

**UNIVERSITÉ DU QUÉBEC À MONTRÉAL**

**LES NEUROSCIENCES ET L'ÉCONOMIE : PERSPECTIVES DE LA  
NEUROÉCONOMIE**

**MÉMOIRE**

**PRÉSENTÉ**

**COMME EXIGENCE PARTIELLE**

**DE LA MAITRISE EN ÉCONOMIQUE**

**PAR**

**B. Y. MARTIN GAUTHIER**

**MAI 2006**

# UNIVERSITÉ DU QUÉBEC À MONTRÉAL

Service des bibliothèques

## Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 -Rév.01-2006). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

## AVANT-PROPOS

J'ai cherché par ce travail à présenter au lecteur quelques-unes des principales recherches rendues possibles par la fusion de l'économie, de la psychologie et des neurosciences. N'ayant pas de formation particulière en psychologie ou en neurosciences, j'ai donc dû, tout au long de mes démarches, expliquer ce qu'est la neuroéconomie et ce qu'elle cherche à faire, de la façon la plus simple possible, en indiquant au lecteur qui désirerait en savoir davantage, où trouver les articles et les publications plus poussés. Comme la neuroéconomie est une discipline émergente qui évolue plutôt rapidement, j'ai parfois été à même de constater certains changements de positions chez des chercheurs, et ceci a représenté un défi additionnel pour mon travail. Lorsque j'ai commencé à m'intéresser au sujet vers l'an 2002, il n'existait aucune définition de la neuroéconomie et très peu de liens Internet abordant ce sujet. L'existence même de la neuroéconomie n'était pas encore évidente. Aujourd'hui (décembre 2005), plus de 136 000 liens apparaissent sur le moteur de recherche Google lorsque l'on tape le terme « neuroeconomics ». En langue française par contre, le terme « neuroéconomique » n'engendre encore qu'une quinzaine de liens. Cette situation s'explique sans doute par le fait que presque toutes les études sont publiées en langue anglaise et que la plupart des centres de recherches sont situés aux États-Unis.

Je profite ici de l'occasion qui m'est donnée pour remercier sincèrement mon directeur de mémoire, Monsieur Robert Leonard, qui a su, tout au long de cette recherche, m'encourager à donner le meilleur de moi-même. Son enthousiasme et ses judicieux conseils m'ont permis d'améliorer le contenu et la forme de cette recherche, et sans son ouverture d'esprit et ses encouragements, ce travail n'aurait jamais pu voir le jour. Je remercie également Martin Renaud qui a eu la gentillesse de réviser le présent document afin de veiller à la qualité de la langue française.

## TABLE DES MATIÈRES

<b>AVANT-PROPOS</b>	<b>ii</b>
<b>LISTE DES FIGURES</b>	<b>vii</b>
<b>LISTE DES TABLEAUX</b>	<b>x</b>
<b>LISTE DES ABRÉVIATIONS</b>	<b>xi</b>
<b>RÉSUMÉ</b>	<b>xii</b>
<b>INTRODUCTION</b>	<b>1</b>
<b>Chapitre I : INTRODUCTION À LA NEUROÉCONOMIE : PROCESSUS SOUS-JACENTS À LA DÉCISION ÉCONOMIQUE...</b>	<b>3</b>
<b>1.0 Introduction .....</b>	<b>3</b>
<b>1.1 L'erreur de Jevons .....</b>	<b>4</b>
<b>1.1.1 Évolution rapide de la neuroéconomie .....</b>	<b>6</b>
<b>1.2 Introduction aux neurosciences .....</b>	<b>9</b>
<b>1.2.1 Le cerveau : éléments de base .....</b>	<b>10</b>
<b>1.2.2 Les outils des neurosciences .....</b>	<b>12</b>

<b>1.3</b>	<b>Processus mentaux : contrôlés et automatiques .....</b>	<b>16</b>
1.3.1	Validations empiriques .....	16
1.3.2	Corollaires physiologiques .....	20
1.3.3	Processus automatiques : quelques implications .....	23
<b>1.4</b>	<b>Approche bidimensionnelle du fonctionnement cérébral ...</b>	<b>29</b>
1.4.1	Interactions entre les systèmes .....	36
1.4.2	Exemple d'interactions à l'intérieur du cadre bidimensionnel ...	40
<b>1.5</b>	<b>Conclusion .....</b>	<b>42</b>
 <b>Chapitre II : LES PRÉFÉRENCES ET LA NEUROÉCONOMIE : SUPPORT ET APPORT À LA THÉORIE ÉCONOMIQUE .....</b>		<b>44</b>
<b>2.0</b>	<b>Introduction .....</b>	<b>44</b>
<b>2.1</b>	<b>Approche holistique de la neuroéconomie .....</b>	<b>45</b>
<b>2.2</b>	<b>Les préférences et l'utilité subjective .....</b>	<b>50</b>
2.2.1	Évaluation et sélection d'objectifs : vers une computation neuronale du concept standard de l'utilité	52
2.2.2	L'utilité et la dopamine .....	62

2.2.3	Les <i>cues</i> environnementaux et la consommation .....	65
2.2.4	Préférences révélées et révision du concept de l'utilité .....	72
2.2.5	Au-delà des préférences révélées .....	78
2.2.6	Homéostasie et <i>alliesthesia</i> .....	80
2.2.7	Utilité directe de la monnaie et de sa source .....	83
2.3	Estimation des probabilités et révision du concept de l'espérance de l'utilité .....	84
2.3.1	Le paradoxe d'Allais .....	85
2.3.2	Le paradoxe d'Ellsberg .....	88
2.3.3	Éléments de la théorie prospective, heuristique et effet de cadrage	91
2.4	Conclusion .....	104
Chapitre III : THÉORIE DES JEUX ET NEUROÉCONOMIE .....		107
3.0	Introduction .....	107
3.1	Théorie des jeux, rationalité et neurosciences .....	108
3.2	Théorie comportementale des jeux et neurosciences .....	117
3.2.1	Stratégie, échange et confiance .....	118

3.2.2	Le jeu de la confiance .....	121
3.2.3	Le dilemme du prisonnier .....	127
3.2.4	Le jeu de l'ultimatum .....	132
3.3	Réciprocité forte et considérations évolutionnistes .....	137
3.3.1	Psychologie évolutionniste et modularité .....	140
3.4	Conclusion .....	144
	CONCLUSION .....	145
	RÉFÉRENCES.....	149

## LISTE DES FIGURES

Figure 1.1	Les lobes du cortex	10
Figure 1.2	Les trois cerveaux de l'humain en coupe transversale	11
Figure 1.3	La barre métallique ayant blessé Phineas P. Gage	13
Figure 1.4	Imagerie à résonance magnétique fonctionnelle et <i>hyperscanning</i>	14
Figure 1.5	Illusion de Müller-Lyer	18
Figure 1.6	Quelques régions pertinentes à la neuroéconomie	22
Figure 1.7	Influence d'un signal subliminal sur la consommation de personnes assoiffées	33
Figure 2.1	Une fonction d'utilité conventionnelle	53
Figure 2.2	Réminiscence d'une fonction d'utilité dans le cerveau	53
Figure 2.3	L'encodage des probabilités de récompense	54
Figure 2.4	Activité neuronale et probabilité de récompense	56
Figure 2.5	Activité neuronale et ampleur du gain	57
Figure 2.6	Anticipation d'un gain monétaire dans le striatum	60
Figure 2.7	Impact neurologique d'un gain ou d'une perte monétaire	61

Figure 2.8	Le circuit de la récompense	65
Figure 2.9	L'encodage de l'erreur d'anticipation	67
Figure 2.10	Courbes d'indifférence avec préférences non-dépaquetées	80
Figure 2.11	Courbes d'indifférence avec préférences dépaquetées	80
Figure 2.12	Le paradoxe d'Ellsberg	89
Figure 2.13	Le cortex préfrontal ventromédial	90
Figure 2.14	La fonction de valeur de la théorie prospective	93
Figure 2.15	Les trois bouliers	101
Figure 2.16	L'hypothèse de comparaisons contrefactuelles sous investigation	103
Figure 3.1	Résultats du jeu de l'inspection et du resquillage	111
Figure 3.2	Comportement humain dans le jeu de l'inspection et du resquillage	112
Figure 3.3	Comportement du singe dans le jeu de l'inspection et du resquillage	113
Figure 3.4	Comportement du singe et activité neuronale	117
Figure 3.5	Jeu de la confiance et de la réciprocité	121
Figure 3.6	Le jeu de la confiance	124
Figure 3.7	Dilemme du prisonnier	128

Figure 3.8	Ronde typique du jeu de l'ultimatum et résultats comportementaux	135
Figure 3.9	Compétition entre raison et émotions pour le contrôle du comportement dans le jeu de l'ultimatum	136
Figure 3.10	Régions activées par une offre inéquitable au jeu de l'ultimatum	137

## LISTE DES TABLEAUX

Tableau		Page
1.1	Principaux centres de recherche de la neuroéconomie	6
1.2	Régions du cerveau associées aux deux systèmes mentaux	22
1.3	Approche bidimensionnelle du fonctionnement cérébral	31
2.1	Les trois niveaux de compréhension d'un dispositif de traitement de l'information selon l'approche de Marr	48
2.2	Les différentes formes d'utilité suggérées par Kahneman	74
3.1	La matrice des paiements du jeu de l'inspection et du resquillage	109
3.2	L'impact de l'ocytocine dans la décision d'investissement	125
3.3	Dilemme du prisonnier et activation dopaminergique dans certaines régions du cerveau	130
3.4	Le problème de Wason	141
3.5	Détection d'une tricherie	142

## LISTE DES ABRÉVIATIONS

BOLD Signal *Blood Oxygen Level Dependent Signal*

BSR *Brain Stimulation Reward* ; stimulus électrique.

EEG Électroencéphalogramme.

FMRI *Functional Magnetic Resonance Imagery* ; imagerie à résonance magnétique fonctionnelle.

TEP *Position Emission Topography* ; tomographie par émission de positrons.

## RÉSUMÉ

Dans la perspective de l'économie comportementale, la neuroéconomie peut informer la théorie économique en utilisant les connaissances des mécanismes du cerveau. Afin de mieux apprécier la contribution de la neuroéconomie, nous présentons de succincts éléments d'introduction à la méthodologie et aux outils des neurosciences. Puis, nous présentons une approche bidimensionnelle du fonctionnement cérébral, non sans avoir préalablement fait le point sur sa pertinence à l'aide de validations empiriques. Cette approche permettra des élaborations concernant l'importance de l'automatisme, des émotions et de l'appariement automatique dans le processus de décision économique, suggérant une moins grande place pour les processus contrôlés de délibérations attentives et un plus grand espace pour les autres processus. Il sera ensuite démontré comment des thèmes comme les préférences, la prise de décision en situation de risque et d'incertitude, les préférences sociales et la théorie comportementale des jeux, peuvent tous bénéficier des avancées de la neuroéconomie. L'enjeu ultime de la neuroéconomie est de développer une nouvelle théorie de la décision qui permettrait de réduire l'écart entre les approches normatives de la théorie économique néoclassique et les approches descriptives de l'économie comportementale. En reliant les causes proximales et ultimes du comportement, elle pourrait ainsi réduire cet écart. L'approche neuroéconomique, pluridisciplinaire, informe tout à la fois les chercheurs des neurosciences, de la psychologie et de l'économie. Pour les chercheurs des neurosciences, elle offre de nouveaux outils conceptuels leur permettant d'étudier le cerveau à l'intérieur d'un autre paradigme; pour les psychologues et les économistes comportementaux, elle est une suite logique de leurs observations empiriques, puisque le comportement observé trouve nécessairement son origine à partir du cerveau et non à partir d'abstractions mathématiques. Enfin, puisque la neuroéconomie utilise le cerveau comme outil d'analyse, elle incorpore inévitablement des considérations évolutionnistes qui pourraient permettre de mieux comprendre pourquoi le comportement s'éloigne parfois de la rationalité substantive, et pourquoi l'être humain semble si souvent tiraillé entre l'ordre interne de l'esprit, l'ordre externe des interrelations sociales et l'ordre étendu des marchés.

Mots clés : neuroéconomie, économie comportementale, décision économique, préférences, théorie de la décision.

## INTRODUCTION

Depuis le début des années 2000, des chercheurs issus de disciplines aussi diverses que l'économie, la psychologie, la biologie et les neurosciences, ont décidé de se regrouper pour travailler ensemble, de telle sorte qu'ils ont initié l'émergence d'une nouvelle science : la neuroéconomie. En reliant la décision à ses substrats neuronaux, la neuroéconomie permet de mieux comprendre le fonctionnement du cerveau et les fondements des processus décisionnels. Ce faisant, la neuroéconomie peut informer toute science du comportement. Dans ce mémoire, nous présentons les recherches pionnières de cette discipline émergente, en soulignant au passage leur pertinence, leur validation ou leur falsification. Au chapitre 1, nous expliquons certains des principaux outils méthodologiques des neurosciences, ainsi que quelques-unes des plus importantes caractéristiques du fonctionnement de notre cerveau; nous montrons qu'à la lumière des connaissances actuelles, la décision économique ne peut être analysée sans tenir compte du rôle joué par les processus automatiques et affectifs. Pour ce faire, nous présentons une approche bidimensionnelle du fonctionnement cérébral, telle que le suggèrent les économistes comportementaux.

