en augmentant le nombre N de 0 à 3 dans une tâche de N-back verbale [29]. Les aires dans lesquelles l'activité est une fonction linéaire croissante de la charge incluent le cortex préfrontal dorso-latéral, le cortex préfrontal ventro-latéral et le cortex pariétal de façon bilatérale, ainsi qu'un certain nombre d'aires motrices et pré-motrices à gauche. D'après les résultats sur le maintien résumés précédemment, les activations du cortex préfrontal ventro-latéral, du pariétal postérieur et motrices reflètent sans doute le réseau d'aires impliquées dans le maintien d'information verbale (par exemple le stockage et a répétition des n lettres les plus récentes). S'il en est ainsi, alors ces résultats impliquent que l'activation bilatérale supplémentaire du cortex préfrontal dorso-latéral sert à la manipulation d'information (par exemple la mise à jour des lettres devant être maintenues).

Dans une autre étude, Smith et collaborateurs [218] rapportent des activation bilatérales du cortex préfrontal dorso-latéral et du cortex frontopolaire dans une tâche de *3-back* verbale et spatiale. Ils observent de plus une légère latéralisation avec davantage d'activation dans le cortex préfrontal dorso-latéral gauche pour l'information verbale et à droite pour l'information spatiale [217]. Ces données suggèrent que les processus de manipulation dans le cortex préfrontal dorso-latéral sont latéralisés gauche/droite pour l'information verbale versus l'information spatiale, comme pour les processus de maintien dans le cortex préfrontal ventro-latéral. Ces études remettent à nouveau en question la dissociation spécifique dorso-ventrale du cortex préfrontal suggérée par Goldman-Rakic.

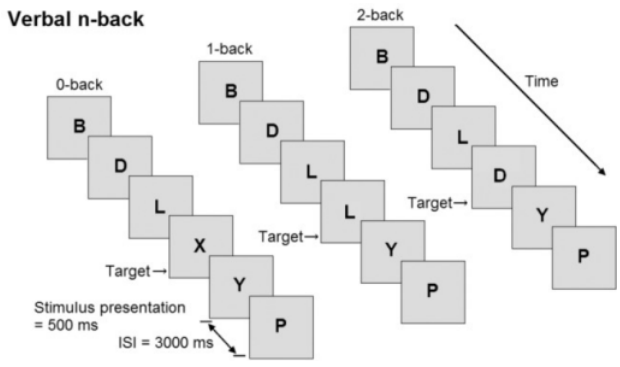


FIGURE 7.1 – **La tâche de *N-back*.**

Pour résumer, l'imagerie fonctionnelle de la mémoire de travail humaine a montré des divisions anatomiques au sein du cortex préfrontal latéral qui sous-tendent différents processus. Le cortex préfrontal ventro-latéral, par exemple, semble plus souvent activé pendant des tâches qui nécessitent du maintien d'information ; alors que le cortex préfrontal dorso-latéral est plus souvent activé pendant des tâches qui nécessitent de la manipulation d'information. Il semble également y avoir une latéralisation des processus du cortex préfrontal selon le type d'information en jeu : la comparaison de tâches verbales et spatiales suggère que le cortex préfrontal ventro-latéral gauche est principalement concerné par le maintien d'information verbale et le cortex préfrontal ventro-latéral droit avec le maintien d'information spatiale. Notons cependant que ces études s'intéressent principalement à une mémoire de type déclarative et non à une mémoire procédurale et à la manipulation d'objets plutôt qu'à la manipulation de *task-sets*.

7.2 Interactions entre le cortex préfrontal et le lobe temporal médian pour la récupération associative

7.2.1 Stockage en mémoire à long terme dans le cortex

La mémoire à long terme est représentée dans de multiples régions du système nerveux. Selon Fuster, il existe deux grandes catégories de mémoire à long terme : perceptuelle et exécutive, qui sont respectivement distribuées dans le cortex postérieur et le cortex frontal (voir figure 7.2) [90]. Dans ces deux secteurs du cortex, la mémoire à long terme est formée sous l'influence de structures du lobe temporal médian, notamment l'hippocampe et l'amygdale [221]. Par ailleurs, ces deux catégories de mémoire corticale, perceptuelle et exécutive, sont organisées de manière hiérarchique :

- La mémoire perceptuelle est acquise via les organes sensoriels. A son niveau le plus basique, elle inclut la mémoire des perceptions sensorielles élémentaires, dans le cortex. A des niveaux plus élevés et plus généraux, la mémoire perceptuelle comprend des mémoires sensorielles multimodales, dans le cortex dit "trans-modal" [158]. A des niveaux encore plus élevés, dans le cortex associatif supérieur, la mémoire perceptuelle inclut les mémoires épisodiques et sémantiques, qui constituent la mémoire dite déclarative. Enfin, au niveau le plus élevé, également dans le cortex associatif postérieur, on trouve la mémoire conceptuelle, la forme la plus générale et la plus abstraite des connaissances perceptuelles.

- La mémoire exécutive est hiérarchisée dans le cortex frontal. Au niveau le plus bas se trouve le cortex moteur primaire (mémoire motrice), qui représente les aspects les plus élémentaires du mouvement, définis par la contraction de muscles définis. Au dessus de cette couche de représentations élémentaires, le cortex prémoteur stocke en mémoire des programmes moteurs définis par un but et une trajectoire. Le cortex préfrontal latéral constitue le niveau le plus élevé

de la hiérarchie corticale frontale pour la représentation de la mémoire exécutive [90]. Cette région corticale semble clairement impliquée dans la représentation de schémas d'action complexes.

Un comportement adapté nécessite l'activation sélective et ordonnée de réseaux corticaux spécifiques. Cette activation sélective et ordonnée est essentielle pour la réalisation de comportements complexes. Elle a lieu dans un système de réseaux qui se chevauchent et s'étendent dans le cortex. Le cortex préfrontal joue un rôle crucial pour assurer la sélectivité et l'ordre de recrutement des réseaux corticaux. C'est ce que Shallice appelle "le contrôle attentionnel superviseur" du cortex préfrontal [166].

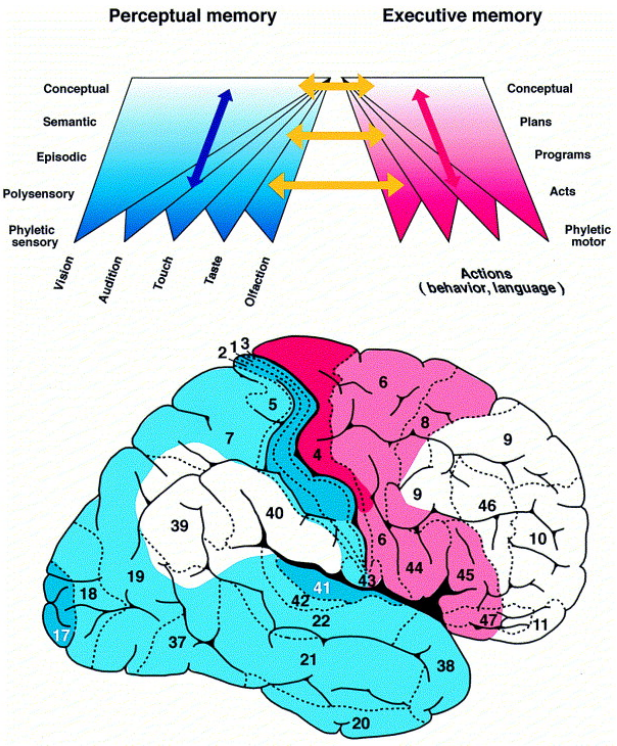


FIGURE 7.2 – **Organization hiérarchique des réseaux de mémoire dans le cortex.** En bleu : mémoire perceptuelle dans le cortex postérieur. En rose : mémoire exécutive dans le cortex frontal. Extrait de [90].

Reste à savoir par quels processus et dans quelles régions cérébrales cette information stockée au niveau cortical peut être récupérée.

7.2.2 Rôle du lobe temporal médian

Rôle général

Le rôle crucial des régions du lobe temporal médian pour la mémoire a été mis en évidence grâce au célèbre patient HM devenu profondément amnésique suite à une opération chirurgicale pour soulager une épilepsie (voir figure 7.3) [208]. Lors de cette opération une grande partie des deux lobes temporaux médians, dont des structures hippocampiques, ont été retirées. Ces structures semblent être impliquées dans l'ensemble des processus mnésiques (encodage, stockage, récupération) pour tout type de mémoire.

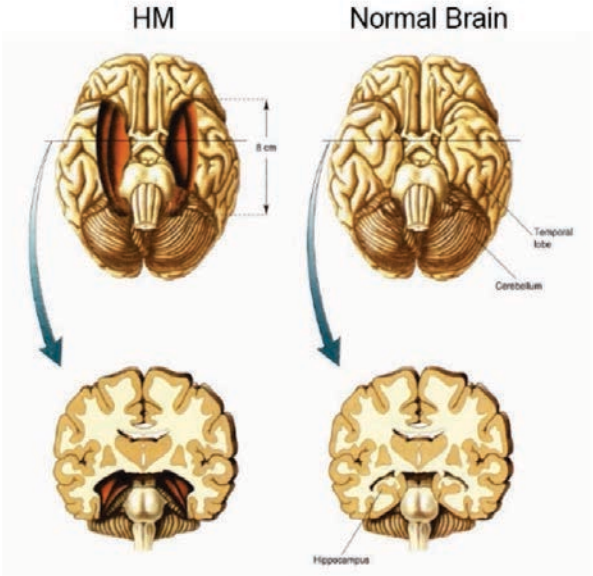


FIGURE 7.3 – **Excision du lobe temporal médian du patient HM. Haut :** vue du dessous ; **Bas :** coupe coronale.

Marr a été le premier à attribuer des opérations aux régions du lobe temporal médian, suggérant que la formation hippocampique classifie et indexe l'information entrante, la stocke temporairement en vue d'un transfert ultérieur aux régions néocorticales où elles sont réorganisées pendant le sommeil [149].

Au moment de la récupération, la présentation d'un sous-ensemble de l'information originelle active l'hippocampe et incite le réseau à réinstaurer le pattern original d'activité par le processus dit de *pattern completion*, c'est-à-dire le processus par lequel une représentation neurale stockée est réactivée par un indice correspondant à un sous-ensemble du pattern, permettant ainsi la récupération d'information stockée.

Ainsi, l'hippocampe ne contiendrait pas le contenu mnésique en lui-même mais serait capable de réactiver des patterns d'activité néocorticaux qui constituent le contenu mnésique [154]. En particulier, Turnock et collaborateurs proposent que cette région joue un rôle de filtre entre le cortex préfrontal et les noyaux sous-corticaux permettant de contrôler, en fonction du contexte, les associations

stimulus-réponses à utiliser [233]. L'hippocampe fournirait ainsi de l'information à d'autres structures qui prennent les décisions [120]. C'est donc un candidat possible pour l'implémentation des associations contexte/*task-set*.

Le modèle de *dual-process* : *recollection* versus *familiarity*

Il est largement admis que le processus mnésique de reconnaissance (*recognition*) comprend deux processus différents : la *recollection* (souvenir) et la $familiarity$ (le caractère familier) [248][243]. Le processus de *recollection* a lieu lorsque la présentation d'un objet à reconnaître suscite également le rappel d'information qualitative au sujet d'un épisode de l'étude. Cette information n'inclut pas seulement l'identité de l'objet étudié, mais également des détails au sujet du contexte dans lequel il a été étudié. Le processus de $familiarity$, quant à lui, correspond à un sentiment de "déjà vu" dénué d'information contextuelle. Ces deux processus permettent d'effectuer de simples jugements de reconnaissance "vieux versus nouveau" à la présentation d'un objet déjà vu ou non. En revanche, les jugements qui incluent un souvenir précis du contenu d'un épisode dépendent largement du processus de *recollection*.

Les caractéristiques fonctionnelles de ces deux processus et leurs substrats neuraux sont un sujet de débat intense. L'une des questions contentieuses est celle de savoir si ces deux processus dépendent de l'hippocampe de façon différentielle [81][220]. Trois théories sont proposées :

L'hippocampe est spécifique au processus de *recollection*. De nombreux auteurs soutiennent que la *recollection* mais pas la $familiarity$ dépend de l'hippocampe. Cette hypothèse est soutenue par plusieurs études chez les patients [1][106][152][118][148], chez l'animal [33] et chez l'humain sain en neuroimagerie fonctionnelle [51][71][82][247].

L'activité de l'hippocampe est corrélée à la "force" du souvenir. Selon une autre interprétation, l'activité de l'hippocampe lors de l'encodage et de la récupération serait le corrélat neuronal non pas du processus de *recollection*, mais d'un souvenir "fort", c'est-à-dire dont on est certain, que cette force soit due à la *recollection* ou à la $familiarity$ ou une combinaison des deux processus mnésiques [220][243].

L'activité de l'hippocampe encode l'information contextuelle. Une troisième interprétation alternative suggère que l'activité de l'hippocampe ne serait sensible ni à la variation de la force d'un signal mnésique, ni au sentiment subjectif de souvenir, mais reflèterait la quantité d'information contextuelle qui est encodée pendant un épisode de l'étude ou qui est récupérée en réponse à un objet test. En soutient de cette interprétation, Rugg et collaborateurs observe une activation de l'hippocampe droit lors de la récupération correcte de la source [195].

Ces deux processus mnésiques peuvent être mis en lien avec le modèle PROBE :

- La *recollection* correspondrait à la récupération par inférence en mémoire de travail. Autrement dit, il y aurait *recollection* lorsqu'un *task-set* possède une fiabilité suffisamment élevée pour être récupérée. Ceci pourrait également se produire lors d'un processus non modélisé dans le cadre du modèle PROBE : la récupération directe depuis la mémoire à long terme lorsqu'aucun *task-set* de la mémoire de travail n'est fiable.

- La *familiarity* correspondrait à la récupération indirecte d'un *task-set* en mémoire à long terme. Plus précisément, lorsque le modèle crée un *probe*, ce dernier est biaisé en faveur du *task-set* le plus fréquent et selon le modèle contextuel. Ainsi, le *probe* créé est déjà "familier" et peut être très vite consolidé vers le *task-set* déjà appris.

7.2.3 Rôle du cortex préfrontal

L'importance des structures du lobe temporal médian pour les processus de mémoire est admise depuis au moins un demi siècle. La preuve de l'implication des régions du lobe frontal est plus récente. Des déficits de mémoire après des lésions du lobe frontal chez l'humain peuvent en effet passer inaperçues parmi des symptômes plus apparents tels que la désinhibition, l'impulsivité et la désorganisation. De plus, les patients fronto-lésés parviennent souvent sans trop de difficulté à se souvenir d'évènements saillants, contrairement aux patients lésés au niveau du lobe temporal médian. Cependant, il est clair que des lésions du lobe frontal peuvent impacter certains aspects de la mémoire, notamment les processus de contrôle qui sont cruciaux lors de la récupération [160][213]. En effet, la grande quantité d'information stockée en mémoire excède très largement l'information spécifique nécessaire à la réalisation d'une tâche en particulier. Les processus de contrôle sont donc nécessaires pour chercher, sélectionner et contrôler l'information lors de la récupération. Ainsi, contrairement à des formes plus pures d'amnésie qui résultent de lésions du lobe temporal médian [208], les déficits après une lésion frontale correspondent davantage à la mauvaise exécution de stratégies de récupération [115]. En d'autres termes, le cortex frontal est requis si une mémoire correcte dépend de l'organisation, la recherche, la sélection et la vérification lors de la récupération d'information stockée.

La théorie du filtrage dynamique

Shimamura a développé un cadre théorique appelé "théorie du filtrage dynamique" pour décrire le rôle du cortex préfrontal lors du processus de récupération [213]. Selon cette vision, le cortex préfrontal surveille et contrôle le traitement de l'information par un mécanisme de filtrage. Quatre aspects du contrôle exécutif peuvent être caractérisés en termes de filtrage dynamique du traitement de l'information : la sélection, le maintien, la mise à jour et le réacheminement.

La sélection correspond à la capacité de centrer son attention sur des caractéristiques perceptuelles ou sur des représentations mnésiques activées ; le maintien se réfère à la capacité de garder active toute information sélectionnée en mémoire de travail ; la mise à jour se réfère aux processus impliqués dans la modulation et le réarrangement de l'activité en mémoire de travail ; le réacheminement se réfère à la capacité de passer d'un processus cognitif ou d'une réponse à un(e) autre.

La théorie du filtrage dynamique suggère que ces quatre aspects du contrôle exécutif peuvent être décrit en terme d'interaction entre des régions du cortex préfrontal et du cortex postérieur. Selon cette vision, les activations corticales postérieures consistent à tout moment en une cacophonie de signaux neuraux en réponse à l'information sensorielle et aux activations en mémoire. Le cortex préfrontal orchestre ces signaux en maintenant certaines activations et en en inhibant d'autres. Ainsi, le cortex préfrontal affine l'activité corticale en augmentant les rapports de signal sur bruit.

Shimamura suggère que les quatre aspects de contrôle exécutif peuvent être interprétés en terme des propriétés de filtrage, tel que la mise en place d'un filtre (sélection), le maintien actif d'un filtre (maintien), le *switch* d'un filtre à un autre (mise à jour ou réacheminement).

Il suppose que ces quatre aspects de contrôle exécutif sont essentiels pour une récupération mnésique réussie. C'est-à-dire, lorsque l'on doit se souvenir d'information, les représentations mnésiques doivent être sélectionnées, maintenue et mises à jour. Etant donné que certaines activation mnésiques deviennent non pertinentes, les activations doivent être inhibées et le processus réacheminé. Ainsi, la recherche stratégique en mémoire impose des demandes significatives au contrôle exécutif.

Spécificités régionales

Cortex préfrontal dorso-latéral. Le cortex préfrontal dorsolatéral semble impliqué dans divers processus de la récupération tels que le traitement des informations contextuelles, la recherche stratégique, la vérification, le monitorage et l'évaluation des représentations qui ont été récupérées depuis la mémoire à long terme et sont maintenues dans le cortex préfrontal ventrolatéral [213][181][87][72][215].

Certains auteurs observent par ailleurs une latéralisation du cortex préfrontal latéral dans la lignée du modèle HERA de Tulving. Fletcher et collaborateurs observent par exemple une spécialisation fonctionnelle du cortex préfrontal latéral droit lors de la récupération depuis la mémoire épisodique [88]. Cette prédominance du cortex préfrontal latéral droit reflèterait un processus d'initiation et de maintien du processus de récupération. D'autres études de neuroimagerie, observent que le cortex préfrontal gauche est particulièrement actif lors de la récupération d'informations contextuelles [194][186].

Quoi qu'il en soit, le cortex préfrontal latéral semble particulièrement impliqué dans le contrôle stratégique de la récupération mnésique, notamment la récupération de *task-sets* [196][34][161].

Cortex frontopolaire. Les lésions du cortex frontopolaire ne causent pas la même amnésie globale qui peut résulter de lésions du lobe temporal médian, mais sont associées à des déficiences dans des tâches de mémoire plus complexes [117].

Cette région pourrait être impliquée dans le processus de récupération [232] [79], bien qu'elle semble plutôt réaliser une fonction de plus haut niveau dans le contrôle mnésique tel que le traitement d'information générée de façon interne [136][48][36]. Autrement dit, le cortex frontopolaire pourrait participer à la récupération intentionnelle d'après des indices internes, en l'absence d'indices externes de récupération.

Cortex préfrontal ventro-médian. Le cortex préfrontal médian est connecté à l'hippocampe ventral et intermédiaire et semble répondre à l'information contextuelle [112]. Une interaction fonctionnelle entre ces deux régions semble exister car les rythmes électrophysiologiques dans ces deux régions sont couplés, en particulier dans la gamme thêta [113][214]. Cette activité synchronisée est associée à la récupération mnésique [19][89]. Des réductions dans le verrouillage de phase entre les décharges du cortex préfrontal médian et ceux de l'hippocampe sont également prédictives des erreurs, ce qui suggère que la synchronie entre le cortex préfrontal médian et l'hippocampe est soit nécessaire pour la récupération correcte, soit reflète la confiance dans la récupération [111].

Pour résumer, le lobe temporal médian peut être vu comme la "mémoire brute", à cause de ses liens étroits et directs avec les processus basiques de mémoire. Le cortex frontal, quant à lui, doit "collaborer" avec la mémoire pour réaliser ses diverses fonctions stratégiques, soit en influençant les entrées du lobe temporal médian soit, en agissant sur les sorties de cette région.

7.2.4 Une vision unifiée

Dans leur revue de littérature de 2003, Simons et Spiers proposent un cadre anatomo-fonctionnel unifié du processus de récupération (voir figures 7.4 et 7.5) [215].

La récupération est définie comme le processus par lequel une information stockée en mémoire à long terme est rendue disponible pour effectuer une opération cognitive. Ce processus s'effectue typiquement via l'utilisation d'indices de récupération. Plusieurs étapes de ce processus impliquent divers modules cognitifs :

- la caractérisation de l'indice de récupération,
- la sélection de représentations stockées en mémoire à long terme avec cet indice,
- la réactivation de l'information stockée,
- le monitorage ou l'évaluation de cette information réactivée.

Plusieurs études de patients amnésiques, de lésions expérimentales chez l'animal et d'imagerie chez l'humain montrent que le cortex préfrontal et le lobe temporal médian interagissent pour permettre la récupération [208][149][221][90][213].

Le lobe temporal médian serait à lui seul responsable du stockage et de l'indexation en mémoire. Par la suite, le cortex préfrontal et le lobe temporal interagiraient de différentes façons à chaque stade de la récupération. Le cortex préfrontal ventrolatéral permettrait de caractériser l'indice de récupération et la nature verbale ou non verbale de l'élément à récupérer. L'indice de récupération serait ensuite utilisé pour interroger le lobe temporal médian afin de chercher de façon stratégique les représentations stockées qui sont en accord avec cet indice, potentiellement par le processus de *pattern completion*. Si l'indice n'est pas suffisamment spécifique ou distinct pour activer une représentation, le processus est répété jusqu'à ce qu'un indice approprié soit trouvé et un souvenir obtenu, ou alors le processus est avorté. Il est également possible qu'un indice active directement le lobe temporal médian si l'indice est hautement spécifique et fortement relié à l'information représentée dans le lobe temporal médian. Il s'agit du processus *associative-cue dependent* [164]. Lorsqu'une ou plusieurs représentations sont identifiées, elles seraient maintenues en mémoire de travail par le cortex préfrontal ventrolatéral. Des opérations de monitorage, effectuées par le cortex préfrontal dorsolatéral, seraient mises en œuvre pour comparer l'information récupérée avec le critère de récupération et permettre la dissociation des représentations ambigües. Le cortex préfrontal antérieur pourrait être recruté si des opérations de récupération particulièrement complexes sont nécessaires, dans le cas par exemple du traitement d'information générée de façon interne, sans indices extérieurs.

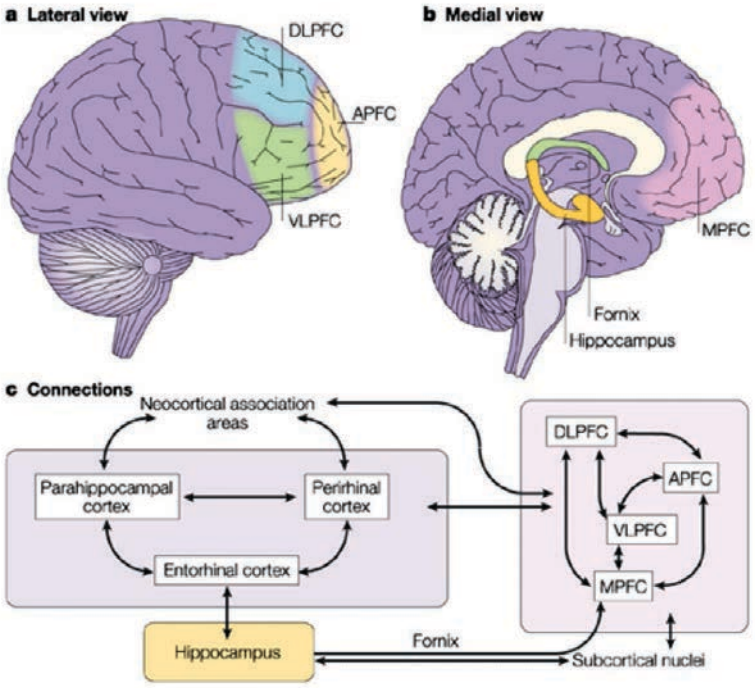


FIGURE 7.4 – **Connectivité entre le cortex préfrontal et le lobe temporal médian. (a,b)** : Le cortex préfrontal (PFC) se subdivise en plusieurs régions : le cortex préfrontal antérieur (APFC, aire de Brodmann (BA) 10), le cortex préfrontal dorsolatéral (DLPFC, BA 46 et 9), le cortex préfrontal ventrolatéral (VLPFC, BA 44, 45 et 47) et le cortex préfrontal médian (MPFC, BA 25 et 32). Les aires de Brodmann 11, 12 et 14 correspondent au cortex orbitofrontal. Le lobe temporal médian comprend l'hippocampe et l'amygdale ainsi que les régions entorhinale, perirhinale and parahippocampique. **(c) Connections.** Il existe des connections réciproques directes entre le cortex préfrontal et le lobe temporal médian. Le cortex orbitofrontal et le cortex dorsolatéral ont de fortes connections réciproques avec le cortex perirhinal et le cortex entorhinal. Il existe davantage de connections du cortex préfrontal vers le cortex perirhinal que dans le sens inverse. Il existe des projections unidirectionnelles de la couche de cellules CA1 vers la région caudale du cortex préfrontal médian. Le complexe subiculaire et les régions temporales médianes ont des connections réciproques avec le cortex préfrontal médian caudal. De plus, le lobe temporal médian reçoit de l'information de plusieurs aires associatives sensorielles unimodales et polymodales. Cette information entre de façon prédominante via le cortex perirhinal et le cortex parahippocampique, qui projettent en retour vers ces régions. Le cortex préfrontal a des connections réciproques avec les cortex associatifs sensoriels dont les régions temporales et pariétales et plusieurs structures sous corticales. D'après [215].

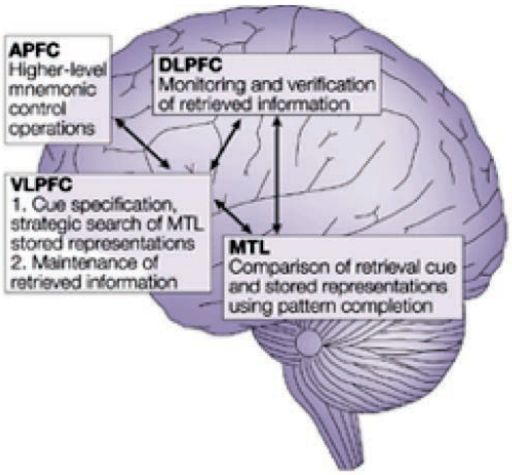


FIGURE 7.5 – **Une vision unifiée du processus de récupération.** Selon la tâche à réaliser, les indices pertinents de récupération sont identifiés dans le cortex préfrontal ventrolatéral (VLPF) avant d'être utilisés pour chercher de façon stratégique des représentations stockées en mémoire à long terme (MTL). Par le processus de *pattern completion*, qui s'effectue en dehors de tout contrôle cognitif, l'indice de récupération est comparé aux représentations stockées jusqu'à l'identification d'une représentation. Cette représentation est récupérée et maintenue dans le cortex préfrontal ventrolatéral pendant que divers processus de monitorage et de vérification sont effectués dans le cortex préfrontal dorsolatéral (DLPFC). L'information récupérée est comparée au critère de récupération spécifié au départ et, si ce critère est atteint, la représentation est rendue disponible pour un traitement conscient. Autrement, une modification de l'indice de récupération peut avoir lieu dans le cortex préfrontal ventrolatéral et d'autres tentatives de récupération sont entreprises. La nature des processus de contrôle mnésique de plus haut niveau, s'ils sont nécessaires, semblent être sous-tendus par le cortex préfrontal antérieur (APFC), et le traitement de l'information pertinente pour le sujet semble engager le cortex préfrontal médian (non figuré). D'après [215].

Troisième partie

Question de thèse

Au cours de l'évolution, le cerveau humain s'est développé de manière à pouvoir s'adapter à des environnements changeants et ouverts. L'être humain est ainsi capable d'apprendre et de mémoriser des tâches complexes, de façon non supervisée, et en association avec le contexte dans lequel elles ont été apprises. Les règles qui gouvernent l'exécution de ses tâches se fondent sur des représentations abstraites des états de l'environnement, les *task-sets*, qui correspondent à une unité hiérarchique du contrôle cognitif.

Le système exécutif est par ailleurs en mesure d'inférer un changement de contingences de l'environnement et de sélectionner en conséquence le *task-set*, et donc le comportement, adapté. Les processus d'inférences peuvent se fonder sur des indices contextuels qui, par association, permettent de récupérer le *task-set* correspondant. Ces inférences sont dites "proactives" car elles ont lieu avant l'exécution de toute action. En l'absence d'indices contextuels, l'inférence peut s'effectuer sur la base de l'observation des résultats obtenus qui ne correspondent plus aux attentes. Ces inférences sont dites "réactives", car elles ont lieu en réaction aux actions effectuées et à leurs conséquences.

L'inférence n'est que la première étape du processus d'adaptation. Elle permet de déduire que le *task-set* courant n'est plus adapté et qu'il est nécessaire d'en changer. Il est possible de quitter le *task-set* courant soit pour en créer un nouveau, soit pour en récupérer un stocké en mémoire. Ainsi, il est nécessaire que le contrôle cognitif et la mémoire interagissent lors de l'initiation d'un nouveau *task-set*.

Le modèle PROBE offre une synthèse des processus computationnels du système exécutif et propose un cadre explicatif intégrant l'inférence et la récupération depuis la mémoire. Des expériences comportementales ont prouvé la capacité de ce modèle à expliquer le comportement de sujets sains dans un protocole expérimental impliquant l'apprentissage et l'arbitrage de *task-sets* multiples, dans des environnements ouverts dont les contingences sont variables et parfois récurrentes [54][74]. Cette aptitude à reproduire et expliquer les données comportementales est un premier argument fort en faveur de ce modèle, et suggère que celui-ci pourrait expliquer certains processus du raisonnement humain. La deuxième étape de la démonstration consiste à découvrir si les processus proposés sont effectivement implémentés dans le cerveau humain.

L'étude de Donoso, Collins et Kœchlin a permis de mettre en évidence les substrats cérébraux de l'inférence réactive dans un protocole où le contexte reste inchangé et non informatif sur les contingences sous-jacentes. Cependant, les substrats cérébraux qui sous-tendent le processus d'inférence proactive et la récupération de règles connues sur la base d'indices contextuels restent à élucider. Il s'agit de la question à laquelle tente de répondre ce travail de thèse.

Au regard de ce que nous avons exploré dans les deux premières parties de cette thèse, nous pouvons énoncer plusieurs prédictions concernant ces substrats cérébraux.

Dans un premier temps, étant donné que notre étude prolonge celle de Donoso, Collins et Kœchlin [74], nous nous attendons à reproduire leurs résultats.

Plus précisément :

Concernant les signaux de fiabilité *ex-post* (μ), c'est-à-dire ceux mis à jour sur la bases des renforcements obtenus en l'absence d'information contextuelle, nous devrions observer :

(i) Une activation du cortex préfrontal ventro-médian en corrélation avec le signal de fiabilité *ex-post* du *task-set* acteur ;

(ii) Une activation du cortex frontopolaire droit en corrélation avec le signal de fiabilité *ex-post* des *task-sets* alternatifs.

Concernant les évènements de *switch*, nous devrions observer :

(i) Une activation du cortex cingulaire antérieur dorsal lors de l'entrée en exploration (création de *probe*) ;

(ii) Une activation du cortex préfrontal dorso-latéral lors de la récupération réactive ;

(iii) Une activation du striatum ventral lors de la consolidation (rejet du *probe*).

Par ailleurs, notre étude intègre une dimension contextuelle qui permet, d'une part, de dissocier théoriquement les signaux de confiance *ex-post* (μ) des signaux de confiance *ex-ante* (λ) ; et, d'autre part, d'observer des évènements spécifiques de récupération proactive.

S'il existe une dissociation anatomo-fonctionnelle des signaux de fiabilité, on pourrait s'attendre à observer une activation du cortex préfrontal dorso-latéral en corrélation avec les signaux de fiabilité *ex-ante*. En effet, nous avons montré que cette région cérébrale est impliquée lors du contrôle contextuel. On peut donc supposer que le cortex préfrontal dorsolatéral soit le siège de l'actualisation des signaux de fiabilité sur la base de l'information contextuelle.

De plus, nous avons vu que le processus de récupération semble nécessiter une collaboration étroite entre le lobe temporal médian et le cortex préfrontal. Ces deux régions, plus précisément l'hippocampe et le cortex préfrontal dorso-latéral, semblent donc de bons candidats pour l'implémentation de la récupération proactive. Nous avons vu par ailleurs que l'hippocampe est fortement couplé au cortex préfrontal ventro-médian. Il est donc possible que cette région soit également impliquée lors des évènements de récupération proactive.

Dans la prochaine partie, nous nous attachons à décrire précisément la méthodologie que nous avons employée afin de tester ces prédictions.

Quatrième partie

Méthodologie

Chapitre 8

Protocole expérimental

« Break the hidden code ! »
Slogan du jeu de société Master Mind, 1972

8.1 Contraintes générales

Le protocole expérimental doit permettre de tester les prédictions du modèle PROBE tout en répondant à des contraintes techniques liées notamment à la méthode d'IRM fonctionnelle. Nous détaillons ci-dessous les prédictions faites par le modèle ainsi que les contraintes techniques qui nous ont guidés dans la conception du protocole expérimental.

8.1.1 Prédictions du modèle PROBE

Le modèle PROBE décrit précédemment rend compte de certaines spécificités du comportement humain et permet de formuler des hypothèses quant au fonctionnement du système exécutif humain. Il décrit notamment l'apprentissage, la sélection, le stockage et la création de *task-sets* en présence d'information contextuelle.

Apprentissage non supervisé par essais et erreurs

L'être humain est capable d'apprendre des associations entre stimuli et actions sur la base des *feedbacks* qu'il reçoit après l'exécution d'une action. Afin de tester cette capacité d'apprentissage par renforcement, notre tâche se doit d'être une tâche d'apprentissage par essais et erreurs où le sujet n'est pas informé de la règle d'association entre stimuli et actions mais peut la déduire par le *feedback* qu'il reçoit après chaque action. Par souci de simplicité, nous nous sommes limités à un *feedbacks* binaire (correct ou incorrect).

Par ailleurs, l'environnement est souvent incertain, ou bruité. Il arrive en

effet qu'un *feedbacks* aléatoire soit contraire à ce qui est attendu. C'est pourquoi on ajoute à notre tâche 10% de bruit : pour chaque action correcte, le sujet reçoit un *feedbacks* correct 90% du temps et incorrect 10% du temps et inversement pour chaque action incorrecte. Ce choix de 10% de bruit est un choix empirique. Des expériences préliminaires ont en effet montré qu'au-delà de cette valeur les sujets ne font que de l'exploration et l'on n'observe plus de *switches*.

Variabilité de l'environnement et contrôle cognitif

Dans un environnement variable (ou changeant, ou volatile), l'être humain est capable de détecter un changement de contingences qui nécessite de modifier son comportement. Plutôt que de modifier pas à pas chaque association stimulus-action indépendamment l'une de l'autre, l'être humain peut directement changer de règle d'association entre l'ensemble des stimuli et l'ensemble des actions. Autrement dit, il peut utiliser la détection d'un changement pour effectuer un *switch* abrupt de comportement, au niveau hiérarchique des *task-sets*, plutôt que de s'adapter localement, au niveau hiérarchique des stimuli.

Notre tâche doit donc permettre la formation de *task-sets*, c'est à dire la construction mentale d'une règle d'association entre stimuli et actions. Pour que cette règle se forme mentalement, elle doit être valide suffisamment longtemps pour que le sujet ait le temps de l'établir. De plus, pour rendre compte de la variabilité de l'environnement et observer des *switches* de comportement, les contingences doivent changer régulièrement et de manière imprévisible.

Par conséquent, il nous faut un nombre suffisant de stimuli et d'actions pour former une collection assez large de règles d'associations. Par ailleurs, il est préférable d'avoir une action possible de plus que le nombre de stimuli. En effet, par simplification, nous nous plaçons dans le cas où une action ne peut être valide que pour un seul stimulus. Dans ce cas, si le nombre d'actions est le même que le nombre de stimuli, la dernière action peut-être déduite par simple élimination.

Récupération d'une stratégie en mémoire de travail

Chaque *task-set* appris peut potentiellement être maintenu en mémoire de travail et éventuellement réutilisé lors de la détection d'un changement dans l'environnement. Ainsi, pour permettre ce phénomène de récupération d'un *task-set* depuis la mémoire de travail, il est nécessaire que certaines contingences réapparaissent régulièrement et de manière suffisamment proches temporellement - soit environ tous les deux à quatre changements - afin que les *task-sets* correspondants ne soient pas encore exclus de la mémoire de travail.

Stockage en mémoire à long terme

Lorsqu'un *task-set* appris est rejeté de la mémoire de travail dont la capacité est limitée, il est stocké en mémoire à long terme. Le contenu de la mémoire à long terme permet d'influencer le réapprentissage d'un *task-set* lorsqu'une contingence réapparaît à une échelle de temps plus large, soit au-delà de quatre ou cinq changements. Pour permettre un apprentissage à long terme et la formation d'une mémoire à long terme biaisée par l'occurrence fréquente de certaines contingences, il est nécessaire de répéter ces contingences, notamment au début de l'expérience.

Création de nouvelles stratégies

Non seulement l'être humain peut exploiter les *task-sets* formés au cours de son expérience passée, mais il peut également explorer de nouvelles stratégies en créant de nouveaux *task-sets*. Pour tester cette possibilité, à tout moment doivent pouvoir apparaître de nouvelles contingences qui nécessitent la formation d'un nouveau *task-set*.

Utilisation de l'information contextuelle

L'être humain est capable de découvrir qu'une entrée sensorielle porte une information contextuelle utile à la sélection a priori d'un *task-set*. Ainsi chaque contingence est associée à un indice contextuel, tel qu'une information visuelle, qui peut être informatif pour prédire la règle en cours.

Au final, notre protocole nécessite une tâche cognitive exigeante qui doit encore répondre à un certain nombre de contraintes techniques.

8.1.2 Contraintes techniques

Notre tâche se doit d'être aussi complète et spécifique que possible au regard des prédictions que nous souhaitons tester, mais doit également être suffisamment simple et d'une durée raisonnable afin d'être réalisable par tout sujet adulte sain.