Au chapitre 2, nous cherchons à montrer les raisons qui ont poussé certains chercheurs des neurosciences à s'intéresser aux théories économiques. Utilisant les notions d'utilité subjective et d'espérance de l'utilité, nous montrons comment ces outils conceptuels leur ont permis de mieux comprendre le fonctionnement du cerveau. Nous soulignons par la suite les similitudes entre le concept de l'utilité subjective et le rôle que joueraient les neurones dopaminergiques dans la décision. Puis, en montrant comment le cerveau parvient à formuler l'anticipation d'une récompense, nous abordons la question de la tentation et revoyons le rôle des processus automatiques dans la

décision économique. Nous montrons ensuite que ces découvertes ont déjà su capter l'attention de certains économistes qui les ont incorporées à leur modèle. Enfin, nous abordons différents thèmes ayant fait l'objet d'études des neurosciences, notamment la théorie des préférences révélées et la révision du concept de l'espérance de l'utilité. Nous cherchons à montrer en quoi la neuroéconomie permet de jeter un nouveau regard sur ces théories.

Au chapitre 3, nous nous intéressons aux recherches qui combinent les neurosciences et la théorie des jeux. Nous montrons comment des notions comme la rationalité de l'agent, la stratégie mixte ou l'équilibre de Nash peuvent servir d'outils conceptuels aux chercheurs issus des neurosciences, et comment ces notions sont d'ailleurs corroborées par certaines de leurs recherches sur le cerveau. Mais les neurosciences peuvent également corroborer d'autres approches, notamment la théorie comportementale des jeux. Nous présentons trois types de jeux qui ont été utilisés dans ce cadre et cherchons à voir si les observations rendues possibles par l'économie expérimentale n'auraient pas des bases biologiques. Nous terminons en montrant comment toutes ces recherches permettent d'incorporer des dimensions évolutionnistes aux interactions humaines.

## CHAPITRE I

### INTRODUCTION À LA NEUROÉCONOMIE : PROCESSUS SOUS-JACENTS À LA DÉCISION ÉCONOMIQUE

#### 1.0) Introduction

De récents développements en neurosciences et en psychologie permettent d'approfondir notre compréhension des mécanismes sous-jacents à la décision économique. Ce chapitre vise à présenter au lecteur certains de ces développements, pour ainsi mieux saisir leur portée et leur influence sur la compréhension des processus décisionnels. Nous cherchons d'abord à comprendre ce qu'est la neuroéconomie et où elle se situe par rapport à la théorie économique. Nous expliquons ensuite certains outils méthodologiques des neurosciences. Il sera alors question des différents mécanismes décisionnels, notamment ceux concernant le rôle des émotions et des processus automatiques. Puis, afin d'élargir la portée des modèles standards de délibération rationnelle et consciente, nous présentons une approche bidimensionnelle et simplifiée du fonctionnement cérébral. Cette approche, qui repose sur certaines évidences issues de la biologie et de la psychologie, sera expliquée à l'aide de validations empiriques fournies par de récentes recherches en neurosciences. Elle servira d'ancrage pour les autres chapitres du présent mémoire.

### 1.1) L'erreur de Jevons

La neuroéconomie est une branche de l'économie comportementale<sup>1</sup>. L'économie comportementale cherche à accroître le degré de réalisme des mécanismes psychologiques de l'analyse économique, sans pour autant rejeter d'emblée le paradigme néoclassique, qui fournit un cadre théorique réfutable, et qui est basé sur la maximisation de l'utilité, les équilibres et l'efficacité (Camerer et Loewenstein, 2002). Un degré plus élevé de réalisme ne nécessite pas nécessairement une scission radicale. Par exemple, au cœur de la théorie néoclassique, il n'est ni spécifiquement indiqué que le risque doit être soupesé de façon linéaire, ni que les valeurs futures doivent être escomptées exponentiellement à un taux constant. Par contre, certaines hypothèses, comme celle de la rationalité substantive, sont remplacées par des hypothèses jugées plus réalistes, afin de tenir compte des limites humaines en matière de capacités de calcul, de volonté et d'égoïsme. C'est dans ce contexte que la neuroéconomie est définie comme étant « l'utilisation de processus cérébraux permettant de trouver de nouvelles fondations pour les théories économiques » (Camerer, 2004)<sup>2</sup>. Dans cette perspective, la neuroéconomie est le prolongement de l'économie comportementale et elle permet avant tout d'accroître le degré de réalisme des théories économiques.

Les premiers économistes néoclassiques voyaient dans le comportement humain la résultante d'un processus décisionnel délibéré où les bénéfices et les coûts des actions étaient soupesés afin de maximiser l'utilité. Au tournant du vingtième siècle, de nombreux économistes espéraient faire de leur discipline une science exacte au même titre que la physique. À l'époque, trouver un secours ou une approche alternative dans la psychologie n'était pas

---

<sup>1</sup> Traduction de « Behavioral Economics ». Le terme économie comportementale est la traduction qui semble la plus appropriée.

<sup>2</sup> Page Web : What is neuroeconomics ? [http://www.hss.caltech.edu/~camerer/web\\_material/n.html](http://www.hss.caltech.edu/~camerer/web_material/n.html)

considéré comme une avenue prometteuse<sup>3</sup>. Malgré tout, l'approche néoclassique ne faisait pas l'unanimité pour autant. Par exemple, Viner remettait en question, dès 1925, le calcul hédonistique rationnel comme base des processus décisionnels :

« Human behaviour, in general, and presumably, therefore, also in the market place, is not under the constant and detailed guidance of careful and accurate hedonic calculations, but is the product of an unstable and unrational (sic) complex of reflex actions, impulses, instincts, habits, customs, fashions and hysteria... »<sup>4</sup>

Malgré ces questionnements, puisque les émotions ne pouvaient être directement quantifiables, les économistes néoclassiques n'empruntèrent jamais cette avenue. Et comme tout portait à croire que l'être humain n'aurait jamais les moyens d'observer directement les émotions ou les préférences, on en était tenu à se contenter, comme le spécifiait Jevons, d'observer les effets quantifiables de celles-ci :

« I hesitate to say that men will ever have the means of measuring directly the feelings of the human heart. It is from the quantitative effects of the feelings that we must estimate their comparative amounts »<sup>5</sup>.

C'est ainsi qu'à partir des années 1940, les concepts d'utilité ordinale et de préférences révélées, ainsi que leurs développements mathématiques ultérieurs, permirent d'une façon commode, d'éviter la complexe réalité sous-jacente à l'utilité. Mais aujourd'hui, d'importantes avancées dans les connaissances que nous avons des processus mentaux ainsi que certains récents développements des neurosciences nous permettent de comprendre que William Stanley Jevons s'était trompé. Certaines émotions peuvent

---

<sup>3</sup> Certains considèrent que la psychologie n'existait pas encore comme discipline scientifiquement reconnue (Camerer et Loewenstein, 2002). Si cette interprétation semble radicale, il n'en demeure pas moins qu'en matière de processus décisionnels, les plus récents développements illustrent à quel point la psychologie a depuis progressé.

<sup>4</sup> Viner, 1925, cité dans Camerer, Loewenstein et Prelec, 2004, p. 555.

<sup>5</sup> Jevons, 1871 cité dans Camerer *et al.*, 2005, p. 9.

effectivement être associées à des observations et à des mesures neurologiques lors de transaction de nature économique.

### 1.1.1) Évolution rapide de la neuroéconomie

La neuroéconomie, qui n'existait pas il y a à peine dix ans, est aujourd'hui une discipline qui évolue très rapidement. Cette émergence récente et a été rendue possible par la collaboration entre chercheurs qui, bien qu'issus de différentes disciplines, tentent tous à leur niveau respectif de mieux comprendre les fondements des processus décisionnels. Le tableau 1.1, qui présente quelques-uns des principaux<sup>6</sup> centres de recherche de la neuroéconomie, ainsi que certains chercheurs affiliés, illustre cette effervescence.

Tableau 1.1 : Principaux centres de recherche de la neuroéconomie

Centre de recherche	Endroit	Chercheurs associés au centre de recherche ou collaborateurs
Centre de recherche pour les sciences neuronales, laboratoire Glimcher de l'Université de New York.	Université de New York (New York)	Paul Glimcher, directeur du laboratoire, chercheur en neurosciences de l'Université de New York.  Michael Platt, chercheur en neurosciences de l'Université Duke (Caroline du Nord).  Michael Dorris, chercheur en neurosciences de l'Université Queen's (Ontario).  Aldo Rustichini, professeur d'économie à l'Université du Minnesota (Minnesota).

<sup>6</sup> Cette liste ne saurait être exhaustive et n'est pas établie selon un quelconque classement. Nous ne cherchons qu'à indiquer au lecteur ce qui nous semble être quelques-uns des principaux centres de recherche, ainsi que certains de leurs chercheurs affiliés. Pour une liste plus élargie, consulter le site [http://neuroeconlab.stanford.edu/necon\\_resources.htm](http://neuroeconlab.stanford.edu/necon_resources.htm)

<p>Laboratoire de neuroéconomie du California Institute of Technology.</p>	<p>California Institute of Technology, Pasadena (Californie)</p>	<p>Colin Camerer, directeur du laboratoire, professeur d'économie et de psychologie cognitive au California Institute of Technology.</p> <p>George Loewenstein, professeur d'économie à l'Université Carnegie Mellon (Pennsylvanie).</p> <p>Drazen Prelec, professeur en sciences de la gestion au MIT (Massachusetts).</p>
<p>Centre de recherche interdisciplinaire pour les sciences économiques.</p> <p>Centre d'étude en neuroéconomie de l'Université George Mason.</p>	<p>Université George Mason (Virginie)</p>	<p>Kevin McCabe, directeur du Centre d'étude de la neuroéconomie, professeur d'économie et de droit à l'Université George Mason, figure pionnière de la neuroéconomie.</p> <p>Vernon Smith, professeur d'économie à l'Université George Mason et figure de proue de l'économie expérimentale.</p> <p>Daniel Houser, professeur d'économie à l'Université George Mason.</p>
<p>Laboratoire Fehr de l'Université de Zurich : Institute for Empirical Research in Economics.</p>	<p>Université de Zurich (canton de Zurich, Suisse)</p>	<p>Ernst Fehr, économiste et directeur de l'institut.</p> <p>Michael Kosfeld, professeur d'économie à l'Université de Zurich.</p> <p>Paul Zak, chercheur en neurosciences, directeur du Center for Neuroeconomics Studies à Claremont (Californie).</p>
<p>Baylor College of Medicine, Human Neuroimaging Laboratory.</p>	<p>Houston (Texas)</p>	<p>Read Montague, chercheur en neurosciences, directeur du laboratoire.</p>

		<p>Brook King-Casas, chercheur en psychologie et économie comportementale, fellow du laboratoire.</p> <p>Greg Berns, chercheur en neurosciences, professeur à l'Université Emory (Atlanta, Georgie), directeur du Computation and Cognitive Neuroscience Lab à l'Université Emory.</p>
Laboratoire de neuroéconomie de l'Université Stanford.	Stanford (Californie)	<p>Antonio Rangel, directeur du laboratoire de neuroéconomie, professeur d'économie à l'Université Stanford.</p> <p>Douglas Bernheim, professeur d'économie à l'Université Stanford.</p> <p>Brian Knutson, chercheur en neurosciences, directeur du Symbiotic Project on Affective Neuroscience (Stanford).</p> <p>Paul Glimcher, Colin Camerer.</p>
Centre d'étude du cerveau, de l'esprit et du comportement de l'Université Princeton : Neuroscience of Cognitive Control Laboratory.	Princeton (New Jersey)	<p>Jonathan Cohen, chercheur en neurosciences, directeur du laboratoire.</p> <p>Joshua Greene, diplômé en philosophie de l'Université Princeton, membre fellow du laboratoire.</p> <p>Jim Rilling, assistant-professeur d'anthropologie à l'Université Emory.</p> <p>Alan Sanfey, assistant-professeur de psychologie à l'Université de l'Arizona.</p>

Autres chercheurs impliqués	Université Cambridge (Royaume-Uni)  Université Harvard (Massachusetts)  Université Concordia (Québec)	Wolfram Shultz, chercheur en neurosciences.  David Laibson, professeur d'économie.  Peter Shizgal, chercheur en neurosciences, directeur du Center for Studies in Behavioral Neurobiology de l'Université Concordia, collaborateur de Daniel Kahneman dans ce qui s'est avéré être l'une des recherches pionnières de la neuroéconomie (Breiter <i>et al.</i> , 2001)
Faits notables	Université du Minnesota (Minnesota)	Première conférence sur les neurosciences et l'économie tenue en octobre 2002.  Le rythme des publications sur la neuroéconomie s'accélère beaucoup depuis 2003.