Une expérience d'IRM fonctionnelle nécessite de travailler en présence d'un fort champ magnétique ce qui implique des règles de sécurité très strictes et une sélection des sujets aptes à y être exposés. De plus, lors de l'expérience IRM, le sujet est confiné dans l'espace réduit de l'imageur et soumis à un bruit d'origine mécanique important pendant l'acquisition des images. A cela s'ajoute la nécessité pour le sujet de rester immobile pendant toute la durée de l'acquisition afin d'éviter les artéfacts liés aux mouvements. Ces contraintes limitent donc la durée d'une session d'IRM (une heure maximum) pour assurer le confort du sujet et l'obtention de données de qualité.

Par ailleurs, le signal enregistré en réponse à un événement étant long (dix secondes environ), chaque événement susceptible de provoquer une réponse se doit d'être suffisamment distant des autres afin d'obtenir un signal analysable. Ceci nécessite d'effectuer un compromis entre la durée et le nombre d'essais afin d'obtenir une quantité suffisante de données de qualité pour les analyses comportementales et fonctionnelles, tout en limitant le temps total d'acquisition.

Ainsi, l'ensemble de ces prédictions et des contraintes techniques nous a amené à concevoir une expérience composée en tout de trois sessions : une session comportementale dite 'd'apprentissage' sur ordinateur et deux sessions dites 'test' en IRM fonctionnelle. Les caractéristiques de ces sessions sont décrites ci-après et récapitulées dans le tableau 8.1. La tâche est illustrée dans la figure 8.1.

L'ensemble de la tâche a été implémenté avec l'outil *Psychtoolbox* (http ://psychtoolbox.org) de Matlab (http ://www.mathworks.fr).

8.2 Structure des essais

8.2.1 Stimulus

A chaque essai, un stimulus (chiffre arabe) est présenté au sujet à l'intérieur de quatre cases noires aux bordures blanches alignées au centre de l'écran. Il y a trois stimuli possibles (1, 2 ou 3) et les quatre cases correspondent aux quatre actions possibles. Les sujets répondent à ces stimuli en actionnant avec les doigts de la main droite l'une des quatre touches [j, k, l, m] d'un clavier azerty (pour la session d'apprentissage sur ordinateur) ou les quatre touches d'un boitier réponse (pour les sessions en IRM) selon la règle suivante :

- j (touche 1) : index
- k (touche 2) : majeur
- l (touche 3) : annulaire
- m (touche 4) : auriculaire

Les doigts sont fixes sur les touches, dans une position naturelle de la main droite, permettant une dissociation motrice des quatre actions possibles. De plus l'expérience peut être réalisée avec une seule main ce qui évite le contrôle des effets de latéralisation.

En plus des stimuli, une information contextuelle, codée par une texture colorée en fond d'écran, est présentée au sujet pendant chaque essai (voir Annexe A pour l'ensemble des textures utilisées).

Le stimulus est affiché en blanc dans toutes les cases pendant 1 seconde. L'information contextuelle reste affichée en continu jusqu'à ce qu'elle change après un nombre variable d'essais.

Le but pour le sujet est de découvrir quelle touche, parmi les quatre à sa disposition, doit être pressée en réponse à chaque stimulus. Un essai gagné correspond à un essai où l'action choisie par le sujet suit l'association correcte. Par exemple, si les contingences sont : [1-k, 2-j, 3-m], si le sujet presse la touche k à l'apparition du stimulus 1, sa réponse est correcte. Cette tâche revient donc pour le sujet à trouver une combinaison de chiffres. Dans notre exemple il s'agit de la combinaison :



Le sujet peut également s'aider de l'information contextuelle pour trouver la combinaison correcte s'il repère une association entre cette information et la combinaison de chiffres.

Etant donné qu'il y a trois stimuli, quatre actions possibles et qu'il n'y a qu'une action correcte par stimulus, nous avons au total 24 associations stimulus-action possibles.

8.2.2 Temps de réponse et intervalle inter-stimulus

Le sujet a 1.3 secondes pour répondre, ce qui correspond au temps d'affichage du stimulus (1 seconde) et 0.3 secondes supplémentaires.

Pour la session d'apprentissage sur ordinateur, le *feedback* est affiché après un intervalle inter-stimulus d'une durée fixe de 0.5 secondes.

Pour les sessions en IRM, le *feedback* est affiché après une durée variable (0.4, 2.15 ou 3.9 secondes réparties sur les essais de manière pseudo-aléatoire). Cette procédure dite de *jittering* permet d'obtenir un signal IRM décomposable en évènements différenciés tout en évitant des essais d'une durée trop longue [1].

8.2.3 Feedback

Si l'essai est gagné (r = 1), le stimulus réapparaît en vert dans la case choisie pendant 1 seconde. Si l'essai est perdu (r = 0), une croix rouge apparaît dans la case choisie pendant 1 seconde également. Si aucune réponse n'a été donnée par

1. Le *jittering* consiste à faire varier l'écart temporel entre le début de l'échantillonnage des images d'un volume cérébral et le début de la présentation du stimulus au sujet. Si le temps de répétition (TR) est de 2 secondes, et que le stimulus est toujours présenté exactement au début d'un TR et toujours pour une durée de 10 secondes, alors on échantillonne systématiquement le même point dans la réponse cérébrale du sujet. On risque donc de manquer des points entre ces points d'échantillonnage. Or ces 'inter points' peuvent aider à caractériser la forme de la fonction de réponse hémodynamique (hrf). Pour pallier ce problème, tout en maintenant le TR à 2 secondes, on peut faire varier le temps entre les essais (l'ISI) aléatoirement entre 0 et 4 secondes. De cette manière on échantillonne davantage de points de la hrf qu'avec un ISI fixe.

Les protocoles avec ISI variable présentent également un avantage d'un point de vue psychologique. En effet, dans ce cas, les essais semblent d'autant plus aléatoires aux yeux du sujet. Lorsque l'ISI est fixe, les sujets peuvent percevoir le rythme de l'expérience et anticiper leur réponse avant l'apparition du stimulus. Selon son degré de variabilité, un ISI variable peut minimiser cet effet.

le sujet, des tirets rouges s'affichent dans les quatre cases. Le sujet est informé au début de l'expérience qu'un essai sans réponse est un essai considéré comme perdu.

Un nouvel essai commence avec un nouveau stimulus après une durée fixe de 0.5 secondes (pour la session d'apprentissage) ou variable (0.1, 1.85 ou 3.6 secondes réparties sur les essais de manière pseudo-aléatoire pour les sessions en IRM).

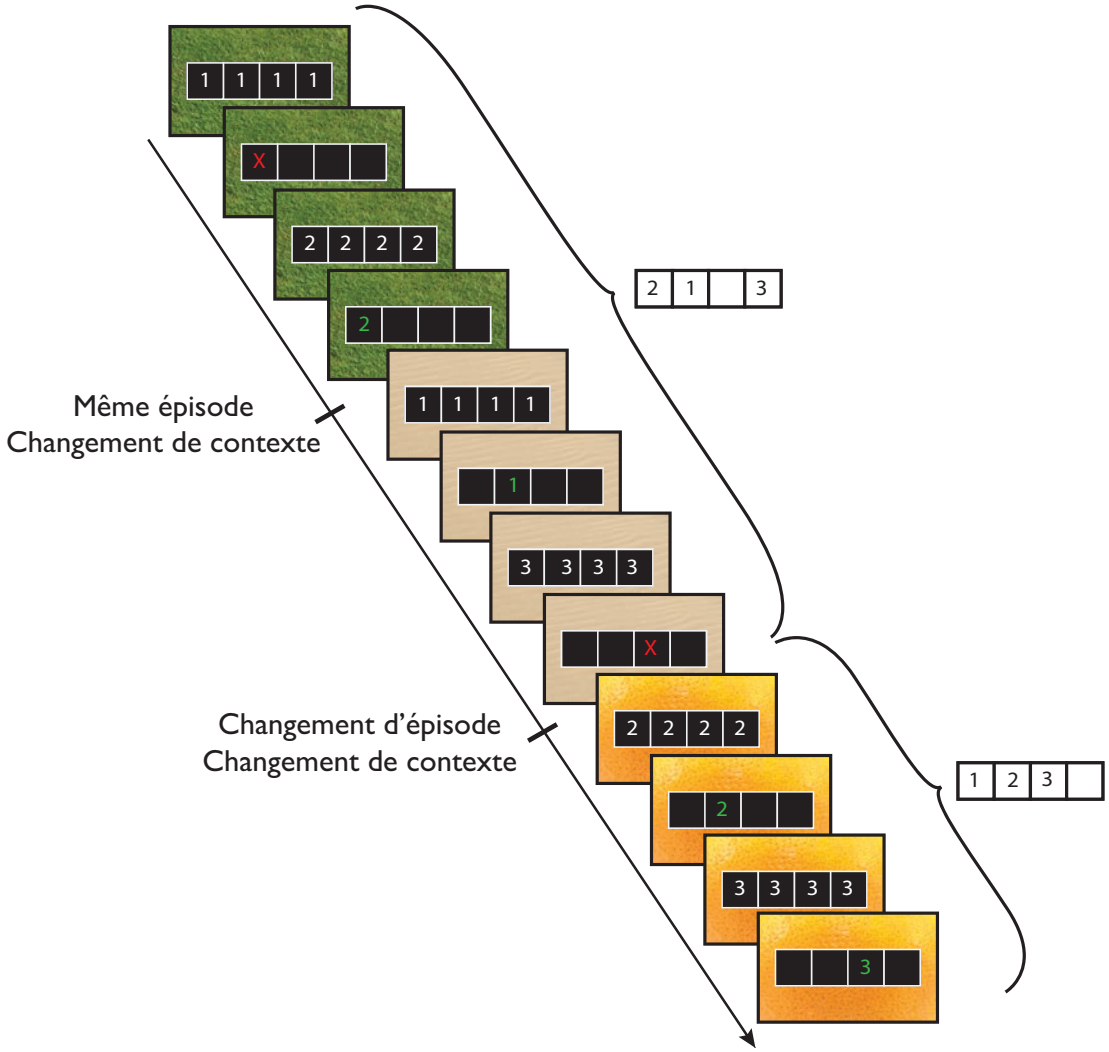


FIGURE 8.1 – **Protocole expérimental**

8.3 Structure des Sessions

8.3.1 La Session d'Apprentissage

Instructions (cf. Annexe B)

L'expérience commence par la lecture des instructions par le sujet seul puis avec l'expérimentateur. Afin d'assurer la motivation du sujet, ce dernier est informé que ses gains dépendent de ses performances. Le sujet est explicitement informé qu'il doit apprendre par lui-même, en fonction des *feedbacks* qu'il reçoit, les associations entre stimulus et réponse, à l'aide ou non des textures. Il reçoit pour consigne de répondre le plus rapidement et surtout le plus correctement possible. Le sujet est également informé que le *feedback* reçu est bruité, ce qui implique que dans un certain nombre de cas, dont le pourcentage n'est pas précisé, un *feedback* positif peut correspondre à une réponse fausse, et un *feedback* négatif à une réponse correcte. Ainsi, face à un *feedback* inattendu, le sujet peut prendre en compte ce bruit et ne pas changer automatiquement de *task-set*. Enfin, le sujet est informé que les combinaisons entre stimuli et actions, autrement dit les contingences de l'environnement, restent constantes pendant un certain nombre d'essais puis changent de manière non prédictible. De même il est informé que les textures en fond d'écran peuvent changer et qu'une même texture peut apparaître plusieurs fois. En revanche il n'est pas informé qu'il existe des associations entre les combinaisons et les textures. C'est à lui d'en faire l'éventuelle observation. Afin d'inciter le sujet à porter attention aux combinaisons et aux textures, il est informé qu'il sera interrogé sur ce qu'il aura observé à la fin de la session.

Entraînement

La tâche cognitive demandée dans l'expérience est assez complexe. Afin de s'y familiariser, le sujet effectue pendant quelques minutes une version simplifiée de la tâche comportant deux stimuli et deux actions seulement. Les stimuli choisis sont un carré et un rond plutôt que des chiffres, afin que les sujets puissent se familiariser avec la tâche sans pour autant faire de transpositions aux contingences qu'ils verront par la suite. Les textures présentées sont également différentes de celles qui apparaîtront par la suite. L'entrainement est constitué de trois épisodes où sont présentées deux combinaisons possibles (carré-rond (■●) et rond-carré (●■)) et trois contextes ($Cont_1$, $Cont_1$', $Cont_2$) dans l'ordre suivant :
1) (■●, $Cont_1$ puis $Cont_1$')
2) (●■, $Cont_2$)
3) (■●, $Cont_1$')

Les actions consistent, comme durant l'expérience, à presser des touches sur un clavier d'ordinateur. L'expérimentateur reste avec le sujet durant cet entrainement et commente la tâche avec lui pour s'assurer qu'il en a bien compris les éléments essentiels et qu'il est suffisamment à l'aise avec le déroulement des essais pour aborder sans difficulté l'expérience proprement dite.

Spécificités de la session d'apprentissage

La session d'apprentissage est la première de l'expérience. Elle se déroule sur ordinateur dans la salle de test du Laboratoire de Neurosciences Cognitives. Elle se divise en 3 *runs* d'une durée d'environ 20 minutes chacun. Entre chaque *run* le sujet est invité à faire une pause de la durée qu'il souhaite (typiquement de quelques secondes à quelques minutes).

Chaque *run* est composé de 9 'épisodes'. On appelle 'épisode' une séquence d'essais durant laquelle les contingences de l'environnement restent constantes, c'est à dire qu'une seule combinaison de chiffres est valide tout au long de l'épisode. Un épisode est constitué de 42, 48 ou 54 essais. Cette variabilité rend la longueur d'un épisode imprévisible pour le sujet et il ne peut ainsi inférer un changement de contingence sur la base de la durée.

Pour cette session, seules trois combinaisons de chiffres ($Comb_1$, $Comb_2$ et $Comb_3$), chacune associée à deux contextes ($Cont_1$ et $Cont_1$', $Cont_2$ et $Cont_2$', $Cont_3$ et $Cont_3$'), sont présentées. Par la suite, nous ferons référence à ces combinaisons de chiffres et contextes comme 'anciens' (indicés par un 'a', par opposition aux combinaisons et contextes 'nouveaux' indicés par un 'n'). A chaque *run*, chaque combinaison de chiffres est présentée trois fois de manière pseudo-randomisée :
- une fois avec son premier contexte associé
- une fois avec son second contexte associé
- une fois avec ses deux contextes associés de manière consécutive, c'est à dire avec un changement de contexte au milieu de l'épisode (il s'agit des épisodes comportant 54 essais).

Chaque changement d'épisode est nécessairement accompagné d'un changement de contexte. En revanche, le changement de contexte au milieu d'un épisode permet d'éviter l'inférence qu'un changement de contexte implique nécessairement un changement de contingences. De plus, deux contingences successives sont incongruentes ou non redondantes : pour chaque stimulus, l'action associée change.

Après chaque pause, l'expérience continue sur la contingence précédente durant six ou neuf essais, appelés essais 'post-pause' afin que le sujet n'interprète pas le redémarrage de l'expérience comme un signe de changement de contingences. La session d'apprentissage comprend au total 1311 essais.

Les séquences de stimuli sont générées aléatoirement pour chaque épisode. Un épisode contient toujours le même nombre de chaque type de stimuli. Les essais 'pièges' ou 'bruités', c'est à dire les essais dont le $feedback$ est inversé, représentent 10% des essais totaux : ils sont générés pseudo-aléatoirement avec les quatre contraintes suivantes : (i) Le premier essai d'un épisode ne peut pas être un essai piège, (ii) chaque épisode contient au moins un essai piège, (iii) chaque *run* contient le même nombre d'essais pièges, (iv) dans chaque *run* on trouve deux essais pièges consécutifs afin que les sujets n'infèrent pas que les essais pièges sont des points isolés, nécessairement suivis par des essais 'normaux'.

Débriefing

A la fin de la session d'apprentissage, le sujet est soumis à un débriefing actif (ou 'post-test') en trois parties :
- **Test des combinaisons**. On présente au sujet six combinaisons de chiffres dans un ordre aléatoire : les trois combinaisons présentées au cours de la session et trois nouvelles.
- **Test des contextes**. On présente au sujet douze contextes dans un ordre aléatoire : les six contextes présentés au cours de l'expérience et six nouveaux.
- **Test des associations**. Chaque combinaison 'ancienne' de chiffres est présentée quatre fois : une fois avec chacun de ses contextes anciens associés et deux fois avec des contextes anciens qui ne lui sont pas associés.

A chaque étape, le sujet est invité à répondre, à l'aide de l'échelle de réponse ci-dessous, s'il se souvient d'avoir observé ces combinaisons, contextes et associations combinaison-contexte.
1 : OUI, je suis sûr(e)
2 : Oui, je crois
3 : Je ne sais pas
4 : Non, je ne crois pas
5 : NON, je suis sûr(e)

Ce post-test permet de juger si le sujet a explicitement retenu les combinaisons et associations qu'il a utilisées. Il permet également d'inciter le sujet à faire d'autant plus attention aux combinaisons et aux associations pour les sessions suivantes. Ce post-test a été réalisé afin de différencier d'éventuels sous-groupes de sujets utilisant des stratégies d'adaptation différentes en suivant la méthode d'Anne Collins et Etienne Kœchlin dans leur article de 2012.

8.3.2 La Session Test

La session Test est découpée en deux sessions d'IRM fonctionnelle passées au Centre National d'Imagerie de Recherche (CENIR) de l'hôpital de la Pitié-Salpêtrière.

Entraînement

Avant chaque session d'IRM, le sujet est invité à faire un entrainement sur ordinateur correspondant à une version courte (environ dix minutes) de la session d'apprentissage avec ajout du temps variable (*jittering*) entre deux évènements. Cet entrainement permet au sujet d'une part de se remémorer les combinaisons, contextes et associations rencontrés lors de la session d'apprentissage ; et d'autre part de se familiariser avec le *jittering*.

Spécificités de la session test

Chaque session d'IRM est divisée en trois *runs* d'une durée d'environ vingt minutes chacun. Entre chaque *run* le sujet est invité à faire une pause de la durée qu'il souhaite (typiquement de quelques secondes à quelques minutes) mais doit rester immobile dans l'IRM.

Chaque *run* est composé de huit 'épisodes'. Il existe quatre types d'épisode correspondant à quatre conditions expérimentales. L'ordre des conditions est pseudo-randomisé.

- **Condition 1** : ancienne combinaison et ancien contexte ($\text{Comb}_a\text{Cont}_a$). Dans deux épisodes par *run*, la combinaison 1 de la session d'apprentissage est valide. Elle est présentée une fois avec l'un de ses deux contextes associés (Cont_1 ou Cont_1') et une fois avec ses deux contextes associés avec un changement de contexte au milieu de l'épisode (Cont_1 puis Cont_1' ou Cont_1' puis Cont_1). Cette condition permet de tester l'acquisition d'un *task-set* et de son association avec les contextes pendant la session d'apprentissage.
- **Condition 2** : ancienne combinaison et nouveau contexte ($\text{Comb}_a\text{Cont}_n$). Dans deux épisodes par *run*, la combinaison 2 de la session d'apprentissage est valide. Elle est présentée une fois avec un nouveau contexte jamais vu par le sujet et une fois avec deux nouveaux contextes consécutifs avec un changement de contexte au milieu de l'épisode. Cette condition permet de tester la capacité des sujets à réutiliser un *task-set* appris dans un nouveau contexte.
- **Condition 3** : nouvelle combinaison et ancien contexte ($\text{Comb}_n\text{Cont}_a$). Dans deux épisodes par *run*, une combinaison nouvelle est valide. Une combinaison nouvelle est présentée une fois avec un des contextes 3 de la session d'apprentissage (Cont_3 ou Cont_3') et une autre combinaison nouvelle est présentée avec les deux contextes 3 de la session d'apprentissage avec un changement de contexte au milieu de l'épisode (Cont_3 puis Cont_3' ou Cont_3' puis Cont_3). Cette condition permet de tester la capacité des sujets à apprendre un nouveau *task-set* dans un contexte connu.
- **Condition 4** : nouvelle combinaison et nouveau contexte ($\text{Comb}_n\text{Cont}_n$). Dans deux épisodes par *run*, une nouvelle combinaison est valide. Une nouvelle combinaison est présentée une fois avec un nouveau contexte et une autre nouvelle combinaison est présentée avec deux nouveaux contextes consécutifs avec un changement de contexte au milieu de l'épisode. Cette condition permet de tester la capacité des sujets à apprendre un nouveau *task-set* dans un contexte nouveau. Elle fournit également une ligne de base pour comparer l'avantage à l'adaptation fourni dans les autres conditions par la connaissance du *task-set* valide ou du contexte.

Ces quatre conditions forment un plan factoriel à deux facteurs qualitatifs : le facteur 'combinaison' à deux niveaux (ancien/nouveau) et le facteur 'contexte' à deux niveaux également (ancien/nouveau).

Un épisode est constitué de 21, 24, 27 ou 30 essais. Les épisodes à 24 et 30 essais sont les épisodes avec changement de contexte au milieu de l'épisode. De

même que lors de la session d'apprentissage, chaque changement d'épisode est nécessairement accompagné d'un changement de contexte mais le contraire est faux. De plus, deux contingences successives sont incongruentes (ou non redondantes). Après chaque pause, l'expérience continue sur la contingence précédente durant six ou neuf essais 'post pause'. Chaque session test comprend 627 essais (soit 1254 essais au total). Les séquences de stimuli et les essais pièges sont générés de la même manière que pour la session d'apprentissage.

Débriefings

Débriefing sur ordinateur. A la fin de l'expérience, c'est à dire à la fin de la deuxième session d'IRM, les sujets sont à nouveau soumis au même post test actif que celui présenté à l'issue de la session d'apprentissage. Ce post test permet d'évaluer ce que le sujet a explicitement retenu (combinaisons et associations) à l'issue de la totalité de l'expérience.

Débriefing informel. En plus du post test sur ordinateur, l'expérimentateur effectue un débriefing informel avec le sujet pour évaluer son ressenti de l'expérience et l'interroger sur ses éventuelles stratégies explicites pour réaliser la tâche. L'expérimentateur répond également aux questions du sujet sur le but de l'expérience (voir Annexe C).

	APPRENTISSAGE	**TEST**
Numéro de session	Session 1	Session 2 | Session 3
Support	Ordinateur MacBook Pro	IRMf (Trio du CENIR)
Nombre de runs par session	3	3
Nombre d'épisodes par run	9 (total = 27 épisodes)	8 (total = 24 épisodes par session)
Nombre d'essais par épisode **Episodes à 2 contextes**	42, 48 ou **54** (moyenne : 48)	21, **24**, 27 ou **30** (moyenne : 25-26)
Nombre total d'essais	1311	1254
Jittering	Non : ISI = 0,5 sec	Oui
Durée moyenne d'un essai	3 secondes	6 secondes
Durée moyenne d'un épisode	48 x 3 = 144 sec = 2,4 min	26 x 6 = 156 sec = 2,6 min
Durée moyenne d'un run	9 x 2,4 = 21,6 min	8 x 2,6 = 20,8 min
Durée moyenne d'une session	3 x 21,6 = 64,8 min = 1h04 min + Entraînement et pauses = 1h30 environ	3 x 20,8 = 62,4 min = 1h02 min + séquence anatomique (8min) = 1h10 environ + entraînement et débriefing = 1h45 environ
Conditions	($Comb_a Cont_a$)	Toutes les 4 conditions
Composition d'un run	3 x [$Comb_i$, $Cont_i$] 3 x [$Comb_i$, $Cont_i$''] 3 x [$Comb_i$, ($Cont_i$ et $Cont_i$'')] i ∈ {1,2,3} 3 $Comb_a$ et 6 $Cont_a$ chaque $Comb_a$ associée à 2 $Cont_a$	1 x [$Comb_1$, $Cont_1$ ou $Cont_1$''] 1 x [$Comb_1$, ($Cont_1$ et $Cont_1$'')] 1 x [$Comb_2$, $Cont_n$] 1 x [$Comb_2$, ($Cont_n$ et $Cont_n$)] 1 x [$Comb_n$, $Cont_3$ ou $Cont_3$''] 1 x [$Comb_n$, ($Cont_3$ et $Cont_3$'')] 1 x [$Comb_n$, $Cont_n$] 1 x [$Comb_n$, ($Cont_n$ et $Cont_n$)]

TABLE 8.1 – **Propriétés générales du dessin expérimental**

8.4 Participants

8.4.1 Expérience comportementale pilote

L'expérience comportementale pilote est identique à l'expérience en IRM mais a été réalisée uniquement sur ordinateur au Laboratoire de Neurosciences Cognitives. Elle a permis de tester et valider le protocole avant de le réaliser en IRM.

Pour cette expérience, nous avons testé 23 sujets volontaires (dont 12 hommes) âgés de 18 à 29 ans. Les sujets on été indemnisés pour leur participation (20 € par session). Parmi les 23 sujets, deux sujets n'ont pas été inclus dans les analyses en raison d'erreurs de manipulation.

8.4.2 Expérience en IRM

Pour l'expérience en IRM, 56 sujets volontaires (dont 26 hommes) âgés de 18 à 28 ans ont été recrutés par annonce via le Relais d'Information sur les Sciences Cognitives (RISC) et les réseaux sociaux. Les sujets ont été soumis à un entretien médical afin de s'assurer qu'ils vérifiaient les conditions de participation à l'expérience :
- Conditions liées à l'IRM : être droitier (contrôle de la latéralisation), ne pas être claustrophobe, ne pas avoir subi d'opérations ayant entrainé la pose d'objets métalliques, absence de tatouages, vision correcte ou corrigée par des lentilles de contact.
- Conditions pour la bonne réalisation de la tâche (capacités cognitives intactes) : être de langue maternelle française (pour la bonne compréhension des instructions), agé de 18 à 30 ans, absence de troubles de l'attention ou de la mémoire, absence d'antécédents neurologiques ou psychiatriques (traumatismes crâniens, épilepsie, etc.), absence de dyslexie, pas de prise de médicaments psychotropes, absence de troubles de la perception des couleurs, pas de consommation de drogue (cannabis ou autre) dans la semaine précédant l'expérience, pas de consommation d'alcool dans les 24 heures précédant l'expérience, ne pas être enceinte, ne pas avoir participé aux expériences précédentes menées par le Laboratoire de Neurosciences Cognitives (afin d'éviter tout biais).

Sur les 56 sujets recrutés :
- Trois sujets se sont désistés avant la fin de l'expérience
- Un sujet a été exclu à la suite de la visite médicale
- Deux sujets ont été exclus pour non réalisation de la tâche (moins de 10% de réponses correctes)
- Un sujet a été exclu en raison d'un problème technique à l'IRM (dysfonctionnement du boitier réponse)
- Un sujet a été exclu en raison de la présence d'un artefact au niveau du lobe temporal droit (présence d'un élément métallique sous cutané)

Au final 48 sujets (dont 24 hommes) ont été inclus dans les analyses pour la totalité de l'expérience. Nous avons fait le choix d'inclure ce nombre important de sujets afin de réaliser d'éventuelles analyses sur des sous-groupes.

Les sujets ont tous été informés des conditions de participation aux expériences au moyen d'une notice écrite d'information légale, validée par le comité d'éthique de l'INSERM et ont signé, en connaissance de cause, un consentement écrit avant leur inclusion dans le protocole.

Les sujets ont été dédommagés à hauteur de 20 € par session comportementale et 80 € par session d'IRM augmenté d'un éventuel bonus de 10 ou 20 € selon les performances (soit du nombre de réponses correctes).

Les différentes sessions de l'expérience ont été espacées de 24h minimum et une semaine maximum.

Chapitre 9

Ajustement des paramètres du modèle sur les données comportementales

9.1 Principe et méthodes d'ajustement de paramètres (ou *fitting*)

9.1.1 Principe général

Un modèle computationnel tel que le modèle PROBE reçoit des variables d'entrée (ou explicatives) et génère des variables de sortie (ou dépendantes). Le calcul des variables de sortie correspond à une ou plusieurs opérations mathématiques qui combinent les variables d'entrée et les paramètres du modèle. Ce sont les valeurs de ces paramètres que nous cherchons à déterminer.

Prenons l'exemple d'une régression linéaire simple pour illustrer notre propos. Considérons N points de données représentés par leurs coordonnées (x_i, y_i), i $\in$ [1, N]. Supposons que ces données peuvent être modélisées par une fonction affine du type $y = ax+b$, où x est une variable d'entrée, y une variable de sortie, a et b sont les paramètres. Ajuster les paramètres de ce modèle revient à déterminer a et b tels que pour chaque x_i, le i correspondant calculé par le modèle soit 'le plus proche possible' du y_i observé. Autrement dit, l'ajustement des paramètres d'un modèle consiste à trouver les valeurs des paramètres qui permettent de prédire au mieux les données observées.

Les méthodes de régression classiques, telle que la régression linéaire, permettent de déterminer les paramètres de façon optimale à partir de calculs sur les données, mais sont inapplicables si la fonction est trop complexe, ce qui est le cas du modèle PROBE. Lorsque l'on ne peut déterminer la solution analytiquement, on peut procéder par essai erreur, c'est-à-dire tester plusieurs combinaisons de paramètres jusqu'à trouver celle qui permet de reproduire au mieux les données. La solution n'est alors pas nécessairement unique.

Pour ajuster les paramètres du modèle PROBE, il faut donc d'une part choisir une méthode de tirage des paramètres et d'autre part définir un critère de la qualité d'ajustement, c'est-à-dire un nombre représentant la différence entre les données et les prédictions du modèle pour tout set de paramètres. Il s'agit en général de la somme des moindres carrés ou du maximum de vraisemblance. L'ajustement des paramètres revient alors à déterminer le set de paramètres qui minimise cette différence.

Ainsi notre objectif est de déterminer, pour chaque sujet, les valeurs des paramètres du modèle PROBE qui lui permettent de reproduire au mieux le comportement observé.

9.1.2 Critères d'ajustements

Il existe plusieurs critères d'ajustement utilisés au sein de la 'communauté des modélisateurs'. Nous détaillons ici ceux que nous avons testés et qui sont les plus courants : la somme des moindres carrés et le maximum de vraisemblance.

La somme des moindres carrés (ou *Least Squares*, LS)

La méthode des moindres carrés a été indépendamment élaborée par Legendre en 1805 et Gauss en 1809.

Reprenons l'exemple d'une régression linéaire simple pour illustrer son principe. Nous avons nos N points de données expérimentales représentés par leurs coordonnées (x_i, y_i), i $\in$ [1, N] et nous cherchons les paramètres a et b tels que la droite affine d'équation $y = ax + b$ s'ajuste le mieux possible aux données. Parmi toutes les droites possibles, on retient celle qui minimise la somme des carrés des écarts entre les valeurs observées (x_i, y_i) et les points (x_i, $\hat{y}_i$) de la droite $\hat{y}_i$=ax_i+b. Autrement dit, celle qui minimise la somme des distances entre chaque point observé et chaque point prédit. Ou plus formellement, si ϵ représente cet écart, appelé aussi résidu, le principe des moindres carrés consiste à choisir les valeurs de a et de b qui minimisent :

$$E = \sum_{i=1}^{N} \epsilon_i^2 = (y_i - (ax_i + b))^2$$

Dans le cas du modèle PROBE, il ne s'agit pas d'une simple régression linéaire mais le principe reste le même. L'écart que nous cherchons à minimiser est celui entre les fréquences de réponses correctes réalisées par le sujet et les probabilités de réponses correctes prédites par le modèle. Plus précisément :
- Pour chaque sujet nous calculons la proportion de réponses correctes à chaque essai après un changement de contingences (typiquement, en moyenne, au bout de cinq essais après un changement de contingences, les sujets sont à 50% de réponses correctes ; au bout de 15 essais le taux de réponses correctes atteint 80%).
- Pour chaque essai, le modèle PROBE calcule la probabilité de sélectionner

chaque action possible. On s'intéresse à la probabilité moyenne de sélectionner l'action correcte à chaque essai après un changement de contingences.
- Nous cherchons à partir de là le set de paramètres qui minimise la somme des carrés des écarts entre les fréquences observées (F) et les probabilités prédites (P) moyennes. Soit plus formellement :

$$LS = \sum_{i=1}^{I} [F_i(correct) - P_i(correct/\theta)]^2$$

Où i ne correspond pas à un essai mais un ensemble d'essais consécutifs de taille variable : de deux essais en début d'épisode (juste après le changement de contingences) jusqu'à 24 essais en fin d'épisode (au delà du 30ème essai). En effet, au delà d'une vingtaine d'essais, les sujets ont atteint l'asymptote de la courbe d'apprentissage. Ces fenêtres variables permettent donc d'accentuer la période d'apprentissage (les 15 premiers essais environ) par rapport à la période asymptotique. Ce critère d'ajustement permet ainsi de prendre en compte la structure de la tâche (les changements de contingences observables).

Le maximum de vraisemblance (ou *log-likelihood*, LLH)

La méthode du maximum de vraisemblance a été introduite en 1912 par le statisticien anglais Ronald Fisher. Il s'agit, appliquée à notre cas, de trouver le set de paramètres qui permet au modèle de prédire au mieux la séquence d'actions effectuées par le sujet. Notons X l'ensemble des actions effectuées par le sujet (X = [a_1, a_2 ... a_T]) et θ le set de paramètres du modèle. A chaque essai, le modèle calcule la probabilité d'effectuer chaque action sachant les actions précédemment effectuées (les actions ne sont pas indépendantes) ainsi que l'historique des stimuli et des récompenses. On peut ainsi calculer la vraisemblance, c'est à dire la probabilité d'obtenir la séquence d'actions du sujet avec les paramètres utilisés. Ou plus formellement :

$$P(X/\theta) = P(a_{1:T}/\theta) = P(a_1/\theta) \prod_{t=T}^{I} P(a_t/\theta, a_{1:t-1})$$

Ainsi, plus cette probabilité est élevée, plus le set de paramètres du modèle est 'vraisemblable' puisqu'il permet de prédire les actions que le sujet a effectuées. On cherche donc à trouver le maximum de cette vraisemblance pour que les probabilités des actions observées soient aussi maximales.

La vraisemblance étant positive et le logarithme népérien une fonction croissante, il est équivalent et souvent plus simple de maximiser le logarithme népérien de la vraisemblance. Le produit se transforme alors en somme, ce qui est plus simple à manipuler. Dans la terminologie anglo-saxonne on parle de *log-likelihood* ou LLH. Soit :

$$LLH = log(P(X/\theta)) = log(P(a_1/\theta)) + \sum_{t=2}^{T} log(P(a_t/\theta, a_{1:t-1}))$$

Le 'meilleur' set de paramètres, ou le plus 'vraisemblable', est donc celui qui maximise la vraisemblance ou la log-vraisemblance [1].

La méthode du maximum de vraisemblance présente l'avantage d'utiliser la probabilité de choix de chaque action effectuée par le sujet, même les actions incorrectes (non récompensées). Par ailleurs, la log-vraisemblance est utilisée pour calculer des critères de comparaison de modèles tels que le pseudo-R^2, le critère d'information bayésien (BIC) ou le critère d'information d'Akaike (AIC).

Cette méthode présente cependant l'inconvénient de donner la même importance à tous les essais sans tenir compte de la structure de la tâche. En effet, les essais n'ont pas nécessairement la même valeur informative : les essais suivants un changement de contingences correspondent à la période d'apprentissage où peuvent se produire des *switches*. A l'inverse, les essais en fin d'épisode correspondent à une période beaucoup plus stable où les contingences sont apprises. Pour cette raison, nous avons préféré la méthode des moindres carrés pour ajuster les paramètres du modèle PROBE car elle tient compte de ce que l'on sait a priori de la structure de la tâche.

Critères de comparaison de modèles : pseudo-R^2 et critères d'information d'Akaike et de Bayes

Afin de comparer différents modèles possibles, il existe un certain nombre de mesures. Nous en présentons trois que nous avons utilisées notamment pour comparer plusieurs variantes du modèle PROBE.

Le pseudo-R^2 (ou R^2 de McFadden) est une mesure de la qualité explicative d'un modèle par rapport au hasard. Il est défini de la manière suivante :

$$Pseudo - R^2 = \frac{LLH_0 - LLH_M}{LLH_0} = 1 - \frac{LLH_M}{LLH_0}$$

Où LLH_0 représente la log-vraisemblance du modèle aléatoire, dont les performances sont au niveau du hasard, et LLH_M représente la log-vraisemblance du modèle que l'on souhaite tester.

Si $\mathrm{LLH}_M = \mathrm{LLH}_0$, le pseudo-$R^2$ est nul ce qui signifie que le modèle ne fait pas mieux que le hasard.

Si $\mathrm{LLH}_M = 0$ (valeur maximale possible d'une log-vraisemblance), le modèle est 'parfait' au sens où il prédit exactement l'ensemble des données.

Le critère d'information d'Akaike (AIC) et le critère d'information bayésien (BIC). Plus un modèle contient de paramètres plus il a de chances de bien s'ajuster aux données. Mais en contrepartie peut se produire un phénomène d'*overfitting* ou 'sur-ajustement'. Un modèle très complexe, même absurde, peut en effet très bien s'ajuster aux données. Ce type de modèles ne peut se généraliser au-delà des données sur lesquelles il a été ajusté. Afin de minimiser ce phénomène,

1. Remarquons qu'étant donné qu'une probabilité est comprise entre 0 et 1, la log-vraisemblance est un nombre négatif.

les différents modèles sont comparés au moyen de critères classiquement utilisés qui pénalisent les modèles dont le nombre de paramètres est important : l'AIC (*Akaike Information Criterion*) et le BIC (*Bayesian Information Criterion*). Ces deux critères ne sont pas interprétables en tant que tel mais uniquement par comparaison (si l'on compare deux modèles, celui dont l'AIC et le BIC sont les plus faibles est le 'meilleur' modèle). Ces deux critères sont définis comme suit :

$$BIC = -2LLH + kln(n)$$
$$AIC = 2k - 2LLH$$

Où LLH correspond à la log-vraisemblance du modèle que l'on teste, k est le nombre de paramètres ajustés du modèle et n le nombre d'observations (c'est-à-dire le nombre d'essais).

Il est à noter que l'AIC et le BIC sont deux critères qui ont été conçus pour comparer des modèles computationnels dont les paramètres n'ont pas nécessairement une signification biologique ou psychophysique. Dans notre cas, les valeurs de ces critères sont à nuancer puisque nous nous intéressons à la modélisation du vivant, et donc à des paramètres qui ne vont pas nécessairement dans le sens de ces critères mais ont une vraie signification biologique ou psychologique.