### 1.2) Introduction aux neurosciences

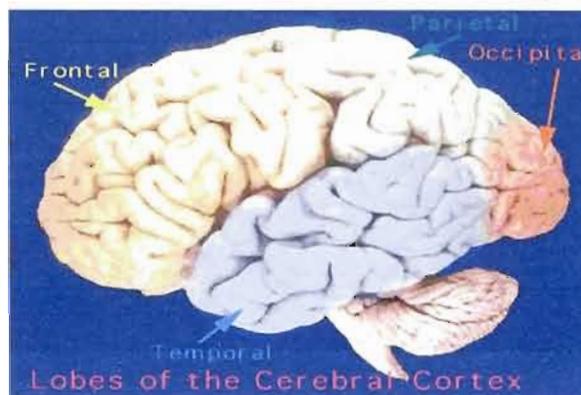
Dans la mesure où elles s'intéressent aux processus décisionnels, des sciences comme l'écologie comportementale, la psychologie ou l'économie sont toutes concernées par les nouvelles découvertes des neurosciences. Quelque processus décisionnel que ce soit doit nécessairement être initié à partir du système nerveux et du cerveau. Une meilleure compréhension du cerveau et de ses mécanismes peut donc nous renseigner sur les fondements du choix économique. Dans la présente section, nous procédons à une brève introduction aux neurosciences.

### 1.2.1) Le cerveau : éléments de base

La figure 1.1 illustre que le cortex du cerveau possède quatre lobes : frontal, pariétal, occipital et temporal. Le lobe frontal est considéré comme le centre de la planification et du contrôle cognitif. Il intègre les inputs des autres régions du cerveau. La région du lobe pariétal gouverne les activités motrices.

Le lobe occipital est là où le décodage visuel s'effectue. Le lobe temporal est important pour la mémoire, la reconnaissance et les émotions.

Figure 1.1 : Les lobes du cortex.

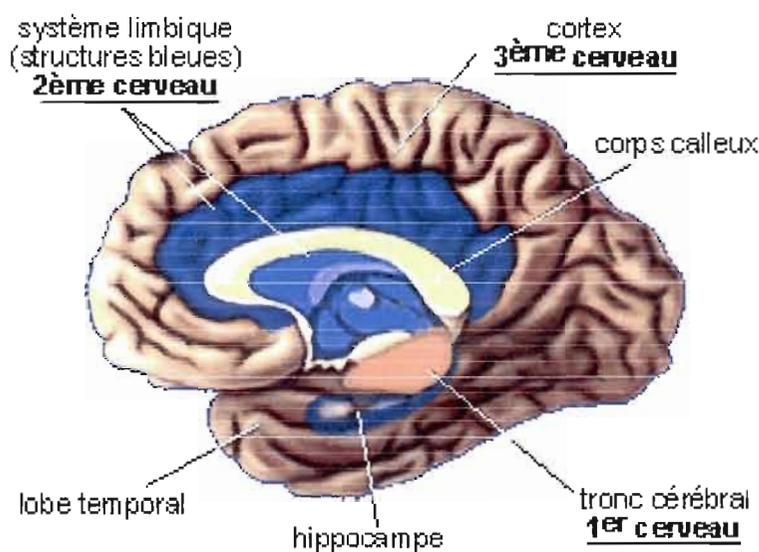


Source : Camerer *et al.* (2003)

Essentiellement, le cerveau humain est un cerveau de mammifère, mais avec un cortex plus développé. Le comportement humain peut donc être considéré comme un compromis entre celui des animaux hautement évolués où les instincts et les émotions jouent un important rôle, et celui où la délibération et la prévoyance sont particulièrement développées (Loewenstein, 1996, cité dans Camerer, Loewenstein et Prelec, 2004). Les humains et les autres mammifères ont des ancêtres communs et leurs cerveaux ont évolué de pair pendant des millions d'années. MacLean (1970) a introduit sa théorie du cerveau tri-unique. Selon cette théorie, le cerveau est passé, au fil du temps, à travers trois stades évolutifs : le reptilien, le paléomammalien et le néomammalien. Chez l'humain et dans une certaine mesure chez les mammifères les plus évolués, les trois cerveaux cohabitent (figure 1.2). Les mammifères les plus primitifs n'ont que les cerveaux paléomammalien et reptilien, alors que les oiseaux, les reptiles, les amphibiens et les poissons n'ont que le cerveau reptilien. Le cerveau reptilien s'occupe des choses les plus élémentaires (boire, manger, copuler). Le cerveau paléomammalien, parfois appelé système limbique, est le siège des sensations et des émotions. Le cerveau néomammalien se caractérise entre

autres choses par la capacité de langage, de raisonnement par association, de contact avec le présent et de projection dans l'avenir. Bien que cette classification (plus particulièrement celle concernant le concept de système limbique) soit aujourd'hui remise en cause (voir Ledoux, 2000), la classification de MacLean demeure intéressante d'un point de vue évolutionniste; les émotions impliquent des modules neuronaux plus primitifs qui furent conservés au travers de l'évolution des mammifères tandis que les processus cognitifs impliquent des circuits plus récents. Comme nous partageons des ancêtres communs avec les autres espèces animales, cette division du cerveau implique que toute étude faite sur des animaux<sup>7</sup> peut fournir de nombreuses pistes quant au fonctionnement cérébral chez l'humain, particulièrement celles concernant le rôle des émotions dans la prise de décision.

Figure 1.2 : Les trois cerveaux de l'humain en coupe transversale.



Source : <http://www.lecerveau.mcgill.ca>

<sup>7</sup> Nous partageons près de 98 % de nos gènes avec certains primates. Un gène est une séquence d'ADN qui est présente dans les chromosomes de toutes les cellules nucléées de l'organisme. Pour le lecteur intéressé, voir Gibbons, 1998, ou encore King et Wilson, 1975.

### 1.2.2) Les outils des neurosciences

Nous présentons ici quatre outils des neurosciences : les dommages cérébraux et les psychopathologies chez l'humain, l'imagerie cérébrale, la mesure d'un seul neurone et les stimuli électriques.

#### 1) Les dommages cérébraux et les psychopathologies chez l'humain

L'une des plus célèbres indications qu'il puisse exister une relation entre certaines fonctions de la personnalité humaine et le cerveau aurait été illustrée la première fois par Phineas P. Gage, qui survécut à une explosion le 13 septembre 1848 à Ludlow (Vermont), alors qu'il supervisait les travaux de construction d'un chemin de fer. Une barre métallique de plus de trois pieds lui passa au travers de la tête (figure 1.3) et détruisit une bonne partie de l'avant de son hémisphère gauche. Étonnamment, il survécut et recommença même à travailler quelques mois plus tard. Par contre, on dénota que sa personnalité avait changé drastiquement ; alors qu'il était reconnu pour ses manières polies avant l'accident, il était devenu exécrable, irrévérencieux et blasphématoire par la suite.<sup>8</sup> Les dommages cérébraux et les psychopathologies chez l'humain permettent d'apprendre sur différentes fonctions cérébrales. La plupart des pathologies (par ex. la schizophrénie, l'autisme, le Parkinson) sont reliées à certaines régions spécifiques du cerveau. Cela permet donc d'identifier les fonctions de ces régions. Par exemple, la maladie de Parkinson commence par affecter une région du cerveau appelée le noyau gris central (*basal ganglia*) avant d'affecter le cortex. Les premiers symptômes de cette maladie nous permettent donc d'apprendre davantage sur ces régions du cerveau (Lieberman, 2000, cité dans Camerer, Loewenstein et Prelec, 2003). Si un patient avec des dommages dans une région spécifique du cerveau ne parvient pas à accomplir certaines tâches qu'une personne normale

---

<sup>8</sup> Rapporté par Malcolm Macmillan de l'école de psychologie de l'Université Deakin (Melbourne, Australie) sur le site [www.deakin.edu.au/hbs/GAGEPAGE/](http://www.deakin.edu.au/hbs/GAGEPAGE/). Voir également Macmillan, 2000.

pourrait effectuer, cela est un indicateur de la fonction de cette région. Particulièrement d'intérêt pour la neuroéconomie, les personnes souffrant de déficits émotionnels (dommages préfrontaux) ; des études ont permis d'illustrer en quoi les émotions peuvent être avantageuses ou désavantageuses lors de certaines décisions économiques.<sup>9</sup>

FIGURE 1.3 : La barre métallique ayant blessé Phineas P. Gage



Source : Harvard's Countway Library of Medicine :  
Warren Anatomical Museum

## 2) L'imagerie cérébrale

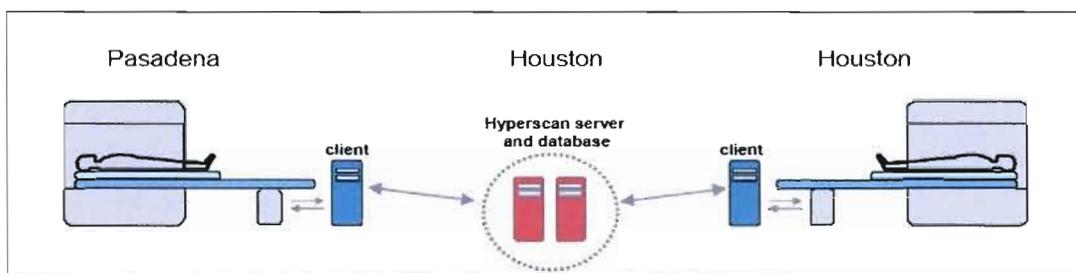
L'imagerie cérébrale permet de visualiser l'activation du cerveau d'un individu effectuant une tâche particulière. On compare ensuite avec l'image obtenue préalablement et on obtient alors une photographie des régions du cerveau activées par la tâche expérimentale. Il existe trois techniques de base. L'encéphalogramme (EEG), qui est effectué en utilisant des électrodes reliées au crâne et qui permet de mesurer l'activité électrique sous-jacente. La tomographie par émissions de positrons (TEP) qui, en mesurant l'influx sanguin dans la région du cerveau associée à la tâche expérimentale, permet une approximation de l'activité neuronale. Enfin, l'imagerie à résonance magnétique fonctionnelle (fMRI), qui fut développée au début des années 1990, enregistre l'activité cérébrale par observation des changements d'oxygénation

---

<sup>9</sup> Nous abordons notamment cette question aux sections 1.4 et 2.3.2.

sanguine.<sup>10</sup> Le signal obtenu est appelé *Bold Signal*.<sup>11</sup> Lors d'une tâche expérimentale, la tête du participant doit demeurer immobile. Ce dernier est informé des différents paiements d'un jeu en observant un écran qui est placé au-dessus de sa tête. Il prend ses décisions en appuyant, avec ses doigts ou avec ses mains, sur une manette ou un bouton. Afin de reproduire ce que les cerveaux font lors d'interactions sociales, des chercheurs du Laboratoire Baylor au Texas ont récemment mis au point un logiciel utilisant des connexions Internet qui permettent de relier plusieurs de ces appareils même si ceux-ci sont situés à milliers de kilomètres les uns des autres<sup>12</sup> (figure 1.4).

Figure 1.4 : Imagerie à résonance magnétique fonctionnelle et *hyperscanning*



Source : King-Casas *et al.* (2005)

<sup>10</sup> L'imagerie à résonance magnétique permet d'obtenir des images très détaillées du cerveau. Un puissant aimant en ligne, vers un champ magnétique, certains atomes (les protons) du cerveau de la personne étendue. On envoie alors une vibration d'ondes radio vers la tête du patient ou du participant. Ces ondes radio sont alors absorbées par les protons qui sortent ainsi de leur alignement. Par la suite, à mesure que les protons reprennent leur alignement, ils émettent des ondes radio qui sont captées et envoyées vers un ordinateur qui construit alors l'image du cerveau. À la base, la technologie fMRI s'appuie sur l'imagerie à résonance magnétique mais elle porte une attention plus particulière aux variations du champ magnétique provoquées par l'arrivée de sang oxygéné dans certaines régions du cerveau.

<sup>11</sup> Pour « Blood Oxygen Level Dependent Signal ».

<sup>12</sup> Cette méthodologie porte le nom d'*hyperscanning* et fut développée au Baylor College of Medicine par Read Montague et ses collègues. Voir Montague *et al.*, 2002. Pour un démo de cette technique, consulter la page suivante : <http://www.hnl.bcm.tmc.edu/hyperScan.html>

### 3) La mesure d'un seul neurone

La mesure d'un seul neurone consiste, comme son nom l'indique, à ne mesurer l'activité que d'un seul neurone à la fois. Cela est possible par l'insertion de minuscules électrodes dans le cerveau. Cet outil est surtout réservé aux animaux car l'insertion de fils peut endommager le cerveau. Néanmoins, comme l'être humain partage avec certains mammifères de nombreuses structures et fonctions cérébrales, la mesure d'un seul neurone permet d'étudier certains processus motivationnels et émotifs communs. C'est cette technique qui a permis de corroborer certaines hypothèses du modèle de l'espérance de l'utilité (voir section 2.2.1).

### 4) Les stimuli électriques

De par leur nature, les stimuli électriques sont surtout réservés aux animaux. En implantant des électrodes dans certaines régions du cerveau, on peut, à l'aide de stimuli électriques, engendrer des types de comportements bien précis tels que manger et boire (Mendelson, 1967 cité dans Camerer *et al.* 2003) ou copuler (Caggiula et Hoebel, 1966 cités dans Camerer *et al.* 2003). Contrairement à des renforçateurs naturels, les stimuli électriques sont inextinguibles et ne permettent pas la satiété. Ils ont un effet directement renforçateur. Pour obtenir de telles stimulations, des rats vont jusqu'à traverser des grillages électriques, sauter par-dessus des haies, refuser des repas alors qu'ils sont affamés, et appuyer sur des leviers pendant des heures jusqu'à la mort (Conover et Shizgal, 2002). C'est en 1954 que James Olds et Peter Milner ont découvert que les rats étaient prêts à exécuter et à apprendre de nouvelles tâches pour obtenir des stimuli électriques. Les premières études associaient les stimuli électriques à une forme d'expérience hédonistique, à une zone de plaisir située dans le cerveau. Si les rats étaient prêts à de tels sacrifices pour obtenir de telles stimulations, n'était-ce pas parce que celles-ci leur procuraient du plaisir ? Cette vision est aujourd'hui

plus nuancée; une action peut être motivée par des éléments non cognitifs prenant naissance en dehors de la conscience et n'est donc pas nécessairement guidée par la recherche du plaisir benthamien (Camerer *et al.* 2003, Shizgal, 1998)<sup>13</sup>. Que ce soit pour l'observation de choix effectués dans un contexte de changement de prix (Green et Rachlin, 1991), pour la mesure de l'ampleur d'une récompense par l'élaboration d'un modèle d'offre de travail (Conover et Shizgal, 2002), pour la recherche des bases neuronales de l'utilité (Shizgal, 1998) ou encore pour la distinction à faire entre le plaisir hédonistique et le choix (Berridge, 1999), les stimuli électriques offrent un grand potentiel pour la neuroéconomie. D'ailleurs, certaines hypothèses standards de l'économie ont été validées à l'aide de cet outil. Par exemple, Conover et Shizgal (1994) démontrèrent le principe de la maximisation au cours d'une étude où des rats avaient à choisir entre deux alternatives : une infusion de solution sucrée ou une stimulation électrique. Conover et son collègue montrèrent que les préférences des rats étaient orientées vers la solution sucrée lorsque les stimuli étaient faibles. Par contre, en augmentant l'ampleur des stimuli, les rats délaissèrent les solutions sucrées pour les stimulations électriques. Les rats se comportaient donc comme s'ils maximisaient la plus large des deux récompenses.

### 1.3) Les processus mentaux : contrôlés et automatiques<sup>14</sup>

#### 1.3.1) Validations empiriques

Savoir jusqu'à quel point les jugements, décisions ou comportements sont sous la gouverne de la pensée consciente demeure l'une des questions les plus fondamentales qui soient. Il est aujourd'hui généralement accepté que nombre de phénomènes psychologiques sont simultanément influencés par des

---

<sup>13</sup> Voir la section 2.2.4 pour plus de détails.

<sup>14</sup> La distinction entre processus contrôlés et automatiques a été faite précédemment par Schneider et Shiffrin, 1977. Depuis, nombre d'auteurs l'ont élaborée, mentionnons ici Sloman, 1996, Lieberman, 2003, Camerer *et al.*, 2003 ainsi que Kahneman, 2003.

processus contrôlés et automatiques<sup>15</sup> (Sloman, 1996). Les débats entourant cette question ont évolué. Initialement, l'existence même de ces processus était questionnée. Aujourd'hui, ce sont les circonstances dans lesquelles l'un ou l'autre de ces processus contrôle l'esprit qui sont débattues (Bargh et Chartrand, 1999). Les processus contrôlés sont définis comme étant des actes mentaux initiés par un acte de volonté, nécessitant l'effort, et que l'on peut contrôler (Bargh et Chartrand, 1999). Parce que ces processus ne peuvent cependant pas opérer en parallèle, ils sont relativement ténus et limités. On ne peut simultanément mémoriser un numéro de téléphone et une adresse électronique; la pensée consciente est fragile et ne peut opérer que de façon séquentielle. Dans une série d'études portant sur les limites de la pensée consciente, il a été démontré qu'un acte nécessitant un effort de volition (comme par exemple, le fait de résister à une tentation) amenuisait ultérieurement la capacité à résoudre une tâche alternative (Baumeister, Bratslavsky, Muraven et Tice, 1998).

Définir les processus automatiques est cependant plus ardu. Ils sont le contraire des processus contrôlés et ils opèrent en parallèle à la conscience (Camerer, Loewenstein et Prelec, 2003). Ils concernent les actes mentaux qui n'interfèrent pas les uns avec les autres, qui ne nécessitent aucun effort, intention ou conscience, et qui sont habituellement vécus comme des perceptions (Lieberman, 2003). Ils correspondent à un système associatif qui infère sur la base de similitudes et de contiguïtés (Sloman, 1996). Bargh et Chartrand (*ibid.*, p. 464) distinguent deux formes de processus mentaux « non conscients ». Les deux opèrent sans efforts mais l'un nécessite la gouverne de la conscience pour initier son action (par exemple conduire un véhicule pour la centième fois), tandis que l'autre évolue complètement en parallèle.

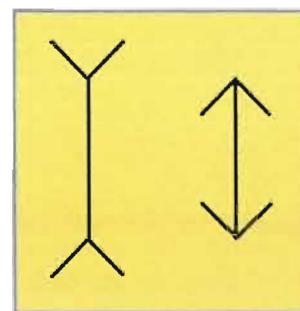
---

<sup>15</sup> Les processus automatiques sont parfois appelés processus non conscients, implicites, intuitifs, tandis que les processus contrôlés sont dits conscients, délibérés ou explicites.

Le mode d'opération des processus contrôlés et automatiques est clairement interactif. Les processus automatiques sont plus rapides et « font toujours valoir leur opinion ». Ils sont le mode par défaut du cerveau (Camerer, Loewenstein et Prelec, 2003) et occupent la plus large part des sources du comportement (Sloman, 1996 ; Bargh et Chartrand, *ibid.*, p. 476 ; Baumeister *et al.*, *ibid.*, p. 1252, 1998). Devant une tâche ou un problème à résoudre, les deux processus peuvent arriver à une solution, et cette solution peut parfois diverger. C'est cela qui a amené Sloman (*ibid.*, 1996) à définir un critère justifiant deux systèmes de raisonnements, l'un clairement associatif, et l'autre basé sur des règles de logique. Ce critère de Sloman est défini par tout ce qui amène une personne à simultanément accepter deux réponses contradictoires, l'une en provenance des processus automatiques et l'autre en provenance des processus contrôlés. Le critère de Sloman est une intéressante façon de démontrer l'existence même des deux processus et leurs potentielles divergences. Pour illustrer, Sloman évoque des formes de croyances qui peuvent être simultanément contradictoires : mentionnons ici l'illusion visuelle et l'erreur de jugement :

Figure 1.5 : Illusion de Müller-Lyer

1) L'illusion de Müller-Lyer illustre le critère de Sloman par l'illusion d'optique (figure 1.5). Les deux lignes sont de même longueur (seul le sens des flèches diffère). La perception visuelle et la compréhension abstraite fournissent pourtant deux réponses différentes. Il semble *a priori* que les deux lignes n'ont pas la même longueur et ce, même après vérification. Seule la connaissance abstraite des règles de la géométrie permet de comprendre que la perception visuelle nous a induit en erreur.



Source : [www.aber.ac.uk/.../](http://www.aber.ac.uk/.../)

2) Une erreur de jugement connue sous le nom de l'erreur de la conjonction<sup>16</sup> (Tversky et Kahneman, 1983) permet également d'illustrer la coexistence des processus contrôlés et automatiques. Un des principes statistiques les plus simples est celui de la conjonction; la probabilité d'une conjonction est inférieure à la probabilité de chacune de ses constituantes, c'est-à-dire :

$$P(A \text{ et } B) < P(A) \text{ et } P(A \text{ et } B) < P(B)$$

Tversky et Kahneman donnèrent le préambule suivant à leurs participants(es) :

Linda a 31 ans. Elle est célibataire, articulée et très intelligente. Elle a obtenu un diplôme en philosophie. Comme étudiante, elle était très concernée par des sujets comme la discrimination et la justice sociale et a également participé à des démonstrations anti-nucléaire.

Ensuite, ils demandèrent aux participants(es) de classer par ordre de probabilité certaines modalités de réponses, dont les suivantes :

Linda est active dans le mouvement féministe. ( A )

Linda est caissière dans une banque. ( B )

Linda est caissière dans une banque et est active dans le mouvement féministe. ( A et B )

La très grande majorité des participants classifièrent les modalités en violation avec la règle de conjonction et ce, même parmi les plus habiles en statistique.<sup>17</sup> Ainsi, 90 % du groupe des répondants ayant des connaissances statistiques firent l'erreur suivante :

$$P(A) > P(A \text{ et } B) > P(B)$$

Sachant pourtant reconnaître la validité de la règle de la conjonction, nombre de participants ne l'appliquèrent tout simplement pas en dehors du contexte

---

<sup>16</sup> Traduction de « Conjunction Fallacy ».

d'un cours de statistique. La coexistence de ces deux formes d'évaluation est une illustration de l'existence de deux processus, un de délibération attentive, et l'autre s'effectuant sur des bases associatives; cela en conformité avec le critère de Sloman.

### 1.3.2) Corollaires physiologiques

Nombre d'études utilisant la technologie fMRI ont permis d'identifier certains des corollaires cérébraux des processus contrôlés et automatiques. L'identification de ces régions n'a pas seulement pour objet de tracer la nomenclature géographique du cerveau et d'ainsi associer celles-ci à différentes formes de comportements; cela ne serait en soit d'aucun intérêt pour l'économie comportementale; ce qui importe, c'est de mieux comprendre ce qui est sous-jacent aux processus décisionnels, ce qui initie l'action ou influence la décision. En sachant quelles sont les composantes contrôlées et automatiques au sein du processus de décision, on peut ensuite regrouper ce qu'on croyait initialement être des phénomènes disparates, et séparer des phénomènes qu'on croyait initialement similaires (Ochsner et Lieberman, 2001, p.727)<sup>18</sup>. Lieberman (2003) classe les processus en deux catégories : le système X et le système C<sup>19</sup>. Les processus contrôlés (système C) sont composés de trois mécanismes neurocognitifs qui oeuvrent de pair : le cortex cingulaire antérieur, le cortex préfrontal et le lobe temporal médial (incluant l'hippocampe). Des circuits neuronaux complexes se trouvant à l'intérieur de ces trois régions permettent la pensée consciente. La région du lobe temporal médial est située à l'intérieur de chaque lobe temporal. Cette région permettrait la mémoire consciente (Squire, Stark et Clark, 2004, cités dans Lieberman, 2003, *ibid.*). Le cortex préfrontal est parfois appelé « la région exécutive », parce qu'il utilise les intrants en provenance d'autres régions du

---

<sup>17</sup> Par exemple, 85 % des doctorants de l'Université Stanford ayant auparavant suivi des cours avancés de statistiques firent l'erreur de la conjonction.

<sup>18</sup> À ce sujet, voir par exemple la section 2.2.5 portant sur les préférences.

cerveau afin de formuler des objectifs à court et long terme, et parce qu'il planifie des actions conformément à ces objectifs (Shallice et Burgess, 1998, cités dans Camerer, Loewenstein et Prelec, 2003). La région connue sous le nom de cortex cingulaire antérieur est notamment activée par différents types de conflits cognitifs, par la douleur physique et par le sentiment de rejet (Eisenberger, Lieberman et Williams, 2003). Cette région peut « alerter » le cortex préfrontal lorsque les processus automatiques ne parviennent pas à atteindre leur objectif (Botvinick *et al.*, 1999, cités dans Lieberman, 2003). Bien qu'il soit admis que cette région soit celle qui fasse clairement appel aux processus contrôlés en cas d'interruption des processus automatiques, on ne peut néanmoins la classer avec trop de rigidité parce que certains processus automatiques y ont également cours (Lieberman, 2003)<sup>20</sup>.