9.1.3 Tirage des paramètres

Il existe plusieurs techniques pour effectuer le tirage initial des paramètres avant d'effectuer une optimisation. La technique la plus simple est la recherche par grille (ou 'grid search'). Cette technique consiste à :

1) Définir un nombre fini de valeurs à tester pour chaque paramètre. Par exemple, pour un taux d'apprentissage, les onze valeurs allant de 0 à 1 avec un pas de 0.1.

2) Tester chaque combinaison de paramètres. Par exemple, si nous décidons de tester nos 9 paramètres avec 10 valeurs chacune, nous ferons tourner le modèle 109 fois (soit un milliard de fois).

3) Enregistrer, pour chaque combinaison de paramètres, la valeur du critère d'ajustement.

4) Choisir la combinaison de paramètres qui minimise (ou maximise selon le critère utilisé) le critère d'ajustement.

Cette technique systématique présente deux inconvénients majeurs. D'une part elle nécessite un temps de calcul très élevé lié au nombre important de paramètres et de valeurs à tester. En effet, l'ajout d'un paramètre augmente exponentiellement le temps de calcul. D'autre part, la plupart des paramètres ne prennent pas des valeurs discrètes mais peuvent prendre n'importe quelle valeur le long d'un continuum et doivent être explorés plus précisément. Or cette méthode de recherche par grille offre un niveau de précision très modéré. Pour ces deux raisons, nous avons utilisé cette technique uniquement pour des explorations préliminaires, notamment pour tester différentes valeurs de la taille de la mémoire

de travail, seul paramètre discret du modèle PROBE.

Nous avons par la suite utilisé les outils d'optimisation de Matlab (`http://www.mathworks.fr/fr/products/optimization/`). L'Optimization Toolbox™ propose des algorithmes permettant de résoudre des problèmes continus d'optimisation non linéaire, en particulier pour effectuer des opérations de minimisation ou de maximisation de fonctions. Pour chaque paramètre que nous souhaitons ajuster, nous sélectionnons dans un premier temps une valeur initiale par tirage aléatoire sur une distribution normale centrée sur la moyenne probable de ce paramètre. Cette moyenne est déterminée lors des explorations préliminaires ou d'après les valeurs obtenues lors d'ajustements sur les données expérimentales d'études précédentes. Ceci permet de contraindre le tirage initial des paramètres autour de valeurs plus probables et d'éviter d'explorer excessivement des valeurs extrêmes moins plausibles. A partir de ces points de départ, les algorithmes d'optimisation réalisent une exploration dite 'par descente (ou montée) de gradient'. C'est à dire qu'ils explorent pas à pas les valeurs aux alentours de la valeur de départ jusqu'à trouver celle qui minimise (ou maximise) le critère d'ajustement. Afin d'éviter de tomber dans des extrema locaux, cette procédure est répétée une centaine de fois avec des points de départ initiaux différents à chaque fois. Cette méthode permet d'explorer tout le continuum de valeurs (avec une précision définie) mais le gain en temps de calcul reste cependant faible et il n'est jamais entièrement garanti d'identifier la solution. La procédure d'ajustement des paramètres a donc représenté une étape particulièrement longue du projet.

9.2 Ajustement des paramètres du modèle PROBE aux données comportementales

« Plus ça rate, plus on a de chances que ça marche »
Devise Shadock

9.2.1 Choix des paramètres à ajuster

Lors du choix des paramètres à ajuster, plusieurs questions se posent, en particulier pour limiter au maximum le temps et la complexité des calculs tout en obtenant le maximum d'information nécessaire.

Une première question consiste à se demander si certains paramètres peuvent ne pas être ajustés mais fixés à une valeur prédéterminée qui soit la même pour tous les sujets. C'est le cas par exemple pour la taille de la mémoire de travail ou la volatilité.

En ce qui concerne la taille de la mémoire de travail, seul paramètre discret du modèle PROBE, par souci de plausibilité, nous limitons sa capacité. En effet, si l'ordinateur a une mémoire parfaite, ce n'est pas le cas des humains. Mais cette limite est-elle la même pour tous les sujets ? Bien que la taille de cette mémoire de travail varie d'un individu à l'autre (de 1 à 5 environ), elle est en moyenne de

trois. Ainsi, et pour simplifier les analyses suivantes, nous avons fait le choix de fixer la taille de la mémoire de travail à trois pour l'ensemble des sujets.

Dans le cas de la volatilité, il est possible de la fixer à sa valeur réelle. En effet, la volatilité correspond à la probabilité qu'à un essai donné se produise un changement de contingences. La volatilité réelle (1/nombre moyen d'essais par épisode) peut donc être calculée. Si on la laisse fixée à cette valeur réelle pour l'ensemble des sujets, on fait l'hypothèse qu'ils perçoivent effectivement cette volatilité réelle. Lors des explorations préliminaires et dans un souci de diminution des temps de calcul, nous avons fait le choix de fixer la volatilité à sa valeur réelle. Par la suite nous l'avons ajustée. En effet, la volatilité perçue par les sujets est souvent supérieure à la volatilité réelle et de plus variable entre les sujets.

Une autre question que l'on peut soulever est celle de la stabilité d'un paramètre pour chaque individu. Autrement dit, un paramètre conserve-t-il la même valeur fixe tout au long de l'expérience ? L'équipe de Rushworth a par exemple proposé un modèle dans lequel le taux d'apprentissage varie au cours d'une même expérience en fonction de la volatilité [18]. Plus la volatilité est élevée plus le taux d'apprentissage augmente et inversement. Il ne s'agit pas d'un ajustement à chaque pas de temps mais par bloc d'essais. Il n'est cependant pas toujours facile de déterminer si un paramètre peut être considéré comme fixe ou s'il doit être vu comme une variable évoluant au fil du temps. Lorsque nous considérons qu'un paramètre est fixe, nous supposons qu'il est le reflet d'une caractéristique psychologique ou neurologique propre, variable entre individus mais constante pour un même individu au cours du temps (ou d'un temps donné). Dans notre cas, même si certains paramètres gagneraient éventuellement à être traités comme des variables, en l'absence d'un modèle capable de prédire leur évolution au cours du temps, nous avons choisi de simplifier le problème et les considérer comme fixes pour chaque sujet au moins au sein d'une même session expérimentale.

Au final, sur les neuf paramètres que comporte le modèle PROBE, nous en avons ajusté huit et maintenu un constant (cf. 2.2.2 pour la description détaillée du modèle PROBE et de ses paramètres) :
1 - N : taille de la mémoire de travail fixée à 3 pour l'ensemble des sujets.
2 - α : taux d'apprentissage du modèle d'apprentissage par renforcement ($0 \leq \alpha \leq 1$)
3 - β : température inverse du softmax (≥ 0)
4 - ϵ : taux d'exploration du softmax ($0 \leq \epsilon \leq 0.5$)
5 - τ : volatilité perçue ($0 \leq \tau \leq 1$)
6 - θ : niveau de confiance initial attribué au probe ($0 \leq \theta \leq 1$)
7 - η : biais d'initialisation de la mémoire à long terme sur le probe ($0 \leq \eta \leq 1$)
8 - κ : poids a priori du modèle contextuel (≥ 1)
9 - δ : biais du contexte pour le *switch* ($0 \leq \delta \leq 1$)

9.2.2 Méthode d'ajustement des paramètres du modèle PROBE

En première approche, nous pouvons envisager d'ajuster un unique ensemble de paramètres pour les trois sessions confondues. Ceci suppose de faire l'hypothèse que les paramètres du sujet ne varient pas d'une session à l'autre. Or, notre protocole expérimental comporte des sessions (apprentissage et test) dont les structures et les conditions diffèrent. De plus, lors d'un changement de session, il est possible qu'une 'réinitialisation' des paramètres ait lieu compte tenu du nouvel environnement. Ceci nous a amené à ajuster séparément un ensemble de paramètres pour la session d'apprentissage et un ensemble de paramètres pour la session test (soit pour les deux sessions d'IRM confondues). Cependant, lors de la session d'apprentissage, certaines combinaisons et associations ont pu être retenues et stockées en mémoire à long terme. Comme on ajuste les paramètres pour chaque session indépendamment les unes des autres, on perd le contenu de la mémoire à long terme. Pour pallier ce problème, nous avons procédé de la manière suivante :

1. On ajuste les paramètres du modèle PROBE sur les données de chaque sujet lors de la session d'apprentissage. On obtient ainsi pour chaque sujet un ensemble de paramètres θ_A.

2. On fait tourner le modèle PROBE avec chaque séquence d'actions des sujets et l'ensemble des paramètres θ_A. On obtient ainsi le contenu de la mémoire à long terme M_A à l'issu de la session. On garde ainsi en mémoire ce qui a été appris lors de la session d'apprentissage.

3. On ajuste les paramètres du modèle PROBE sur les données de la session test en initialisant la mémoire à long terme par la mémoire à long terme M_a. On obtient ainsi l'ensemble de paramètres θ_T pour la session test.

4. On effectue une réinitialisation des variables du modèle au moment du changement de session IRM.

Chapitre 10

L'Imagerie par Résonance Magnétique fonctionnelle (IRMf)

« fMRI is not and will never be a mind reader »
Nikos Logothetis

10.1 Principe et acquisition des données

10.1.1 Un bref historique

L'Imagerie par Résonance Magnétique (IRM) représente l'une des avancées les plus importante dans le domaine de l'imagerie du vivant depuis la découverte des rayons X par Conrad Röntgen en 1895 [201]. En 1973, le chimiste Paul Lauterbur est le premier à décrire les bases de cette nouvelle technique [143], ce qui lui vaut un prix Nobel trente ans plus tard. Il faut cependant attendre encore vingt ans avant que Seiji Ogawa, un chercheur des laboratoires Bell, décrive comment tirer partie des propriétés magnétiques de l'hémoglobine [173]. C'est grâce à cette découverte qu'en 1991 Bellivau et collaborateurs parviennent à réaliser les première images d'un cerveau en fonctionnement à l'aide d'une machine IRM [101]. Avant l'avènement des techniques d'imagerie cérébrale telle que l'IRM, les chercheurs devaient se contenter de tirer des conclusions sur les fonctions cérébrales à partir d'observations de patients avec des lésions cérébrales ou par des études invasives chez l'animal. L'IRM permet désormais d'étudier un cerveau humain sain en activité.

Cependant, même si les "régions allumées" du cerveau créent une fascination populaire et beaucoup de spéculations dans la presse, il faut garder en tête les limites de cette technique et de son interprétation.

10.1.2 Le signal *BOLD* (*Blood Oxygen Level Dependant*)

Plusieurs techniques d'IRMf sont utilisées de nos jours, mais la plus commune est celle fondée sur le signal de détection du niveau d'oxygène dans le sang (signal *BOLD*). Ce signal est lié aux propriétés magnétiques de l'hémoglobine contenue dans les globules rouges du sang. L'hémoglobine se trouve sous deux formes :

- Les globules rouges oxygénés par les poumons contiennent de l'oxyhémoglobine, une molécule non active en IRM ;

- Les globules rouges désoxygénés au niveau des tissus contiennent de la désoxyhémoglobine qui présente des propriétés magnétiques : sous l'effet d'un champ magnétique (produit par l'IRM), elle acquière une aimantation dirigée dans le même sens que le champ magnétique, détectable par l'IRM.

L'activité cérébrale entraine une petite augmentation de la consommation d'oxygène par les neurones et donc des modification hémodynamiques. C'est le phénomène de couplage neuro-vasculaire. Dans les zones cérébrales activées, c'est-à-dire où les neurones sont stimulés, on observe à la fois une augmentation de la consommation en oxygène et une augmentation du débit sanguin cérébral local. Mais l'augmentation du débit sanguin cérébral est proportionnellement plus importante que la consommation d'oxygène. Le ratio consommation sur débit diminue donc. Ceci entraine une diminution relative de la désoxyhémoglobine qui peut être enregistrée par l'IRM. Cette variation est très faible (de l'ordre de 1%) et nécessite des méthodes statistiques puissantes pour être mise en évidence.

L'IRM fonctionnelle est donc une mesure indirecte de l'activité neurale. De plus, étant donné que les modifications hémodynamiques se produisent plusieurs secondes après l'activation neurale, la résolution temporelle de l'analyse est limitée. Cependant, la résolution anatomique est bien meilleure qu'avec d'autres techniques.

10.1.3 Paramètres d'acquisition

Dans une session d'IRM classique, le sujet est placé dans un fort champ magnétique (en général 1.5 ou 3 Tesla ; en comparaison, le champ magnétique terrestre est de 0.00005 Tesla). Dans le cas de notre étude, les images d'IRM fonctionnelle ont été acquises avec un scanner 3 Tesla Siemens (Erlangen, Allemagne) Trio du Centre de Neuroimagerie de Recherche (CENIR) à l'hôpital de la Pitié-Salpêtrière.

Durant chaque session d'IRM, des cartes de champ (*field maps*) sont acquises, ainsi que trois runs d'images fonctionnelles ou EPI (*Echo Planar Imaging*) : en combinant les deux sessions, chaque sujet réalise donc un total de six runs en IRM, chacun d'une durée de vingt minutes en moyenne. Durant la première session, une image anatomique (T1) est également acquise.

Le temps d'acquisition d'un volume entier du cerveau (ou temps de répétition) est de deux secondes. Chaque volume fonctionnel est composé de 37 coupes acquises de façon entrelacée (acquisition d'abord des coupes impaires puis des

coupes paires).

L'unité de volume de référence est le voxel, un cube d'acquisition couvrant une petite région du cerveau dans laquelle l'activité peut être mesurée : ici, notre volume d'acquisition représente un total de 64 x 64 x 37 voxels, chaque voxel mesurant 3 mm d'arête.

Une fois acquises, les données son traitées avec le logiciel SPM8 (*Statistical Parametric Mapping*, *Wellcome Trust Centre for Neuroimaging*, *London*, `http://www.fil.ion.ucl.ac.uk/spm`).

10.2 Prétraitement du signal (*preprocessing*)

Une série d'opérations, le prétraitement (*preprocessing*), permet de préparer les images fonctionnelles pour les analyses statistiques. Une chaine de prétraitements comprend plusieurs étapes décrites ci-dessous. L'ordre dans lequel s'effectuent ces étapes dépend de l'ordre d'acquisition des coupes d'un volume fonctionnel : le mode séquentiel (les coupes sont acquises les unes après les autres du haut vers le bas ou du bas vers le haut) ou le mode entrelacé (acquisition d'abord des coupes impaires puis des coupes paires, du haut vers le bas ou du bas vers le haut). Dans notre expérience, nous avons choisi le mode entrelacé avec une acquisition du bas vers le haut de 37 coupes.

10.2.1 La correction du décalage d'acquisition entre coupes (*slice timing*)

Lors d'une acquisition fonctionnelle, les différentes coupes d'un même volume ne sont pas acquises simultanément, mais successivement pendant une durée égale au temps de répétition. Dans notre expérience, le temps de répétition est d'une durée de deux secondes, ce qui signifie que la dernière coupe est acquise deux secondes après la première.

La correction de ce décalage consiste à "ramener", par interpolation temporelle, l'instant d'acquisition de toutes les coupes à un instant commun, qui est l'instant d'acquisition d'une des coupes du volume (la coupe de référence). Nous avons choisi comme coupe de référence la coupe 37, soit celle acquise au milieu de la séquence (la dernière coupe d'acquisition des coupes impaires). Lors des analyses statistiques ultérieures, on considère que toutes les coupes du volume ont été acquises simultanément.

10.2.2 La correction du mouvement (*realign*)

L'objectif de cette étape est de corriger les artéfacts dus aux mouvements de la tête du sujet. Ces mouvements sont inévitables pendant une acquisition qui peut durer plusieurs dizaines de minutes. Si ces mouvements ne sont pas corrigés, il peut en résulter des "faux positifs" dans les cartes d'activation, c'est-à-dire des

voxels qui sont considérés comme activés après seuillage alors qu'ils ne le sont pas en réalité.

Le principe du recalage consiste à choisir une image de référence au sein de la série temporelle acquise et à corriger le déplacement des autres images de la série par rapport à cette image de référence (dans notre cas il s'agit de la même image de référence que pour la correction de décalage d'acquisition). On suppose que le déplacement est composé uniquement de trois rotations et de trois translations.

10.2.3 Le recalage (*coregister*)

Cette étape permet de mettre en correspondance des images issues d'une même modalité ou de modalités différentes. Typiquement, on l'utilise pour recaler les images fonctionnelles et l'image anatomique d'un même sujet.

10.2.4 La normalisation spatiale (*normalize*)

L'objectif de cette étape est de placer toutes les images anatomiques et fonctionnelles des sujets dans un espace commun (ou *template*) : le repère *MNI* (*Montreal Neurological Institute*). Ceci permet par la suite de déterminer les coordonnées des zones activées dans un repère commun, et donc de comparer des activations entre les différents sujets en contournant les problèmes posés par la variabilité anatomo-fonctionnelle interindividuelle.

10.2.5 Le lissage spatial (*smoothing*)

Les données d'IRM fonctionnelle présentent des corrélations spatiales : le signal acquis dans un voxel n'est pas indépendant du signal acquis dans les voxels voisins. Les caractéristiques de ces corrélations spatiales sont difficiles à estimer. En pratique, on lisse les données avec un filtre passe-bas (qui coupe les hautes fréquences) de façon à ce que, dans les images lissées, les caractéristiques des corrélations spatiales soient connues et imposées par les propriétés du filtre utilisé.

La méthode de seuillage que l'on applique par la suite aux cartes d'activation s'appuie sur la théorie des champs aléatoires gaussiens. Pour que les hypothèses sous-jacentes à cette théorie soient vérifiées, on lisse spatialement les données avec un filtre gaussien (dans notre cas un filtre de 6 mm).

10.3 L'approche *model-based*

10.3.1 Principe de la méthode *model-based*

Afin d'assurer notre survie, l'évolution a façonné notre cerveau pour lui permettre d'optimiser la sélection et l'exécution d'actions. Ceci nécessite de :

1. Traiter l'information sensorielle perçue
2. Evaluer la pertinence de cette information en fonction de l'environnement extérieur dans lequel on se trouve et de notre état interne
3. Choisir une action adaptée parmi plusieurs alternatives
4. Exécuter le mouvement approprié
5. Evaluer la pertinence du choix selon le résultat observé

Souvent, ces processus reposent sur des variables internes qui ne sont pas directement observables par l'expérimentateur, telles que l'expérience passée, la valeur subjective ou la motivation. Même si ces variables peuvent être influencées par des conditions expérimentales, elles ne peuvent être mesurées directement ou contrôlées précisément. Cependant une étude exhaustive des processus neuraux qui sous-tendent le contrôle cognitif nécessite de prendre ces variables en compte. Or, les méthodes expérimentales traditionnelles ne sont souvent pas capables d'élucider ces processus computationnels [57].

Comment alors peut-on étudier ces processus dynamiques non mesurables directement ? Pour comprendre leur implémentation, il est nécessaire d'élaborer des modèles formels qui proposent des algorithmes de calcul de ces variables internes. Il est possible ensuite de confronter ces variables théoriques aux données comportementales et neurales. L'approche *model-based*, permet ainsi d'effectuer des descriptions mécanistiques précises et des prédictions quantitatives au sujet des états cachés que notre cerveau représente mais qui ne sont pas directement mesurables. Cette approche permet d'énoncer des hypothèses sur la manière dont le cerveau pourrait élaborer différents processus cognitifs : ainsi, en plus de répondre à la question "où sont implémentés ces processus ?", elle apporte des éléments de réponse à la question "comment sont-ils implémentés ?" [73].

Cette méthode, appliquée à l'IRM fonctionnelle, a été décrite par John P. O'Doherty, Alan Hampton et Hackjin Kim [172] et est de plus en plus utilisée dans l'étude des bases neurales du contrôle cognitif et de la prise de décisions.

10.3.2 Etapes de la méthode

Elaboration d'un modèle computationnel décrivant un processus cognitif

L'approche *model-based* repose sur un modèle computationnel qui décrit quantitativement les liens entre les stimuli, les actions et les résultats obtenus. Ce modèle doit être en mesure de réaliser la tâche effectuée par les sujets en décrivant la transformation des stimuli en réponses comportementales observables. Il doit par ailleurs contenir les variables non directement observables qui affectent

cette transformation. En règle générale, il n'existe pas un unique modèle mais un ensemble de modèles susceptibles de convenir. Une comparaison entre modèle est nécessaire pour déterminer celui qui est le mieux à même d'expliquer les données.

Ajustement des paramètres du modèle sur les données comportementales

Les paramètres doivent être ajustés à une certaine valeur afin que le modèle soit en mesure de faire des prédictions. Cet ajustement se fait sur les données comportementales de manière à minimiser la différence entre celles-ci et les prédictions du modèle, et a lieu avant que le modèle ne soit utilisé pour prédire les fluctuations des variables cachées essai par essai.

Pour une description détaillée de la méthode, voir chapitre 9.

Comparaison du signal théorique et du signal observé

Après ajustement des paramètres, les variables générées par le modèle sont utilisées comme facteurs explicatifs du signal *BOLD* (*Blood Oxygenation Level-Dependent*), en tant que régresseurs du Modèle Linéaire Généralisé (pour une description plus détaillée, voir 10.4). Autrement dit, les variables sont entrées en tant que série temporelle et convoluées avec la fonction de réponse hémodynamique canonique (*hrf*). La réponse théorique ainsi obtenue est comparée au signal *BOLD* observé par régression linéaire afin d'identifier les régions cérébrales dont l'activation varie en corrélation avec les variables du modèle.

10.3.3 Un exemple historique : erreur de prédiction et dopamine

Le modèle de Rescorla et Wagner

Les exemples les plus simples de modèles théoriques utilisés en psychologie sont ceux qui ont été développés pour décrire le processus de conditionnement pavlovien (voir figure 10.1) [179]. Le modèle de Rescorla et Wagner est le plus souvent cité [239]. Dans ce modèle, deux variables représentent le processus par lequel un stimulus conditionné (le son d'une cloche) déclenche une réponse conditionnée (la salivation au moment du son de cloche) :

- v : l'intensité de la réponse conditionnée (le degré de salivation au moment du son de cloche), ou, de manière plus abstraite, la valeur du stimulus conditionné (à quel point le son de la cloche prédit une récompense),

- u : la valeur du stimulus non conditionné (la nourriture / récompense).

Pour que le conditionnement opère, la variable v (qui peut initialement valoir zéro lorsque le stimulus conditionné n'est associé à aucune récompense) doit

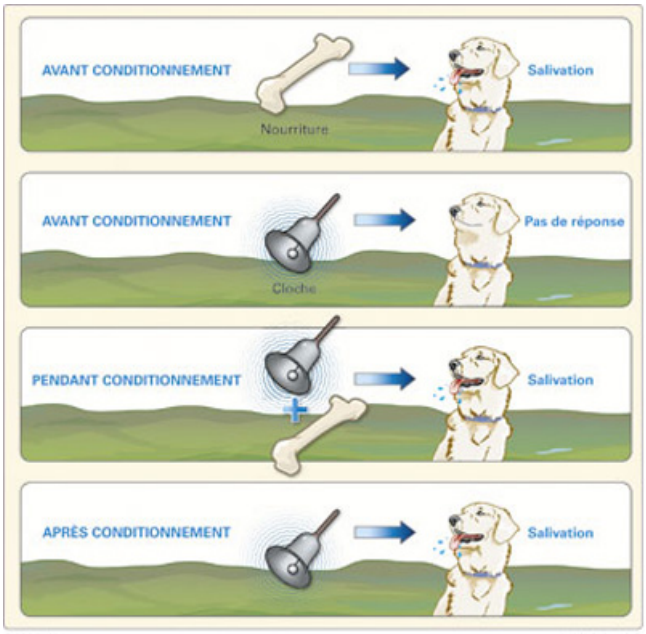


FIGURE 10.1 – **Conditionnement pavlovien**. Avant conditionnement, un stimulus inconditionné (de la nourriture) déclenche une réponse inconditionnée de manière réflexe, sans apprentissage nécessaire (la salivation). Par ailleurs, un stimulus neutre (son de cloche) ne déclenche aucune réponse spécifique. Lors du conditionnement, les stimuli inconditionné et neutre sont présentés simultanément. Après le conditionnement, le stimulus initialement neutre déclenche une réponse conditionnée. Le stimulus neutre devient un stimulus conditionné.

converger vers la valeur de u. C'est à dire qu'à la fin du conditionnement, le stimulus conditionné prédit entièrement la récompense et a donc la même valeur.

Plus formellement, à chaque essai, le modèle de Rescorla et Wagner fait une prédiction au sujet de v à l'essai t (v_t) et met à jour cette estimation pour l'essai suivant, v_{t+1}, en fonction de l'erreur de prédiction, δ_t :

$$v_{t+1} = v_t + \alpha \times \delta_t$$

Où α est le paramètre de taux d'apprentissage, et l'erreur de prédiction à t, δ_t, est définie comme la différence entre la récompense réellement obtenue u_t et la récompense attendue v_t : $\delta_t = u_t - v_t$.

La valeur de v est ainsi mise à jour en proportion de δ, de manière à ce qu'au cours de l'apprentissage, v converge vers u et δ tend vers 0. Ce modèle simple comprend trois variables d'intérêt, v, δ et u. En psychologie comportementale traditionnelle, la seule variable directement observable ici est v, parce qu'elle peut être mesurée en évaluant l'intensité de la réponse conditionnée chez le sujet (animal ou humain) effectuant la tâche de conditionnement. Les autres variables, telles que δ, ne sont pas observables directement mais peuvent être inférée grâce au modèle.

Ainsi, l'erreur de prédiction, non mesurable, peut être étudiée grâce à cet algorithme pour une tâche de conditionnement pavlovien, au cours duquel un stimulus arbitraire, initialement neutre affectivement, endosse une valeur appétitive

par sa présentation répétée avec un stimulus appétitif. La règle de Rescorla et Wagner s'est montrée efficace pour rendre compte de nombreux phénomènes de conditionnement classique.

Enregistrements de neurones dopaminergiques chez le singe

Schultz, Dayan et Montague ont proposé en 1997 une expérience de conditionnement effectuée sur des singes [207]. Les singes voient une lumière (le stimulus conditionné) et doivent appuyer sur un levier pour obtenir une récompense (jus, stimulus non conditionné) après un certain délai. L'activité des neurones dopaminergiques de l'aire tegmentale ventrale est mesurée. En début d'apprentissage, ces neurones répondent de manière phasique à la récompense (figure 10.2, en haut). Cependant, au cours de l'apprentissage, cette activation phasique se produit de plus en plus tôt avant la récompense, tandis qu'on ne l'observe plus au moment de la récompense. En fin d'apprentissage, l'activation phasique a lieu au moment de l'apparition de l'indice prédictif de la récompense (le stimulus conditionné) (figure 10.2, au milieu). Si la récompense est omise, on observe une désactivation phasique des neurones dopaminergiques à l'instant prédit de la récompense (figure 10.2, en bas). Ce pattern d'activations est très bien expliqué par l'hypothèse que la dopamine code une erreur de prédiction. Plus précisément, au début de l'expérience aucun renforcement n'est attendu (le stimulus conditionné n'est pas prédictif d'une récompense). Par conséquent, à la réception de la récompense, on a une erreur de prédiction élevée (quasiment égale à la valeur de la récompense), qui correspond au pic d'activation lors de la récompense. Au fur et à mesure de l'apprentissage, le singe apprend que le stimulus conditionné prédit une récompense ce qui diminue l'erreur de prédiction (jusqu'à l'annuler) et correspond à la disparition du pic à la présentation de la récompense. L'erreur de prédiction se propage en arrière jusqu'à être reportée au stimulus conditionné. Ainsi, quand le stimulus conditionné est présenté en fin d'apprentissage, on a une erreur de prédiction positive, ce qui correspond au pic d'activation observé pour le stimulus conditionné. Enfin, si la récompense n'est pas présentée une fois le comportement appris, on a une erreur de prédiction négative, ce qui correspond à la désactivation observée en cas d'omission de la récompense. Cette étude suggère donc que la dopamine code pour une erreur de prédiction.

Notons que le pattern de réponse de ces neurones dopaminergiques ne correspond pas à une simple règle de Rescorla et Wagner mais plutôt à une extension en temps réel de cette règle appelé *temporal difference learning* dans laquelle les prédictions des récompenses futures sont calculées à des intervalles de temps discrets au sein de chaque essai.

Cette étude est une des premières à avoir corrélé une activité neuronale avec une variable cachée calculée par un modèle théorique. Elle a posé les bases de l'approche *model-based*.

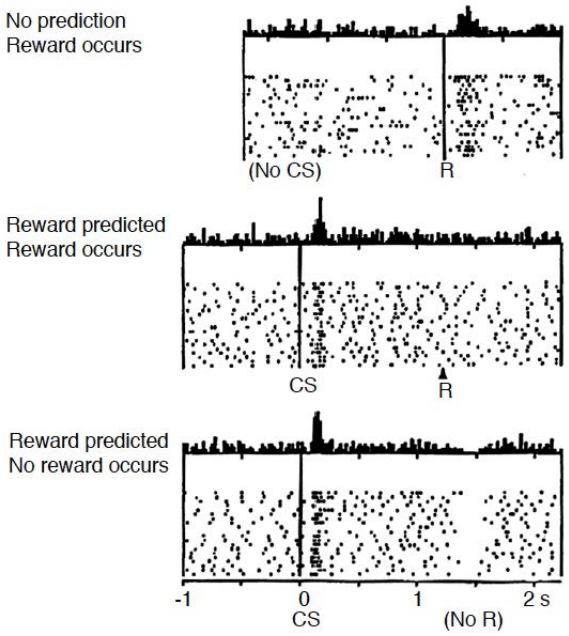


FIGURE 10.2 – Activité des neurones dopaminergiques suite à une récompense R non prédite (en haut), suite à un stimulus conditionné (CS) suivi de la récompense prédite (milieu) et suite à un stimulus conditionné avec omission de la récompense prédite (bas). Reproduit de [207].

IRM fonctionnelle chez l'humain

Pour tester l'existence d'un signal d'erreur de prédiction dans le cerveau humain, O'Doherty et ses collègues ont proposé un protocole d'IRM fonctionnelle utilisant une tâche classique de conditionnement. On présente aux sujets humains des stimuli visuels arbitraires (des fractales). Un stimulus est suivi d'une récompense (une solution sucrée), alors qu'un autre stimulus est non récompensé. Dans certains essais peu fréquents, le stimulus récompensant est présenté mais la récompense est omise, et inversement le stimulus non récompensant est présenté et une récompense inattendue est délivrée. La séquence d'essais que chaque sujet a expérimentée est incluse dans un modèle de *temporal difference learning* afin de générer une série temporelle spécifiant l'erreur de prédiction à différents temps donnés [171]. Cette série temporelle est ensuite convoluée avec une fonction de réponse hémodynamique canonique pour générer un signal théorique. Ce signal théorique est comparé au signal *BOLD* enregistré pour chaque sujet afin d'identifier les régions cérébrales où les signaux corrèlent. Des corrélations significatives ont été observées dans plusieurs régions cérébrales, et en particulier dans le striatum ventral et le cortex orbitofrontal, tous deux des régions cibles de neurones dopaminergiques. Ces résultats suggèrent donc que les signaux d'erreur de prédiction sont présents dans le cerveau humain et que leurs variations sont en accord avec le modèle computationnel spécifique de *temporal difference learning*.

D'autres études ont permis de reproduire ce résultat [155] et d'étendre l'analyse *model-based* à des tâches et des modèles plus complexes. Cette approche a

été employée dans plusieurs domaines tels que l'apprentissage associatif [68], les interactions sociales [17] ou le contrôle cognitif [32][249].

10.3.4 Limites de la méthode

Notons qu'obtenir des corrélations entre les variables d'un modèle théorique et des activations cérébrales ne démontre pas de manière concluante que la région activée implémente le modèle considéré. Utiliser la méthode *model-based* revient à tester des signaux cérébraux très spécifiques correspondant à des hypothèses computationnelles précises. En contrepartie, cette approche limite la possibilité de détecter des évènements inattendus qui ne correspondent pas à des hypothèses a priori. De plus, la méthode *model-based* n'exclut pas tous les inconvénients intrinsèques de l'IRM fonctionnelle (mesure indirecte de l'activité neuronale, mauvaise résolution spatio-temporelle, etc.)[150].

10.4 Modèle linéaire général

10.4.1 Spécification du modèle

Principe général

Pour chaque voxel du cerveau, notons $\mathbf{Y}$ le signal *BOLD* acquis au cours du temps, c'est-à-dire la réponse hémodynamique mesurée dans ce voxel. $\mathbf{Y}$ est une fonction du temps comprenant autant de mesures que de volumes d'acquisition (notons q ce nombre de mesures). $\mathbf{Y}$ ne représente pas le signal mesuré brut mais les données prétraitées.

On fait classiquement l'hypothèse que la réponse hémodynamique est linéaire. Cela signifie que si le protocole expérimental comprend n types d'évènements (par exemple l'apparition d'un stimulus), si la réponse mesurée dans un voxel pour les événements de type j est Y_j, alors le signal $\mathbf{Y}$ acquis dans ce voxel est : $Y = Y_1 + Y_2 + ... + Y_n$.

Spécifier un modèle signifie traduire sous forme mathématique les hypothèses quant à la forme des réponses Y_1, Y_2,...,Y_n associées à chacun des types d'évènements. On connaît donc les valeurs du signal $\mathbf{Y}$ en q points, et spécifier le modèle consiste à définir la forme théorique de la réponse hémodynamique pour ces mêmes q points.

Il faut donc choisir pour chaque type d'évènement j une forme théorique pour la réponse Y_j. Cette forme est modélisée par une ou plusieurs fonctions temporelles, qu'on appelle des régresseurs, et qui sont, comme $\mathbf{Y}$, des vecteurs de taille q. Ces régresseurs sont regroupés dans une matrice $\mathbf{X}$ qu'on appelle la matrice de dessin expérimental. $\mathbf{X}$ comporte q lignes et autant de colonnes que de régresseurs.

Pour un seul événement de durée très courte, quasi instantanée (par exemple

la visualisation d'un bref stimulus visuel), on peut considérer que le signal *BOLD* a une forme proche de la fonction hémodynamique dite canonique (*hrf*). La *hrf* atteint son maximum 5 secondes après l'événement qui l'a induite.

Si on suppose que le signal *BOLD* est plus complexe (par exemple, le maximum n'est pas atteint exactement en 5 secondes dans tous les voxels), on peut considérer que la forme du signal est plus complexe que la *hrf* seule, et est plus proche de la combinaison linéaire de plusieurs fonctions, par exemple une *hrf* et sa dérivée. On choisit alors de modéliser la condition avec deux "fonctions de base".

Dans notre cas, nous avons choisi de modéliser le signal avec la *hrf* seule.

Par ailleurs, dans le cas où un type d'évènement est répété plusieurs fois à différents instants du paradigme, on fait l'hypothèse que, à chaque occurrence, ce type d'événement apporte une contribution identique au signal *BOLD* mesuré. Autrement dit, la réponse au stimulus est invariante dans le temps.

Chaque régresseur est donc une fonction temporelle constituée d'une suite de fonctions élémentaires *hrf* synchronisées avec les instants de présentation des évènements. Ils traduisent ainsi la forme attendue de la variation du signal *BOLD* induite par chaque type évènement.

SPM construit alors un modèle statistique appelé modèle linéaire général, qui est une généralisation du modèle de régression linéaire, et qui a pour but de décomposer le signal mesuré $\mathbf{Y}$ comme une combinaison linéaire de régresseurs. Le même modèle est appliqué à tous les voxels du cerveau.

Le modèle statistique décompose le signal mesuré comme suit :

$$\mathbf{Y} = \mathrm{M} + \beta_1 \mathrm{R}_1 + \beta_2 \mathrm{R}_2 + \ldots + \beta_n \mathrm{R}_n + \mathrm{E}$$

E est appelée l'erreur résiduelle, et correspond à la partie des données non expliquée par le modèle. C'est un vecteur de taille q. M est la valeur moyenne du signal, qu'on décrit classiquement sous la forme $\mathrm{M} = \beta_0 \mathrm{R}_0$, où R_0 est un vecteur qui vaut "1" pour tous les points de mesure (on l'appelle le régresseur "constant", il est défini automatiquement).

Autrement dit, SPM caractérise la réponse Y_j (contribution des évènements de type j au signal $\mathbf{Y}$ total mesuré dans le voxel) par :

- sa forme théorique, représentée par le régresseur R_j ;
- un coefficient multiplicatif associé β, que l'on appelle coefficient de régression ou paramètre.

Régresseurs

L'ensemble des régresseurs que nous avons modélisés est récapitulé dans le tableau 10.1.

Numéro	Nom	Onset	Durée
1 à 40	Contexte	Apparition du contexte	box-car (toute la durée d'affichage)
41	Essai non-réponse	Stimulus	event
42	Essai non-réponse	Feedback	event
43	3 premiers essais d'un run	Stimulus	event
44	3 premiers essais d'un run	Feedback	event
45	**Apparition d'un contexte ancien**	Stimulus	event
46	**Apparition d'un contexte nouveau**	Stimulus	event
47	Effet de l'apparition d'un contexte ancien	Stimulus	boxcar (durée d'un essai)
48	Effet de l'apparition d'un contexte nouveau	Stimulus	boxcar (durée d'un essai)
49	**Exploitation**	Stimulus	event
50	**Exploration**	Stimulus	event
51	Creation de *probe* proactive	Stimulus	event
52	**Creation de *probe* réactive**	Stimulus	event
53	Creation de *probe* secondaire	Stimulus	event
54	Récupération réactive, contexte ancien	Stimulus	event
55	**Récupération réactive, contexte nouveau**	Stimulus	event
56	Récupération réactive secondaire	Stimulus	event
57	**Récupération proactive, contexte ancien**	Stimulus	event
58	Récupération proactive, contexte nouveau	Stimulus	event
59	Récupération directe a posteriori	Stimulus	event
60	Récupération directe a posteriori secondaire	Stimulus	event
61	**Consolidation**	Stimulus	event
62	Consolidation secondaire	Stimulus	event
63	Stimulus	Stimulus	event
64	Feedback positif	Feedback	event
65	Feedback négatif	Feedback	event
66	Stimulus au feedback	Stimulus	box-car (du stimulus au feedback)
	Modulations paramétriques :		
67	Valeur de l'action		
68	*dummy* λ		
69	**λ de la meilleure alternative**		
70	**λ de la seconde meilleure alternative**		
71	Feedback au stimulus	Feedback	box-car (du feedback au stimulus)
	Modulations paramétriques :		
72	γ acteur		
73	γ de la meilleure alternative		
74	γ de la seconde meilleure alternative		
75	*dummy* μ		
76	**μ de la meilleure alternative**		
77	**μ de la seconde meilleure alternative**		
78 à 82	Runs		boxcar (durée d'un run)
83 à 88	Mouvements		

TABLE 10.1 – **Régresseurs**. Les régresseurs en vert correspondent à ceux qui apparaissent par la suite dans les résultats présentés.