Lieberman<sup>21</sup> associe trois structures neuronales au système X : l'amygdale, le noyau gris central (*basal ganglia*) et le cortex latéral temporal. L'amygdale réagirait spontanément à des stimuli associables à la menace sous différentes formes (serpents, expressions de peur) et initierait différents mouvements moteurs de « fuite ou affrontement ».

Le noyau gris central répondrait à différentes formes de récompenses anticipées et le cortex latéral temporal permettrait la reconnaissance de l'identité, attributs et comportements d'objets sociaux et non sociaux. Le tableau 1.1 illustre les corollaires physiologiques des systèmes X et C.

---

<sup>19</sup> Le système X pour *reflexive* correspond aux processus automatiques tandis que le système C provient de *reflective* et réfère aux processus contrôlés.

<sup>20</sup> Une certaine prudence est de mise pour toutes les autres régions du cerveau qu'on associe aux systèmes X et C. Il s'agit d'une classification valable, mais imparfaite.

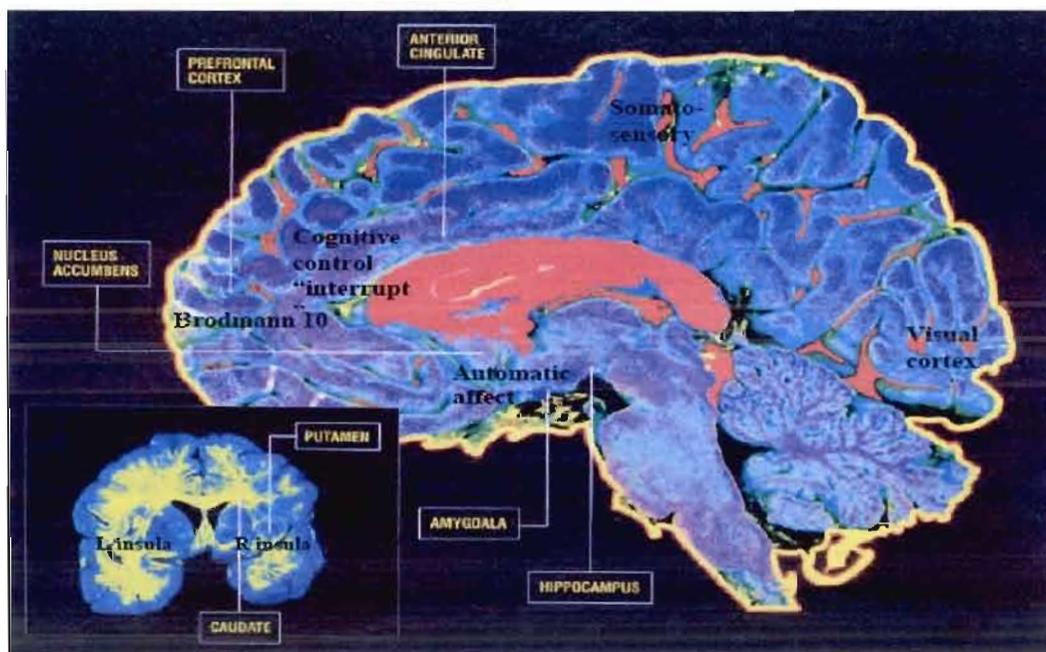
<sup>21</sup> Pour une revue plus récente, voir Lieberman, M. (sous presse). « The X and C systems : The Neural Basis of Reflexive and Reflective Social Cognition ». In : Harmon-Jones, E. et P. Winkelman (éd). (À paraître). *Fundamentals of Social Neuroscience*, New York : Guilford.

Tableau 1.2 : Régions du cerveau associées aux deux systèmes mentaux

Systeme X	Systeme C
Amygdale	Cortex préfrontal
Région du noyau gris central	Cortex cingulaire antérieur
Cortex latéral temporal	Lobe temporal médial (incluant l'hippocampe)

La figure 1.6 présente certaines des régions du cerveau qui sont particulièrement pertinentes à la neuroéconomie. Mentionnons ici : les régions du *noyau accumbens* (striatum ventral), de *putamen* et de *caudate* (noyau caudé) qui font toutes partie de la région du noyau gris central. L'économiste Kevin McCabe et ses collègues (2001) ont notamment observé que la région Brodmann 10 serait active lors d'un choix coopératif.

Figure 1.6 : Quelques régions pertinentes à la neuroéconomie



Source : Camerer, Loewenstein et Prelec (2003)

### 1.3.3) Processus automatiques : quelques implications

Comme les processus automatiques exercent continuellement leur influence sur la prise de décision, nous présentons ici certaines de leurs plus importantes caractéristiques : le parallélisme, la plasticité, la modularité et la spécialisation. Puisque nous ferons implicitement ou explicitement continuellement allusion à ces notions tout au long du présent mémoire, nous prenons le temps de les expliciter.

#### Le parallélisme

L'essentiel de l'activité cérébrale s'effectue de manière parallèle et en dehors de notre conscience. C'est le principe du parallélisme. Le parallélisme donne au cerveau un extraordinaire pouvoir parce qu'il facilite des réponses rapides et obtenues sans efforts. L'identification visuelle en est un bel exemple. En général, les processus parallèles ne sont pas accessibles à la conscience. Des études faites à l'aide d'imagerie cérébrale ont montré que le visage humain était automatiquement décodé (en dehors de la pensée consciente) par l'amygdale et que celle-ci était particulièrement activée par des visages effrayants (Öhman, 2002). Reinders et ses collègues (2003) interprètent cet automatisme comme étant le fruit d'un long processus évolutionniste associant les chances de survie et de reproduction de nos ancêtres à la capacité de détection rapide de menaces. Tout au long du présent mémoire, nous reviendrons souvent sur la notion du parallélisme parce que plusieurs recherches en neurosciences indiquent que celle-ci semble exercer une continue influence sur la décision et le comportement.

#### La plasticité

Dès qu'une activité neurologique a lieu, des changements physiologiques s'effectuent dans le cerveau, c'est le principe de la plasticité. Lorsqu'un signal

est envoyé de façon répétée d'un neurone à un autre, la connexion se renforce. Le cerveau change donc physiologiquement selon les opérations qu'il effectue. Par exemple, la région du cerveau associée à la conduite et à la mémoire spatiale (l'hippocampe) est plus développée chez les chauffeurs de taxi (Camerer *et al.*, 2003). L'environnement économique peut donc façonner les cerveaux. Une implication majeure de la plasticité est que l'encodage d'information est irréversible parce que les processus physiologiques qui produisent l'apprentissage sont eux-mêmes irréversibles. Un message aura donc tendance à être cru s'il est répété plusieurs fois, même en sachant que le message est faux (Gilbert et Gill, 2000 cités dans Camerer *et al.*, 2003); quand les gens se font des croyances qui sont par la suite discréditées pour de bon, les croyances persistent (Ross, Lepper et Hubbard, 1975 cités dans Camerer *et al.*, 2003); et quand ils savent qu'une chose est vraie ou fausse, ils surestiment le degré de connaissance qu'ont les autres de cette même chose (Camerer, Loewenstein et Weber, 1989). Les théories standards de l'information présupposent que l'on peut ignorer des messages inutiles. Pourtant, plusieurs recherches ont démontré le contraire. La plasticité est donc en lien avec la publicité, la propagande et l'information. Par exemple, une information où l'on a préalablement vanté les mérites d'un titre (pourtant surévalué) pourrait (quand même) engendrer son achat.

### La modularité

La *tabula rasa* de John Locke (1690) et toute ses réfutations de l'innéisme, à savoir qu'aucune connaissance ou disposition ne sont présentes à la naissance, ont graduellement fait place à l'hypothèse de modularité au sein de la psychologie. Pour accomplir certaines tâches spécifiques, des modules spécialisés sont appelés, c'est le principe de modularité. Selon cette théorie, le cerveau est une collection de différents dispositifs spécialisés. Les neurones que l'on retrouve dans les diverses parties du cerveau ont différentes formes et structures, différentes propriétés fonctionnelles et peuvent s'assembler en des

modules relativement discontinus et spécialisés. La modularité implique que certaines régions du cerveau accomplissent certaines tâches spécifiques. Les recherches faites sur des patients souffrant de déconnexion calleuse<sup>22</sup> (*split brain*) ont permis de découvrir que l'hémisphère gauche du cerveau est en charge des principales activités cognitives (pour une revue de ces recherches, voir Gazzaniga, 1998). Par exemple, les zones de Broca et de Wernicke sont respectivement responsables de la production et de la compréhension du langage. L'existence ou non de certains types de modules spécialisés est d'intérêt pour les économistes. Ainsi, avec la notion d'équilibre bayésien, on prend pour acquis que les gens peuvent facilement faire des inférences sur les préférences et les croyances des autres. Or, procéder à de telles inférences n'est pas si facile. De récentes recherches en neurosciences ont cherché à démontrer l'existence d'un module spécialisé dans la reconnaissance des aspirations des autres (Rilling *et al.*, 2004) et il ne serait pas étonnant que des personnes normales puissent différer dans leurs habiletés à procéder à de telles inférences. Une étude de Baron-Cohen (1995) suggère l'existence d'un tel module d'inférence. On montre une boîte de friandises à un enfant autiste. On remplace par des stylos et on lui demande ce qu'un autre enfant pensera qu'il y a dans la boîte. La plupart du temps, il répondra : « Des stylos ». Les enfants normaux résolvent ce problème vers l'âge de 3 ou 4 ans. Les enfants autistes souffrant du syndrome d'Asperger n'y parviennent que vers l'âge de 8 à 12 ans et avec beaucoup de difficultés. Les autistes ont des lacunes en matière d'inférence sur autrui. La région du *superior temporal sulcus*, serait spécialisée dans l'identification du comportement des autres (Lieberman *et al.* 2002). Elle serait particulièrement importante pour pouvoir faire des inférences sur le comportement d'autrui et serait peu fonctionnelle chez les autistes (Baron-Cohen, 1995 cité dans Lieberman *et al.*, 2002). Les autistes adultes, lorsque face à un problème d'inférence, peuvent parvenir à atténuer cette lacune. Pour ce faire, leur cortex préfrontal ventral (normalement associable au raisonnement) doit par contre compenser. C'est donc dire qu'ils

---

<sup>22</sup> La déconnexion calleuse implique l'absence de transferts d'informations interhémisphériques.

utilisent leurs processus délibérés ou leur raison, à un coût neurologique élevé, pour résoudre ce problème. Mais cela se fait non sans difficultés.

### La spécialisation

Le cerveau tente d'accomplir des tâches en utilisant des connexions neuronales déjà existantes, le plus efficacement possible; c'est le principe de spécialisation. Le cerveau ne représente environ que 2 % du poids du corps humain et pourtant, il mobilise en permanence environ 20 % du sang et de l'oxygène de notre organisme (<http://www.lecerveau.mcgill.ca>). Lorsqu'une nouvelle tâche est effectuée, plusieurs régions du cerveau sont activées à un coût neurologique très élevé. À mesure que l'apprentissage de la tâche s'effectue, le coût neurologique décroît et l'activité cérébrale associée à la tâche se canalise dans des sites très précis. Lo et Repin (2002) ont comparé des mesures psychophysiologiques de cambistes. Ils arrivent à la conclusion que les réactions biologiques des cambistes sont négativement corrélées avec leur expérience. Au fur et à mesure que ces derniers apprennent à effectuer leur tâche, les processus automatiques sont alors appelés à jouer un rôle de plus en plus important. La spécialisation implique que les tâches que l'on doit accomplir s'effectueront avec ce que le cerveau a « entre les mains ». Ainsi, lorsque appelé à prendre une décision, le cerveau, plutôt que de soupeser de façon explicite les gains et pertes, utilisera fréquemment des mesures d'« appariement automatiques »<sup>23</sup> (Leboeuf, 2002 cité dans Camerer *et al.* 2003; Medin et Bazerman, 1999) qui sont moins coûteuses en terme neurologique. Les implications de ces mesures d'appariement automatiques sur les processus décisionnels sont donc fondamentales. Elles sous-tendent que le comportement ou la décision peuvent être fortement tributaires du contexte comme le suggèrent Gazzaniga, Ivry et Mangun, 1998 (cités dans Smith, 2004, p.3):

---

<sup>23</sup> Traduction du terme « Pattern Matching ».

« Two decision tasks with the same underlying logical structure may lead to different responses because they are embedded in different contexts and invoke different memory experiences. ».