Régresseurs de non intérêt.
- *Runs* : ce régresseur modélise les six *runs* d'IRM fonctionnelle (trois *runs* par session).
- Mouvements : ce régresseur modélise les mouvements du sujet dans le scanner dans les trois rotations possibles de la tête.
- Non réponse : ce régresseur modélise tous les essais où aucune réponse n'a été enregistrée.
- Premiers essais d'un *run* : nous avons modélisé séparément les premiers essais d'un *run* qui correspondent à la reprise de la tâche juste après une pause, lorsque le signal n'est pas encore stabilisé.

Effets du contexte. Nous avons modélisé individuellement chacun des 40 contextes (4 anciens et 36 nouveaux) qui apparaissent lors des sessions d'IRM afin de retirer leurs effets spécifiques non désirés (saillance visuelle, valeur émotionnelle, etc.). Nous avons également modélisé les changements de contextes en dissociant l'apparition des contextes anciens et nouveaux. Ces changements sont modélisés à la fois en *event*, modélisant ainsi l'effet brusque et saillant du changement de contexte ; et en *box-car* afin de modéliser un potentiel effet durable du changement (tel que l'attention visuelle du sujet).

Régresseurs définis par le modèle PROBE. A chaque essai, le modèle PROBE reçoit, maintient ou actualise plusieurs variables : le stimulus présenté, l'action réalisée, la récompense obtenues, les valeurs de chaque action, les prédictions de récompense, les signaux de fiabilité, les transitions d'état (exploitation et exploration). Nous avons modélisé l'ensemble des évènements que nous souhaitons étudier.

Les essais d'exploitation correspondent aux derniers essais de chaque épisode où le sujet, selon la prédiction du modèle, est en état d'exploitation du *task-set* courant. Les essais d'exploration correspondent quant à eux aux essais au milieu de la première période d'exploration de chaque épisode. Ces deux régresseurs représentent des évènements *stay* et nous fournissent une ligne de base pour comparer avec les évènements de transition d'état.

On s'intéresse aux premiers évènements de transition d'état de chaque épisode. C'est pourquoi on modélise les premiers évènements d'une part et les évènements secondaires d'autre part, pour chacune des transitions d'état possibles (création de *probe*, récupération et consolidation). On dissocie également les différents types de transition en fonction de leur moment d'apparition : soit dès le premier essai de l'épisode (au moment du changement de contexte), c'est à dire de façon proactive ; soit un à plusieurs essais après le changement de contexte, de façon réactive. On dissocie également en fonction du type de contexte (ancien ou nouveau). Ces diverses dissociations nous permettent de dégager des évènements "purs". Par exemple, dans les épisodes où le contexte est nouveau, les récupérations sont nécessairement fondées sur les $feedbacks$. On est donc pratiquement certains qu'ils s'agit de récupérations réactives où le contexte n'a pas d'effet.

Enfin, les régresseurs "stimulus" et "feedback" possèdent des modulations paramétriques issues des variables du modèle PROBE.

Au moment du stimulus, on modélise : la valeur de l'action choisie (selon le module sélectif), la fiabilité du *task-set* aléatoire (*dummy*), la fiabilité *ex-ante* (λ) du meilleur *task-set* alternatif dans la mémoire de travail et la fiabilité *ex-ante* du second meilleur *task-set* alternatif.

Au moment du $feedback$, on modélise : la prédiction (γ) de la récompense pour le *task-set* acteur (selon le module prédictif), la prédiction de la récompense pour le meilleur et second meilleur *task-sets* alternatifs et la fiabilité *ex-post* (μ) pour le meilleur et second meilleur *task-set* alternatif.

Notons que ces régresseurs sont modélisés en *box-car*. En effet, il est difficile de déterminer avec précision le moment de mise à jour de ces variables. Pour les capturer, cette modélisation s'est avérée préférable à une modélisation en *event*.

Par ailleurs, nous n'avons pas pu modéliser les signaux de confiance du *task-set* acteur avec les alternatives car, par construction, leur corrélation est trop élevée. En effet, dans le modèle PROBE, lorsque la fiabilité de l'acteur chute, les fiabilités des *task-sets* alternatifs ont tendance à augmenter simultanément. Ces interdépendances entre variables sont le fruit de la structure même du modèle et peuvent difficilement être évitées. Nous les avons donc modélisés dans des modèles séparés. Les résultats présentés dans cette thèse sont ceux du GLM incluant les alternatives uniquement [1].

Enfin, il est à noter que SPM implémente par défaut une orthogonalisation des régresseurs dans l'ordre défini par l'utilisateur. Le premier régresseur intègre toute la variance commune, le deuxième est orthogonalisé par rapport au premier et exclut donc la variance commune avec le premier. Le troisième est orthogonalisé par rapport au deuxième et exclut donc la variance des deux premiers, et ainsi de suite. Pour le dernier régresseur dans l'ordre d'orthogonalisation, SPM ne calcule que la variance spécifique, en soustrayant toute variance partagée avec les autres. Nous avons donc veillé à déclarer nos régresseurs d'intérêt en fin de chaîne d'orthogonalisation afin de ne capturer que leur effet individuel, en excluant tous les effets communs.

10.4.2 Estimation des paramètres

L'étape d'estimation des paramètres du modèle a pour objectif d'estimer les valeurs des coefficients de corrélation β, de sorte que le modèle "modélise au mieux les données" ou, en termes mathématiques, que la variance de l'erreur résiduelle E soit minimale. Ceci revient à effectuer, pour chaque voxel, une régression linéaire des données sur l'ensemble des régresseurs théoriques définis dans la matrice $\mathbf{X}$, afin d'estimer les paramètres β qui minimisent la variance de l'erreur résiduelle. Pour se faire, on utilise la méthode de résolution des "moindres carrés". Cette

1. La corrélation des régresseurs se mesure au moyen du *Variance Inflation Factor* (*VIF*). Il permet de déterminer si une variable donnée est suffisamment décorrélée des autres variables pour pouvoir être intégrée sans problème dans un GLM. Lorsqu'il dépasse la valeur de 5, la corrélation est considérée comme étant élevée.

méthode permet d'obtenir :

- une estimation des paramètres dans chaque voxel ;

- le signal estimé dans chaque voxel, ou signal "reconstruit" à partir des régresseurs théoriques et des paramètres estimés ;

- l'erreur résiduelle dans chaque voxel, c'est-à-dire la différence entre les données et le signal estimé.

10.4.3 Définition et estimation de contrastes

Une fois que l'on a estimé les paramètres β, on les utilise lors de l'étape de définition et d'estimation de contrastes pour comparer les "activations" associées aux différents types d'évènements. Par exemple, on considère qu'un voxel est plus "activé" par les événements de type 1 que par les événements de type 2 si, dans ce voxel, un test statistique permet de conclure que β_1 est significativement plus élevé que β_2. On peut ensuite visualiser, par exemple, la localisation des voxels effectivement plus activés par les événements de type 1 que les événements de type 2.

Chaque paramètre β reflète en quelque sorte la contribution du régresseur associé à la variation du signal *BOLD* dans le voxel considéré : les β significativement différents de 0 correspondent aux régresseurs (et donc aux types d'évènements) qui contribuent significativement à expliquer le signal dans le voxel. L'étape d'estimation de contrastes consiste à effectuer des tests statistiques sur les valeurs des paramètres β estimés.

10.4.4 Analyses de groupe

Dans le cadre de notre expérience, nous ne nous intéressons pas au fonctionnement du cerveau de chaque sujet, mais au fonctionnement du cerveau humain en général. On cherche donc à généraliser à une population entière des résultats obtenus chez quelques sujets. Pour cela, on analyse ensemble les données obtenues chez les différents sujets. Nous avons effectué une analyse de groupe dite "à effets aléatoires" (*random effects*).

L'analyse à effet aléatoire s'utilise après que l'on a effectué des analyses individuelles chez chacun des sujets du groupe. Elle s'applique aux images de contraste calculées lors des analyses individuelles. Elle prend en compte à la fois la variabilité intra-individuelle (*within subject variance*) puisqu'elle utilise les résultats issus d'analyses individuelles, mais aussi la variabilité interindividuelle (*between subject variance*). A ce titre, elle permet en théorie d'inférer des conclusions généralisables à l'ensemble de la population de laquelle est tiré le groupe de sujets sur lequel est conduite l'étude. L'analyse à effets aléatoires renseigne donc sur les régions qui sont activées de la même manière chez tous les sujets.

10.5 Analyse des régions d'intérêt

Les cartes statistiques obtenues au niveau du groupe permettent d'identifier des régions dont l'activité corrèle avec l'événement considéré. Il est possible par la suite d'analyser plus précisément l'activité de ces régions, notamment en extrayant les valeurs des coefficients de corrélations (β). Cependant, extraire les β des voxels concernés à partir de ces cartes revient à faire une analyse circulaire. En effet, ces cartes sont établies en fonctions des valeurs des β obtenus pour chaque sujet inclus dans l'analyse. Par conséquent, extraire les β à nouveau revient à extraire des valeurs déjà présélectionnées. D'où la circularité.

Pour contourner ce biais, nous avons utilisé la procédure *leave one out* : nous établissons une carte statistique sur n-1 sujets, où n est le nombre total de sujets. Autrement dit nous excluons un sujet de l'analyse statistique de groupe. Les régions d'intérêt sont définies à partir de cette nouvelle carte. On en extrait ensuite les valeurs des β pour le sujet exclu de l'analyse de groupe. Ainsi, les β sont extraits de régions définies indépendamment. Cette procédure est répétée autant de fois qu'il y a de sujets dans le groupe.

10.6 Histogrammes péri-stimulus

10.6.1 Réponse en impulsions finies (*Finite Impulse Response, FIR*)

Dans le modèle linéaire généralisé présenté précédemment, nous faisons l'hypothèse que, dans chaque voxel, il est possible de modéliser le signal BOLD par la fonction de réponse hémodynamique (hrf). Cette modélisation, qui est la plus courante dans la communauté, s'appuie sur des études qui justifient cette approximation [40]. Cependant, il est possible que les voxels présentent différentes formes de signal. Le modèle canonique est donc parfois incorrect. Il devient alors nécessaire d'estimer la forme du signal pour chaque voxel. Il existe plusieurs façons d'estimer le signal. Nous détaillons ici le modèle de réponse en impulsions finies.

L'idée est de modéliser le signal avec une combinaison de fonctions simples. Dans le cas du modèle en impulsions, il s'agit d'une série de N impulsions distinctes, de magnitude égale à un, espacées régulièrement. Autrement dit, à chaque instant t, une impulsion unitaire a lieu, le reste du temps la fonction est nulle. On pondère ensuite chaque impulsion de manière à se rapprocher de la valeur réelle du signal à chaque instant. Ainsi ce modèle ne fait aucune hypothèse sur la forme du signal : le poids appliqué à chaque impulsion peut prendre n'importe quelle valeur, ce qui permet au modèle de capturer une large gamme de signaux. Cette méthode permet par ailleurs d'observer le décours temporel du signal.

Cette modélisation présente cependant l'inconvénient d'être parfois trop flexible. En effet, dans les cas où le signal est très bruité, cette méthode aura tendance à modéliser le bruit au détriment du signal d'intérêt.

Nous avons utilisé ce type de modèle pour explorer la forme du signal. Cependant les histogrammes présentés dans cette thèse correspondent à ceux réalisés avec la méthode suivante.

10.6.2 Résidus

Lors de l'étape d'estimation du modèle linéaire généralisé, SPM génère des images contenant les "résidus", c'est-à-dire la partie du signal non expliquée par le modèle pour chaque image fonctionnelle (il y a donc autant d'image de résidus que d'images fonctionnelles). SPM efface ensuite ces images une fois l'estimation terminée car elles sont très volumineuses et nécessitent un grand espace de mémoire pour les conserver. Les conserver risque de très vite saturer la mémoire du disque. Cependant, ces résidus peuvent fournir une information intéressante. Nous avons ainsi effectué des GLM en conservant ces images (en désactivant la commande du code SPM qui les supprime par défaut).

Pour chaque régresseur d'intérêt, à chaque événement, nous récupérons la valeur des résidus autour de l'évènement (quelques secondes avant jusqu'à quelques secondes après). Puis nous moyennons sur l'ensemble des événements de ce régresseur, ce qui nous donne une valeur moyenne des résidus. Autrement dit la part moyenne de signal non expliqué. Nous ajoutons ce décours à une fonction *hrf* canonique pondérée par le β estimé pour le régresseur correspondant. Chacune de ces courbes est ensuite moyennée sur l'ensemble des sujets. Ainsi, si notre modèle capture bien l'ensemble du signal, les résidus devraient être quasiment nuls et ne pas altérer le signal estimé (β x *hrf*). Autrement dit, avec cette méthode, on observe la déviation par rapport à la *hrf* théorique.

Cinquième partie

Résultats

Chapitre 11

Résultats comportementaux

11.1 Classification des essais

Le protocole expérimental est construit de manière à assurer que d'un épisode à l'autre, deux contingences successives sont rigoureusement incongruentes. Une association stimulus-action correcte lors d'un épisode donné est donc toujours fausse à l'épisode suivant. Lorsqu'un stimulus est présenté, parmi les quatre actions possibles, l'une d'entre elles est nécessairement correcte, une action est incorrecte pour l'épisode courant mais était correcte à l'épisode précédent, et deux actions sont incorrectes à la fois pour l'épisode courant et l'épisode précédent. En d'autres termes, on dénombre une action correcte, une action de persévération et deux actions exploratoires. Les proportions respectives des ces quatre actions varient au cours de l'épisode.

Nous définissons ainsi la persévération comme la mesure, pour le nième essai de chaque épisode, de la proportion de réponse de persévération : autrement dit, la proportion d'actions qui auraient été correctes durant l'épisode précédent. On peut s'attendre à ce que cette persévération soit plus importante en début d'épisode. Nous pouvons en effet supposer que le sujet, après avoir fructueusement utilisé une association stimulus-action pendant un certain temps, aura tendance à continuer à l'utiliser au début de l'épisode suivant, à moins que l'indice contextuel l'incite à changer.

Nous définissons la performance, quant à elle, comme la mesure de la proportion de réponses correctes pour le nième essai de chaque épisode : il s'agit donc de la proportion d'actions qui sont effectivement correctes pour l'épisode en cours. Cette mesure devrait suivre le chemin inverse de la persévération, c'est-à-dire être minimale en début d'épisode, puis commencer à croitre au fur et à mesure que les nouvelles contingences sont apprises ou retrouvées.

Nous définissons enfin l'exploration comme la mesure de la proportion de réponses exploratoires, donc des réponses qui ne sont ni correctes, ni de la persévération, pour le nième essai de chaque bloc. Il s'agit typiquement d'actions que le sujets réalise en début d'épisode, lorsqu'il teste différentes possibilités pour tenter d'apprendre les actions correctes.

11.2 Comportement des sujets

11.2.1 Session d'apprentissage

Performance et exploration

La performance moyenne lors de la session d'apprentissage s'élève à 80,2 % de réponses correctes (sem. 0,84). La proportion d'essais sans réponse est de 0,6% (sem. 0,08).

En début d'épisode, le taux de réponses correctes est d'environ 5% puis atteint un plateau d'environ 90% au bout d'une vingtaine d'essais (niveau de la chance : 25%).

En début d'épisode, le taux de réponses exploratoires est d'environ 15%, il atteint un pic d'environ 40% après 3 essais puis retourne à 10% (niveau de la chance : 50%).

Ajout de l'information contextuelle. Dans la figure 11.1, on compare les données comportementales de la session d'apprentissage (courbes rouges) avec les données comportementales de la session dite "récurrente" de l'expérience de Donoso, Collins et Kœchlin (courbes bleues) [74]. Dans ces deux sessions, trois combinaisons seulement sont présentées au sujet. Ce qui les différencie est la présence d'information contextuelle dans notre protocole. Cet ajout ne modifie pas les performances asymptotiques : à la fin de l'épisode, les sujets ont appris la combinaison (les deux courbes atteignent la même asymptote). On observe en revanche une différence significative entre les deux courbes pendant environ les dix premiers essais après un changement d'épisode (et donc un changement de contexte). Plus précisément, on peut observer, sur la courbe d'apprentissage de l'expérience sans contexte, un plateau en tout début d'épisode. Ce plateau traduit une période de persévération systématique qui est absente dans l'expérience avec contextes. Cette absence de persévération se traduit également par une exploration immédiate que l'on peut observer sur les courbes d'exploration : dès le premier essais de l'épisode, l'exploration est relativement élevée.

Ces résultats confirment que, contexte présent ou non, les sujets sont capables d'apprendre correctement un *task-set* au sein d'un épisode. Si le contexte apporte un effet, il est logique qu'il soit limité au début de l'épisode, et non à la fin, où la performance asymptotique est atteinte dans tous les cas. Ces données suggèrent ainsi que les sujets utilisent l'information contextuelle a priori pour guider leurs choix.

Effet du changement de contexte. Par ailleurs, si l'on observe le comportement des sujets lors d'un changement de contexte au milieu d'un épisode (au 27ème essai), alors que la combinaison n'a pas changé, on observe une légère diminution de la performance (voir flèches noires sur les figures 11.4 et 11.5).

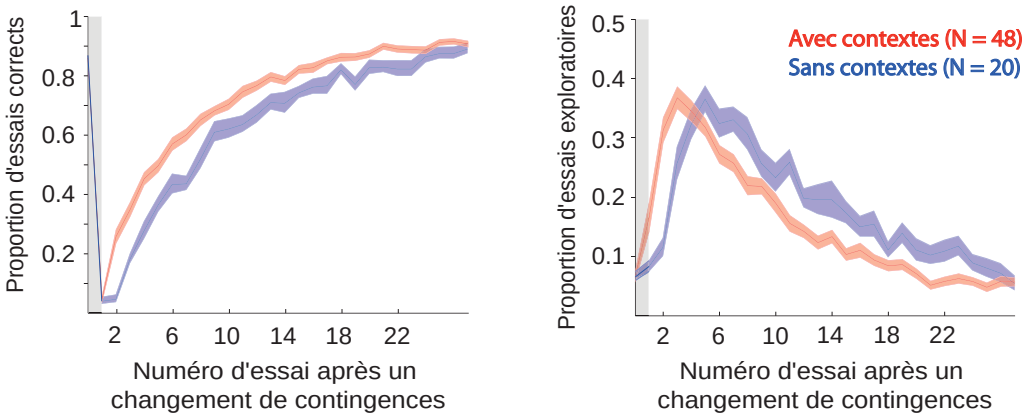


FIGURE 11.1 – **Résultats comportementaux pour la session d'apprentissage.** Comparaison avec l'expérience sans contextes de Donoso, Collins et Kœchlin [74]. Les zones ombragées correspondent à l'erreur moyenne standard. La zone grisée correspond aux contingences précédentes.

Ce phénomène suggère un déplacement de l'attention vers le contexte, au détriment des performances, et/ou l'existence d'un biais de sensibilité au contexte, autrement dit une tendance à changer de comportement lors d'un changement de contexte (cet effet est traduit par le paramètre de biais de contexte δ du modèle PROBE).

Apprentissage à long terme

La figure 11.2 présente la proportion de réponses correctes par épisode en fonction du numéro d'épisode. On observe une augmentation significative de la performance au cours de la session d'apprentissage, passant d'environ 75% de bonnes réponses à presque 85% de bonnes réponses ($r = 0.7$, $F = 61$, $p < 5.10^{-8}$). Cette évolution traduit un effet d'apprentissage à long terme. Etant donné que cet effet n'est pas observé dans l'expérience sans contextes [53], on peut en déduire que, dans notre expérience, les sujets ont appris certaines associations entre les contextes et les combinaisons correspondantes. En fin de session ils peuvent utiliser cette information pour retrouver plus rapidement la combinaison correcte.

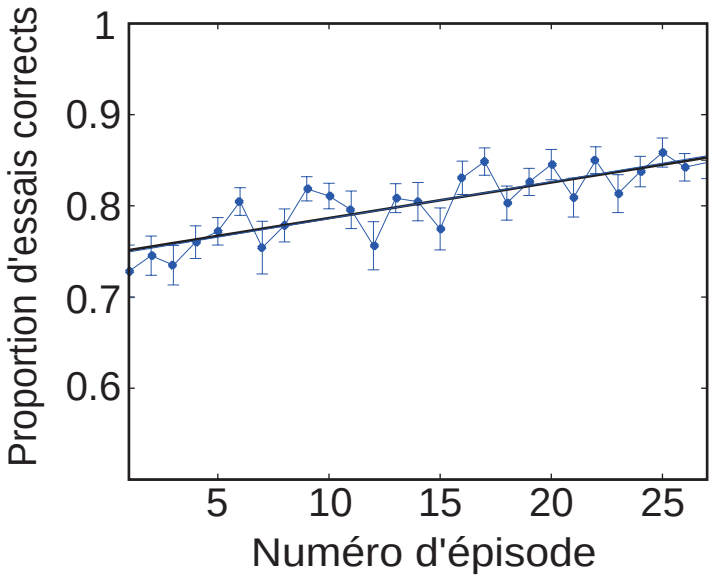


FIGURE 11.2 – **Apprentissage à long terme lors de la session d'apprentissage.** Les barres d'erreur correspondent à l'erreur moyenne standard.

Débriefing

Le débriefing reflète une variabilité comportementale très forte entre les sujets. Après la session d'apprentissage, ce qu'on perçu les sujets de l'expérience va d'un extrême où ils ont correctement identifié trois combinaisons, six contextes et leurs associations, à l'autre extrême où ils ne se sont pas rendu compte que seul un nombre limité de combinaisons était utilisé ni que les textures en fond d'écran étaient informatives pour l'expérience.

Dans leur étude de 2012, Collins et Kœchlin parviennent à distinguer trois groupes de sujets sur la base des réponses au débriefing : un groupe d'exploiteurs "proactifs" qui ont retenu l'ensemble des associations et utilisent l'information contextuelle a priori pour réutiliser les *task-sets* appris ; un groupe d'exploiteurs "rétroactifs" qui n'ont pas retenu les associations entre les contextes et les combinaisons mais sont en mesure de réutiliser les *task-sets* appris sur la base des renforcements ; et enfin un groupe d'explorateurs qui n'ont observé ni le nombre limité de combinaisons ni l'information portée par les contextes. Ces trois groupes sont également corrélés avec des valeurs de paramètres du modèle PROBE.

Dans le cas de notre expérience, nous avons observé, plutôt que des "catégories" de sujets, un continuum de comportements avec toutes les combinaisons possibles. Par conséquent, par la suite, nous n'avons pas effectué d'analyse de groupe sur la base de ce débriefing.

11.2.2 Session test

La performance moyenne lors de la session test (les deux sessions d'IRM confondues) s'élève à 64,5% de réponses correctes (sem. 1,16). La proportion d'essais sans réponse est de 2 % (sem. 0,29). La diminution importante des performances dans cette session par rapport à la session d'apprentissage peut s'expliquer d'une part par les changements de conditions expérimentales (expérience sur ordinateur dans le cas de la session d'apprentissage et dans l'IRM dans le cas de la session test), et d'autre part par la complexité accrue de la session test (davantage de combinaisons et de contextes, plusieurs conditions expérimentales).

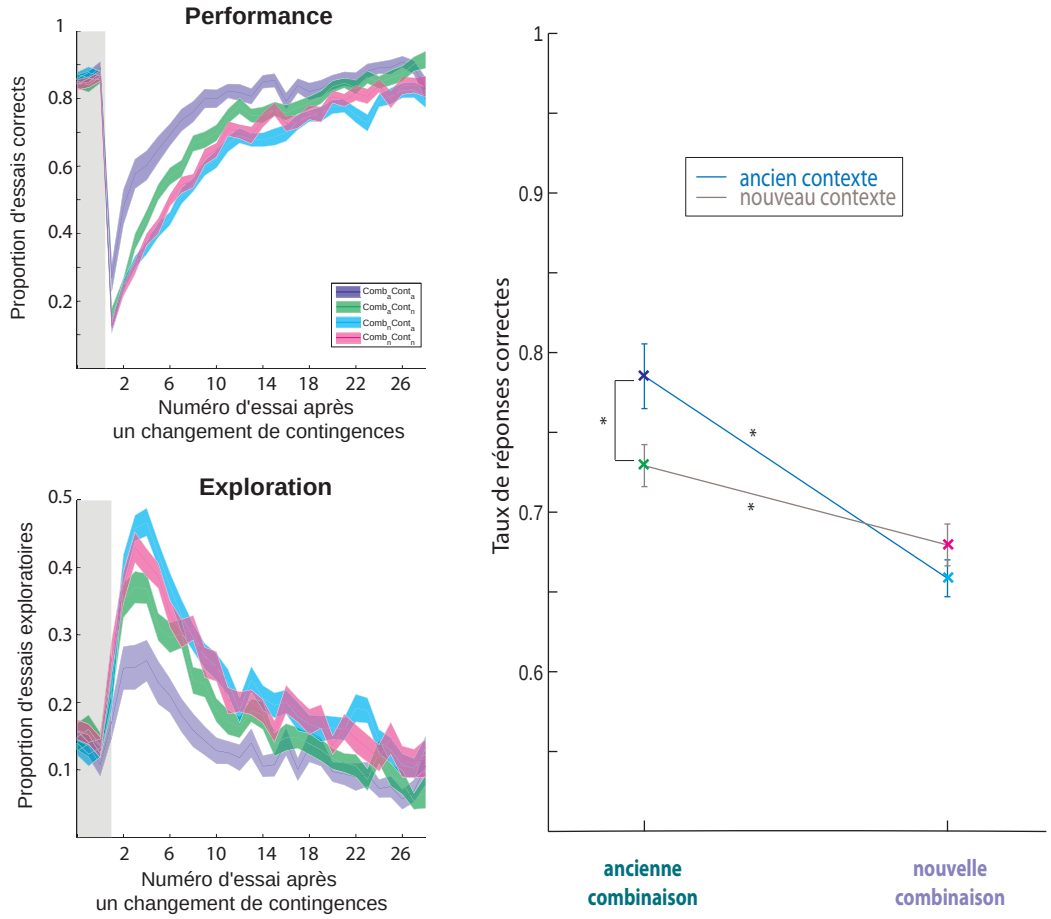


FIGURE 11.3 – **Résultats comportementaux pour la session test.** Les zones ombragées et les barres d'erreur correspondent à l'erreur moyenne standard. La zone grisée correspond aux contingences précédentes. * $p < 0.05$.

D'après le modèles PROBE, les sujets peuvent utiliser deux types de stratégies : la récupération ou la création de *task-sets*, et l'on peut distinguer trois comportements adaptatifs distincts :

- La récupération "proactive", c'est-à-dire la réutilisation de *task-sets* précédemment appris en association avec leur indice contextuel dès la présentation de cet indice ;

- La récupération "rétroactive", c'est-à-dire la réutilisation de *task-sets* précédemment appris mais sans l'association avec leur indice contextuel, uniquement sur la base des renforcements reçus.

- L'exploration, c'est-à-dire la création d'un nouveau *task-set*, quel que soit le contexte, dès la détection d'un changement de contingences.

Ainsi, les quatre conditions de la session test nous permettent de tester la capacité des sujets à réutiliser des *task-sets* déjà appris ou d'en créer de nouveaux, dans des contextes nouveaux ou anciens. Les résultats obtenus nous permettent de dégager deux effets (voir figure 11.3) :

- Un effet combinaison : la performance est moins bonne dans les conditions où la combinaison est nouvelle (le *task-set* doit être créé, courbes roses et cyans) que dans les conditions où la combinaison est connue (le *task-set* peut être réutilisé, courbes bleues et vertes). L'exploration est également plus importante dans les conditions où le *task-set* doit être créé.

- Un effet contexte lors de la réutilisation de *task-sets* connus : en présence d'un contexte connu, les performances sont meilleures que lorsque le contexte est nouveau.

11.3 Comportement du modèle PROBE

11.3.1 Qualité de l'ajustement du modèle sur les données comportementales

Session d'apprentissage

Les résultats de l'ajustement du modèle PROBE sur les données comportementales de la session d'apprentissage sont illustrés dans la figure 11.4.

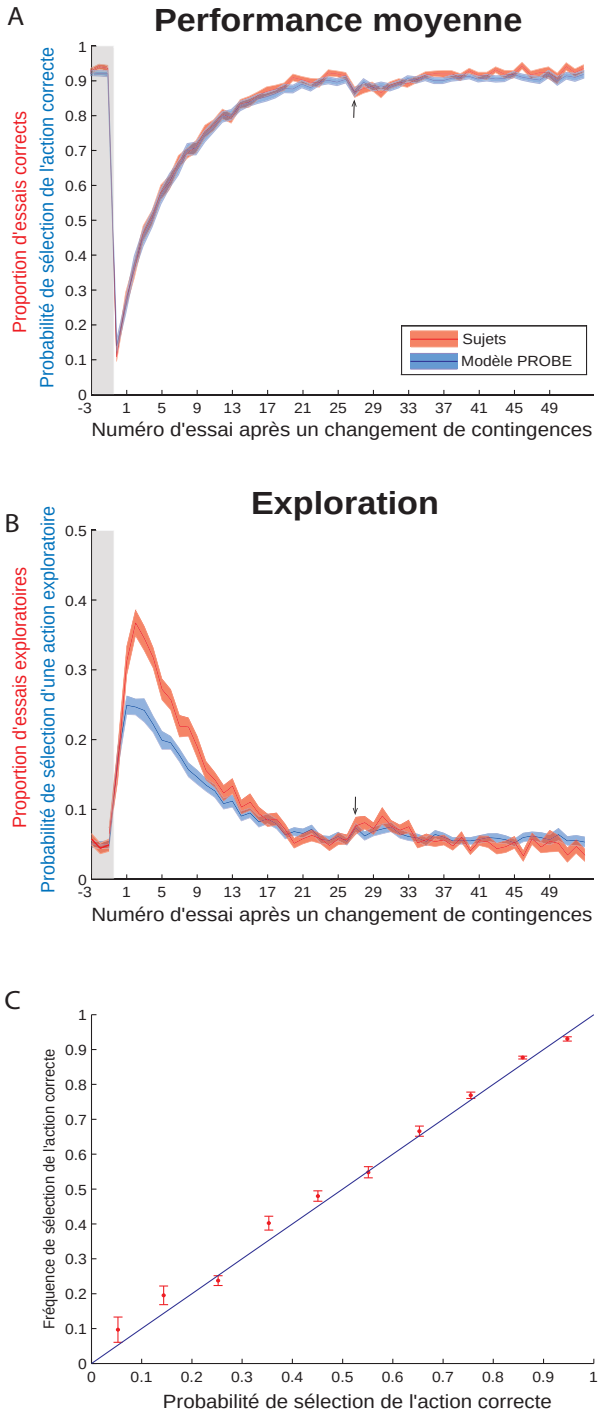


FIGURE 11.4 – **Ajustement du modèle PROBE sur les données comportementales de la session d'apprentissage.** Les flèches noires indiquent un changement de contexte au milieu d'un épisode (sans changement de contingences). Les zones ombragées et les barres d'erreur correspondent à l'erreur moyenne standard. La zone grisée correspond aux contingences précédentes.

A : Courbes d'apprentissage. Les prédictions du modèle PROBE (courbe bleue) semblent refléter fidèlement le comportement des sujets (courbe rouge). Le modèle rend même compte de la légère baisse de performance lors d'un changement de contexte au milieu de l'épisode (27ème essai, flèche noire).

B : Courbes d'exploration. Le modèle PROBE reproduit la forme globale de la courbe d'exploration des sujets. On observe cependant un écart de près de 10% entre les deux courbes : le modèle semble explorer significativement moins que les sujets lors des premiers essais de l'épisode. Ceci peut être expliqué par le phénomène de "réaction d'inhibition" (*inhibition of return*). Les sujets, lorsqu'ils quittent le *task-set* courant, inhibent la réutilisation de ce *task-set*, c'est-à-dire l'ensemble des associations stimulus-réponse qui lui correspondent. Ce qui n'est pas le cas du modèle (cette inhibition n'est pas implémentée). Les sujets auront donc davantage tendance à sélectionner des essais exploratoires que le modèle qui peut continuer à sélectionner des essais considérés comme de la persévération.

Notons que cet effet diminue au cours de la session (non montré). En effet, en début de session, lorsque le modèle entre en exploration, il crée un *task-set* probe. Ce *task-set* est initialement uniforme : toutes les actions ont la même probabilité d'être sélectionnées et le modèle contextuel a peu d'influence. En fin d'épisode, l'initialisation d'un *probe* est biaisée par deux effets :

(i) un effet de fréquence : le probe est initialement plus proche des *task-sets* qui ont été appris de façon récurrente. Certaines associations stimulus-action sont donc plus probables que d'autres ;

(ii) un effet du contexte : à la fin de a session, les associations entre contexte et combinaison ont pu être apprises ce qui biaise à nouveau le probe vers le *task-set* correspondant au contexte.

C : Graphique probabilités / fréquences. Ce graphique est une autre manière de représenter la qualité de l'ajustement du modèle. On trie les actions correctes en fonction de leur probabilité d'être sélectionnée tel que prédit par le modèle. Puis, pour chaque fenêtre de probabilité, on regarde la fréquence effective de sélection de l'action correcte. Par exemple, pour tous les essais où le modèle prédit que la sélection de l'action correcte est comprise entre 0 et 0.1, on regarde le nombre de fois où, pour ces essais, le sujet a effectivement sélectionné l'action correcte. Ainsi, si les prédictions du modèle sont correctes, les probabilités et les fréquences observées devrait être similaires. Les points de la courbe devraient donc s'aligner sur la diagonale principale, ce que nous observons effectivement.

paramètre	α	β	ϵ	τ	θ	η	κ	δ
min	0.01	6	0.0001	0.01	-1.35	0.0003	5	0.21
max	0.92	94	0.2	0.49	2.53	1	829	1
moyenne	0.41	52	0.04	0.23	0.81	0.38	146	0.66
erreur standard	0.03	3	0.01	0.02	0.12	0.04	26	0.03

TABLE 11.1 – **Valeurs des paramètres du modèle PROBE ajustés sur les données de la session d'apprentissage.**

Session test

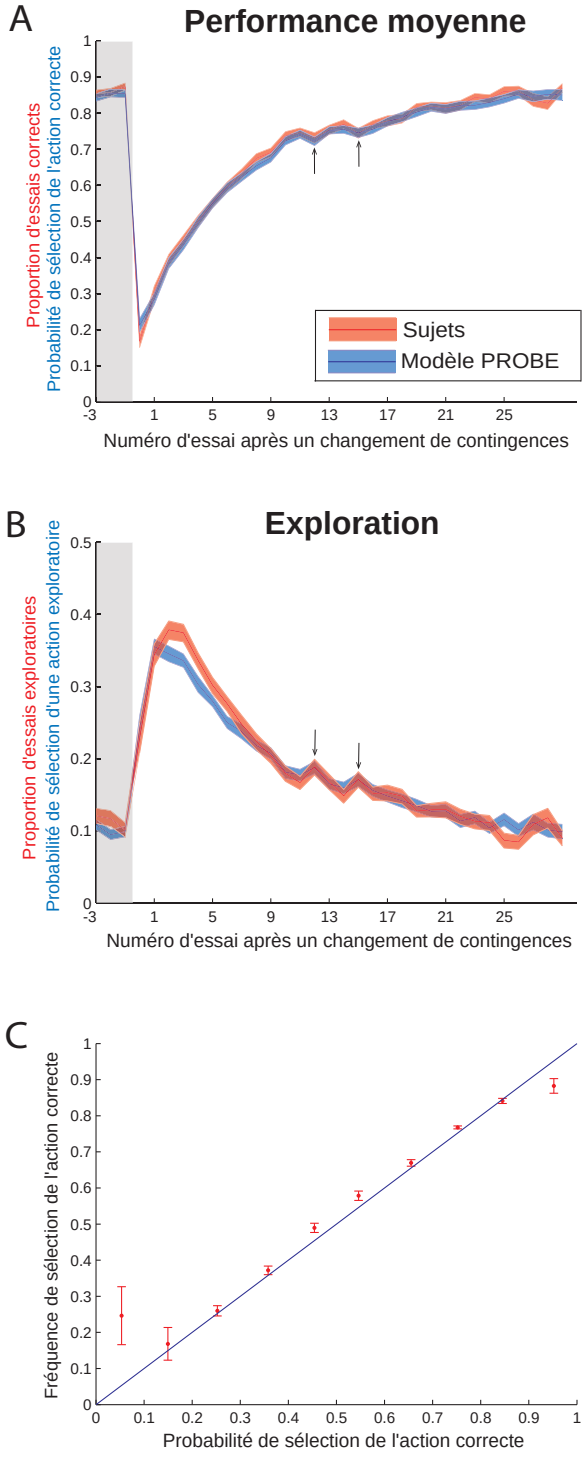


FIGURE 11.5 – **Ajustement du modèle PROBE sur les données comportementales de la session test.** Les flèches noires indiquent un changement de contexte au milieu d'un épisode (sans changement de contingences). Les zones ombragées correspondent à l'erreur moyenne standard. La zone grisée correspond aux contingences précédentes.

La figure 11.5 représente les résultats de l'ajustement du modèle PROBE sur les données comportementales de la session test de la même façon que pour la session d'apprentissage.

Notons que l'effet "d'inhibition retour" sur les courbes d'apprentissage est moins fort que pour la session test. Ceci est sans doute lié au fait que la mémoire à long terme n'est pas vierge dès le début de la session ; et au nombre bien plus important de combinaisons et de contextes au cours de cette session.

Par ailleurs, sur la courbe de probabilités / fréquences, on peut observer que les prédictions du modèles semblent moins bonnes pour les probabilités très faibles et les probabilités très élevées. Ceci peut s'expliquer par les différentes conditions expérimentales de la session. Notons dans un premier temps que le modèle est meilleur que les sujets pour retenir et utiliser les associations entre contexte et *task-sets*, comme nous le verrons par la suite. Par conséquent :

- Dans la condition $\text{Comb}_n\text{Cont}_a$, le modèle prédit une forte probabilité de sélection des actions correspondant au *task-set* ancien qu'il a associé au contexte ancien. Or ici le *task-set* est tout autre. Le modèle prédit donc une très faible probabilité de sélection de l'action correcte. Les sujets, quant à eux, sont moins conservateurs et peuvent sélectionner par hasard l'action correcte plus fréquemment que ce que prédit le modèle. Cependant, notons que ces cas sont assez rares, ce qui s'illustre d'ailleurs par la grande barre d'erreur du premier point du graphique.