L'effet de contexte est donc vécu comme une expérience de perception et une perception est le produit de la mémoire; les processus par lesquels nous apprenons sur l'environnement externe impliquent une relation entre les conditions présentes et nos propres expériences de conditions similaires (Hayek, 1952 cité par Smith, 2001).<sup>24</sup> Supposer que l'évaluation de prospects se fasse principalement par l'évaluation des probabilités d'événements est dès lors remis en cause par certaines recherches issues des neurosciences et de la psychologie ; le problème de « la maladie asiatique » (Tversky et Kahneman, 1981) dans lequel les gens modifient leur choix selon le contexte en est un célèbre exemple.<sup>25</sup>

#### L'appariement automatique: deux exemples

Les sciences économiques suggèrent que les gens ont généralement des capacités cognitives qui peuvent être adaptées à n'importe quel type de problème et que les résultats seront équivalents si la structure est la même. Les processus automatiques, particulièrement si le contexte change, suggèrent au contraire que les résultats seront tributaires de la capacité d'un module cérébral à répondre à une tâche particulière. Comme les processus

---

<sup>24</sup> Il est intéressant de noter que Hayek a exercé une réelle influence sur les neurosciences et sur certains préceptes de la psychologie évolutionniste (Smith, 2001). À ce sujet, voir Fuster, 1999.

<sup>25</sup> Tversky et Kahneman (1981) ont décrit ce problème. Devant la menace d'une maladie asiatique qui va tuer 600 personnes si rien n'est fait, deux alternatives sont proposées. Si le programme A est adopté, exactement 200 personnes seront sauvées. Si le programme B est adopté, il y a une probabilité de 1/3 que 600 personnes soient sauvées et une probabilité de 2/3 que personne ne soit sauvé. Placés devant ce choix, la plupart des gens choisissent le programme A. Lorsqu'on reformule la question en suggérant de choisir entre le programme C (exactement 400 personnes vont mourir) ou le programme D (une probabilité de 1/3 que personne ne meure et une probabilité de 2/3 que 600 personnes meurent), les gens choisissent alors généralement le programme D. Pourtant, en choisissant A initialement, ils auraient dû choisir C dans le second contexte (les couples A-C et B-D représentent les mêmes choix, seule la formulation diffère). Lorsque le cadrage est présenté dans un contexte de gain, les gens sont averses au risque et choisissent l'option A. Lorsque le cadrage est présenté dans un contexte de perte, les gens préfèrent l'option la plus risquée. Nous revenons sur ce phénomène à la section 2.3.3.

automatiques sont moins coûteux en terme neurologique et qu'ils interviennent donc plus rapidement que les processus contrôlés, leur influence est fondamentale :

« The struggle between rapid unconscious pattern-detection processes and their slow, effortful adjustment by deliberation is not a fair fight ; so automatic impressions will influence behavior much of the time. »<sup>26</sup>

Deux exemples permettront d'illustrer l'importance de cet appariement automatique sur les processus décisionnels. Gobet et Simon (1996) ont montré que des joueurs d'échecs experts sont capables de mémoriser davantage de positions que des joueurs novices seulement si ces positions correspondent à des situations de matchs. Ainsi, dans les cas où les pièces sont disposées aléatoirement et ne correspondent pas à une situation plausible, la différence entre experts et novices n'est pas significative. Lorsque le contexte change, la performance ou capacité des experts n'est plus équivalente. Pourtant, la structure du problème demeure la même (mémoriser comment des pièces sont distribuées sur un échiquier). Un autre exemple provient de Gneezy et Rustichini (2000) : les garderies qui instaurent une pénalité monétaire aux parents retardataires vont subséquemment, contrairement à toute logique économique, faire face à davantage de retards. L'appariement automatique peut également expliquer ce phénomène. Alors qu'initialement la ponctualité est vue comme une forme d'obligation morale, les retards, après l'introduction de la pénalité, sont vus comme un choix impliquant un prix monétaire. Cette situation étrange peut s'expliquer par le fait que le contexte lui-même devient une variable explicative. Si cette interprétation est la bonne, l'hypothèse *ceteris paribus* devient alors très forte et la relation entre le prix et les quantités demandées peut étrangement être positive. Camerer *et al.* (2003) suggèrent que de tels processus décisionnels sont généralisés et ont donc souvent un impact sur la décision économique. La prochaine section permettra d'approfondir cette question.

---

<sup>26</sup> Camerer *et al.*, 2003, p. 27.

#### 1.4) Approche bidimensionnelle du fonctionnement cérébral

Encore tout récemment, les économistes considéraient le cerveau comme étant une boîte noire qu'on pouvait tout au mieux décrire par des équations mathématiques simplificatrices. Mais aujourd'hui, les récentes avancées en neurosciences, notamment les nouvelles techniques d'imagerie cérébrale, permettent l'observation des différentes parties du cerveau qui sont actives lorsque les gens prennent des décisions de nature économique. Une meilleure compréhension de ce que notre cerveau cherche à faire peut permettre, en utilisant des hypothèses qui partent d'une meilleure description neurologique, non seulement d'améliorer les modèles mathématiques du comportement économique, mais aussi de mieux comprendre tout ce qui concerne les processus décisionnels.

C'est dans cette optique, afin d'illustrer les bases du processus décisionnel, que Camerer *et al.*<sup>27</sup> (2003, 2005) présentent un schéma simple d'approche bidimensionnelle du fonctionnement cérébral (tableau 1.3). La première dimension du modèle concerne les processus automatiques et contrôlés. Nous savons que les processus contrôlés opèrent un à la fois, nécessitent un effort, sont délibérés et permettent un accès introspectif partiel tandis que les processus automatiques opèrent en parallèle, s'effectuent sans impression d'effort, sont indirects et ne permettent pas d'accès introspectifs. La deuxième dimension du fonctionnement neurologique est représentée par les processus affectifs et cognitifs. Une telle distinction est importante en psychologie et remonte aux anciens Grecs depuis l'époque où Platon comparait l'esprit humain à des chariots tirés par deux chevaux : les émotions et la raison. Les

---

<sup>27</sup> Colin Camerer enseigne la psychologie cognitive et l'économie au California Institute of Technology. Diplômé de l'Université de Chicago en 1981 (doctorat en théorie décisionnelle à l'âge de 22 ans), il est l'une des figures dominantes de la neuroéconomie. Pour ce qui s'avère une contribution fondamentale par la publication de l'article « Neuroeconomics : How Neuroscience Can Inform Economics », publié dans le *Journal of Economic Literature* en 2005, il a notamment collaboré avec George Loewenstein de l'Université Carnegie Mellon (doctorat en économie à Yale en 1985 et théoricien des choix intertemporels) et avec Drazen Prelec, qui enseigne les sciences de la gestion au MIT.

processus cognitifs concernent la connaissance, le calcul délibéré et la raison. Buck (1999) (cité dans Camerer *et al.*, 2003) définit les processus affectifs comme étant reliés aux émotions telles la colère, la tristesse ou la honte, ainsi qu'à des éléments de considérations biologiques de base comme la faim, la douleur ou le désir sexuel. Il est important de souligner que notre espèce partage beaucoup de mécanismes neuronaux avec les autres mammifères et que ces mécanismes concernent davantage les émotions que la cognition; des conclusions issues d'études faites sur des animaux s'appliquent également à notre espèce. Certaines dimensions biologiques de la décision remettent en cause des hypothèses standards du comportement économique. Par exemple, la perspective économique considère la maximisation du bien-être comme étant à la base des processus décisionnels. Partant du précepte que l'homme n'a pas évolué pour être heureux mais pour s'adapter, survivre et se reproduire, la perspective des neurosciences, elle, implique que le sentiment de bien-être et le plaisir sont plutôt considérés comme des signaux homéostatiques parmi tant d'autres.<sup>28</sup>

Le tableau 1.3 permet d'enrichir l'explication du comportement économique. Le quadrant I, qui correspond à une démarche cognitive consciente, est en action lorsque, par exemple, on planifie un budget ou on effectue un calcul mathématique. Le quadrant II est certainement le moins sollicité d'entre les quatre. Pour expliquer que des processus affectifs puissent être volitifs, Camerer et ses collègues donnent l'exemple d'un acteur faisant appel à des émotions antérieurement vécues pour jouer un rôle. Le quadrant III est en action lorsque l'on fait une action machinale, comme par exemple attacher ses chaussures. Le quadrant IV intervient lorsque l'on sursaute, suite à une peur soudaine par exemple.

---

<sup>28</sup> L'axiome de transitivité des préférences est remis en cause par la notion d'homéostasie. La section 2.2.6 approfondit ce sujet.

Tableau 1.3: Approche bidimensionnelle du fonctionnement cérébral <sup>29</sup>

	Cognitif	Affectif
Processus contrôlés <ul style="list-style-type: none"> <li>- un à la fois</li> <li>- nécessitent un effort</li> <li>- délibérés</li> <li>- accès introspectif limité</li> </ul>	<b>I</b>	<b>II</b>
Processus automatiques <ul style="list-style-type: none"> <li>- en parallèle</li> <li>- sans effort</li> <li>- indirects</li> <li>- pas d'accès introspectif</li> </ul>	<b>III</b>	<b>IV</b>

Source : Camerer, Loewenstein et Prelec

Les modèles économiques standards sont construits à partir de l'implicite hypothèse que la décision économique émane principalement du quadrant I. En incorporant d'autres dimensions, le tableau 1.3 plaide pour un plus grand réalisme; une décision économique nécessite une interaction constante entre les différents quadrants avec, comme mode par défaut, les processus automatiques des quadrants III et IV. Les processus contrôlés sont relativement peu sollicités et sont l'exception plutôt que la règle, c'est-à-dire qu'ils sont le mode d'interruption du cerveau. Cela a fait dire aux économistes David Laibson,<sup>30</sup> de l'Université Harvard, et Andrew Caplin,<sup>31</sup> de l'Université de

<sup>29</sup> Ce schéma n'est pas propre aux neurosciences parce que trop simplificateur. Malgré cela, Camerer et ses collègues considèrent qu'il peut nuancer et améliorer une approche unidimensionnelle de maximisation de l'utilité.

<sup>30</sup> David Laibson est un économiste connu pour avoir formalisé un modèle de consommation incorporant des notions issues des neurosciences. Pour plus de détails, voir Laibson, 2001.

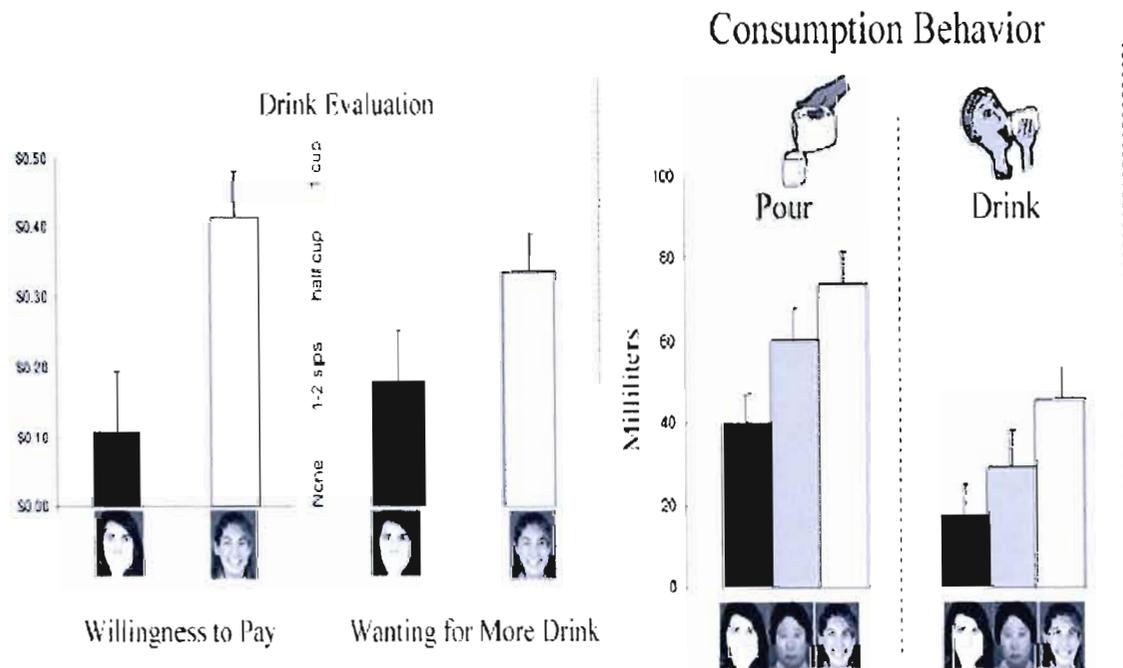
New York, que la théorie économique est érigée à partir du mode d'interruption. En fait, même lorsque les gens croient agir en toute conscience, et que les processus contrôlés semblent être à l'origine d'un comportement, ce sont souvent les processus automatiques qui ont initié l'action par le biais de signaux subtils (Bargh, 1997 cité dans Baumeister *et al.*, 1998).

L'influence des processus automatiques sur la décision économique est démontrée par Winkielman, Berridge et Wilbarger (2005) qui montrent qu'une présentation subliminale de visages à multiples humeurs (pendant 16 ms) peut avoir une incidence sur la consommation et sur l'évaluation de l'utilité d'une boisson. Les présentations subliminales minimisent la contribution des processus cognitifs les plus complexes et réduisent la réponse stratégique (Bargh, 1996 cité dans Winkielman *et al.*, 2005). Lorsqu'on présente un visage de façon subliminale, la brièveté de la présentation est telle qu'il est impossible d'avoir consciemment aperçu un visage et son humeur. Comme préalablement souligné à la section 1.3.3, des études faites à l'aide de la technique d'imagerie cérébrale ont montré que les visages agressifs pouvaient être enregistrées par l'amygdale en dehors de la conscience (Öhman, 2002). Au cours d'une série d'expérimentations, Winkielman et ses collègues montrèrent que la présentation subliminale de visages heureux moussait la consommation et la volonté de payer plus cher pour une boisson (figure 1.7). Leur étude montre que le cerveau enregistre donc des signaux qui ne sont pas nécessairement accessibles à la conscience (système C) mais qui influencent néanmoins la décision.

---

<sup>31</sup> Andrew Caplin est co-directeur du Centre expérimental de sciences sociales de l'Université de New York. Il croit fermement que les frontières entre les disciplines (psychologie et économie) ne survivront plus encore très longtemps. Pour plus de détails, voir <http://www.econ.nyu.edu/user/caplina/>

Figure 1.7 : Influence d'un signal subliminal sur la consommation de personnes assoiffées



Source : Winkielman, Berridge et Wilbarger (2005)

La partie gauche de la figure 1.7 (*Drink Evaluation*) illustre que des personnes assoiffées désirent en moyenne davantage de jus et sont prêtes à payer plus cher pour une consommation lorsque le visage subliminal est joyeux<sup>32</sup>. Une conclusion contraire s'applique pour un visage malheureux. La partie droite de la figure 1.7 (*Consumption Behavior*) montre la relation entre les quantités versées et bues en fonction de l'humeur des visages<sup>33</sup> (de gauche à droite, l'humeur des visages est mauvaise, neutre et bonne). Les travaux de

<sup>32</sup> Cette étude fut faite auprès de 29 étudiants en psychologie (6 garçons et 23 filles) de premier cycle universitaire. En premier, on présenta huit fois un visage subliminal neutre afin d'établir une base comparative. Puis, on présenta huit visages heureux ou agressifs avant de demander aux participants d'évaluer une boisson qu'on leur avait servie.

Winkielman et de ses collègues montrent comment les processus affectifs et automatiques du quadrant IV peuvent avoir une incidence sur la consommation et sur les prix. Le fait que des *cues* subliminaux puissent être enregistrés par le cerveau (sans faire appel au quadrant I) est difficilement réconciliable avec l'idée que la décision économique est principalement le fruit d'un calcul précis et délibéré. En fait, non seulement les processus automatiques et affectifs sont-ils fréquemment sollicités, mais en plus, ils sont parfois essentiels à l'efficacité même d'une décision. Bechara, Damasio, Tranel et Damasio (1997) montrèrent que la décision stratégique, que l'on croit à tort uniquement reliée au raisonnement délibéré et logique, peut être précédée de biais non conscients ayant une incidence avantageuse sur la décision. Pour illustrer, Bechara et ses collègues ont créé une expérimentation où l'objectif est de maximiser les gains en sélectionnant des cartes à partir de quatre paquets<sup>34</sup>. L'expérimentation est conçue de telle sorte que deux de ces paquets engendrent de forts gains à court terme mais aussi de fortes pertes et une espérance de gains négative à long terme. Les deux autres paquets peuvent sembler moins alléchants à cause de leurs modestes gains mais ils affichent néanmoins une espérance positive à long terme et ils correspondent au choix optimal de cette expérimentation. Bechara et ses collègues demandèrent à des participants normaux ainsi qu'à des patients souffrant de certains types de dommages préfrontaux de choisir entre ces différentes options risquées. Ces patients ont la particularité de ne plus pouvoir traiter correctement les informations émotionnelles; ils ont toutes sortes de déficiences en matière de

---

<sup>33</sup> Cette étude fut faite auprès de 39 étudiants (14 garçons et 25 filles) de premier cycle universitaire. Après la présentation de 8 visages subliminaux (agressifs, neutres ou heureux), on demanda aux participants de se verser du jus et d'en boire.

<sup>34</sup> Au cours de cette expérimentation connue sous le nom de *Iowa Gambling Task*, Bechara et ses collègues ont demandé à des participants d'effectuer des choix à partir de quatre paquets : A, B, C et D afin de maximiser les gains et d'éviter les pertes. En tournant chaque carte, les participants obtiennent une récompense immédiate (100 \$ dans les paquets A et B et 50 \$ dans les paquets C et D). Les rendements à court terme sont plus élevés dans les paquets A et B. Cependant, tourner une carte peut également engendrer une perte, et les pertes à long terme sont plus grandes dans les paquets A et B mais les participants ne le savent pas. Jouer à partir des paquets C et D est la meilleure stratégie et permet des gains à long terme. En tout, les participants doivent faire 100 sélections de cartes mais ils ne le savent pas non plus.

décisions sociales et ont beaucoup de difficultés à choisir des amis, partenaires ou activités (Bechara, Damasio, *et al.*, 2000; Bechara, Tranel et Damasio, 2002, Bechara, 2004); ils affichent également une insensibilité aux conséquences futures de leurs choix. Malgré tout, ces personnes ont la particularité de performer normalement au test de quotient intellectuel et la plupart de leurs fonctions cognitives sont demeurées intactes (Bechara, 2004).

Les résultats sont révélateurs : Bechara et ses collègues observèrent que les personnes normales, contrairement aux patients, commencèrent à choisir les options avantageuses avant même d'avoir pu explicitement réaliser laquelle des stratégies était la meilleure. En soi, cela est une autre indication que le quadrant I ne fait pas le travail à lui tout seul et que les processus affectifs se doivent d'être fonctionnels pour permettre la meilleure prise de décision stratégique. En outre, on observa que chez les personnes normales, les choix les plus risqués engendraient des mesures physiologiques (sueur de la peau) qui ne furent pas observées chez les patients. Ce qui est frappant est que ces mesures furent observées alors que les participants ignoraient encore laquelle des stratégies était la plus efficace. Chez les patients, jamais ces mesures ne furent enregistrées, même après qu'ils aient compris lesquels des paquets étaient les plus risqués. Comme ces patients ont souvent des difficultés à évaluer les conséquences futures de leur choix, ils continuèrent de sélectionner les mauvais paquets même après avoir explicitement identifié la stratégie la plus efficace. Cette étude suggère que les processus délibérés seuls, lorsqu'ils ne sont pas sous l'influence de processus parallèles à la conscience, ne suffisent pas pour assurer la prise de décision optimale. Les résultats de l'expérimentation montrèrent que les personnes souffrant de déficits émotionnels (patients) finirent par perdre de l'argent contrairement aux

personnes normales. Il semble que l'usage seul du processus cognitif et délibéré (quadrant I) est insuffisant pour assurer la décision optimale.<sup>35</sup>

#### 1.4.1) Interaction entre les systèmes

La décision, qu'elle soit économique ou d'une autre nature, est le fruit d'une continue interaction entre les quatre différents quadrants du tableau 1.3. Trois aspects de cette interaction sont notables : la collaboration, la compétition et l'auto-inférence. Une décision ou un choix appropriés doivent être le fruit d'une efficace *collaboration* entre les quadrants ; taper une lettre à l'ordinateur nécessite à la fois les processus contrôlés (réfléchir à ce qu'on veut écrire) et les processus automatiques (taper la lettre). Comme nous l'avons souligné à la section précédente, la collaboration entre les processus cognitifs et affectifs est essentielle à la prise de décision; en certaines circonstances, la raison seule (quadrant I) ne suffit pas pour qu'une décision soit adéquate; les processus automatiques et affectifs importent aussi. Mais il est également vrai que trop d'émotions peuvent altérer le jugement. Par exemple, la colère peut diminuer la perception que l'on a du risque alors qu'un excès de tristesse l'amplifiera (Lerner et Keltner, 2001 cités dans Camerer *et al.*, 2003). Et sous forte influence des émotions, les gens vont souvent faire le contraire de ce qu'ils savent être le mieux pour eux (Loewenstein, 1996 cité dans Cohen et Blum, 2002). On peut néanmoins affirmer que la plupart du temps, nos choix sont le fruit d'une collaboration appropriée entre nos différents processus mentaux.

La *compétition* entre les différents quadrants implique que ceux-ci puissent parfois ne pas collaborer efficacement. La compétition entre les processus automatiques et contrôlés a préalablement été démontrée par le critère de Sloman (voir section 1.3.1); nous avons souligné à la section 1.3.1 que les

<sup>35</sup> Mais il arrive néanmoins que l'usage presque exclusif du quadrant I puisse effectivement permettre le choix optimal. Par exemple, Gneezy et Potters, 1997, cités dans Camerer *et al.*, 2003 ont montré que la

processus automatiques et contrôlés peuvent se faire compétition l'un et l'autre : devant la résolution d'un problème donné, la compétition implique que l'on puisse obtenir deux réponses différentes, l'une en provenance des processus automatiques, et l'autre relevant des processus contrôlés. On sait qu'une impression (initiée par les processus automatiques) peut parfois avoir trop d'influence et nous induire en erreur comme l'illustre l'erreur de la conjonction. Par contre, grâce à la rapidité de ces mêmes processus, on peut prendre des décisions à des coûts neurologiques relativement plus faibles. Par ailleurs, la compétition entre les systèmes affectifs et cognitifs s'illustre naturellement par le phénomène de la tentation. Nous abordons cette question à la section 2.2.3. Mentionnons ici que l'apparition de certains *cues* peut initier le désir de consommation (quadrant IV) et que les processus cognitifs (quadrant I) peuvent le contenir. La connaissance des conséquences ultérieures de certains types de consommation (achats compulsifs, drogues, sucreries, etc.) peut en effet réduire l'influence du quadrant IV. La compétition entre le système affectif et cognitif n'est pas encore très bien comprise. Il semble que le striatum, faisant partie d'une plus large région appelée noyau gris central, jouerait un rôle important. L'ampleur de l'intensité des émotions implique différents résultats : à un niveau d'intensité faible, les émotions serviront « d'avertisseurs ». À un niveau intermédiaire, la personne prend conscience d'un conflit entre le cognitif et l'affectif. À un niveau supérieur, les circuits neuronaux favoriseront la victoire des émotions sur la raison (Carter, 1999 cité dans Camerer *et al.*, 2003). Corriger une première impression par le biais des processus délibérés n'est donc pas une tâche facile. Cela nécessite de véritables efforts. La lutte entre les processus automatiques (rapides) et les processus délibérés (lents) est un combat inégal et les impressions automatiques auront l'avantage la plupart du temps (Camerer *et al.*, 2003).

L'*auto inférence* ou *sense-making*<sup>36</sup> correspond aux inférences que l'on fait sur ce qui nous entoure, sur les fondements de nos propres croyances, agissements ou préférences. Le quadrant I, qui est le centre de la pensée consciente, est particulièrement important pour l'accomplissement de cette tâche. Cependant, on sait que le quadrant I ne permet pas toujours un très bon accès introspectif. Il n'est donc pas étonnant qu'il attribuera parfois les causes d'un comportement donné à un processus conscient et délibéré alors que ce n'est pas le cas. Les quadrants III et IV opèrent en parallèle. Il est donc fréquent de croire qu'un comportement origine d'un processus conscient et délibéré alors que ce n'est pas le cas<sup>37</sup>. Certaines recherches effectuées à l'aide d'électroencéphalogrammes ont démontré que ce qu'on croit être un mouvement issu de la conscience, peut être en fait initié bien avant le moment précis où la pensée consciente n'intervienne :

« Because the behavior reliably follows the intention, we feel that this freely willed intention is causing the action, but in fact, both the sensation of intention and the overt action are caused by prior neural events which are inaccessible to consciousness. »<sup>38</sup>

Certaines recherches suggèrent que c'est l'hémisphère gauche du cerveau humain qui aurait la fonction de donner un sens aux événements, qui en serait « l'interprète ». Ce dernier serait constamment occupé à trouver une signification à ce qui nous entoure, même lorsqu'il n'y en a aucune; il nous pousserait donc à continuellement commettre des erreurs interprétatives (Gazzaniga, 1998). Dans une expérimentation rapportée par Gazzaniga (1998), on demanda à des personnes de prédire correctement la direction que

<sup>36</sup> Nous avons traduit *sense-making* par le terme « auto-inférence ». La portée du terme *sense-making* est un peu plus large puisqu'elle déborde de l'interprétation plus limitée du terme « auto » et implique qu'on puisse chercher à donner un sens aux événements qui nous entourent « en dehors de nous-mêmes », que l'on en soit l'instigateur ou non. Même si nous utilisons le terme « auto inférence », nous nous référons à l'interprétation que Camerer *et al.* font du terme *sense-making*.