- Dans la condition $\text{Comb}_a\text{Cont}_a$, le modèle prédit à nouveau, dès la présentation du contexte, une très forte probabilité de sélection des actions correspondant au *task-set* ancien associé, qui, cette fois, est le *task-set* correct. Le modèle aura donc tendance à sélectionner l'action correcte dès le premier essai de l'épisode. Les sujets, en revanche, fournissent plus rarement une réponse correcte dès le premier essai. D'où un dernier point sur la courbe qui paraît être une moins bonne prédiction.

Notons cependant que ces distorsions ne sont pas significatives.

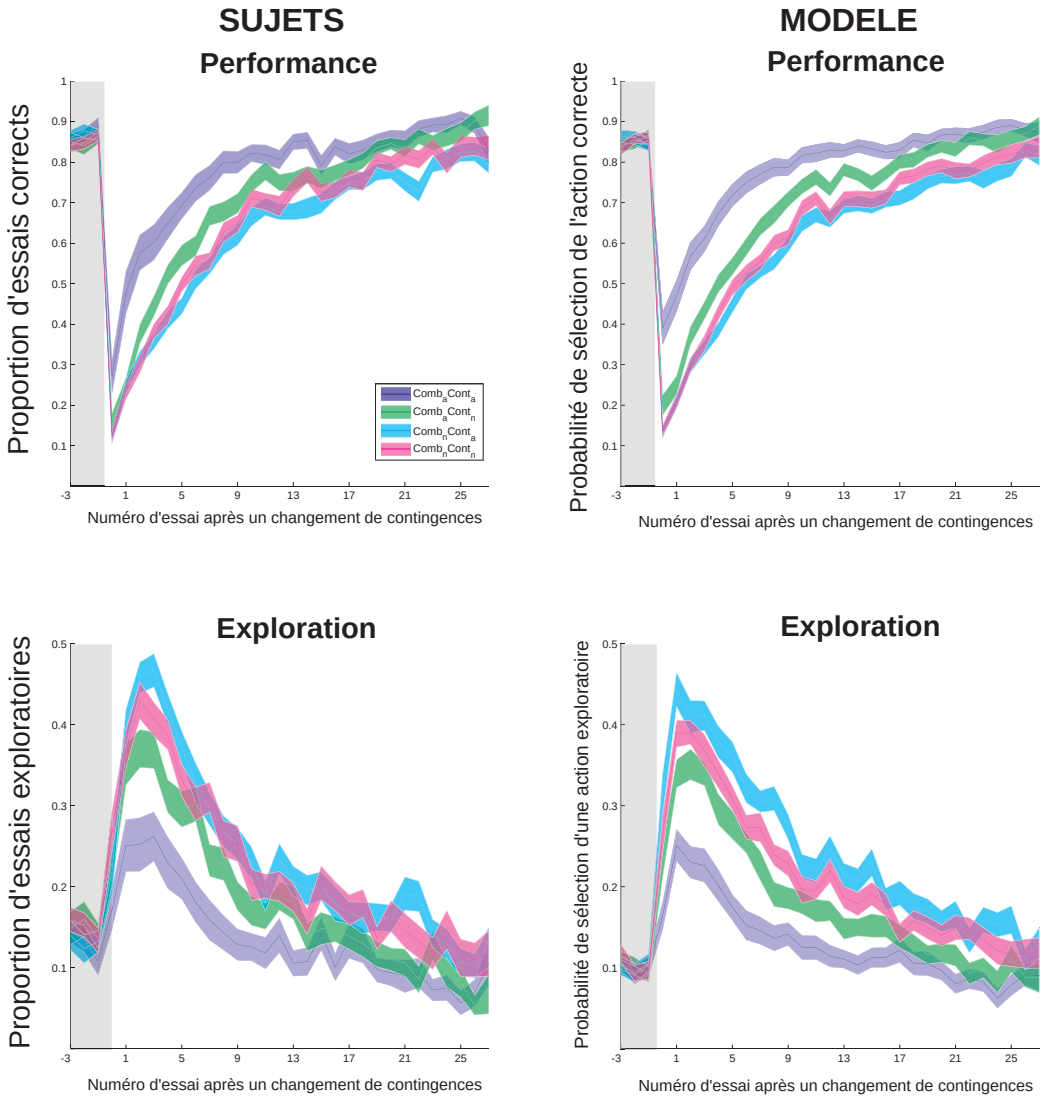


FIGURE 11.6 – **Résultat de l'ajustement du modèle PROBE pour les quatre conditions expérimentales de la session test.** Les zones ombragées et les barres d'erreur correspondent à l'erreur moyenne standard. La zone grisée correspond aux contingences précédentes.

La figure 11.6 présente les résultats d'ajustement du modèle en dissociant les quatre conditions de la session test. Notons que l'ajustement du modèle a été effectué sur les courbes d'apprentissage moyennes, toutes conditions confondues. Il est donc logique d'observer un bon ajustement sur les courbes moyennes. Cependant, rien ne prédit a priori que ce type d'ajustement permet de reproduire les effets des différentes conditions expérimentales. C'est toutefois ce que nous observons. Ceci permet de justifier de la qualité de l'ajustement du modèle et de sa pertinence.

Notons cependant que le modèle est légèrement meilleur que les sujets dans la condition $\text{Comb}_a\text{Cont}_a$ (le taux de réponses correctes lors du premier essai est

de l'ordre de 10% au dessus des sujets). Par ailleurs, sur les courbes d'exploration, le pic d'exploration apparait pour le modèle quasiment toujours un essai avant les sujets. Ces observations suggèrent que le modèle est légèrement plus rapide que les sujets. Cet effet peut s'expliquer par les contraintes du protocole expérimental. En effet, lors d'un changement d'épisode, le stimulus du premier essai apparaît en même temps que le changement de contexte. Il est possible que les sujets, contrairement au modèle, ne traitent pas ces deux informations simultanément. Il est possible qu'ils privilégient de traiter le stimulus en priorité pour fournir une réponse et prennent en compte le contexte, qui reste affiché tout au long de l'épisode, seulement au bout du deuxième essai ou plus.

paramètre	α	β	ϵ	τ	θ	η	κ	δ
min	0.06	7	0.0001	0.05	-0.37	0.0003	2	0.13
max	1	92	0.34	0.45	2.34	0.97	352	0.94
moyenne	0.45	58	0.05	0.22	0.91	0.45	43	0.6
erreur standard	0.03	3	0.01	0.01	0.09	0.04	12	0.03

TABLE 11.2 – **Valeurs des paramètres du modèle PROBE ajustés sur les données de la session test.**

11.3.2 Points de transition prédits par le modèle

Les courbes d'apprentissage sont des moyennes sur plusieurs épisodes. D'où leur aspect continu qui cache les évènements discrets et abrupts de *switch*. En effet, un changement de contingences (d'épisode) ne correspond pas nécessairement au moment où le sujet change de stratégie (*switch* de *task-set*). Le changement d'épisode est défini par l'expérimentateur, en revanche le *switch* n'est pas directement observable. Comment savoir à quel essai le sujet a effectivement repéré les changements de contingences et s'est adapté en conséquence ? Pour répondre à cette question, on s'appuie sur les prédictions du modèle PROBE afin de localiser la position théorique de ces évènements de *switches*.

Notons que nous ne considérons que les premiers points de transition de chaque épisode afin de nous concentrer sur la phase initiale d'apprentissage ou de récupération du *task-set*, même si parfois le modèle prédit des transitions supplémentaires en cours d'épisode.

Par souci de clarté et de concision, nous montrons ici uniquement les courbes de persévération et d'exploration de la session test, qui correspondent aux événements considérés par la suite pour les analyses des données d'IRM. Dans chaque graphique, le point 0 correspond à l'essai *switch* (le premier essai avec le *task-set* sélectionné). La barre verticale correspond au changement de contingences (changement d'épisode). Les barres d'erreur correspondent à l'erreur standard.

Création de *probe* proactive (figure 11.7)

La création de *probe* proactive correspond à la création de *probe* (l'entrée en exploration) dès le premier essai d'un nouvel épisode, donc dès le changement de contexte. D'après le modèle, ces évènements représentent en moyenne 18 essais sur la session (sem : 1.9 ; médiane : 17.5) ; ils apparaissent donc dans environ 38% des épisodes avec cependant une forte variabilité interindividuelle. La transition abrupte de persévération et d'exploration est bien visible sur les courbes du modèle (figure 11.7 à droite). Ces transitions sont beaucoup plus graduelles chez les sujets (figure 11.7 à gauche). Les sujets atteignent en effet les taux prédits par le modèle seulement au bout de trois essais. Les sujets semblent donc intégrer l'information de changement de contexte plus tardivement. Etant donné qu'il est difficile d'identifier exactement si et à quel moment les sujets intègrent cette information, nous n'avons pas considéré ces évènements par la suite.

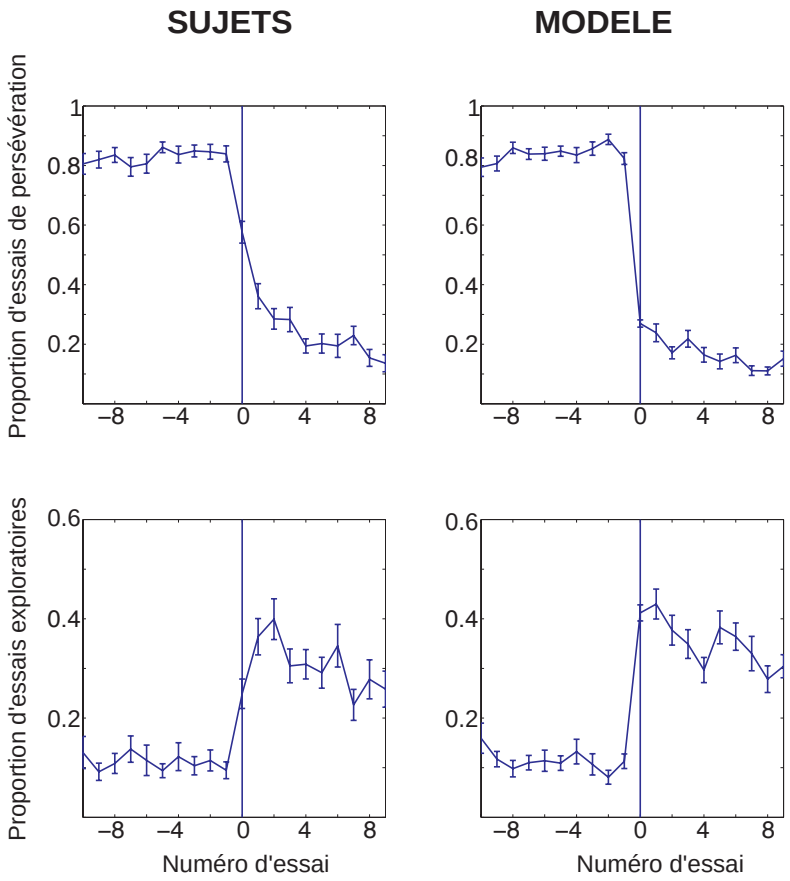


FIGURE 11.7 – **Creation de probe proactive**

Création de *probe* réactive (figure 11.8)

La création de *probe* réactive correspond à la création de probe (l'entrée en exploration) après le changement de contexte, donc après avoir reçu un ou plusieurs $feedbacks$. D'après le modèle, ces évènements représentent en moyenne 16 essais sur la session (sem : 2.1 ; médiane : 7) ; ils apparaissent donc dans environ 33% des épisodes avec à nouveau une forte variabilité interindividuelle. Cette transition abrupte est bien visible sur les courbes du modèle : chute abrupte du taux de persévération et augmentation abrupte du taux d'exploration au bout de 1-2 essais. Ces effets sont également observables chez les sujets. Le modèle semble donc prédire correctement ces évènements.

Notons cependant que ces évènements sont particulièrement rapides. Il est fort probable que le contexte ait également une influence. C'est pourquoi, pour éliminer une partie de l'effet du contexte sur ces évènements, nous avons, pour les analyses IRM, considéré uniquement les créations de *probe* apparaissant au bout d'au moins deux essais.

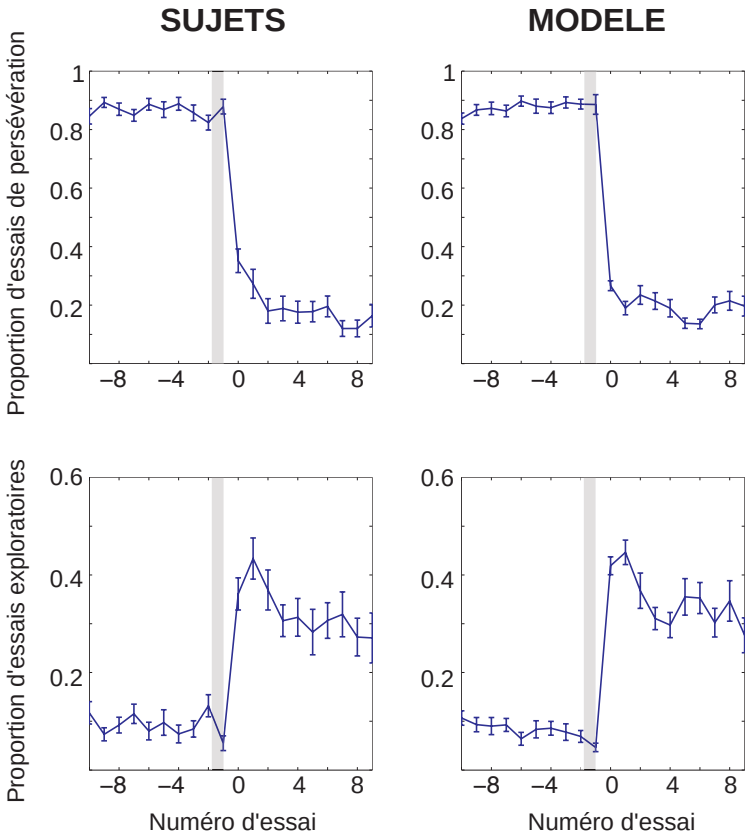


FIGURE 11.8 – **Creation de probe réactive.** La zone grisée correspond à l'erreur moyenne standard du changement de contingences.

Récupération proactive (figure 11.9)

La récupération proactive correspond à la récupération d'un *task-set* connu dès le premier essai de l'épisode, c'est-à-dire dès l'apparition du contexte connu. Seuls 32 sujets présentent ce type d'évènements. D'après le modèle, ces évènements représentent en moyenne 4 essais sur la session (sem : 0.5 ; médiane : 3) ; ils apparaissent donc dans environ 17% des épisodes dont le contexte est connu. Là encore, le modèle semble prédire correctement le comportement des sujets bien que ces derniers présentent une transition légèrement moins abruptes (deux essais de transition environ). Notons par ailleurs la différence entre la condition $\text{Comb}_a\text{Cont}_a$ et la condition $\text{Comb}_n\text{Cont}_a$ sur les courbes d'exploration (encarts). En effet, dans la condition $\text{Comb}_a\text{Cont}_a$, la récupération est la plus part du temps correcte. Le taux d'exploration est donc très faible. En revanche, dans la condition $\text{Comb}_n\text{Cont}_a$, la récupération est la plupart du temps incorrecte puisque la combinaison est toujours nouvelle et donc différente de celle qui a été associée au contexte. Par conséquent, les réponses effectuées correspondent à des réponses exploratoires.

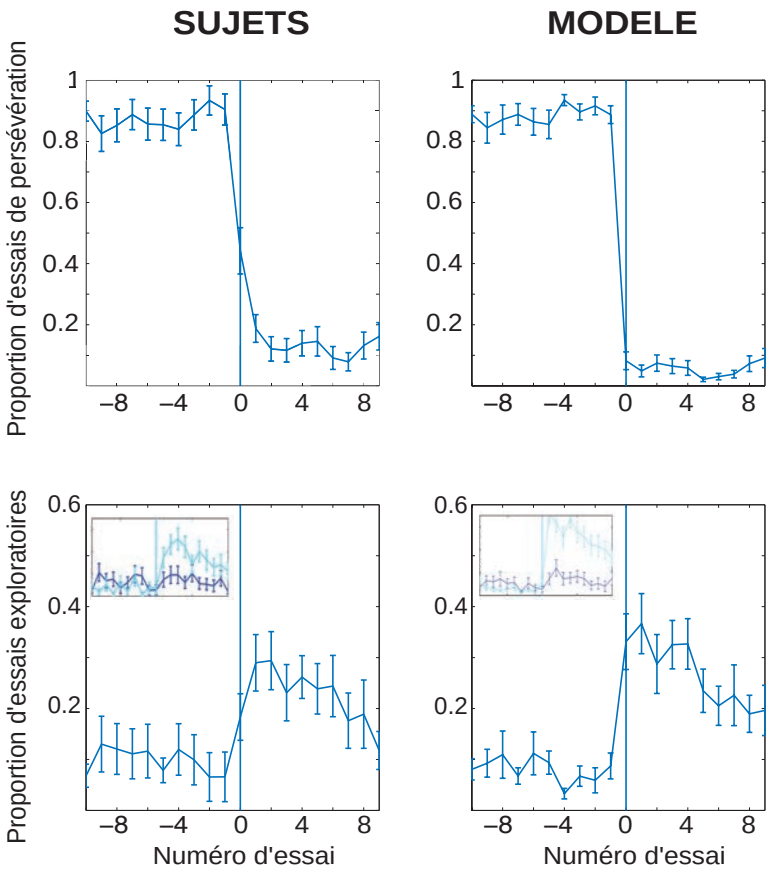


FIGURE 11.9 – **Récupération proactive**

Récupération réactive et consolidation (figure 11.10)

La récupération réactive correspond à la récupération d'un *task-set* connu après plusieurs essais suivant le changement d'épisode où le contexte est inconnu (donc non informatif), c'est-à-dire sur la base des *feedbacks* obtenus. 42 sujets présentent ce type d'évènements. D'après le modèle, ces évènements représentent en moyenne 6 essais sur la session (sem : 0.6 ; médiane : 6) ; ils apparaissent donc dans environ 26% des épisodes dont le contexte est nouveau.

La consolidation correspond à la confirmation du *probe* après plusieurs essais. D'après le modèle, ces évènements représentent en moyenne 26 essais sur la session (sem : 1.4 ; médiane : 25) ; ils apparaissent donc dans environ 54% des épisodes.

Les essais de récupération réactive et de consolidation correspondent tous deux à une transition d'un état d'exploration à un état d'exploitation. Cependant, comme l'illustre la figure 11.10, ces transitions présentent des particularités :

- la récupération réactive (courbe rouge) est une transition abrupte. Elle survient environ 7-8 essais après un changement d'épisode. Nous représentons ici uniquement les récupérations réactives correctes. En effet, elles seules permettent d'observer la transition abrupte de performance. Les récupérations incorrectes sont assez fréquentes d'une part à cause de la condition $\text{Comb}_n\text{Cont}_a$ qui, si elle entraîne une récupération, est nécessairement incorrecte ; et d'autres part à cause du nombre élevé de combinaisons présentées qui sont en partie congruentes. Il est donc possible qu'un *task-set* congruent avec le *task-set* effectivement correct soit présent dans la mémoire de travail et récupéré.

- la consolidation (courbe jaune) est un processus continu. Elle survient environ 6-7 essais après un changement d'épisode.

D'après le modèle, ces deux processus sont en compétition pendant la période d'exploration. Dans le premier cas il s'agit de se fonder sur les résultats obtenus pour éventuellement récupérer un *task-set* alternatif en mémoire de travail. Dans le second, il s'agit d'apprendre le *probe*. En début d'exploration, la fiabilité du *probe* peut être plus élevée que le *task-set* correct stocké en mémoire. En effet, le *task-set* correct est doté d'une fiabilité très faible à l'épisode précédent car les combinaisons de chiffres d'un épisode à l'autre sont strictement incongruentes. Par conséquent, si par chance le *probe* produit des réponses correctes dès le début de la période d'exploration, sa fiabilité a d'autant plus de chance d'atteindre le seuil de 0.5 avant un *task-set* en mémoire, aboutissant ainsi à une consolidation plutôt qu'une récupération. Cet effet est visible sur la figure 11.10 : les réponses correctes sont significativement plus fréquentes dans la période précédant une consolidation (courbe jaune). De plus, les évènements de consolidation apparaissent en moyenne plus tôt que les évènements de récupération (barres verticales jaunes et rouges).

Par ailleurs, les évènements de consolidation peuvent avoir un effet délétère sur la récupération de *task-sets* corrects. En effet, du fait de la capacité limitée de la mémoire de travail, la consolidation d'un *probe* peut provoquer le rejet d'un *task-set* correct hors de la mémoire de travail. Auquel cas, la mémoire de travail peut contenir deux *task-sets* redondants : le *probe* consolidé et le *task-set* cor-

rect qui n'a pas été récupéré. Le *task-set* rejeté n'est quant à lui plus en mesure d'être récupéré dans les épisodes suivants. Il est nécessaire de le réapprendre en créant et confirmant à nouveau un *probe*. Au final, ces effets expliquent pourquoi les évènements de récupération sont globalement moins nombreux que ceux de consolidation.

Notons que nous reproduisons ici les résultats comportementaux obtenus par Donoso, Collins et Kœchlin [74].

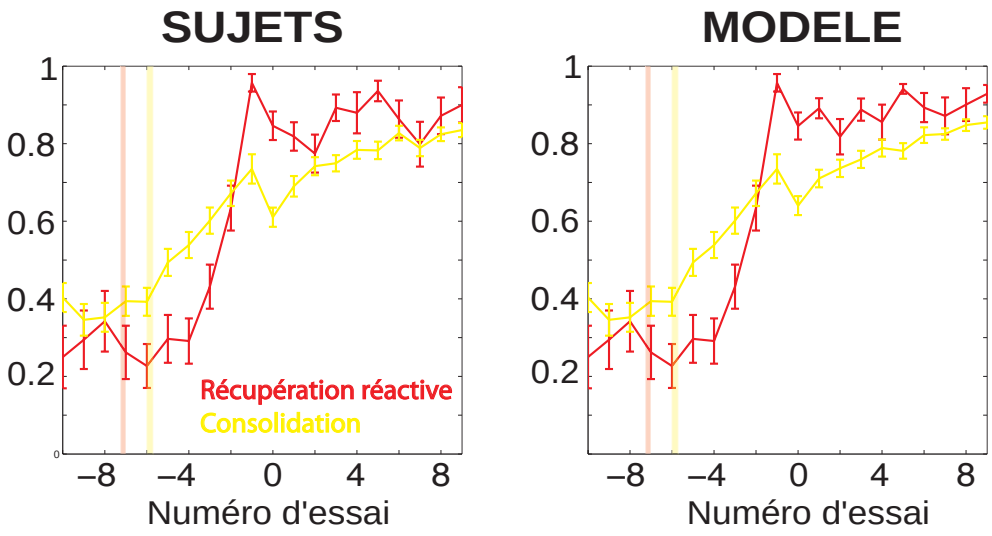


FIGURE 11.10 – **Récupération réactive et consolidation.**

En conclusion, notre expérience permet de valider les prédictions du modèle. Elle montre que nous pouvons apprendre un répertoire de *task-sets* et en tirer profit pour les réutiliser en cas de besoin, même dans un nouveau contexte. Elle montre que nous sommes capables de découvrir qu'une dimension sensorielle apporte de l'information contextuelle, et d'apprendre à utiliser cette information contextuelle a priori, pour la sélection des *task-sets*.

Chapitre 12

Résultats d'imagerie

Dans cette partie nous présentons l'ensemble des activations que nous avons observées dans le lobe frontal et les structures sous-corticales pour les signaux de fiabilité et les différents évènements de *switch*. Par souci de concision, nous ne présentons pas les activations postérieures (régions pariétales, temporales ou occipitales).

Nous décrivons les régions d'une taille minimale de 20 voxels à un seuil de significativité de p = 0.05 pour correction multiple sur l'ensemble des voxels du lobe frontal.

Pour les régions d'intérêt, nous présentons également les histogrammes péristimulus. Nous comparons l'effet des essais *switch* (création de *probe*, récupérations et consolidation) par rapport aux essais *stay* (exploration ou exploitation) précédents. Autrement dit, nous comparons d'une part la création de *probe* et la récupération proactive à l'exploitation ; et d'autre part la récupération réactive et la consolidation à l'exploration.

12.1 Les signaux de fiabilité

Dans leur étude de 2014, Donoso, Collins et Kœchlin ont identifié deux régions dont l'activité corrèle avec les signaux de fiabilité [74] :

- le cortex préfrontal ventro-médian encode le signal de fiabilité du *task-set* acteur ;

- le cortex frontopolaire droit encode les signaux de fiabilité des deux *task-sets* alternatifs.

Dans le cas de notre expérience, on s'attend donc à :

(i) Retrouver l'activation du cortex préfrontal ventro-médian ;

(ii) Reproduire le second résultat pour la fiabilité *ex-post* (μ) qui correspond au signal de fiabilité sans l'intégration de l'information contextuelle ;

(iii) Observer une région dont l'activité corrèle avec la fiabilité *ex-ante* (λ), c'est à dire la fiabilité intégrant l'information contextuelle. On peut supposer :

Soit que cette région est la même que celle des μ. Tous les signaux de fiabi-

lité seraient ainsi actualisés dans une région commune.

Soit qu'il existe une région spécifique d'intégration de l'information contextuelle pour la mise à jour des fiabilités.

12.1.1 Signal de fiabilité du *task-set* acteur et cortex préfrontal ventro-médian

Dans un premier temps, nous avons cherché à reproduire l'activation ventro-médiale observée par Donoso, Collins et Kœchlin. Dans un autre GLM que celui présenté dans cette thèse, nous avons inclus les fiabilités du *task-set* acteur (sans les fiabilités des *task-sets* alternatifs). Nous observons effectivement une activation du cortex préfrontal ventro-médian qui corrèle avec la fiabilité du *task-set* acteur.

Sachant que nous nous intéressons spécifiquement aux substrats cérébraux de la récupération de *task-sets* alternatifs, nous présentons ici uniquement les résultats du GLM incluant les fiabilités des *task-sets* alternatifs sans l'acteur.

12.1.2 Signaux de fiabilité *ex-post*, cortex frontopolaire et préfrontal dorso-latéral

Dans un deuxième temps, nous avons analysé les signaux de fiabilité des deux *task-sets* alternatifs.

Nous observons des activations qui corrèlent avec les signaux de fiabilité *ex-post* dans trois régions préfrontales (figure 12.1) : le cortex frontopolaire (pic voxel aux coordonnées MNI [34, 62, -5]), une région dorso-latérale antérieure (pic voxel aux coordonnées MNI [46, 32, 32]) et une région dorso-latérale postérieure (pic voxel aux coordonnées MNI [30, 8, 61]).

Notons que nous répliquons bien le résultat de Donoso, Collins et Kœchlin (activation du cortex frontopolaire droit). Cependant, comme le montre les valeurs des coefficients de corrélation β (diagramme du bas dans la figure 12.1), cette activation n'est pas spécifique des signaux de fiabilité *ex-post* μ. En effet, l'activité de cette région corrèle également avec le signal de fiabilité *ex-ante* du meilleur *task-set* alternatif (λ_1).

De plus, nous observons deux régions du cortex préfrontal dorso-latéral dont l'activité corrèle avec les quatre signaux de fiabilité λ_1, λ_2, μ_1 et μ_2 (diagrammes du haut dans la figure 12.1).

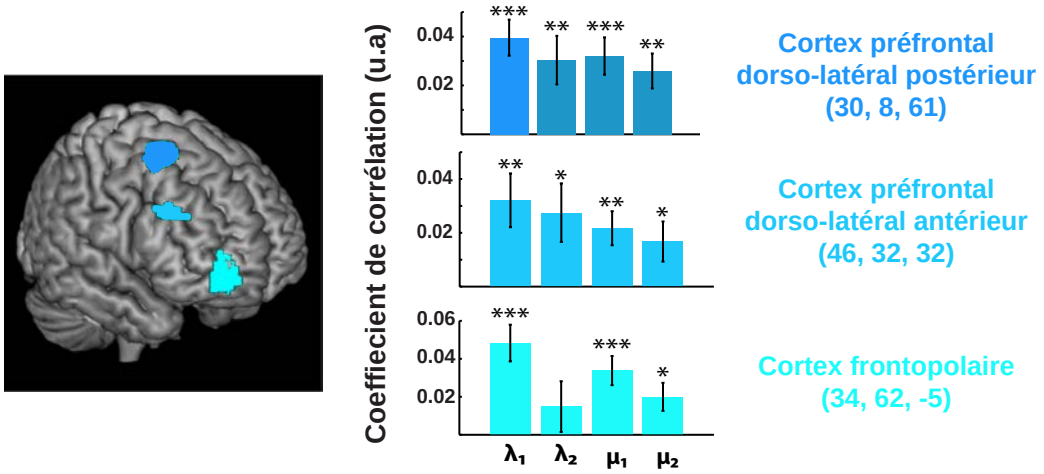


FIGURE 12.1 – **Activations cérébrales associées au signal de fiabilité *ex-post* (μ).** A gauche : intersection des cartes statistiques des signaux de fiabilité *ex-post* du meilleur (μ_1) et second meilleur (μ_2) *task-sets* alternatifs (au seuil p = 0.05 pour correction multiple). A droite : valeur des coefficients de corrélation (β) pour les quatre signaux de fiabilité (λ_1, λ_2, μ_1 et μ_2). Les coordonnées MNI des pics voxels sont indiqués entre parenthèse. u.a : unité arbitraire. Les barres d'erreur correspondent à l'erreur standard calculée sur les 48 sujets. *p<0.05, **p<0.005, ***p<0.0005.

12.1.3 Signaux de fiabilité *ex-ante* et cortex préfrontal dorso-latéral

Enfin, pour les signaux de fiabilité *ex-ante* (λ), nous observons deux régions préfrontales (figure 12.2) : le cortex préfrontal dorso-latéral droit (pic voxel aux coordonnées MNI [34, 6, 64]) et le cortex préfrontal dorso-latéral gauche (pic voxel aux coordonnées MNI [-48, 27, 37]). La région de droite est plus antérieure que celle de gauche.

les coefficients de corrélation β indiquent que ces deux régions ne sont pas spécifiques aux signaux de fiabilité *ex-ante*. En effet, l'activité de la région du cortex préfrontal dorso-latéral droit corrèle avec l'ensemble des signaux de fiabilité (diagramme de gauche sur la figure 12.2) ; et l'activité de la région du cortex préfrontal dorso-latéral gauche corrèle également avec le signal de fiabilité *ex-post* du meilleur *task-set* alternatif μ_1 (diagramme de droite sur la figure 12.2).

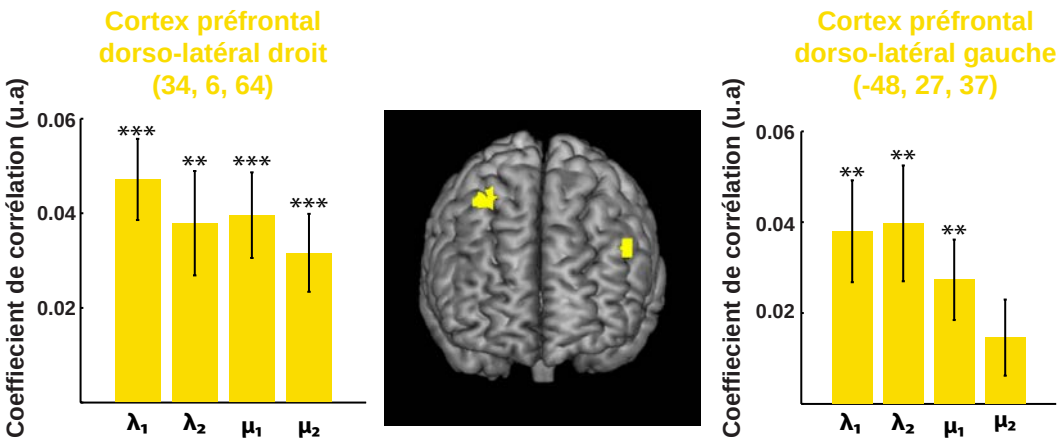


FIGURE 12.2 – **Activations cérébrales associées au signal de fiabilité *ex-ante* (λ).** Au centre : intersection des cartes statistiques des signaux de fiabilité *ex-ante* du meilleur (λ_1) et second meilleur (λ_2) *task-sets* alternatifs (au seuil p = 0.05 pour correction multiple). A gauche et à droite : valeur des coefficients de corrélation (β) pour les quatre signaux de fiabilité (λ_1, λ_2, μ_1 et μ_2) pour la région du cortex préfrontal dorso-latéral droit et la région du cortex préfrontal dorso-latéral gauche, respectivement. Les coordonnées MNI des pics voxels sont indiqués entre parenthèse. u.a : unité arbitraire. Les barres d'erreur correspondent à l'erreur standard calculée sur les 48 sujets. **p<0.005, ***p<0.0005.

12.2 Abandon du *task-set* acteur et pré-SMA

Donoso, Collins et Kœchlin observent une activation spécifique du cortex cingulaire antérieur dorsal et d'une partie de la pré-SMA lors de la création d'un *probe* (entrée en exploration). Ils suggèrent que cette région détecte quand le *task-set* acteur devient non fiable afin de déclencher la création d'un *probe*.

Dans notre expérience, lors de la création d'un *probe*, nous observons également une activation médiale dans une région plus petite que la région observée dans l'étude de Donoso, Collins et Kœchlin (figure 12.3, pic voxel aux coordonnées MNI [-12, 8, 55]). Les histogrammes péri-stimulus présentent également un pattern d'activité similaire : les évènements de création de *probe* engendrent une activité cérébrale supérieure aux évènements d'exploitation (courbe bleue dans la figure 12.3, $T > 2.59$, $p < 0.05$), alors que les évènements de récupération réactive et de consolidation engendre une activation non significative (courbes rouge et jaune dans la figure 12.3, $p > 0.7$ pour les deux types d'évènements). Nous observons de plus une activation signification en corrélation avec les évènements de récupération proactive (courbe cyan sur la figure 12.3, $T > 2.76$, $p < 0.01$).

De façon symétrique, nous observons une région médiale dans la carte d'activation des évènements de récupération proactive (figure 12.4) dont le pattern d'activité est identique (pic voxel aux coordonnées MNI [-3, 8, 43], $T > 1.95$ et $p < 0.05$ pour la création de *probe* (bleu) et la récupération proactive (cyan) ; $p > 0.5$ pour la récupération réactive (rouge) et la consolidation (jaune)).

Dans notre expérience, les évènements de création de *probe* et de récupération proactive ont en commun le fait de quitter une période d'exploitation, c'est-à-dire un *task-set* acteur fiable. Il semble donc logique d'observer une région commune qui indique le moment de quitter ce *task-set* acteur.

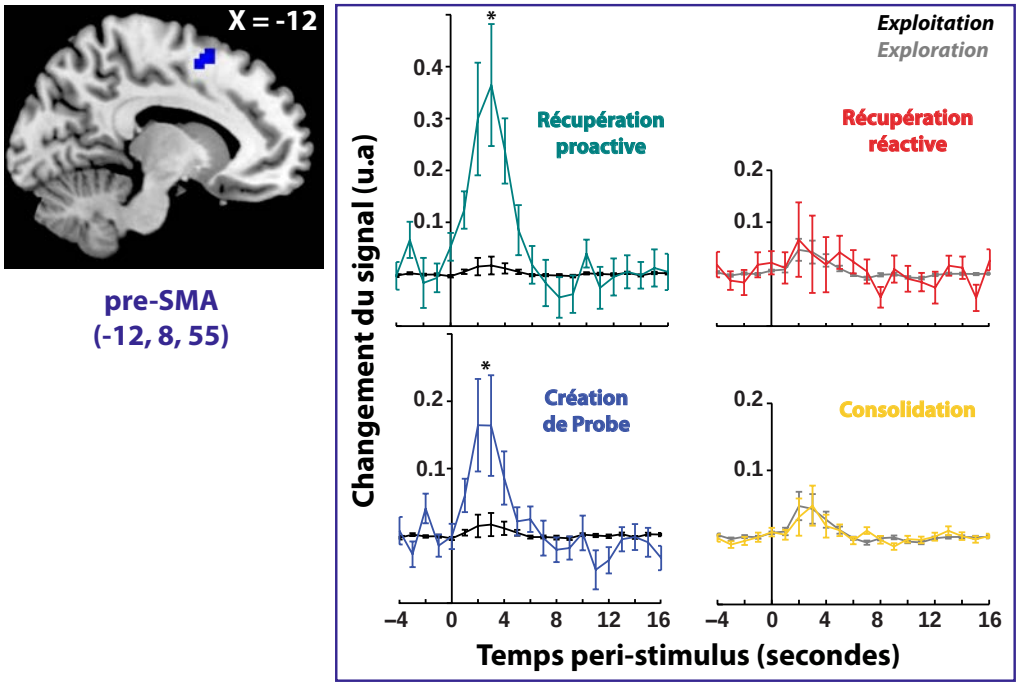


FIGURE 12.3 – **Activations cérébrales dans la pré-SMA (Région de la création de *probe*).** A gauche : Région de la pré-SMA observée sur la carte statistique de la création de *probe* (au seuil p = 0.05 pour correction multiple). Les coordonnées MNI des pics voxels sont indiqués entre parenthèse. A droite : Histogrammes péri-stimulus pour les quatre types d'évènements *switch* (création de *probe* (bleu), récupération proactive (cyan), récupération réactive (rouge) et consolidation (jaune)) comparés aux évènements *stay* (exploitation (noir) et exploration (gris)). u.a : unité arbitraire. Le point 0 correspond à l'apparition du stimulus au premier essai suivant le *switch*. Les barres d'erreur correspondent à l'erreur standard calculée sur l'ensemble des sujets présentant le type d'évènement. *p<0.05.

Lors des évènements de création de *probe*, nous avons également observé des activations dans l'aire motrice supplémentaire (pic voxel aux coordonnées MNI [0, -13, 64]), dans les aires de Brodmann 44 et 47 (pics voxels aux coordonnées MNI [-45, 35, -5] et [-51, 17, 13]), dans le cortex préfrontal dorso-latéral (pic voxel aux coordonnées MNI [42, 5, 37]) et dans le striatum (pic voxel aux coordonnées MNI [30, -13, 1]). Ce réseau correspond à celui observé par Donoso, Collins et Kœchlin. Par conséquent, par souci de concision, nous avons choisi de ne pas détailler l'ensemble des activations dans ces régions.

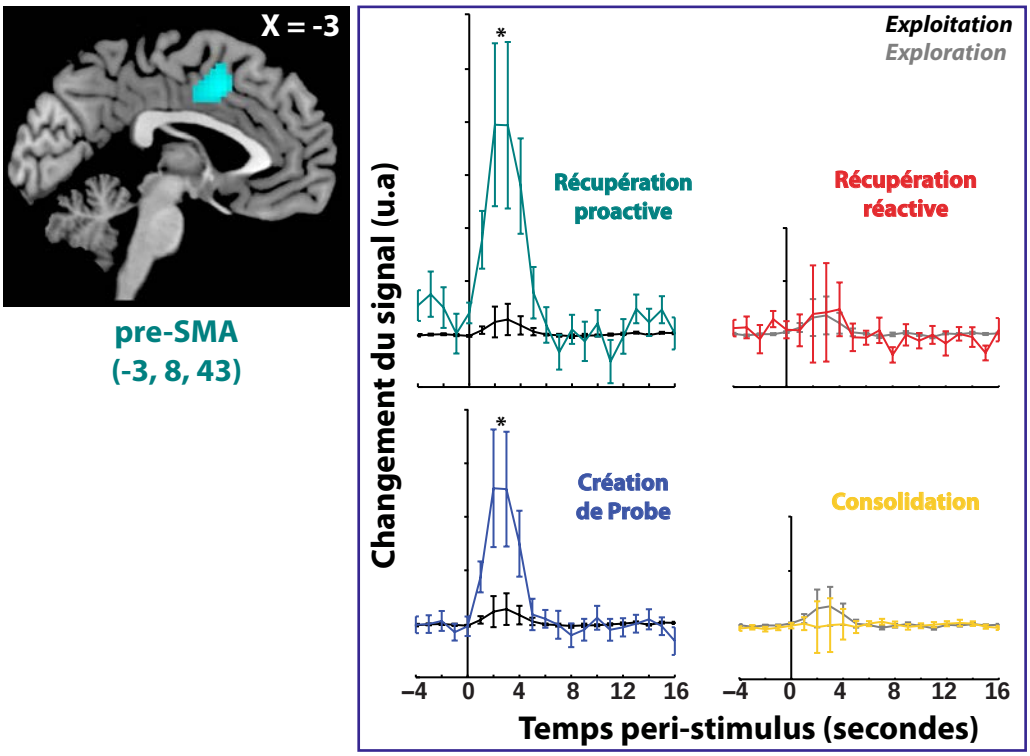


FIGURE 12.4 – **Activations cérébrales dans la pré-SMA (Région de la récupération proactive).** A gauche : Région de la pré-SMA observée sur la carte statistique de la récupération proactive (au seuil p = 0.05 pour correction multiple). Les coordonnées MNI des pics voxels sont indiqués entre parenthèse. A droite : Histogrammes péri-stimulus pour les quatre types d'évènements *switch* (création de *probe* (bleu), récupération proactive (cyan), récupération réactive (rouge) et consolidation (jaune)) comparés aux évènements *stay* (exploitation (noir) et exploration (gris)). u.a : unité arbitraire. Le point 0 correspond à l'apparition du stimulus au premier essai suivant le *switch*. Les barres d'erreur correspondent à l'erreur standard calculée sur l'ensemble des sujets présentant ce type d'évènements. *p<0.05.

12.3 Récupération et cortex préfrontal dorso-latéral

Donoso, Collins et Kœchlin observent une activation du cortex préfrontal dorso-latéral (BA 45 et cortex préfrontal latéral médian) pour les événements de rejet du *probe* (donc de récupération réactive). Ils suggèrent que cette région détecte lorsqu'une une stratégie alternative maintenue dans le cortex frontopolaire devient fiable pour la récupérer comme acteur.

Nous observons également une activation du cortex préfrontal dorso-latéral pour les récupérations proactives et réactives avec une latéralisation : à gauche pour la récupération proactive et à droite pour la récupération réactive (figure 12.3, pics voxels aux coordonnées [-33, 44, 28] pour la région de gauche et [54, 14, 34] pour la région de droite). Les histogrammes péri-stimulus montrent que ces ré-

gions sont spécifiques :

- Dans la région de gauche, l'activité qui corrèle avec la récupération proactive est significative par rapport à l'exploitation (courbe cyan à gauche, $T > 2.85$, $p < 0.01$) alors que l'activité des autres types de *switch* ne l'est pas ($p > 0.1$ pour la création de *probe* (bleu), la récupération réactive (rouge) et la consolidation (jaune)).

- Dans la région de droite, l'activité qui corrèle avec la récupération réactive est significative par rapport à l'exploration (courbe rouge à droite, $T > 1.90$, $p < 0.05$) alors que l'activité des autres types de *switch* ne l'est pas ($p > 0.1$ pour la création de *probe* (bleu), la récupération proactive (cyan) et la consolidation (jaune)).

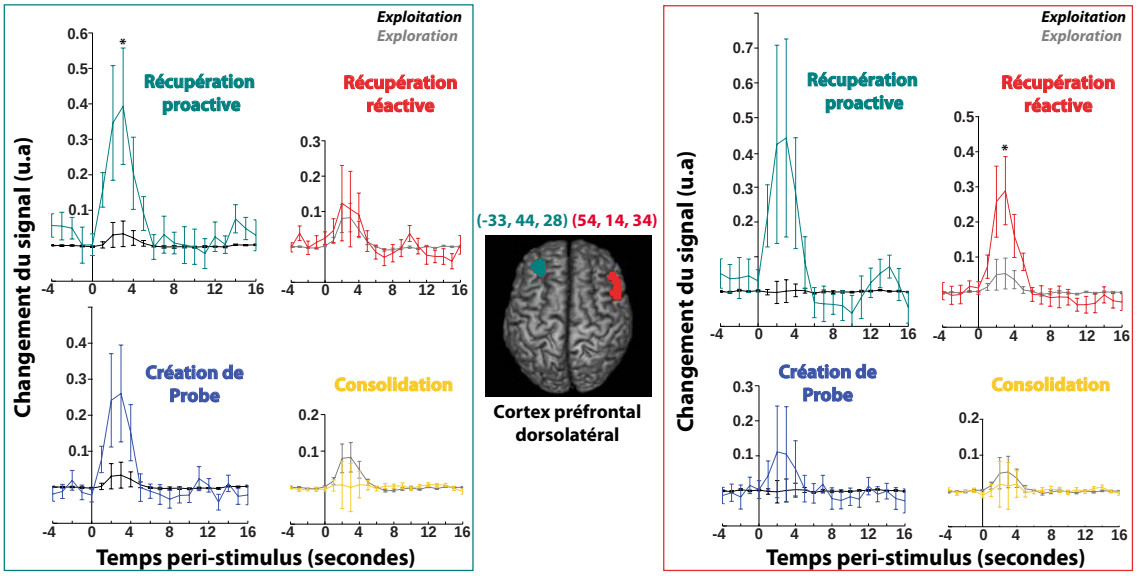


FIGURE 12.5 – **Cortex préfrontal dorso-latéral et récupération.** Au centre : superposition des régions préfrontales dorso-latérales des cartes statistiques de la récupération proactive (cyan) et de la récupération réactive (rouge) (au seuil $p = 0.05$ pour correction multiple). Les coordonnées MNI des pics voxels sont indiqués entre parenthèse. A gauche et à droite : Histogrammes péri-stimulus pour les quatre types d'évènements *switch* (création de *probe* (bleu), récupération proactive (cyan), récupération réactive (rouge) et consolidation (jaune)) comparés aux évènements *stay* (exploitation (noir) et exploration (gris)) pour la région de gauche et la région de droite respectivement. u.a : unité arbitraire. Le point 0 correspond à l'apparition du stimulus au premier essai suivant le *switch*. Les barres d'erreur correspondent à l'erreur standard calculée sur l'ensemble sujets présentant ce type d'évènements. *$p<0.05$.

12.4 Spécificités des récupérations proactives et réactives

12.4.1 Récupération réactive et cortex frontopolaire

En plus du cortex préfrontal dorso-latéral droit, nous observons lors de la récupération réactive une activation bilatérale du cortex frontopolaire (figure 12.4, pics voxels aux coordonnées MNI [33, 59, -2] et [-45, 47, -2]). Les histogrammes péri-stimulus révèlent que ces régions sont spécifiques de la récupération réactive ($T > 2.09$, $p < 0.05$). Les trois autres évènements de *switch* présentent une activité non significative ($p > 0.2$).

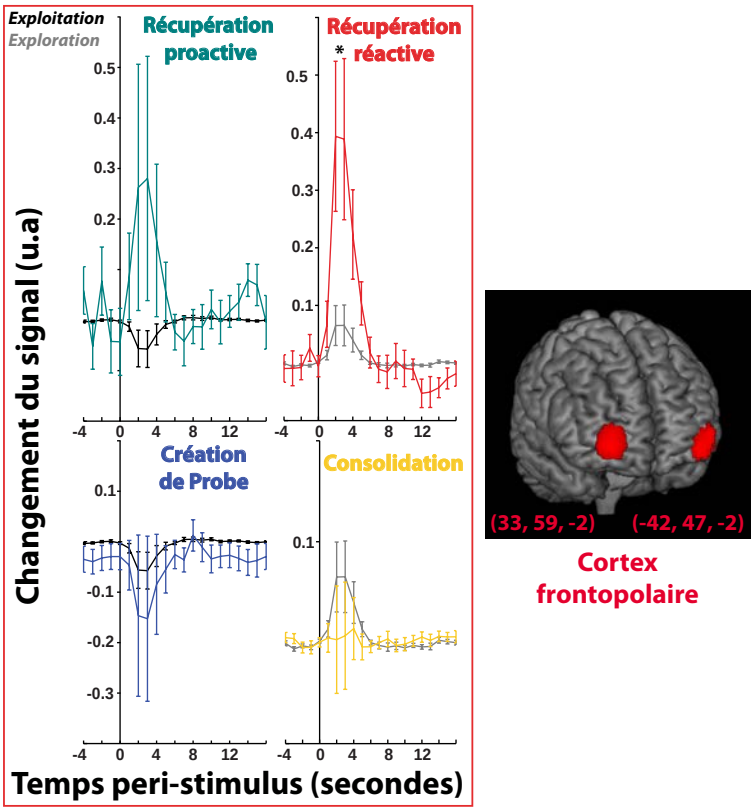


FIGURE 12.6 – **Régions spécifiques de la récupération réactive.** A droite : régions frontopolaires observées sur la carte statistique de la récupération réactive (au seuil $p = 0.05$ pour correction multiple). Les coordonnées MNI des pics voxels sont indiqués entre parenthèses. A gauche : Histogrammes péri-stimulus pour les quatre types d'évènements *switch* (création de *probe* (bleu), récupération proactive (cyan), récupération réactive (rouge) et consolidation (jaune)) comparés aux évènements *stay* (exploitation (noir) et exploration (gris)). u.a : unité arbitraire. Le point 0 correspond à l'apparition du stimulus au premier essai suivant le *switch*. Les barres d'erreur correspondent à l'erreur standard calculée sur les 42 sujets présentant ce type d'évènements. *p<0.05.

12.4.2 Récupération proactive : hippocampe, striatum ventral et cortex préfrontal ventro-médian

Lors de la récupération proactive, en plus du cortex préfrontal dorso-latéral gauche et de la pré-SMA, nous observons un réseau ventral très spécifique incluant l'hippocampe droit (pic voxel aux coordonnées MNI [24, -28, -11]), le striatum ventral droit (pic voxel aux coordonnées MNI [21, 2, -11]) et le cortex préfrontal ventro-médian (pic voxel aux coordonnées MNI [0, 35, -8]).

Activation hippocampique (figure 12.7)

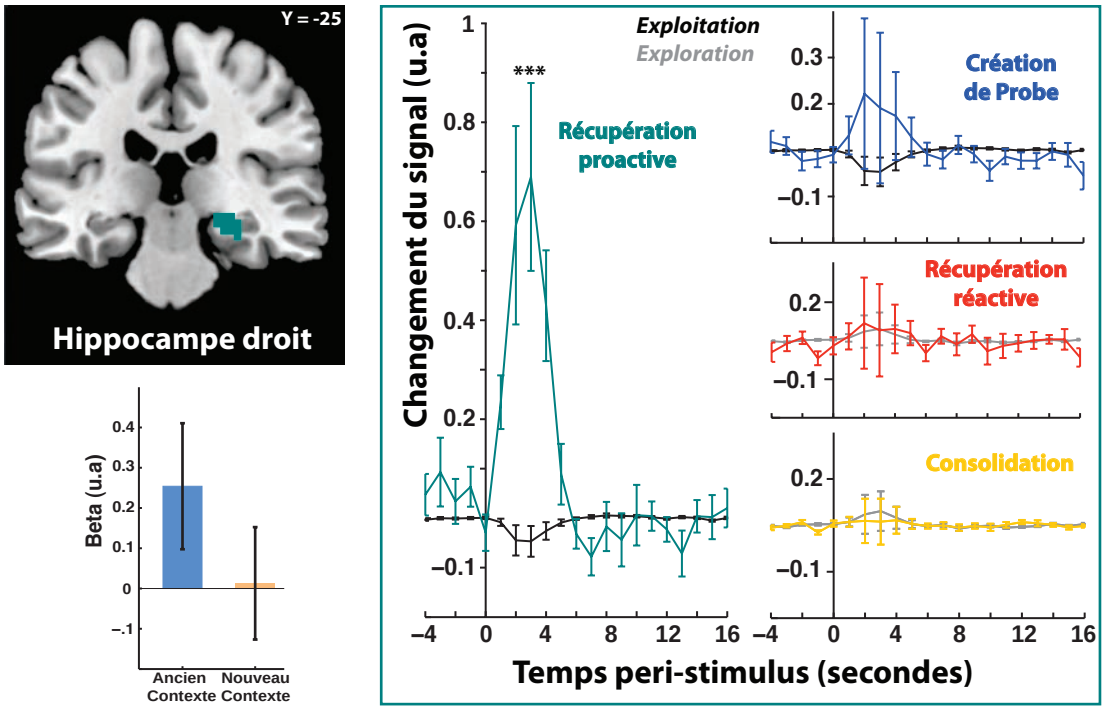


FIGURE 12.7 – **Région spécifique de la récupération proactive (1) : l'hippocampe droit.** En haut à gauche : région de l'hippocampe droit observé sur la carte statistique de la récupération proactive (au seuil p = 0.05 pour correction multiple). Pic voxel aux coordonnées MNI [24, -28, -11]. En bas à gauche : valeur des coefficients de corrélation β pour l'apparition d'un contexte ancien (bleu) ou nouveau (orange). A droite : Histogrammes péri-stimulus pour les quatre types d'évènements *switch* (récupération proactive (cyan), création de *probe* (bleu), récupération réactive (rouge) et consolidation (jaune)) comparés aux évènements *stay* (exploitation (noir) et exploration (gris)). u.a : unité arbitraire. Le point 0 correspond à l'apparition du stimulus au premier essai suivant le *switch*. Les barres d'erreur correspondent à l'erreur standard calculée sur l'ensemble des sujets présentant ce type d'évènements. ***p<0.0005.

Les histogrammes péri-stimulus extraits de l'hippocampe droit confirment

sa spécificité pour les évènements de récupération proactive (courbe cyan, T > 4.13, p < 0.0005). Aucun des autres types de *switch* (récupération réactive, création de *probe* et consolidation) n'engendre d'activité significative (p > 0.2).

Notons par ailleurs que, comme le montre les valeurs des coefficients de corrélation β, cette région ne corrèle pas avec le type de contexte (nouveau ou ancien) (p > 0.1).

Activation du striatum ventral (figure 12.8)

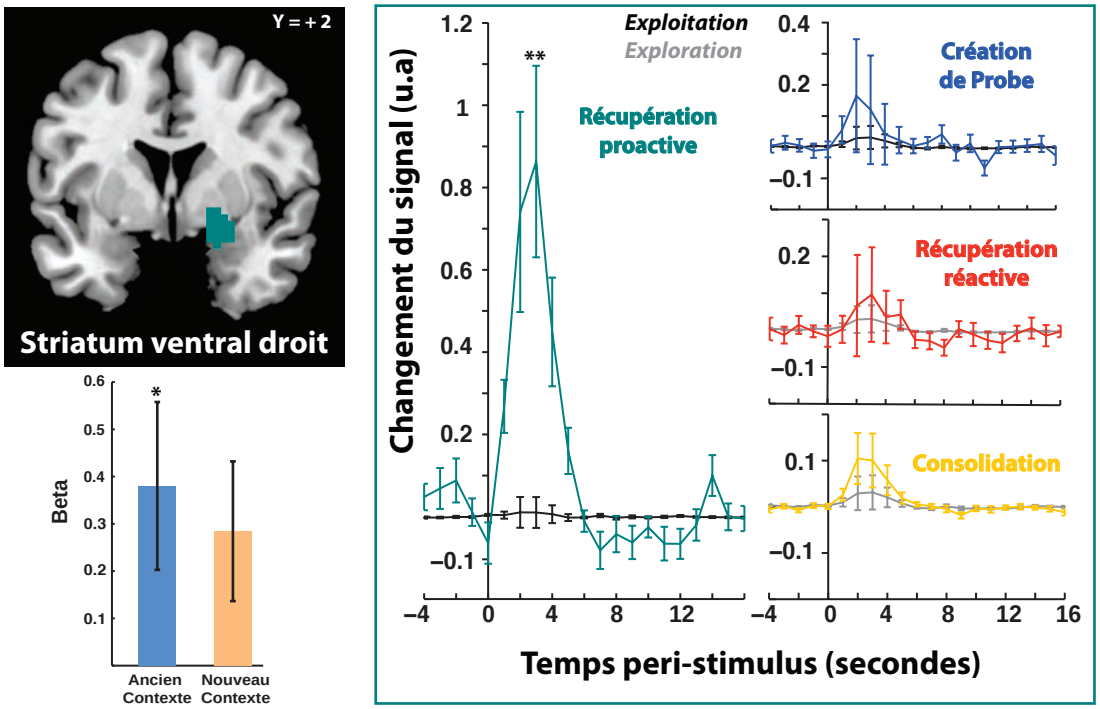


FIGURE 12.8 – **Région spécifique de la récupération proactive (2) : le striatum ventral droit.** En haut à gauche : région du striatum ventral droit observé sur la carte statistique de la récupération proactive (au seuil p = 0.05 pour correction multiple). Pic voxel aux coordonnées MNI [21, 2, -11]. En bas à gauche : valeur des coefficients de corrélation β pour l'apparition d'un contexte ancien (bleu) ou nouveau (orange). A droite : Histogrammes péri-stimulus pour les quatre types d'évènements *switch* (récupération proactive (cyan), création de *probe* (bleu), récupération réactive (rouge) et consolidation (jaune)) comparés aux évènements *stay* (exploitation (noir) et exploration (gris)). u.a : unité arbitraire. Le point 0 correspond à l'apparition du stimulus au premier essai suivant le *switch*. Les barres d'erreur correspondent à l'erreur standard calculée sur l'ensemble des sujets présentant ce type d'évènements. **p<0.005.

Les histogrammes péri-stimulus extraits du striatum ventral droit confirment sa spécificité pour les évènements de récupération proactive (courbe cyan, T > 2.99, p < 0.005). Aucun des autres types de *switch* (récupération réactive, création de *probe* et consolidation) n'engendre d'activité significative (p > 0.2).

Les valeurs des coefficients de corrélation β montrent que cette région corrèle avec l'apparition d'un contexte ancien ($T > 2.16$, $p < 0.05$) mais pas de façon significativement différente que pour l'apparition d'un contexte nouveau ($p > 0.5$).

Activation du cortex préfrontal ventro-médian (figure 12.9)

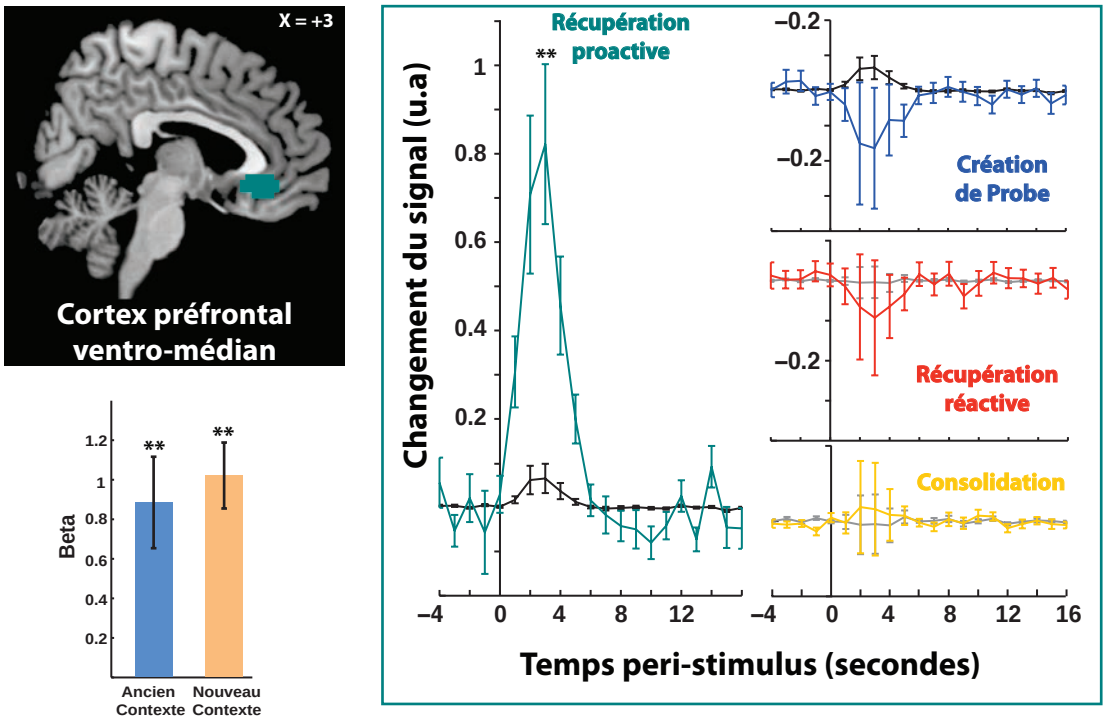


FIGURE 12.9 – **Région spécifique de la récupération proactive (3) : le cortex préfrontal ventro-médian.** En haut à gauche : région du cortex préfrontal ventro-médian observé sur la carte statistique de la récupération proactive (au seuil $p = 0.05$ pour correction multiple). Pic voxel aux coordonnées MNI [0, 35, -8]. En bas à gauche : valeur des coefficients de corrélation β pour l'apparition d'un contexte ancien (bleu) ou nouveau (orange). A droite : Histogrammes péri-stimulus pour les quatre types d'évènements *switch* (récupération proactive (cyan), création de *probe* (bleu), récupération réactive (rouge) et consolidation (jaune)) comparés aux évènements *stay* (exploitation (noir) et exploration (gris)). u.a : unité arbitraire. Le point 0 correspond à l'apparition du stimulus au premier essai suivant le *switch*. Les barres d'erreur correspondent à l'erreur standard calculée sur l'ensemble des sujets présentant ce type d'évènements. **$p<0.005$.

Les histogrammes péri-stimulus extraits du cortex préfrontal ventro-médian confirment sa spécificité pour les évènements de récupération proactive (courbe cyan, $T > 3.5$, $p < 0.005$). Aucun des autres types de *switch* (récupération réactive, création de *probe* et consolidation) n'engendre d'activité significative ($p > 0.2$).

Les valeurs des coefficients de corrélation β montrent que cette région corrèle avec le changement de contexte, qu'il soit ancien ou nouveau ($T > 3.86$, $p < 0.0005$, t-test apparié : $p > 0.3$).

12.5 Consolidation : amygdale et striatum

Ainsi que l'observent Donoso, Collins et Kœchlin, nous observons une activation du striatum ventral lors de la consolidation (figure 12. 10, pic voxel aux coordonnées MNI [18, 17, -8]). Selon ces auteurs, cette région détecterait lorsque un *task-set* nouvellement créé devient fiable pour confirmer son stockage en mémoire. Les histogrammes péri-stimulus dans cette région montrent une activation significative pour les évènements de consolidation et de récupération réactive (courbes jaune et rouge, $T > 2.27$, $p < 0.05$ pour les deux) et une activation non significative pour les évènements de récupération proactive et de création de *probe* (courbes cyan et bleu, $p > 0.3$).

Notons que la région du striatum ventral observé ici est anatomiquement différente de celle observée lors des événements de récupération proactive.

Nous observons par ailleurs une légère activation de l'amygdale lors de ces mêmes évènements (pic voxel aux coordonnées MNI [-18, -4, -23]). Cependant, aucun histogramme péri-stimulus ne révèle une activation significative dans cette région (non montré).

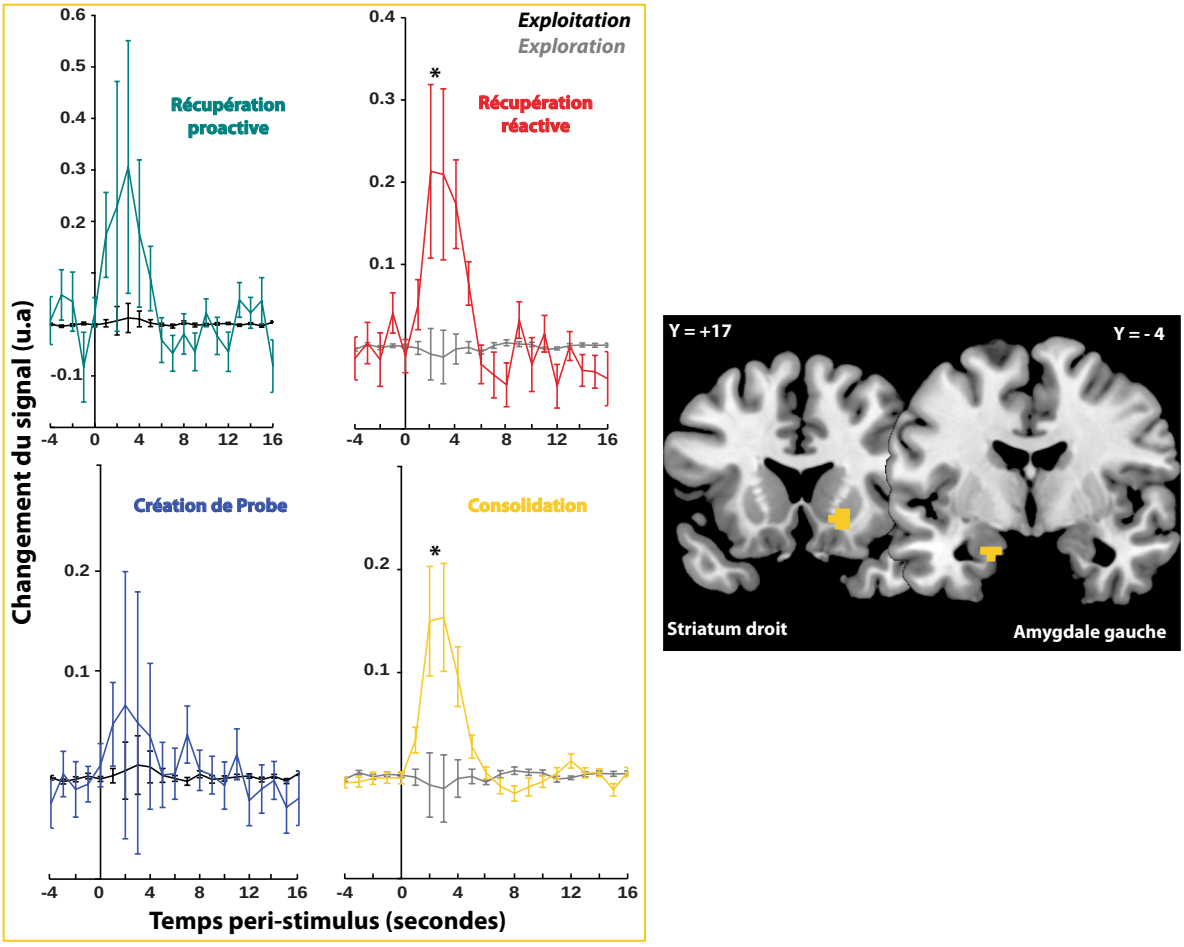


FIGURE 12.10 – **Régions de la consolidation : striatum ventral et amygdale.** A droite : striatum ventral droit et amygdale gauche observés sur la carte statistique de la consolidation (au seuil p = 0.05 pour correction multiple). Pics voxels aux coordonnées MNI [18, 17, -8] et [-18, -4, -23] respectivement. A gauche : Histogrammes péri-stimulus pour les quatre types d'évènements *switch* (récupération proactive (cyan), création de *probe* (bleu), récupération réactive (rouge) et consolidation (jaune)) comparés aux évènements *stay* (exploitation (noir) et exploration (gris)). u.a : unité arbitraire. Le point 0 correspond à l'apparition du stimulus au premier essai suivant le *switch*. Les barres d'erreur correspondent à l'erreur standard calculée sur l'ensemble des sujets présentant ce type d'évènements. *p<0.05.

Sixième partie

Discussion

Chapitre 13

Dissociation des processus réactifs et proactifs dans les lobes frontaux

13.1 Manifestations comportementales de l'apprentissage et de la récupération de *task-sets*

Les résultats comportementaux de notre expérience confirment que les humains peuvent apprendre, de façon non supervisée, un répertoire de *task-sets* et en tirer profit pour les réutiliser en cas de besoin. Ils montrent que nous sommes capables de découvrir qu'une dimension sensorielle apporte de l'information contextuelle, et d'apprendre à utiliser cette information a priori, pour la sélection des *task-sets*.

D'après le modèles PROBE, il existe trois types de transitions entre *task-sets* : la création de *probe*, la récupération et la consolidation.

La création de *probe* correspond à l'entrée en exploration. Elle a lieu lorsque le *task-set* acteur n'est plus fiable et se traduit par une chute abrupte de la persévération.

La récupération est soit proactive, soit réactive. La récupération proactive correspond à l'utilisation a priori de l'information contextuelle pour récupérer le *task-set* correspondant en mémoire de travail. Elle se traduit également par une chute abrupte de la persévération dès l'apparition de l'indice contextuel. La récupération réactive correspond quant à elle à l'abandon d'un *probe* (c'est-à-dire la sortie d'exploration), en fonction des *feedbacks* obtenus, pour récupérer un *task-set* fiable en mémoire de travail. Elle se traduit par une augmentation rapide des performances lorsque la récupération est correcte.

Enfin, la consolidation correspond à la validation du *probe* et se traduit par une augmentation graduelle de la performance.

Ces différents types de transitions prédits par le modèle sont observables chez les sujets et présentent des corrélats cérébraux spécifiques.

13.2 Réseaux cérébraux des processus d'inférences

Ainsi que l'on observé Donoso, Collins et Kœchlin, nous retrouvons dans notre étude deux réseaux d'inférence dans les lobes frontaux : un réseau médian et un réseau latéral [74].

13.2.1 Réseau médian de l'inférence sur le *task-set* acteur

Ce réseau comprend le cortex préfrontal ventro-médian, la pré-SMA et le striatum ventral. Il implémente des inférences sur le *task-set* acteur, c'est-à-dire celui qui sélectionne les actions.

L'activité du cortex préfrontal ventro-médian corrèle avec la fiabilité du *task-set* acteur. Ainsi, comme nous l'avions suggéré (voir 5.1), cette région semble représenter le *task-set* acteur.

Donoso, Collins et Kœchlin suggèrent que le cortex cingulaire antérieur (et la pré-SMA) détecte le moment où le *task-set* acteur devient non fiable et qu'il est nécessaire de l'abandonner pour entamer une période d'exploration (création de *probe*). Dans notre cas, seule la pré-SMA semble jouer cette fonction. Plus précisément, étant donné que nous observons cette région pour la création de *probe* et la récupération proactive (les deux seuls cas où l'on quitte un *task-set* acteur), plutôt que pour l'entrée en exploration, cette région semble nécessaire pour inhiber le *task-set* acteur. Cette interprétation est d'ailleurs cohérente avec le fait que la pré-SMA est impliquée dans l'inhibition de réponses comportementales établies ou automatiques [116]. Cette inhibition vient de l'inférence que le *task-set* acteur n'est plus fiable. Il est abandonné et devient un *task-set* alternatif.

La fiabilité du *task-set* acteur peut chuter, d'une part, parce que les récompenses qu'il prédit n'apparaissent plus (les *feedbacks* sont davantage négatifs que positifs) et, d'autre part, à l'apparition d'un contexte qui ne lui est pas associé. Deux cas peuvent alors se produire :

(i) L'apparition d'un contexte connu provoque la diminution brusque de la fiabilité du *task-set* acteur en cours d'utilisation et la récupération proactive du *task-set* associé à ce contexte ;

(ii) Le *task-set* acteur devient non fiable mais aucune alternative ne semble correspondre aux nouvelles contingences de l'environnement. Il s'ensuit donc une période d'exploration, c'est-à-dire la formation d'un *task-set probe* pondéré par l'ensemble des *task-sets* stockés en mémoire à long terme.

Dans le cas de la création d'un *probe*, le striatum ventral détecte lorsque ce *task-set probe* devient fiable, ce qui achève la période d'exploration et permet la consolidation en mémoire de ce nouveau *task-set*. Notons que l'événement de consolidation peut être considéré comme un renforcement positif. En effet, il correspond au moment où, d'après le modèle, le sujet semble avoir trouvé la combinaison correcte, ce qui assimilable à une récompense. De plus, lors de ces évène-

ments, nous avons observé une activation de l'amygdale gauche. La confirmation d'un *probe* s'accompagne donc peut-être d'une émotion positive. Autrement dit, peut-être observons nous ici le réseau cérébral lié au sentiment de "Eurêka ! J'ai trouvé !".

13.2.2 Réseau latéral de l'inférence sur les *task-sets* alternatifs

Ce réseau latéral comprend le cortex frontopolaire et le cortex préfrontal dorso-latéral. Ces régions encodent les signaux de fiabilité associés aux *task-sets* alternatifs stockés en mémoire de travail et détectent lorsque l'un d'eux devient fiable afin de le récupérer.

En accord avec les résultats de Donoso, Collins et Kœchlin, nous observons une activation du cortex frontopolaire en corrélation avec les signaux de confiance des *task-sets* alternatifs. Ceci est cohérent avec le rôle de cette région dans le maintien d'options alternatives [139][23].

Nous n'observons pas de dissociation entre les signaux de fiabilité *ex-post* (μ) actualisés en fonction des *feedbacks* reçus et les signaux de fiabilité *ex-ante* (λ) actualisés en fonction de l'information contextuelle. Notons que, bien que nous n'observions pas de dissociation claire, il n'est pas impossible qu'elle existe. Notre protocole n'est cependant pas en mesure de la capturer. En effet, étant donné que, d'une part, le modèle PROBE n'inclut pas de notion de temporalité ou de durée d'actualisation des signaux de fiabilité et que, d'autre part, dans notre protocole, l'information contextuelle reste affichée sur une longue durée, il est difficile de déterminer à quel moment précis les signaux de confiance sont actualisés, et donc de dissocier les effets des *feedbacks* des effets du contexte (voir 14.3). Nos résultats suggèrent cependant que le cortex frontopolaire est légèrement plus spécifique des signaux de confiance *ex-post* μ et que le cortex préfrontal dorso-latéral (à gauche notamment) est légèrement plus spécifique de signaux de confiance *ex-ante* λ. Cette interprétation va dans le sens des rôles respectifs du cortex frontopolaire et du cortex préfrontal dorso-latéral pour la mise à jour des signaux de fiabilité en fonctions des *feedbacks* [74] et l'intégration de l'information contextuelle [132].

Le cortex préfrontal dorso-latéral apparaît quant à lui pour tous les évènements de récupération (proactive et réactive), c'est-à-dire les mécanismes qui permettent, en se fondant sur les signaux de fiabilité, de rechercher en mémoire de travail un *task-set* correspondant aux contingences externes. Nous avons déjà mentionné le rôle du cortex préfrontal dorso-latéral pour le contrôle cognitif et la sélection d'éléments maintenus en mémoire de travail. Cette région semble donc responsable de la détection et de la sélection d'un *task-set* alternatif devenu fiable afin de le récupérer, c'est-à-dire l'implémenter dans le système exécutif.

Notons que nous observons une latéralisation des processus de récupération : à gauche pour la récupération proactive et à droite pour la récupération réactive. Il est délicat de s'avancer sur l'interprétation de cette latéralisation qui

ne traduit pas nécessairement une véritable dissociation anatomo-fonctionnelle de ces processus. En effet :

- D'une part, l'activité du cortex-préfrontal dorso-latéral droit pour la récupération réactive n'est pas significativement différente de celle pour la récupération réactive ;

- D'autre part, Donoso, Collins et Kœchlin observent une activation latéralisée à gauche pour les évènements de récupération réactive [74].

13.3 Réseaux cérébraux des processus de récupération

Les réseaux des processus d'inférence décrits dans la partie précédente correspondent, en grande partie, à la reproduction des résultats de l'étude de Donoso, Collins et Kœchlin [74]. Notre étude permet donc de confirmer ces résultats tout en y intégrant la dimension contextuelle.

En revanche, jusqu'à présent, aucune étude n'a tenté d'identifier les réseaux spécifiques des récupérations réactive et proactive. Nos résultats montrent que ces deux processus, bien que de nature similaire, dépendent de réseaux distincts.

Rappelons que la récupération se définit comme le processus par lequel une information stockée en mémoire est rendue disponible pour effectuer une opération cognitive. Ce processus s'effectue typiquement via l'utilisation d'un indice de récupération qui permet d'identifier la représentation à récupérer.

13.3.1 Récupération réactive

La récupération réactive correspond à la récupération d'un *task-set* en mémoire de travail d'après les *feedbacks* observés. Autrement dit, aucun indice extérieur ne permet de déterminer le *task-set* correct. L'observation des *feedbacks* permet d'accumuler des preuves en faveur d'un éventuel *task-set* stocké en mémoire. Ainsi, on peut considérer que l'indice de récupération correspond à la quantité cumulée de preuves en faveur de ce *task-set*, quantité qui se reflète dans la valeur du signal de fiabilité. L'indice de récupération est donc généré de façon interne. Or, nous avons vu (voir 7.2.3) que le cortex frontopolaire est impliqué dans le traitement d'information générée de façon interne et semble lié au processus de récupération. L'activation frontopolaire que nous observons lors des évènements de récupération réactive est donc cohérente. Ce résultat rejoint d'ailleurs le rôle de cette région dans le monitorage des options alternatives et notamment de la fiabilité *ex-post* actualisée selon les récompenses obtenues. Autrement dit, ce type de récupération est intimement lié aux processus d'inférence du système exécutif.

13.3.2 Récupération proactive

Lors de la récupération proactive, l'indice de récupération (le contexte) est donné dans l'environnement. Il est directement observable. Il est donc légitime d'observer un réseau distinct du précédent pour ce type de récupération. Le réseau que nous observons est un réseau ventral très spécifique qui inclut l'hippocampe, le striatum ventral et le cortex préfrontal ventro-médian. Nous proposons ici une interprétation quant au rôle de chacune de ces régions.

Réactivation du *task-set* dans l'hippocampe

Comme nous l'avons vu (voir 7.2.2), l'hippocampe est une région clé des processus mnésiques. Il ne contient pas le contenu mnésique en lui même mais, à la présentation de l'indice de récupération, il permet la réactivation du contenu associé à cet indice. Ainsi, lors de la présentation de l'indice contextuel connu, le *task-set* correspondant pourrait émerger par mémoire associative via l'hippocampe.

L'hippocampe entretient par ailleurs des connexions avec le cortex préfrontal ventro-médian et le striatum [250]. Ainsi l'hippocampe pourrait transmettre des représentations mnésiques à ces modules d'estimation que nous observons également [120].

Il est possible que la représentation réactivée par l'hippocampe corresponde à la combinaison de chiffres plutôt qu'à l'ensemble des associations stimulus-réponse du *task-set* correspondant. Autrement dit, il ne s'agirait pas d'une représentation de mémoire procédurale (d'un *task-set*) mais d'une représentation de mémoire déclarative. Cela dit, comme le soulève Oberauer [168], ces deux types de représentations sont intimement liés puisque toute représentation déclarative est "traduite" en représentation procédurale pour permettre la sélection d'une action.

Nous pouvons également relier ce résultat au modèle de *dual-process* (voir 14.1). Rappelons qu'il existe trois interprétations quant au rôle de l'hippocampe dans ce modèle :

1) Il pourrait être spécifique du processus de *recollection* ;

2) Il pourrait être corrélé à la "force" du souvenir ;

3) Il encoderait uniquement l'information contextuelle.

Dans le cadre du modèle PROBE, nous envisageons la récupération (proactive ou réactive) en mémoire de travail comme un processus de *recollection*.

Nous observons une activation de l'hippocampe uniquement lors des évènements de récupération proactive. Nous pouvons donc supposer que l'hippocampe est spécifique du processus de *recollection* lorsqu'un indice de récupération est présent dans l'environnement. Or ce processus dépend également d'une "force d'activation" (la valeur du signal de fiabilité) qui elle-même, pour la récupération proactive, dépend de la force d'association avec les indices contextuels. Autrement dit, les trois interprétations concernant le rôle de l'hippocampe dans le modèle de *dual-process* semblent compatibles, au moins dans le cadre de notre étude.

Validation dans le striatum ventral

Le striatum ventral est connecté à l'hippocampe et au cortex préfrontal. Il est donc au carrefour des informations provenant de ces deux régions. On pourrait ainsi supposer que, lors de la récupération proactive, le striatum reçoit d'une part une représentation *bottom-up* provenant de l'hippocampe ; et d'autre part une représentation *top-down* sélectionnée par le cortex préfrontal dorso-latéral lors du processus d'inférence proactive. L'adéquation de ces deux représentations pourrait être interprétée comme un signal de renforcement positif. Autrement dit, la représentation qui émerge de façon spontanée depuis l'hippocampe et celle sélectionnée de façon contrôlée par le cortex préfrontal dorso-latéral sont identiques, ce qui permet la validation de cette représentation par le striatum ventral.

On peut se demander quel est l'intérêt de ce dédoublement du processus de récupération. Lorsqu'un *task-set* est fortement associé à un contexte donné (par exemple parce que leur occurrence simultanée est très fréquente), il semble avantageux que la récupération ce de *task-set* devienne quasiment automatique et qu'il émerge de façon *bottom-up* sans passer par le système exécutif. Cependant, il est tout à fait possible que ce processus de récupération nécessite un certain contrôle qu'effectuerait le cortex préfrontal dorso-latéral. Prenons un exemple pour illustrer notre propos. Admettons que je sois chez moi et que le téléphone sonne. A priori, dès que je suis dans le contexte "à la maison", je récupère automatiquement le *task-set* "habitudes à la maison". Par conséquent, au stimulus "sonnerie du téléphone" je devrais associer l'action de "répondre". Cependant, supposons qu'un soir je reçois des amis à dîner. Si le téléphone sonne, par souci de bienséance, je me retiendrai de répondre. Ainsi, on pourrait supposer que dans ce cas, la représentation provenant de l'hippocampe nécessite d'être inhibée car une information contextuelle supplémentaire (la présence de mes amis), intégrée par le cortex préfrontal dorso-latéral, induit un autre comportement.

Ainsi, contrairement à la récupération réactive, la récupération proactive semble partiellement indépendante des processus d'inférence. Elle reste cependant sous son contrôle.

Prédiction dans le cortex préfrontal ventro-médian

Ainsi que nous l'avons déjà mentionné, le cortex préfrontal ventro-médian semble représenter le *task-set* acteur. On peut donc penser qu'une fois qu'un *task-set* alternatif a été sélectionné et validé par les régions précédemment décrites, le cortex préfrontal ventro-médian intègre l'ensemble des informations et implémente le nouveau *task-set* acteur.

Remarquons que nous n'observons pas d'activation dans le cortex préfrontal ventro-médian lors de la récupération réactive. Ceci pourrait s'expliquer par le fait que lors de la récupération proactive, on passe directement d'un *task-set* acteur fiable à un autre *task-set* acteur fiable. La récupération proactive correspond

au seul cas où ce type de transition se produit. On change donc entièrement de prédiction. En revanche, pour la récupération réactive, on ne fait "que" rejeter un *probe* non fiable. Ceci pourrait expliquer la forte activation du cortex préfrontal ventro-médian lors de la récupération proactive et non lors de la récupération réactive.

Afin de confirmer ces interprétations, il serait nécessaire d'effectuer des analyses de connectivité et/ou d'utiliser une méthode dont la résolution temporelle est meilleure que celle de l'IRM fonctionnelle. En effet, ces interprétations soulèvent des hypothèses quant à l'ordre de recrutement de ces diverses régions (hippocampe et cortex préfrontal dorso-latéral dans un premier temps, puis striatum ventral, et enfin cortex préfrontal ventro-médian).

Chapitre 14

Limites et perspectives

14.1 Limites du modèle PROBE et du protocole expérimental

Quel que soit le critère considéré, le modèle PROBE s'avère quantitativement et qualitativement convaincant pour expliquer le comportement des sujets dans des tâches qui nécessitent l'apprentissage, le stockage et la sélection de *task-sets* dans des environnements variables et ouverts. Il présente cependant des limites, tout comme le protocole élaboré pour le tester.

14.1.1 Schéma de renforcement

Dans notre protocole nous nous sommes limités à un schéma de renforcement binaire basique tel qu'il est implémenté dans le modèle PROBE. Cependant, il est fréquent que, plutôt qu'un simple "gagné" ou "perdu", le retour de l'environnement soit continu et complexe (nombre de points, argent, quantité de plaisir, etc.).

Le point clé du modèle PROBE est l'estimation bayésienne du signal de confiance dans les *task-sets*. Pour cela chaque *task-set* est équipé d'un modèle prédictif, lui permettant de prédire (lorsque ce *task-set* est adapté aux contingences extérieures) les conséquences de la sélection d'une action face à un stimulus.

Dans le cadre de renforcements binaires, le modèle peut estimer simplement, après chaque essai, quelle était la probabilité de la conséquence observée. Cependant, un modèle prédictif pourrait également être implémenté dans le cadre d'une fonction de renforcement plus complexe. Le choix du renforcement binaire est donc davantage un choix de simplification qu'une limitation théorique.

14.1.2 Modélisation de la mémoire

Nous construisons probablement, au cours de notre vie, un nombre très important de *task-sets*. Il semble peu crédible que l'ensemble de ces *task-sets* soient mis en concurrence et évalués à tout instant où du contrôle cognitif est nécessaire. Il semble plus probable que selon le contexte, seul un petit nombre de *task-sets* est considéré par le cerveau. En effet, la plupart des *task-sets* disponibles dans notre répertoire ne seront pas appropriés dans un contexte donné, ce qui implique que leur confiance a priori est extrêmement faible. On peut ainsi supposer que seul un nombre limité de *task-sets* émergent et est stocké en mémoire de travail où la sélection a lieu. C'est ce que modélise le modèle PROBE.

Cependant, si aucun *task-set* présent en mémoire de travail n'est fiable, on devrait pouvoir récupérer une représentation directement en mémoire à long terme (sans passer par la mémoire de travail) sur la base du modèle contextuel. Ainsi on pourrait considérer une troisième façon de sortir d'exploration qui n'est pas implémentée dans le modèle : la récupération d'un *task-set* fiable qui émerge directement de la mémoire à long terme grâce à une forte association avec un indice contextuel. Cette possibilité est impossible en l'absence de contexte. Elle n'a pas été modélisée dans le modèle PROBE car il ne s'agit pas vraiment d'un processus exécutif, mais plutôt d'un processus purement mnésique. Cependant, si le modèle exécutif n'a rien à proposer, il devrait considérer ce qui émerge potentiellement de la mémoire à long terme. L'interface entre le système exécutif et la mémoire reste donc à explorer plus précisément.

14.1.3 Temporalité des évènements

Le modèle PROBE ne fait aucune hypothèse sur le temps nécessaire pour réaliser les calculs d'inférence, ni sur le moment précis où ces évènements ont lieu. Il est en mesure d'intégrer simultanément les informations telles que les stimuli et les contextes. Or nos données suggèrent que les sujets, lorsque le protocole le permet, traitent les informations de façon sérielle. Dans notre cas par exemple, le contexte reste affiché tout au long de l'épisode. Les sujets semblent donc prioriser l'information en traitant en premier lieu le stimulus pour lequel ils doivent fournir une réponse rapidement, puis intègrent l'information du contexte. C'est probablement pour cette raison que, d'une part, nous ne parvenons pas à dissocier les signaux de confiance *ex-post* et *ex-ante* et, d'autre part, que nous observons des créations de *probe* proactives moins nettes chez les sujets que ce que prédit le modèle, ainsi qu'un nombre de récupérations proactives assez faible. Il aurait sans doute été préférable, par exemple, d'afficher le contexte quelques secondes avant le premier stimulus pour que les sujets aient le temps de traiter cette information avant d'avoir à fournir une réponse. Nous aurions dans ce cas observé sans doute davantage de *switches* proactifs.

14.2 Neurones et neuromodulateurs

14.2.1 Implémentation neuronale

Le modèle PROBE propose une théorie algorithmique des processus d'inférence et de récupération au niveau du contrôle cognitif. Nous avons identifié, en continuité du travail de Donoso, Collins et Kœchlin [74], les substrats cérébraux de ces processus. Cependant, si nous supposons que le modèle PROBE modélise effectivement le fonctionnement du cerveau humain, son implémentation au niveau neuronal doit être possible.

Plusieurs études ont montré que des transferts d'informations semblables à des processus d'inférence bayésienne sont représentés dans des réseaux de neurones biologiquement précis.

Kœchlin et collaborateurs ont élaboré un réseau de neurones permettant d'encoder une inférence bayésienne dans le cadre de la modélisation de la perception du mouvement [135]. Les auteurs montrent que ce modèle, qui repose sur un principe d'inférence bayésienne, permet d'expliquer des propriétés connues des neurones de l'aire visuelle MT, responsables de la perception du mouvement. Ils en déduisent qu'il est possible que ces neurones effectuent effectivement ce type d'inférence bayésienne.

Bien que ce modèle soit fondé sur un réseau de neurones, il reste assez loin d'une implémentation biologiquement plausible d'un calcul bayésien puisqu'il est fondé sur des neurones ponctuels.

D'autres études montrent que, comme nécessaire pour une inférence bayésienne, les populations de neurones représentent des distributions de probabilités [146][16]. Ils montrent ainsi qu'on peut modéliser, avec des neurones respectant des propriétés biologiques connues, des calculs d'inférence bayésienne.

Ainsi, bien que la plupart de ces études soient limitées à des zones sensorielles du cortex, et non associatives ou préfrontales, plusieurs modèles semblent montrer que des calculs d'inférence bayésienne peuvent être représentés par des neurones.

Par ailleurs, des données neuronales ont confirmés que les neurones possèdent des propriétés qui leur permettent de réaliser ce type de calculs [131][43].

14.2.2 Neuromodulateurs

Si l'implémentation du modèle PROBE doit trouver des corrélats au niveau cellulaire, il devrait également être possible d'en trouver au niveau moléculaire.

Neuromodulateurs et modélisation

Les avancées récentes dans les techniques de biologie moléculaire ont permis d'obtenir des données abondantes sur la localisation spatiale et les effets physiolo-

giques de différents neuromédiateurs et de leurs récepteurs. Les aires préfrontales sont les cibles de projections majeures de monoamines (dopamine, noradrénaline, sérotonine et acétylcholine). La dopamine et la noradrénaline sont deux systèmes majeurs qui ont été très étudiés dans le contexte de l'adaptation comportementale.

Doya propose par exemple différents rôles spécifiques pour les neuromédiateurs [76]. Selon lui, les neuromédiateurs seraient un moyen de signaler des variables et des paramètres spécifiques qui régulent les modules d'apprentissage répartis dans le cerveau. Plus spécifiquement :

- La dopamine, comme nous l'avons déjà mentionné, encoderait le signal d'erreur de prédiction [207] [155] ;

- Le de taux sérotonine encoderait le facteur de dévaluation d'un modèle de *temporal difference learning*, c'est-à-dire le ratio entre la prédiction à long terme et la prédiction à court terme de l'obtention d'une récompense ;

- La noradrénaline encoderait l'aléa lors de la sélection de l'action, c'est-à-dire le paramètre de température inverse du *softmax* [5] ;

- L'acétylcholine serait responsable de la vitesse de mise à jour de la mémoire et corrélerait donc avec le taux d'apprentissage d'un modèle d'apprentissage par renforcement.

Neuromodulateurs et variabilité interindividuelle

Il semble pertinent de considérer la possibilité que les différences comportementales observées entre les sujets puissent être dues à des différences dans le niveau individuel des divers neurotransmetteurs [55][56].

Une série d'études menée par Cools et collaborateurs confirme cette possibilité dans le cas particulier de la dopamine [55][56]. Par exemple, dans l'une de leurs expériences, ces auteurs on soumis des sujets à une tâche de renversement (*reversal learning*). Ils ont mesuré en parallèle, par la technique de tomographie par émission de positrons, leur niveau individuel de production de dopamine dans le striatum. Ils montrent une corrélation forte entre le niveau de dopamine striatale et la performance des sujets : les sujets ayant un fort niveau de dopamine striatale apprennent mieux à partir des récompenses, tandis que ceux ayant un faible niveau de dopamine striatale apprennent mieux à partir des punitions.

Ainsi, les niveaux individuels de neuromodulateurs pourraient être des causes de variabilité comportementale entre les sujets. Il est possible que ces différences de taux de neuromodulateurs soient corrélée à des différences génétiques.

14.3 Des processus spécifiques à l'humain ?

Tous les mammifères semblent posséder un cortex préfrontal [235]. Cependant, chez les rongeurs, il se réduit principalement à des régions cérébrales limbiques incluant le cortex orbitofrontal et le cortex cingulaire antérieur [235]. Chez

les primates, le cortex préfrontal a évolué avec le développement des régions préfrontales latérales [92]. Chez les humains, ces régions ont évolué davantage, avec notamment le développement bilatéral de sa portion la plus antérieure, le cortex frontopolaire latéral [210], qui n'a apparemment pas d'homologue chez les singes [133].

Kœchlin propose que chacune de ces étapes évolutives s'accompagne d'une augmentation graduelle des capacités d'inférences, avec dans un premier temps l'émergence des inférences réactives dans les régions préfrontales limbiques chez les rongeurs ; puis l'apparition des inférences proactives dans les régions préfrontales latérales chez les primates ; et enfin l'apparition des inférences contrefactuelles réactives et proactives dans les régions frontopolaires humaines (on appelle inférence contrefactuelle une inférence sur une stratégie alternative, c'est-à-dire autre que l'acteur) [134].

Ainsi, les régions préfrontales limbiques des petits mammifères permettraient d'inférer un changement de contingences en se fondant sur l'écart entre les résultats observés et ceux prédits par la stratégie actrice. Les régions du cortex préfrontal latéral chez les primates permettrait quant à elles d'intégrer l'information contextuelle a priori afin d'effectuer des inférences proactives. Enfin, le cortex frontopolaire permettrait d'inférer qu'une stratégie alternative est désormais fiable. Ce dernier processus dote les êtres humains de la capacité, non seulement d'inférer un changement de contingences, mais également de déterminer quelle stratégie choisir suite à ce changement.

Conclusion

Les décisions, l'ultime expression de la volonté, doivent leur véritable existence à des processus cachés dans les replis du cerveau. La compréhension des mécanismes cérébraux qui sous-tendent la prise de décision chez l'humain est en plein essor. La multitude de facteurs qui contribuent à la prise de décision, et la multitude de comportements qui en découlent, en font un problème complexe. Malgré tout, l'approche multidisciplinaire des sciences cognitives, qui intègrent des domaines tels que la neuroimagerie, l'étude de patients, l'électrophysiologie et la modélisation mathématique, est en mesure de révéler nos représentations mentales, et ainsi de faire le lien entre nos perceptions et notre comportement.

En effet, c'est grâce aux outils de modélisation et d'imagerie cérébrale que nous avons mis en évidence un réseau cérébral spécifique de la récupération proactive, réseau inconnu jusqu'à présent.

Dans notre protocole, les sujets devaient apprendre une collection de *task-sets* en association avec des indices contextuels, puis les réutiliser. Nous avons montré que les sujets sont capables, dès l'apparition de l'indice contextuel, de récupérer le *task-set* correspondant. Le modèle PROBE, sur lequel se fonde notre étude, prédit formellement ces évènements. Toutefois, les bases cérébrales de la récupération proactive restaient à déterminer.

Récupérer proactivement un *task-set* nécessite une collaboration étroite entre le système exécutif et le système mnésique. Nous avons révélé que les bases cérébrales associées à ces évènements impliquent effectivement le cortex préfrontal dorso-latéral et un réseau ventral incluant l'hippocampe, reliant ainsi prise de décision et mémoire.

Bien que le cadre expérimental de cette thèse soit assez peu écologique, il paraît toutefois évident que nos résultats s'appliquent à de nombreuses situations quotidiennes. En effet, nous sommes en permanence confrontés à la nécessité de passer d'un comportement à un autre, la plupart du temps dans des contextes connus. Et, dans un contexte connu, il est pertinent de faire appel à nos apprentissages passés pour les réutiliser. Ceci nécessite cependant un certain contrôle afin de maintenir un comportement suffisamment flexible et s'adapter à tout changement. Les systèmes exécutif et mnésique sont ainsi essentiels pour faire de nous ce que nous sommes : des êtres capables de s'adapter à un monde à la fois stable et en constante évolution.

Annexes

Annexe A

Les contextes

LES CONTEXTES

Les contextes correspondent à des textures colorées en fond d'écran. Les images ont toutes été modifiées au format 800*600 pixels, longueur = 28.22 cm et largeur = 21.17 cm, avec une résolution de 200 DPI. Certaines ont été homogénéisées sur iPhoto.

Les contextes ont été sélectionnés de telle sorte qu'ils soient discernables les uns des autres. La distribution des contextes par couleur/catégorie est récapitulée dans le tableau ci-dessous.

Les contextes et leur numéro correspondant

	Animal	Minéral	Aliment	Nature	Tissus	Manufacturé
Vert	1. Crocodile	8. Émeraude	15. Salade	22. Pelouse	29. Coussin	36. Tapisserie
Rouge	2. Cuir	9. Pierre	16. Épices	23. Fleurs	30. Drap	37. Taule
Jaune	3. Léopard	10. Terre peinte	17. Citron	24. Fleurs	31. Rideau	38. Éponge
Bleu	4. Écailles	11. Azurite	18. Eau	25. Ciel	32. Jean	39. Carrelage
Brun	5. Pelage	12. Terre	19. Café	26. Feuilles	33. Tissu	40. Chocolat
Beige	6. Peau humaine	13. Sable	20. Mie	27. Graines	34. Filet	41. Papier froissé
Gris	7. Fourrure	14. Mur pavé	21. Riz	28. Écorce	35. Laine	42. Aluminium

Les 42 contextes



Annexe B

Les instructions

REGLES DU JEU

Cette expérience est un jeu où il faut mettre des chiffres dans des cases. Votre gain dépendra de vos performances.

I. ENTRAÎNEMENT

Pendant l'entraînement, un stimulus s'affiche à l'écran à intervalles réguliers : il s'agit soit d'une paire de ronds, soit d'une paire de carrés. Ce stimulus s'affiche sur une texture en fond d'écran qui peut changer au cours de l'entraînement. Une même texture peut apparaître plusieurs fois.

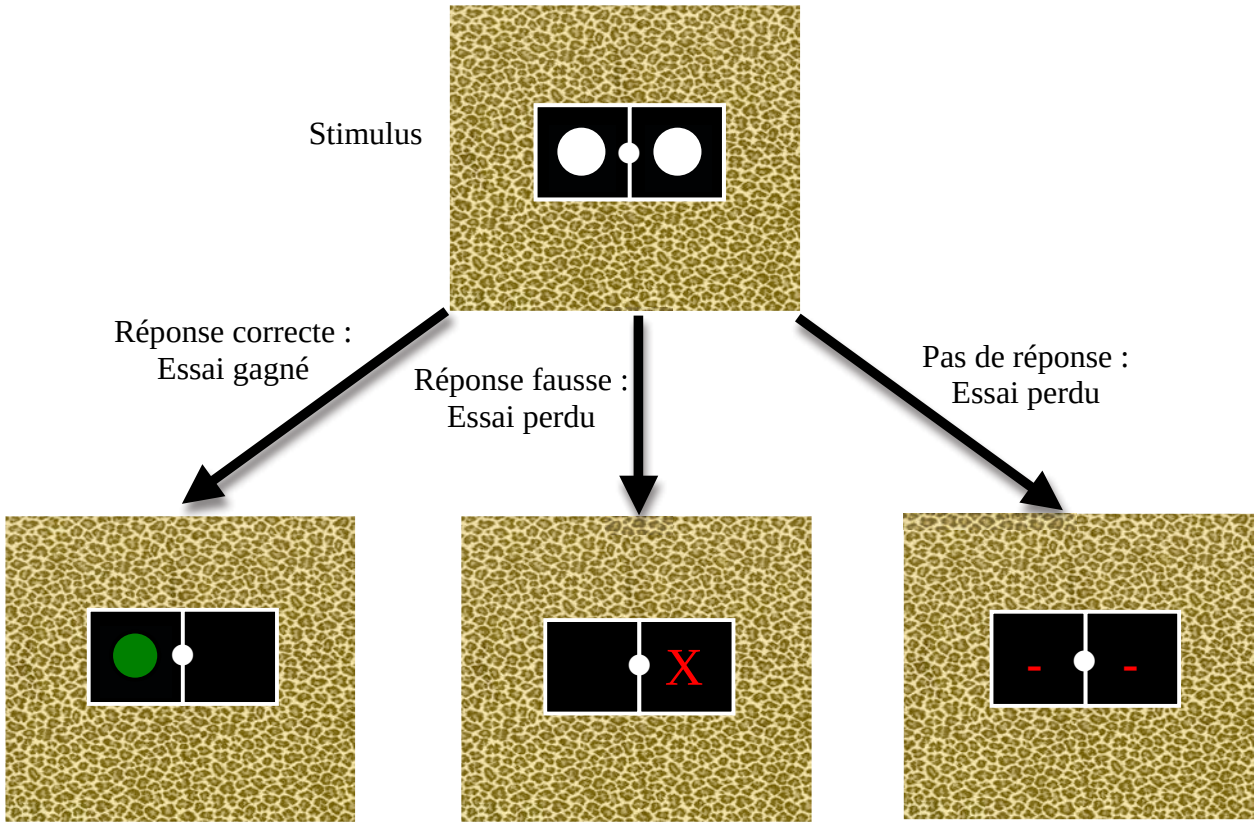
Vous devez répondre en appuyant sur *une* des 2 touches j ou k du clavier en utilisant l'index et le majeur de la main droite. Sur l'écran deux cases représentent les 2 touches. Si vous avez gagné, la forme prend une **couleur verte** et apparaît dans la case correspondant à la touche. Si vous avez perdu, une croix de **couleur rouge** apparaît dans la case correspondant à la touche.

C'est à vous d'apprendre comment associer les formes aux touches à l'aide ou non des textures. Le seul critère est d'essayer de gagner le plus possible. Il vous est demandé de répondre le plus rapidement et correctement possible.

Si vous ne répondez pas dans le temps imparti, l'essai est considéré comme perdu. N'appuyez sur une touche qu'une seule fois par stimulus : seule la première réponse est prise en compte.

Vous allez apprendre une 'configuration correcte' d'association entre formes et touches. Vous devez observer pendant l'entraînement :

- que la réponse de l'ordinateur n'est pas 100% fiable. Il peut arriver que vous appuyiez sur la touche correcte mais que vous perdiez à l'essai. Inversement, il peut arriver que vous appuyiez sur une touche incorrecte mais que vous gagniez à l'essai.
- que la configuration correcte peut changer au cours de l'entraînement.



II. **EXPÉRIENCE** (3 sessions)

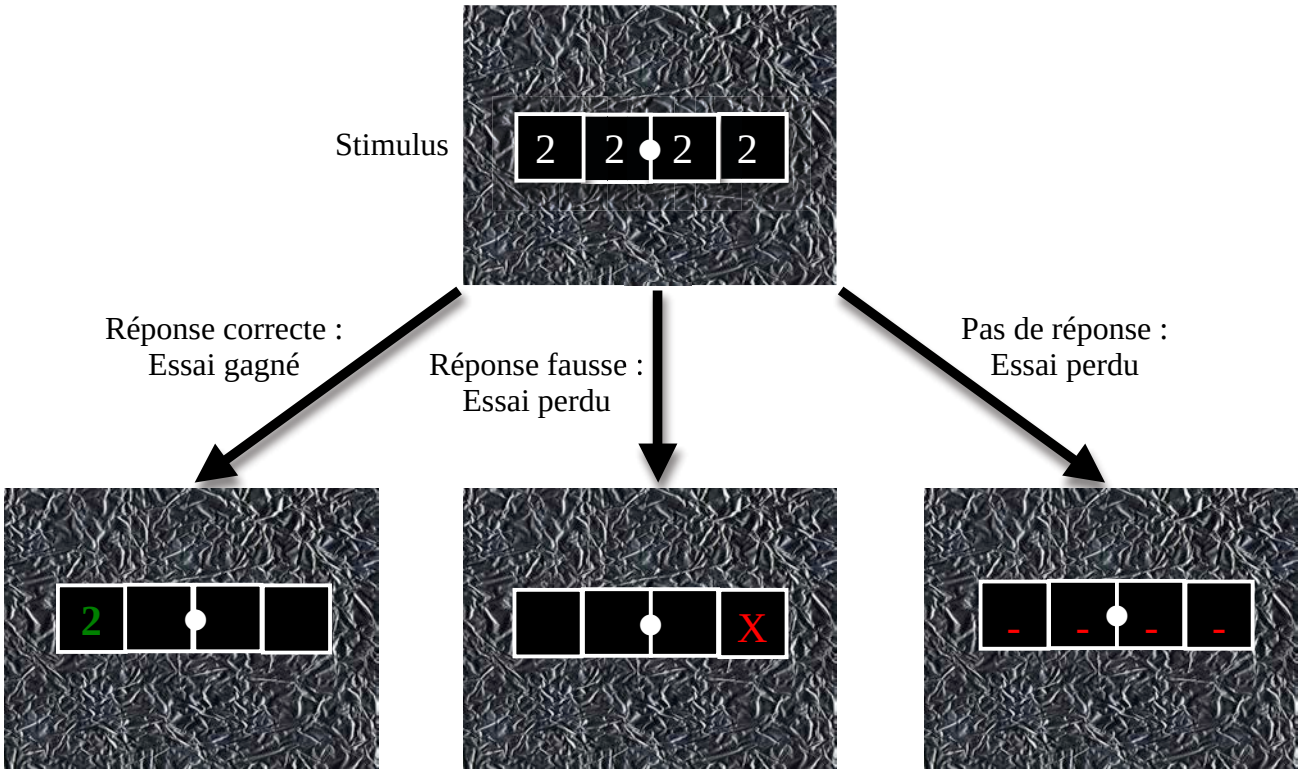
Pendant l'expérience, un stimulus s'affiche à l'écran à intervalles irréguliers. Les stimuli sont des quadruplets d'un même chiffre (1, 2 ou 3). Ce stimulus s'affiche sur une texture en fond d'écran qui peut changer au cours de l'expérience. Une même texture peut apparaître plusieurs fois.

Vous devez répondre en appuyant sur *une* des 4 touches j, k, l ou m du clavier en utilisant les doigts de la main droite. Sur l'écran 4 carrés représentent les 4 touches. Si vous avez gagné, le chiffre prend une **couleur verte** et apparaît dans la case correspondant à la touche. Si vous avez perdu, une croix de **couleur rouge** apparaît dans la case correspondant à la touche.

C'est à vous d'apprendre comment associer les chiffres aux touches à l'aide ou non des textures. Le seul critère est d'essayer de gagner le plus possible.

Il vous est demandé de répondre le plus rapidement et correctement possible. Si vous ne répondez pas dans le temps imparti, l'essai est considéré comme perdu. N'appuyez sur une touche qu'une seule fois par stimulus : seule la première réponse est prise en compte.

Au cours de l'expérience il y aura 2 pauses. Après une pause, l'expérience reprend son cours, donc n'oubliez pas ce que vous faisiez avant la pause.



A la fin de chaque session, nous vous testerons sur les configurations et les textures présentées.

III. **REMUNERATION**

La participation aux sessions 2 et 3 est conditionnée par une participation correcte à la session 1. A la fin de l'expérience, 10% des essais seront tirés au hasard et votre performance sera calculée sur ces essais. Selon votre taux de réussite, vous recevrez une indemnisation pouvant aller de 20 à 200 €.

Bon courage !

Annexe C

Le débriefing informel

Debriefing

Numéro du sujet :

Qu'avez-vous pensé de l'expérience ? Avez-vous rencontré un problème particulier ?

Était-ce :

- Difficile / facile ?
- Long / correct ?
- Fatigant ? OUI / NON
- Rapide / suffisamment lent ?
- Autre ?

Avez-vous réussi à être concentré pendant toute l'expérience ? OUI / NON

Aviez-vous une stratégie ? Si oui, laquelle ?

Pouvez-vous nommer des textures que vous avez vues ?

Avez-vous d'autres commentaires à faire sur l'expérience ?

IMPORTANT :

Merci de ne pas parler de l'expérience à des personnes qui la passeront après vous : il est important de ne pas savoir à l'avance la structure de l'expérience pour que nos données expérimentales soient valides.

Bibliographie

[1] J. P. Aggleton, S. D. Vann, C. Denby, S. Dix, A. R. Mayes, N. Roberts, and A. P. Yonelinas. Sparing of the familiarity component of recognition memory in a patient with hippocampal pathology. *Neuropsychologia*, 43(12) :1810–1823, Jan. 2005.

[2] W. H. Alexander and J. W. Brown. Medial prefrontal cortex as an action-outcome predictor. *Nature Neuroscience*, 14(10) :1338–1344, 2011.

[3] Alvarez and Cavanagh. The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, 15(2) :106–111, 2004.

[4] Anderson and Lebiere. *The atomic components of thought*. Hillsdale, NJ : Lawrence Erlbaum Associates. Psychology Press, 1998.

[5] G. Aston-Jones and J. D. Cohen. Adaptive gain and the role of the locus coeruleus-norepinephrine system in optimal performance. *The Journal of Comparative Neurology*, 493(1) :99–110, Dec. 2005.

[6] A. Baddeley. Working memory : The interface between memory and cognition. *Journal of cognitive neuroscience*, 4(3) :281–288, 1992.

[7] A. Baddeley. The episodic buffer : a new component of working memory ? *Trends in cognitive sciences*, 4(11) :417–423, 2000.

[8] A. Baddeley and G. Hitch. Working memory. *Psychology of Learning and Motivation*, 8 :47–89, 1974.

[9] D. Badre. Cognitive control, hierarchy, and the rostro–caudal organization of the frontal lobes. *Trends in Cognitive Sciences*, 12(5) :193–200, May 2008.

[10] D. Badre and M. D'Esposito. Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex. *Cognitive Neuroscience, Journal of*, 19(12) :2082–2099, 2007.

[11] D. Badre, J. Hoffman, J. W. Cooney, and M. D'Esposito. Hierarchical cognitive control deficits following damage to the human frontal lobe. *Nature Neuroscience*, 12(4) :515–522, Apr. 2009.

[12] D. Badre, A. S. Kayser, and M. D'Esposito. Frontal cortex and the discovery of abstract action rules. *Neuron*, 66(2) :315–326, Apr. 2010.

[13] D. Badre and A. D. Wagner. Computational and neurobiological mechanisms underlying cognitive flexibility. *Proceedings of the National Academy of Sciences*, 103(18) :7186–7191, 2006.

[14] H. Barbas, H. Ghashghaei, S. Dombrowski, and N. Rempel-Clower. Medial prefrontal cortices are unified by common connections with superior temporal cortices and distinguished by input from memory-related areas in the rhesus monkey. *Journal of Computational Neurology*, 410(3) :343–367, 1999.

[15] A. Bechara. The role of emotion in decision-making : Evidence from neurological patients with orbitofrontal damage. *Brain and Cognition*, 55(1) :30–40, June 2004.

[16] J. M. Beck and A. Pouget. Exact inferences in a neural implementation of a hidden markov model. *Neural Computation*, 19(5) :1344–1361, 2007.

[17] T. E. J. Behrens, L. T. Hunt, M. W. Woolrich, and M. F. S. Rushworth. Associative learning of social value. *Nature*, 456(7219) :245–249, Nov. 2008.

[18] T. E. J. Behrens, M. W. Woolrich, M. E. Walton, and M. F. S. Rushworth. Learning the value of information in an uncertain world. *Nature Neuroscience*, 10(9) :1214–1221, Sept. 2007.

[19] K. Benchenane, A. Peyrache, M. Khamassi, P. L. Tierney, Y. Gioanni, F. P. Battaglia, and S. I. Wiener. Coherent theta oscillations and reorganization of spike timing in the hippocampal- prefrontal network upon learning. *Neuron*, 66(6) :921–936, June 2010.

[20] M. Bertin, N. Schweighofer, and K. Doya. Multiple model-based reinforcement learning explains dopamine neuronal activity. *Neural Networks*, 20(6) :668–675, Aug. 2007.

[21] J. R. Binder, R. H. Desai, W. W. Graves, and L. L. Conant. Where is the semantic system ? a critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, 19(12) :2767–2796, Dec. 2009.

[22] R. M. Bonelli and J. L. Cummings. Frontal-subcortical circuitry and behavior. *Dialogues in clinical neuroscience*, 9(2) :141, 2007.

[23] E. D. Boorman, T. E. Behrens, M. W. Woolrich, and M. F. Rushworth. How green is the grass on the other side ? frontopolar cortex and the evidence in favor of alternative courses of action. *Neuron*, 62(5) :733–743, June 2009.

[24] M. M. Botvinick. Conflict monitoring and decision making : reconciling two perspectives on anterior cingulate function. *Cognitive, Affective, & Behavioral Neuroscience*, 7(4) :356–366, 2007.

[25] M. M. Botvinick. Hierarchical models of behavior and prefrontal function. *Trends in Cognitive Sciences*, 12(5) :201–208, May 2008.

[26] M. M. Botvinick, T. S. Braver, D. Barch, C. Carter, and J. D. Cohen. Conflict monitoring and cognitive control. *Psychological Review*, 108(3) :624–652, 2001.

[27] M. Braine. *Mental Logic*. Lawrence Erlbaum Associates, 1998.

[28] T. S. Braver and S. R. Bongiolatti. The role of frontopolar cortex in subgoal processing during working memory. *NeuroImage*, 15(3) :523–536, Mar. 2002.

[29] T. S. Braver, J. D. Cohen, L. E. Nystrom, J. Jonides, E. E. Smith, and D. C. Noll. A parametric study of prefrontal cortex involvement in human working memory. *Neuroimage*, 5(1) :49–62, 1997.

[30] K. Brodmann and L. J. Gary. *Brodmann's localization in the cerebral cortex the principles of comparative localisation in the cerebral cortex based on cytoarchitectonics*. Springer, [New York, NY], 2006.

[31] J. W. Brown. Learned predictions of error likelihood in the anterior cingulate cortex. *Science*, 307(5712) :1118–1121, Feb. 2005.

[32] J. W. Brown. Learned predictions of error likelihood in the anterior cingulate cortex. *Science*, 307(5712) :1118–1121, Feb. 2005.

[33] M. W. Brown and J. P. Aggleton. Recognition memory : what are the roles of the perirhinal cortex and hippocampus ? *Nature Reviews Neuroscience*, 2(1) :51–61, 2001.

[34] R. L. Buckner. Functional–anatomic correlates of control processes in memory. *The journal of neuroscience*, 23(10) :3999–4004, 2003.

[35] R. L. Buckner and D. C. Carroll. Self-projection and the brain. *Trends in Cognitive Sciences*, 11(2) :49–57, Feb. 2007.

[36] P. W. Burgess, S. K. Scott, and C. D. Frith. The role of the rostral frontal cortex (area 10) in prospective memory : a lateral versus medial dissociation. *Neuropsychologia*, 41(8) :906–918, Jan. 2003.

[37] P. W. Burgess, E. Veitch, A. de Lacy Costello, and T. Shallice. The cognitive and neuroanatomical correlates of multitasking. *Neuropsychologia*, 38(6) :848–863, 2000.

[38] B. G. Burton, V. Hok, E. Save, and B. Poucet. Lesion of the ventral and intermediate hippocampus abolishes anticipatory activity in the medial prefrontal cortex of the rat. *Behavioural brain research*, 199(2) :222–234, 2009.

[39] T. Buschman, E. Denovellis, C. Diogo, D. Bullock, and E. Miller. Synchronous oscillatory neural ensembles for rules in the prefrontal cortex. *Neuron*, 76(4) :838–846, Nov. 2012.

[40] R. Buxton, E. Wong, and M. Frank. Dynamics of blood flow and oxygenation changes during brain activation : the balloon model. *Magn Reson Med*, 39(6) :855–364, 1998.

[41] I. G. Cameron, M. Watanabe, G. Pari, and D. P. Munoz. Executive impairment in parkinson's disease : Response automaticity and task switching. *Neuropsychologia*, 48(7) :1948–1957, June 2010.

[42] N. Camille. The involvement of the orbitofrontal cortex in the experience of regret. *Science*, 304(5674) :1167–1170, May 2004.

[43] M. Carandini and D. J. Heeger. Summation and division by neurons in primate visual cortex. *Science*, 264(5163) :1333–1336, 1994.

[44] S. Carmichael and J. Price. Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. *J Comp Neurol*, 363(4) :642–664, 1995.

[45] S. Charron and E. Koechlin. Divided representation of concurrent goals in the human frontal lobes. *Science*, 328(5976) :360–363, Apr. 2010.

[46] Z. Chen and N. Cowan. Chunk limits and length limits in immediate recall : A reconciliation. *Journal of Experimental Psychology : Learning, Memory, and Cognition*, 31(6) :1235–1249, 2005.

[47] V. S. Chib, A. Rangel, S. Shimojo, and J. P. O'Doherty. Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *Journal of Neuroscience*, 29(39) :12315–12320, Sept. 2009.

[48] K. Christoff, V. Prabhakaran, J. Dorfman, Z. Zhao, J. K. Kroger, K. J. Holyoak, and J. D. Gabrieli. Rostrolateral prefrontal cortex involvement in relational integration during reasoning. *NeuroImage*, 14(5) :1136–1149, Nov. 2001.

[49] J. D. Cohen, K. Dunbar, and J. McClelland. On the control of automatic processes : A parallel distributed processing account of the stroop effect. *Psychological Review*, 97(3) :332–361, 1990.

[50] J. D. Cohen, S. M. McClure, and A. J. Yu. 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 B : Biological Sciences*, 362(1481) :933–942, May 2007.

[51] M. Cohn, M. Moscovitch, A. Lahat, and M. P. McAndrews. Recollection versus strength as the primary determinant of hippocampal engagement at retrieval. *Proceedings of the National Academy of Sciences*, 106(52) :22451–22455, 2009.

[52] M. W. Cole, T. Yarkoni, G. Repovs, A. Anticevic, and T. S. Braver. Global connectivity of prefrontal cortex predicts cognitive control and intelligence. *Journal of Neuroscience*, 32(26) :8988–8999, June 2012.

[53] A. Collins. *Apprentissage et contrôle cognitif : une théorie computationnelle de la fonction exécutive préfontale humaine*. PhD thesis, Université Pierre et Marie Curie-Paris VI, 2010.

[54] A. Collins and E. Koechlin. Reasoning, learning, and creativity : Frontal lobe function and human decision-making. *PLoS Biology*, 10(3) :e1001293, Mar. 2012.

[55] R. Cools. Differential responses in human striatum and prefrontal cortex to changes in object and rule relevance. *Journal of Neuroscience*, 24(5) :1129–1135, Feb. 2004.

[56] R. Cools, M. J. Frank, S. E. Gibbs, A. Miyakawa, W. Jagust, and M. D'Esposito. Striatal dopamine predicts outcome-specific reversal learning and its sensitivity to dopaminergic drug administration. *Journal of Neuroscience*, 29(5) :1538–1543, Feb. 2009.

[57] G. Corrado and K. Doya. Understanding neural coding through the model-based analysis of decision making. *Journal of Neuroscience*, 27(31) :8178–8180, Aug. 2007.

[58] N. Cowan. Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, 104(2) :163–191, 1988.

[59] N. Cowan. The magical number 4 in short-term memory : A reconsideration of mental storage capacity. *Behavioral and brain sciences*, 24(01) :154–176, 2000.

[60] N. Cowan. The magical mystery four : How is working memory capacity limited, and why ? *Current Directions in Psychological Science*, 19(1) :51–57, Feb. 2010.

[61] N. Cowan, E. M. Elliott, J. Scott Saults, C. C. Morey, S. Mattox, A. Hismjatullina, and A. R. Conway. On the capacity of attention : Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, 51(1) :42–100, Aug. 2005.

[62] H. D. Critchley, Y. Nagai, M. A. Gray, and C. J. Mathias. Dissecting axes of autonomic control in humans : Insights from neuroimaging. *Autonomic Neuroscience*, 161(1-2) :34–42, Apr. 2011.

[63] E. A. Crone, S. E. Donohue, R. Honomichl, C. Wendelken, and S. A. Bunge. Brain regions mediating flexible rule use during development. *Journal of Neuroscience*, 26(43) :11239–11247, Oct. 2006.

[64] J. Cummings. Anatomic and behavioral aspects of frontal-subcortical circuits. *Annals of the New York Academy of Sciences*, 769 :1–14, 1995.

[65] H. Damasio, T. Grabowski, R. Frank, A. M. Galaburda, A. R. Damasio, and others. The return of phineas gage : clues about the brain from the skull of a famous patient. *Science*, 264(5162) :1102–1105, 1994.

[66] N. D. Daw, J. P. O'Doherty, P. Dayan, B. Seymour, and R. J. Dolan. Cortical substrates for exploratory decisions in humans. *Nature*, 441(7095) :876–879, June 2006.

[67] B. De Martino, S. M. Fleming, N. Garrett, and R. J. Dolan. Confidence in value-based choice. *Nature Neuroscience*, 16(1) :105–110, Dec. 2012.

[68] H. E. M. den Ouden, K. J. Friston, N. D. Daw, A. R. McIntosh, and K. E. Stephan. A dual role for prediction error in associative learning. *Cerebral Cortex*, 19(5) :1175–1185, May 2009.

[69] R. Desimone. Neural mechanisms for visual memory and their role in attention. *Proceedings of the National Academy of Sciences*, 93(24) :13494–13499, 1996.

[70] M. D'Esposito, J. A. Detre, D. C. Alsop, R. K. Shin, S. Atlas, and M. Grossman. The neural basis of the central executive system of working memory. *Nature*, 378 :279–281, 1995.

[71] R. A. Diana, A. P. Yonelinas, and C. Ranganath. Medial temporal lobe activity during source retrieval reflects information type, not memory strength. *Journal of cognitive neuroscience*, 22(8) :1808–1818, 2010.

[72] I. G. Dobbins, H. Foley, D. L. Schacter, and A. D. Wagner. Executive control during episodic retrieval : multiple prefrontal processes subserve source memory. *Neuron*, 35(5) :989–996, 2002.

[73] R. Dolan. Neuroimaging of cognition : Past, present, and future. *Neuron*, 60(3) :496–502, Nov. 2008.

[74] M. Donoso, A. G. E. Collins, and E. Koechlin. Foundations of human reasoning in the prefrontal cortex. *Science*, 344(6191) :1481–1486, June 2014.

[75] N. U. Dosenbach, K. M. Visscher, E. D. Palmer, F. M. Miezin, K. K. Wenger, H. C. Kang, E. D. Burgund, A. L. Grimes, B. L. Schlaggar, and S. E. Petersen. A core system for the implementation of task sets. *Neuron*, 50(5) :799–812, June 2006.

[76] K. Doya. Metalearning and neuromodulation. *Neural Networks*, 15(4) :495–506, 2002.

[77] K. Doya, K. Samejima, K.-i. Katagiri, and M. Kawato. Multiple model-based reinforcement learning. *Neural Computation*, 14(6) :1347–1369, 2002.

[78] J.-C. Dreher, E. Koechlin, S. O. Ali, and J. Grafman. The roles of timing and task order during task switching. *NeuroImage*, 17(1) :95–109, Sept. 2002.

[79] E. Düzel, R. Cabeza, T. W. Picton, A. P. Yonelinas, H. Scheich, H.-J. Heinze, and E. Tulving. Task-related and item-related brain processes of memory retrieval. *Proceedings of the National Academy of Sciences*, 96(4) :1794–1799, 1999.

[80] M. D'Esposito and B. R. Postle. The dependence of span and delayed-response performance on prefrontal cortex. *Neuropsychologia*, 37(11) :1303–1315, 1999.

[81] H. Eichenbaum, A. Yonelinas, and C. Ranganath. The medial temporal lobe and recognition memory. *Annual Review of Neuroscience*, 30(1) :123–152, July 2007.

[82] L. L. Eldridge, B. J. Knowlton, C. S. Furmanski, S. Y. Bookheimer, and S. A. Engel. Remembering episodes : a selective role for the hippocampus during retrieval. *Nature neuroscience*, 3(11) :1149–1152, 2000.

[83] Eriksen and J. St. James. Visual attention within and around the field of focal attention : A zoom lens model. *Perception & Psychophysics*, 40(4) :225–240, 1986.

[84] A. Etkin, T. Egner, and R. Kalisch. Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends in Cognitive Sciences*, 15(2) :85–93, Feb. 2011.

[85] J. S. B. Evans. The mental model theory of conditional reasoning : critical appraisal and revision. *Cognition*, 48(1) :1–20, 1993.

[86] M. Falkenstein, J. Hoormann, S. Christ, and J. Hohnsbein. ERP components on reaction errors and their functional significance : a tutorial. *Biological psychology*, 51(2) :87–107, 2000.

[87] P. Fletcher and R. Henson. Frontal lobes and human memory. insights from functional neuroimaging. *Brain research bulletin*, 124 :849–881, 2001.

[88] P. C. Fletcher, T. Shallice, C. D. Frith, R. S. Frackowiak, and R. J. Dolan. The functional roles of prefrontal cortex in episodic memory. II. retrieval. *Brain*, 121(7) :1249–1256, 1998.

[89] S. Fujisawa and G. Buzsáki. A 4 hz oscillation adaptively synchronizes prefrontal, VTA, and hippocampal activities. *Neuron*, 72(1) :153–165, Oct. 2011.

[90] J. M. Fuster. Prefrontal neurons in networks of executive memory. *Brain research bulletin*, 52(5) :331–336, 2000.

[91] J. M. Fuster. The prefrontal cortex-an update-time is of the essence. *Neuron*, 30(2) :319–333, 2001.

[92] J. M. Fuster. *The Prefrontal Cortex*. New York, Raven Press, 2 edition, 2008.

[93] W. J. Gehring and R. T. Knight. Prefrontal–cingulate interactions in action monitoring. *Nature neuroscience*, 3(5) :516–520, 2000.

[94] J. Glascher, A. N. Hampton, and J. P. O'Doherty. Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. *Cerebral Cortex*, 19(2) :483–495, Feb. 2009.

[95] J. Gläscher, R. Adolphs, H. Damasio, A. Bechara, D. Rudrauf, M. Calamia, L. K. Paul, and D. Tranel. Lesion mapping of cognitive control and value-based decision making in the prefrontal cortex. *Proceedings of the National Academy of Sciences*, 109(36) :14681–14686, 2012.

[96] O. Godefroy. Frontal syndrome and disorders of executive functions. *Journal of Neurology*, 250(1) :1–6, Jan. 2003.

[97] O. Godefroy, M. Cabaret, V. Petit-Chenal, J.-P. Pruvo, and M. Rousseaux. Control functions of the frontal lobes. modularity of the central-supervisory system ? *Cortex*, 35 :1–20, 1999.

[98] P. S. Goldman and H. E. Rosvold. Localization of function within the dorsolateral prefrontal cortex of the rhesus monkey. *Experimental neurology*, 27(2) :291–304, 1970.

[99] P. Goldman-Rakic. Development of cortical circuitry and cognitive function. *Child development*, 58(3) :601–622, 1987.

[100] P. Goldman-Rakic, A. Puve, J. Krystal, J. Gore, A. Belger, and G. McCarthy. Dissociation of mnemonic and perceptual processes during spatial and nonspatial working memory using fMRI. *Hum Brain Mapp*, 6(1) :14–32, 1998.

[101] P. T. Golumbek, A. J. Lazenby, H. I. Levitsky, L. M. Jaffee, H. Karasuyama, M. Baker, and D. M. Pardoll. Treatment of established renal cancer by tumor cells engineered to secrete interleukin-4. *Science*, 254(5032) :713–716, 1991.

[102] A. N. Hampton. The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *Journal of Neuroscience*, 26(32) :8360–8367, Aug. 2006.

[103] A. N. Hampton. The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *Journal of Neuroscience*, 26(32) :8360–8367, Aug. 2006.

[104] H. Harlow. The formation of learning sets. *Psychological Review*, 56(1) :51–65, 1949.

[105] M. Haruno. A neural correlate of reward-based behavioral learning in caudate nucleus : A functional magnetic resonance imaging study of a stochastic decision task. *Journal of Neuroscience*, 24(7) :1660–1665, Feb. 2004.

[106] J. Holdstock, A. Mayes, Q. Gong, N. Roberts, and N. Kapur. Item recognition is less impaired than recall and associative recognition in a patient with selective hippocampal damage. *Hippocampus*, 15(2) :203–215, 2005.

[107] C. B. Holroyd, J. Dien, and M. G. Coles. Error-related scalp potentials elicited by hand and foot movements : evidence for an output-independent error-processing system in humans. *Neuroscience letters*, 242(2) :65–68, 1998.

[108] C. B. Holroyd, S. Nieuwenhuis, N. Yeung, L. Nystrom, R. B. Mars, M. G. H. Coles, and J. D. Cohen. Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. *Nature Neuroscience*, 7(5) :497–498, May 2004.

[109] B. Hommel, J. Gehrk, and L. Knuf. Hierarchical coding in the perception and memory of spatial layouts. *Psychological Research*, 60 :1–10, 2000.

[110] A. Hyafil, C. Summerfield, and E. Koechlin. Two mechanisms for task switching in the prefrontal cortex. *Journal of Neuroscience*, 29(16) :5135–5142, Apr. 2009.

[111] J. M. Hyman, M. E. Hasselmo, and J. K. Seamans. What is the functional relevance of prefrontal cortex entrainment to hippocampal theta rhythms ? *Frontiers in Neuroscience*, 5, 2011.

[112] J. M. Hyman, L. Ma, E. Balaguer-Ballester, D. Durstewitz, and J. K. Seamans. Contextual encoding by ensembles of medial prefrontal cortex neurons. *Proceedings of the National Academy of Sciences*, 109(13) :5086–5091, Mar. 2012.

[113] J. M. Hyman, E. A. Zilli, A. M. Paley, and M. E. Hasselmo. Medial prefrontal cortex cells show dynamic modulation with the hippocampal theta rhythm dependent on behavior. *Hippocampus*, 15(6) :739–749, 2005.

[114] H. Imamizu, N. Sugimoto, R. Osu, K. Tsutsui, K. Sugiyama, Y. Wada, and M. Kawato. Explicit contextual information selectively contributes to predictive switching of internal models. *Experimental Brain Research*, 181(3) :395–408, July 2007.

[115] A. Incisa della Rocchetta and B. Milner. Strategic search and retrieval inhibition : the role of the frontal lobes. *Neuropsychologia*, 31(6) :503–524, 1993.

[116] M. Isoda and O. Hikosaka. Switching from automatic to controlled action by monkey medial frontal cortex. *Nature Neuroscience*, 10(2) :240–248, Feb. 2007.

[117] J. Janowsky, A. P. Shimamura, and L. R. Squire. Source memory impairment in patients with frontal lobe lesions. *Neuropsychologia*, 27(8) :1043–1056, 1989.

[118] A. Jeneson, C. B. Kirwan, R. O. Hopkins, J. T. Wixted, and L. R. Squire. Recognition memory and the hippocampus : A test of the hippocampal contribution to recollection and familiarity. *Learning & Memory*, 17(1) :63–70, Feb. 2010.

[119] Jerison. *Evolution of the Brain and Intelligence*. New York : Academic Press. Academic Press, 2 edition, 1994.

[120] A. Johnson, M. A. van der Meer, and A. D. Redish. Integrating hippocampus and striatum in decision-making. *Current Opinion in Neurobiology*, 17(6) :692–697, Dec. 2007.

[121] Johnson-Laird, P. Legrenzi, and M. Sonino Legrenzi. Reasoning and a sense of reality. *British Journal of Psychology*, 63 :3, 1972.

[122] K. Johnston, H. M. Levin, M. J. Koval, and S. Everling. Top-down control-signal dynamics in anterior cingulate and prefrontal cortex neurons following task switching. *Neuron*, 53(3) :453–462, Feb. 2007.

[123] J. Jonides, E. E. Smith, R. A. Koeppe, E. Awh, S. Minoshima, and M. A. Mintun. Spatial working memory in humans as revealed by PET. *Nature*, 363(6430) :623–625, 1993.

[124] M. Just and P. Carpenter. A capacity theory of comprehension : Individual differences in working memory. *Psychological Review*, 99(1) :122–149, 1992.

[125] J. W. Kable and P. W. Glimcher. The neural correlates of subjective value during intertemporal choice. *Nature Neuroscience*, 10(12) :1625–1633, Dec. 2007.

[126] D. Kahneman and A. Tversky. Choices, values, and frames. *American psychologist*, 39(4) :341, 1984.

[127] M. P. Karlsson, D. G. R. Tervo, and A. Y. Karpova. Network resets in medial prefrontal cortex mark the onset of behavioral uncertainty. *Science*, 338(6103) :135–139, Oct. 2012.

[128] S. W. Kennerley, M. E. Walton, T. E. J. Behrens, M. J. Buckley, and M. F. S. Rushworth. Optimal decision making and the anterior cingulate cortex. *Nature Neuroscience*, 9(7) :940–947, July 2006.

[129] J. G. Kerns. Anterior cingulate conflict monitoring and adjustments in control. *Science*, 303(5660) :1023–1026, Feb. 2004.

[130] K. A. Kiehl, P. F. Liddle, and J. B. Hopfinger. Error processing and the rostral anterior cingulate : An event-related fMRI study. *Psychophysiology*, 37(2) :216–223, 2000.

[131] C. Koch and I. Segev. The role of single neurons in information processing. *nature neuroscience*, 3 :1171–1177, 2000.

[132] E. Koechlin. The architecture of cognitive control in the human prefrontal cortex. *Science*, 302(5648) :1181–1185, Nov. 2003.

[133] E. Koechlin. Frontal pole function : what is specifically human ? *Trends in Cognitive Sciences*, 15(6) :241, June 2011.

[134] E. Koechlin. An evolutionary computational theory of prefrontal executive function in decision-making. *Philosophical Transactions of the Royal Society B : Biological Sciences*, 2014.

[135] E. Koechlin, J. Anton, and Y. Burnod. Bayesian inference in populations of cortical neurons : a model of motion integration and segmentation in area MT. *Biol Cybern*, 80(1) :25–44, 1999.

[136] E. Koechlin, G. Basso, P. Pietrini, S. Panzer, and J. Grafman. The role of the anterior prefrontal cortex in human cognition. *Development*, 1 :157–167, 1991.

[137] E. Koechlin, G. Basso, P. Pietrini, S. Panzer, and J. Grafman. The role of the anterior prefrontal cortex in human cognition. *Nature*, 399 :148–151, 1999.

[138] E. Koechlin, G. Corrado, P. Pietrini, and J. Grafman. Dissociating the role of the medial and lateral anterior prefrontal cortex in human planning. *Proceedings of the National Academy of Sciences*, 97(13) :7651–7656, 2000.

[139] E. Koechlin and A. Hyafil. Anterior prefrontal function and the limits of human decision-making. *Science*, 318(5850) :594–598, Oct. 2007.

[140] E. Koechlin and C. Summerfield. An information theoretical approach to prefrontal executive function. *Trends in Cognitive Sciences*, 11(6) :229–235, June 2007.

[141] N. Kolling, T. E. J. Behrens, R. B. Mars, and M. F. S. Rushworth. Neural mechanisms of foraging. *Science*, 336(6077) :95–98, Apr. 2012.

[142] K. S. LaBar, D. R. Gitelman, T. B. Parrish, M. Mesulam, and others. Neuroanatomic overlap of working memory and spatial attention networks : a functional MRI comparison within subjects. *Neuroimage*, 10(6) :695–704, 1999.

[143] P. Lauterbur. Image formation by induced local interactions : Examples employing nuclear magnetic resonance, 1973.

[144] F. Lhermitte. Human autonomy and the frontal lobes. part II : patient behavior in complex and social situations : the "environmental dependency syndrome". *Annals of neurology*, 19(4) :335–343, 1986.

[145] F. Lhermitte, B. Pillon, and M. Serdaru. Human autonomy and the frontal lobes. part i : Imitation and utilization behavior : a neuropsychological study of 75 patients. *Annals of neurology*, 19(4) :326–334, 1986.

[146] W. J. Ma, J. M. Beck, P. E. Latham, and A. Pouget. Bayesian inference with probabilistic population codes. *Nature Neuroscience*, 9(11) :1432–1438, Nov. 2006.

[147] A. W. MacDonald. Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, 288(5472) :1835–1838, June 2000.

[148] J. R. Manns, R. O. Hopkins, J. M. Reed, E. G. Kitchener, and L. R. Squire. Recognition memory and the human hippocampus. *Neuron*, 37(1) :171–180, 2003.

[149] D. Marr. Simple memory : A theory for archicortex. *Philos Trans R Soc Lond B Biol Sci.*, 262(841) :23–81, 1971.

[150] R. B. Mars, N. J. Shea, N. Kolling, and M. F. S. Rushworth. Model-based analyses : Promises, pitfalls, and example applications to the study of cognitive control. *The Quarterly Journal of Experimental Psychology*, 65(2) :252–267, Feb. 2012.

[151] M. Matsumoto, K. Matsumoto, H. Abe, and K. Tanaka. Medial prefrontal cell activity signaling prediction errors of action values. *Nature Neuroscience*, 10(5) :647–656, May 2007.

[152] A. Mayes, J. Holdstock, C. Isaac, D. Montaldi, J. Grigor, A. Gummer, P. Cariga, J. Downes, D. Tsivilis, D. Gaffan, Q. Gong, and K. Norman. Associative recognition in a patient with selective hippocampal lesions and relatively normal item recognition. *Hippocampus*, 14(6) :763–784, 2004.

[153] U. Mayr and R. Kliegl. Task-set switching and long-term memory retrieval. *Journal of Experimental Psychology : Learning, Memory, and Cognition*, 26(5) :1124–1140, 2000.

[154] J. L. McClelland, B. L. McNaughton, and R. C. O'Reilly. 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, 1995.

[155] S. M. McClure, G. S. Berns, and P. R. Montague. Temporal prediction errors in a passive learning task activate human striatum. *Neuron*, 38(2) :339–346, 2003.

[156] A. N. McCoy and M. L. Platt. Expectations and outcomes : decision-making in the primate brain. *Journal of Comparative Physiology A*, 191(3) :201–211, Mar. 2005.

[157] N. Meiran, Z. Chorev, and A. Sapir. Component processes in task switching. *Cognitive Psychology*, 41(3) :211–253, Nov. 2000.

[158] M.-M. Mesulam. From sensation to cognition. *Brain*, 121(6) :1013–1052, 1998.

[159] E. K. Miller and J. D. Cohen. An integrative theory of prefrontal cortex function. *Annual review of neuroscience*, 24(1) :167–202, 2001.

[160] B. Milner, M. Petrides, and M. Smith. Frontal lobes and the temporal organization of memory. *Hum Neurobiol*, 4(3) :137–142, 1985.

[161] Y. Miyashita. Cognitive memory : Cellular and network machineries and their top-down control. *Science*, 306(5695) :435–440, Oct. 2004.

[162] S. Monsell. Task switching. *Trends in Cognitive Sciences*, 7(3) :134–140, Mar. 2003.

[163] P. Morgane, J. Galler, and D. Mokler. A review of systems and networks of the limbic forebrain/limbic midbrain. *Progress in Neurobiology*, 75(2) :143–160, Feb. 2005.

[164] M. Moscovitch. Memory and working with memory : Evaluation of a component process model and comparisons with other models. *Memory systems*, 1994 :369–394, 1994.

[165] S. Nieuwenhuis. Sensitivity of electrophysiological activity from medial frontal cortex to utilitarian and performance feedback. *Cerebral Cortex*, 14(7) :741–747, Mar. 2004.

[166] D. Norman and T. Shallice. Attention to action, 1986.

[167] M. Oaksford and N. Chater. The probabilistic approach to human reasoning. *Trends in Cognitive Sciences*, 5(8) :349–357, 2001.

[168] K. Oberauer. Declarative and procedural working memory : Common principles, common capacity limits ? *Psychologica Belgica*, 50 :277–308, 2010.

[169] K. Oberauer, A. S. Souza, M. D. Druey, and M. Gade. Analogous mechanisms of selection and updating in declarative and procedural working memory : Experiments and a computational model. *Cognitive Psychology*, 66(2) :157–211, Mar. 2013.

[170] K. Oberauer, H.-M. Sü\s s, R. Schulze, O. Wilhelm, and W. W. Wittmann. Working memory capacity—facets of a cognitive ability construct. *Personality and Individual Differences*, 29(6) :1017–1045, 2000.

[171] J. P. O'Doherty, P. Dayan, K. Friston, H. Critchley, and R. J. Dolan. Temporal difference models and reward-related learning in the human brain. *Neuron*, 38(2) :329–337, 2003.

[172] J. P. O'Doherty, A. Hampton, and H. Kim. Model-based fMRI and its application to reward learning and decision making. *Annals of the New York Academy of Sciences*, 1104(1) :35–53, Apr. 2007.

[173] S. Ogawa, T. M. Lee, A. R. Kay, and D. W. Tank. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences*, 87(24) :9868–9872, 1990.

[174] A. Owen. The functional organization of working memory processes within human lateral frontal cortex : the contribution of functional neuroimaging. *Eur J Neurosci*, 9(7) :1329–1339, 1997.

[175] A. M. Owen, J. J. Downes, B. J. Sahakian, C. E. Polkey, and T. W. Robbins. Planning and spatial working memory following frontal lobe lesions in man. *Neuropsychologia*, 28(10) :1021–1034, 1990.

[176] J. X. O'Reilly, S. Jbabdi, and T. E. J. Behrens. How can a bayesian approach inform neuroscience ? : How can a bayesian approach inform neuroscience ? *European Journal of Neuroscience*, 35(7) :1169–1179, Apr. 2012.

[177] R. Passingham. Functional specialization of the supplementary motor area in monkeys and humans. *Advance Neurology*, 70 :105–116, 1996.

[178] E. Paulesu, C. Frith, and R. Frackowiak. The neural correlates of the verbal component of working memory. *Nature*, 362 :342–345, 1993.

[179] I. P. Pavlov. *Conditioned reflexes*. Courier Dover Publications, 2003.

[180] M. Petrides. Frontal lobes and behaviour. *Curr Opin Neurobiol*, 4(2) :207–211, 1994.

[181] M. Petrides, B. Alivisatos, E. Meyer, and A. C. Evans. Functional activation of the human frontal cortex during the performance of verbal working

memory tasks. *Proceedings of the National Academy of Sciences*, 90(3) :878–882, 1993.

[182] M. Petrides and B. Milner. Deficits on subject-ordered tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia*, 20(3) :249–262, 1982.

[183] M. Petrides and D. N. Pandya. Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. *European Journal of Neuroscience*, 16(2) :291–310, 2002.

[184] N. Picard and P. L. Strick. Motor areas of the medial wall : a review of their location and functional activation. *Cerebral cortex*, 6(3) :342–353, 1996.

[185] R. Quilodran, M. Rothé, and E. Procyk. Behavioral shifts and action valuation in the anterior cingulate cortex. *Neuron*, 57(2) :314–325, Jan. 2008.

[186] C. Ranganath, M. K. Johnson, M. D'Esposito, and others. Left anterior prefrontal activation increases with demands to recall specific perceptual information. *Journal of Neuroscience*, 20(22) :1–5, 2000.

[187] A. Rangel and T. Hare. Neural computations associated with goal-directed choice. *Current Opinion in Neurobiology*, 20(2) :262–270, Apr. 2010.

[188] K. R. Ridderinkhof. The role of the medial frontal cortex in cognitive control. *Science*, 306(5695) :443–447, Oct. 2004.

[189] L. Rips. *The Psychology of Proof : Deductive Reasoning in Human Thinking*. A Bradford Book, 1994.

[190] E. T. Rolls. On the brain and emotion. *Behavioral and brain sciences*, 23(02) :219–228, 2000.

[191] M. Roy, D. Shohamy, and T. D. Wager. Ventromedial prefrontal-subcortical systems and the generation of affective meaning. *Trends in Cognitive Sciences*, 16(3) :147–156, Mar. 2012.

[192] O. Rubin and N. Meiran. On the origins of the task mixing cost in the cuing task-switching paradigm. *Journal of Experimental Psychology : Learning, Memory, and Cognition*, 31(6) :1477–1491, 2005.

[193] J. Rubinstein, D. Meyer, and j. Evans. Executive control of cognitive processes in task switching. *Psychological Review*, 2001.

[194] M. D. Rugg, P. C. Fletcher, P. M. Chua, and R. J. Dolan. The role of the prefrontal cortex in recognition memory and memory for source : An fMRI study. *Neuroimage*, 10(5) :520–529, 1999.

[195] M. D. Rugg, K. L. Vilberg, J. T. Mattson, S. S. Yu, J. D. Johnson, and M. Suzuki. Item memory, context memory and the hippocampus : fMRI evidence. *Neuropsychologia*, 50(13) :3070–3079, Nov. 2012.

[196] M. D. Rugg and E. L. Wilding. Retrieval processing and episodic memory. *Trends in cognitive sciences*, 4(3) :108–115, 2000.

[197] M. Rushworth, M. Walton, S. Kennerley, and D. Bannerman. Action sets and decisions in the medial frontal cortex. *Trends in Cognitive Sciences*, 8(9) :410–417, Sept. 2004.

[198] M. Rushworth, M. Walton, S. Kennerley, and D. Bannerman. Action sets and decisions in the medial frontal cortex. *Trends in Cognitive Sciences*, 8(9) :410–417, Sept. 2004.

[199] M. F. Rushworth, M. J. Buckley, T. E. Behrens, M. E. Walton, and D. M. Bannerman. Functional organization of the medial frontal cortex. *Current Opinion in Neurobiology*, 17(2) :220–227, Apr. 2007.

[200] M. F. S. Rushworth and T. E. J. Behrens. Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature Neuroscience*, 11(4) :389–397, Apr. 2008.

[201] W. Röntgen. *Über eine neue Art von Strahlen.* Wu¨rzburg, Stahel, 1895.

[202] K. Sakai. Task set and prefrontal cortex. *Annual Review of Neuroscience*, 31(1) :219–245, July 2008.

[203] J. Sallet, R. Quilodran, M. Rothé, J. Vezoli, J.-P. Joseph, and E. Procyk. Expectations, gains, and losses in the anterior cingulate cortex. *Cogn Affect Behav Neurosci*, 7(4) :327–336, 2008.

[204] A. Sanborn, T. Griffiths, and D. Navarro. Rational approximations to rational models : Alternative algorithms for category learning. *Psychological Review*, 117(4) :1144–1167, 2010.

[205] M. Scheffers, M. Coles, P. Bernstein, W. Gehring, and E. Donchin. Event-related brain potentials and error-related processing : an analysis of incorrect responses to go and no-go stimuli. *Psychophysiology*, 33(1) :42–53, 1996.

[206] D. Schiller and M. R. Delgado. Overlapping neural systems mediating extinction, reversal and regulation of fear. *Trends in Cognitive Sciences*, 14(6) :268–276, June 2010.

[207] W. Schultz. A neural substrate of prediction and reward. *Science*, 275(5306) :1593–1599, Mar. 1997.

[208] W. B. Scoville and B. Milner. Loss of recent memory after bilateral hippocampal lesions. *J. Neurol. Neurosurg. Psychiat.*, 20(11) :11–21, 1957.

[209] L. D. Selemon and P. S. Goldman-Rakic. Common cortical and subcortical targets of the dorsolateral prefrontal and posterior parietal cortices in the rhesus monkey : evidence for a distributed neural network subserving spatially guided behavior. *The Journal of Neuroscience*, 8(11) :4049–4068, 1988.

[210] K. Semendeferi, E. Armstrong, A. Schleicher, K. Zilles, G. W. Van Hoesen, and others. Prefrontal cortex in humans and apes : a comparative study of area 10. *American Journal of Physical Anthropology*, 114(3) :224–241, 2001.

[211] K. Semendeferi, H. Damasio, R. Frank, and G. W. Van Hoesen. The evolution of the frontal lobes : a volumetric analysis based on three-dimensional reconstructions of magnetic resonance scans of human and ape brains. *Journal of Human Evolution*, 32(4) :375–388, 1997.

[212] K. Semendeferi, A. Lu, N. Schenker, and H. Damasio. Humans and great apes share a large frontal cortex. *Nature Neuroscience*, 5(3) :272–276, Mar. 2002.

[213] A. P. Shimamura. The role of the prefrontal cortex in dynamic filtering. *Psychobiology*, 28(2) :207–218, 2000.

[214] A. G. Siapas, E. V. Lubenov, and M. A. Wilson. Prefrontal phase locking to hippocampal theta oscillations. *Neuron*, 46(1) :141–151, Apr. 2005.

[215] J. S. Simons and H. J. Spiers. Prefrontal and medial temporal lobe interactions in long-term memory. *Nature Reviews Neuroscience*, 4(8) :637–648, Aug. 2003.

[216] E. E. Smith. Storage and executive processes in the frontal lobes. *Science*, 283(5408) :1657–1661, Mar. 1999.

[217] E. E. Smith and J. Jonides. Neuroimaging analyses of human working memory. *Proceedings of the National Academy of Sciences*, 95(20) :12061–12068, 1998.

[218] E. E. Smith, J. Jonides, and R. A. Koeppe. Dissociating verbal and spatial working memory using PET. *Cerebral Cortex*, 6(1) :11–20, 1996.

[219] E. Sowell, P. Thompson, C. Holmes, T. Jernigan, and A. Toga. In vivo evidence for post-adolescent brain maturation in frontal and striatal regions. *Nature Neuroscience*, 2(10) :859–861, 1999.

[220] L. R. Squire, J. T. Wixted, and R. E. Clark. Recognition memory and the medial temporal lobe : a new perspective. *Nature Reviews Neuroscience*, 8(11) :872–883, Nov. 2007.

[221] L. R. Squire and S. Zola-Morgan. Memory : brain systems and behavior. *TINS*, 11(4) :170–175, 1988.

[222] K. E. Stanovich and R. F. West. Advancing the rationality debate. *Behavioral and brain sciences*, 23(05) :701–717, 2000.

[223] G. Stanton, C. Bruce, and M. Goldberg. Topography of projections to posterior cortical areas from the macaque frontal eye fields. *J Comp Neurol*, 353(2) :291–305, 1995.

[224] E. Stein. *Without Good Reason : The Rationality Debate in Philosophy and Cognitive Science*. Oxford University Press, 1996.

[225] S. P. Stich. Could man be an irrational animal ? *Synthese*, 64(1) :115–135, 1985.

[226] A. Stocco, C. Labiere, and J. Anderson. Conditional routing of information to the cortex : A model of the basal ganglia's role in cognitive coordination. *Psychological Review*, 117(2) :541–574, 2010.

[227] R. Sutton and A. Barto. *Reinforcement Learning : An Introduction*. MIT Press, 1998.

[228] D. Swick and U. Turken. Dissociation between conflict detection and error monitoring in the human anterior cingulate cortex. *Proceedings of the National Academy of Sciences*, 99(25) :16354–16359, 2002.

[229] Y. W. Teh, M. Jordan, M. Beal, and D. Blei. Hierarchical dirichlet processes, 2006.

[230] J. N. Towse, G. J. Hitch, and U. Hutton. On the interpretation of working memory span in adults. *Memory & Cognition*, 28(3) :341–348, 2000.

[231] A. Treisman and S. Gormican. Feature analysis in early vision : Evdence from search asymmetries. *Psychological Review*, 95(1) :15–48, 1988.

[232] E. Tulving. Episodic memory : from mind to brain. *Annual review of psychology*, 53(1) :1–25, 2002.

[233] M. Turnock and S. Becker. A neural network model of hippocampal–striatal–prefrontal interactions in contextual conditioning. *Brain Research*, 1202 :87–98, Apr. 2008.

[234] L. Ungerleider and J. Haxby. 'what' and 'where' in the human brain. *Curr Opin Neurobiol*, 4(2) :157–165, 1994.

[235] H. B. Uylings, H. J. Groenewegen, and B. Kolb. Do rats have a prefrontal cortex ? *Behavioural Brain Research*, 146(1-2) :3–17, Nov. 2003.

[236] H. T. van Schie, R. B. Mars, M. G. H. Coles, and H. Bekkering. Modulation of activity in medial frontal and motor cortices during error observation. *Nature Neuroscience*, 7(5) :549–554, May 2004.

[237] A. Vandierendonck, M. Deschuyteneer, A. Depoorter, and D. Drieghe. Input monitoring and response selection as components of executive control in pro-saccades and anti-saccades. *Psychological Research*, 72 :1–11, 2008.

[238] E. Volle, R. Beato, R. Levy, and B. Dubois. Forced collectionism after orbitofrontal damage. *Neurology*, 58 :488–490, 2002.

[239] A. R. WAGNER. ROBERT a. RESCORLA ALLAN r. WAGNER. In *Classical conditioning : a symposium*, page 64. New York, Appleton-Century-Crofts, 1965.

[240] M. E. Walton, P. L. Croxson, T. E. Behrens, S. W. Kennerley, and M. F. Rushworth. Adaptive decision making and value in the anterior cingulate cortex. *NeuroImage*, 36 :T142–T154, Jan. 2007.

[241] M. E. Walton, J. T. Devlin, and M. F. S. Rushworth. Interactions between decision making and performance monitoring within prefrontal cortex. *Nature Neuroscience*, 7(11) :1259–1265, 2004.

[242] F. Waszak, B. Hommel, and A. Allport. Task-switching and long-term priming : Role of episodic stimulus–task bindings in task-shift costs. *Cognitive Psychology*, 46(4) :361–413, June 2003.

[243] J. T. Wixted, L. Mickes, and L. R. Squire. Measuring recollection and familiarity in the medial temporal lobe. *Hippocampus*, 20(11) :1195–1205, Nov. 2010.

[244] K. Wunderlich, A. Rangel, and J. P. O'Doherty. Neural computations underlying action-based decision making in the human brain. *Proceedings of the National Academy of Sciences*, 106(40) :17199–17204, Oct. 2009.

[245] Wylie and Allport. Task switching and the measurement of "switch costs". *Psychol Res*, pages 212–33, 2000.

[246] E. Yehene, N. Meiran, and N. Soroker. Basal ganglia play a unique role in task switching within the frontal-subcortical circuits : evidence from patients with focal lesions. *Journal of Cognitive Neuroscience*, 20(6) :1079–1093, 2008.

[247] A. P. Yonelinas. Separating the brain regions involved in recollection and familiarity in recognition memory. *Journal of Neuroscience*, 25(11) :3002–3008, Mar. 2005.

[248] A. P. Yonelinas and B. Levy. Dissociating familiarity from recollection in human recognition memory : different rates of forgetting over short retention intervals. *Psychon Bull Rev*, 9(3) :575–582, 2002.

[249] A. J. Yu and P. Dayan. Uncertainty, neuromodulation, and attention. *Neuron*, 46(4) :681–692, May 2005.

[250] D. Öngür, A. T. Ferry, and J. L. Price. Architectonic subdivision of the human orbital and medial prefrontal cortex. *The Journal of Comparative Neurology*, 460(3) :425–449, June 2003.