<sup>37</sup> Cette idée n'est pas nouvelle en économie comme en fait foi Hayek : « ... Il n'est pas possible pour le cerveau humain de toujours s'expliquer pleinement lui-même... Il n'est pas possible pour lui de rendre compte de – ou de prédire – l'interaction d'un grand nombre de cerveaux humains... » Chap. 5 : La vanité fatale in *La présomption fatale*, page 109 (1988, p. 109 et 1952, 8.66/8.86).

<sup>38</sup> Camerer *et al.*, 2003.

prendrait un faisceau lumineux (vers le haut ou vers le bas d'un écran). Le faisceau apparaissait, d'une façon aléatoire, 80 pour cent des fois vers le haut de l'écran. Alors qu'ils auraient simplement dû sélectionner le bouton du haut afin de maximiser leur chance, les gens persistèrent à penser qu'il existait une séquence qu'ils pourraient trouver, ce qui au bout du compte ne leur donna la bonne réponse que dans seulement 68 pour cent des cas. Pourtant, des expériences similaires faites sur des rats ont démontré que ceux-ci obtenaient de meilleurs résultats en choisissant tout simplement et continuellement le bouton du haut, maximisant ainsi l'espérance de gain. Pour montrer le rôle « interprétatif » que joue l'hémisphère gauche, Wolford, Miller et Gazzaniga (2000) demandèrent à des patients souffrant d'une déconnexion calleuse<sup>39</sup> (*split brain*) de jouer à un jeu similaire. Ils observèrent que les résultats de l'hémisphère gauche de ces patients se comparaient aux résultats des personnes normales, c'est-à-dire qu'ils cherchaient des séquences alors qu'il n'y en avait pas et que leur performance était ainsi moins bonne que celle des rats.<sup>40</sup>

Par contre, Wolford et ses collègues constatèrent que les réponses données par l'hémisphère droit se rapprochaient de celles des rats, un peu comme si l'hémisphère droit ne se contentait que de maximiser les gains sans chercher à interpréter quoi que ce soit. Un autre exemple illustrant les limites de l'auto-inférence est donné par Fischman et Foltin (1992)<sup>41</sup>; à partir de deux leviers reliant leur bras à une intraveineuse, des drogués avaient la possibilité de s'administrer des doses de cocaïne. Pour ce faire, ceux-ci n'avaient qu'à appuyer fréquemment en sélectionnant le levier de leur choix. Selon le jour, chacun des deux leviers pouvaient masquer différentes combinaisons variables

<sup>39</sup> On sait qu'en ne regardant que de l'œil gauche, toute l'information captée se dirige vers l'hémisphère droit du cerveau. Et l'information captée par l'œil droit est dirigée vers l'hémisphère gauche. En fermant l'œil gauche de ces patients, on peut ainsi voir ce que cherche à faire l'hémisphère gauche.

<sup>40</sup> Il est intéressant de constater que cette caractéristique humaine à tout vouloir expliquer est bien connue de certains investisseurs : « When science responds by saying there is no known answer, or worse yet, there is no answer, people are not satisfied... In such circumstances people are perfectly willing to make up answers, or even to pay others to make them up for them (Loewenstein, 1996, p. 316).

de récompenses (allant d'un placebo à de plus fortes doses de 8 à 15 mg de cocaïne). Au cours d'une séance particulière, on plaça un placebo dans un levier et une très petite dose de cocaïne (4mg) dans l'autre. Lorsque interrogés, les participants crurent à tort qu'il s'agissait de deux placebos. Ils se déclarèrent indifférents entre l'un ou l'autre des leviers. Pourtant, lorsqu'on fit le décompte des choix de levier en fin de journée, on observa que le levier qui masquait de petites doses de cocaïne avait été significativement préféré à l'autre. Outre que cet exemple montre que les préférences des drogués n'étaient pas initiées par des processus conscients, il montre également comment les capacités de *sense-making* sont limitées.

En somme, les erreurs d'auto-inférence illustrent les limites de nos capacités cognitives; selon Camerer *et al.*, ces erreurs plaident pour des hypothèses économiques plus réalistes<sup>42</sup> :

« Because people lack introspective access to the processes that produce such bias, they are unable to correct for them even when they are motivated to make impartial judgments and decisions... Even though much of the brain's activity is "cognitively inaccessible", we have the illusion that we are able to make sense of it, and we tend to make sense of it in terms of Quadrant I processes. This is exactly why we need to rely on neuroscience rather than intuitive sense making to figure out what causes behavior. ».<sup>43</sup>

#### 1.4.2) Exemple d'interactions à l'intérieur du cadre bidimensionnel

##### Le plat de sushi

Comme nous l'avons vu, toute décision ou action implique une interaction entre ces différents quadrants. Afin d'illustrer cette interaction, Camerer et ses collègues donnent l'exemple d'une personne à qui l'on offre un plat de sushi ;

<sup>41</sup> Berridge, 1999, p. 531 en référence à Fischman et Foltin, 1992.

<sup>42</sup> Par exemple, Camerer *et al.*, 2003 suggèrent que certains modèles économiques qui présupposent que la discrimination du marché du travail contre les minorités correspond à un goût ou à une croyance particulière pourraient être améliorés par l'apport des neurosciences.

<sup>43</sup> Camerer *et al.*, 2003, p. 25.

aussitôt, le quadrant III s'active et le nerf optique permet la transmission de l'information vers le cortex occipital où elle sera décodée en des lignes et en des formes. En aval, cette information sera utilisée dans le cortex visuel primaire où elle sera intégrée avec un ensemble de représentations d'objets qui permettront de l'identifier comme étant un sushi. Cette étape est extrêmement complexe et donne toutes sortes de difficultés aux chercheurs en intelligence artificielle qui tentent de la recréer sur des ordinateurs. Le quadrant IV s'activera à partir du moment où le cortex visuel primaire et d'autres régions comme le cortex orbitofrontal permettront d'identifier l'appréciation que la personne se fait de cette nourriture. L'histoire qu'a cette personne avec les sushis impliquera différentes réactions cérébrales. Si la personne a déjà été malade et qu'elle en a gardé un souvenir désagréable (aversion nauséuse), la région amygdalienne sera particulièrement active puisque cette région est souvent associée au dégoût et à la peur. Par ailleurs, si la personne a très faim, les neurones de l'hypothalamus s'activeront plus rapidement. En supposant qu'elle n'a rien contre l'idée de manger un sushi, le cortex moteur de la personne s'activera (quadrant III) lorsqu'elle tendra la main et prendra cette nourriture. Le quadrant IV entrera en action selon son appréciation de la collation. Aidé de cet exemple, on constate que les quadrants I et II ne se sont jamais activés. La consommation peut être initiée par des processus ne faisant pas appel à la délibération consciente. C'est un cas possible si la personne a fait son geste machinalement tout en parlant aux autres convives. Les quadrants I et II seraient sollicités si, par exemple, la personne a récemment lu que le sushi est un mauvais plat pour la santé. Dès lors, il est possible que les quadrants entrent en compétition (la personne aime manger du sushi mais doit considérer les conséquences de son geste).

### 1.5) Conclusion

La neuroéconomie est le prolongement naturel de l'économie comportementale car elle permet d'augmenter le degré de réalisme des fondements de la décision économique. Des outils tels que les psychopathologies chez l'humain, les dommages cérébraux, l'imagerie cérébrale, la mesure d'un seul neurone et les stimuli électriques permettent tous une meilleure compréhension de ce que le cerveau cherche à accomplir. Comme nous partageons des ancêtres communs avec les autres espèces animales, la division du cerveau suggérée par MacLean implique que toute étude faite sur des animaux peut nous renseigner sur le fonctionnement du cerveau humain; cela élargit en soi les frontières de la neuroéconomie et lui incorpore nécessairement une dimension évolutionniste. D'importantes avancées dans les connaissances que nous avons des processus mentaux ainsi que de récents développements dans les neurosciences devraient permettre de reconsidérer certaines hypothèses et de tenter de nouveaux développements dans les sciences économiques. L'approche bidimensionnelle du fonctionnement cérébral et certaines recherches suggèrent que les modèles économiques uniquement basés sur les processus de délibérations conscientes et rationnelles sont incomplets. Il est aujourd'hui accepté que nombre de phénomènes psychologiques sont à la fois volitifs et non délibérés. Des chercheurs issus des neurosciences, de la psychologie et de l'économie, ont entrepris d'établir les circonstances dans lesquels l'un ou l'autre de ces processus entre en action et ainsi d'identifier leurs corollaires cérébraux. Aidé d'outils des neurosciences, il est donc possible de mieux comprendre ce qui initie l'action ou la décision, de regrouper des phénomènes que l'on croyait différents et d'en séparer d'autres qu'on croyait identiques. On sait que différents processus influencent la décision et chacun ont leur rôle et leur importance. Les processus automatiques sont le mode par défaut de notre cerveau. Ils ont une incidence directe sur notre choix de consommation et opèrent le plus souvent en parallèle à notre conscience. Sans ceux-ci, nous sommes lents et inefficaces. Le fait que notre cerveau soit affecté par la

plasticité, la modularité et la spécialisation, pourrait avoir des répercussions fondamentales sur notre façon d'appréhender la décision de nature économique. Ces processus impliquent qu'un choix pourrait changer selon le contexte, même si la structure du problème est équivalente. L'importance des processus automatiques est donc fondamentale. En fait, même la décision stratégique optimale, que l'on croit uniquement associée aux processus contrôlés, nécessite l'apport des processus automatiques et affectifs. Les émotions ont une influence continue sur la décision et comme les recherches faites à l'aide de patients souffrant de déficits émotionnels l'ont montré, elles sont même absolument nécessaires à la prise de décision optimale. Par contre, selon le contexte, elles peuvent également être aveuglantes au point d'inciter quelqu'un à choisir un paiement qui va contre son intérêt propre. Nous reviendrons sur cette question au chapitre 3.

La connaissance des différents processus mentaux et de leurs corollaires physiologiques nous permet de jeter un nouveau regard sur certains fondements du comportement. Aidé des différentes connaissances et techniques des neurosciences, il est aujourd'hui possible d'observer l'activation cérébrale au moment de la décision économique; le lieu et le degré d'activation nous renseignant sur les processus guidant l'action. De récents développements issus des neurosciences permettent d'observer ce que Jevons n'aurait jamais cru possible : l'observation, la mesure et l'influence d'une émotion lors d'une transaction économique. Nous aborderons cette question aux prochains chapitres. L'approche bidimensionnelle du fonctionnement cérébral s'inscrit dans une perspective descriptive du comportement économique et se veut une suite logique à l'économie comportementale. Les neurosciences peuvent donc informer la théorie économique et cette transmission de connaissances est nécessairement mutuelle. Nous abordons maintenant ces considérations en poursuivant avec le thème des préférences.

## CHAPITRE II

### LES PRÉFÉRENCES ET LA NEUROÉCONOMIE : SUPPORT ET APPORT À LA THÉORIE ÉCONOMIQUE.

#### 2.0) Introduction

Dans ce chapitre, nous chercherons à comprendre les raisons pour lesquelles des chercheurs issus des neurosciences, de l'économie et de la psychologie se sont récemment unis pour initier l'émergence de la neuroéconomie (section 2.1). Nous verrons comment des concepts de base de la théorie économique, tels que l'utilité subjective et l'espérance de l'utilité peuvent être d'intéressants outils pour les chercheurs issus des neurosciences, et comment un retour d'ascenseur est rendu possible par l'utilisation du cadre d'analyse de la théorie prospective pour interpréter certains résultats (sections 2.2 à 2.2.1).

Nous tenterons de montrer les liens entre le concept de l'utilité et le rôle joué par la dopamine dans la prise de décision (section 2.2.2). Nous verrons en quoi les neurones dopaminergiques, et leurs répercussions automatiques, peuvent influencer la décision de consommation et en quoi cela est d'intérêt pour les économistes (section 2.2.3). Nous mentionnerons d'ailleurs au passage comment certaines théories économiques ont déjà été fortement influencées par des découvertes de la neuroéconomie. Nous aborderons ensuite la question des préférences révélées et nous situerons cette théorie par rapport aux récentes avancées de la psychologie et des neurosciences (section 2.2.4). Nous verrons par la suite en quoi une approche neuroéconomique, qui incorpore la notion de préférences implicites, permet de dépaqueter le concept de préférence (section 2.2.5). Après avoir abordé la question des processus homéostatiques et de l'utilité subjective engendrée par le gain monétaire

(sections 2.2.6 et 2.2.7), nous tenterons d'établir le lien entre la neuroéconomie et quelques falsifications de l'espérance de l'utilité. Nous terminerons en présentant les travaux de recherches des neurosciences et de la neuroéconomie en lien avec certaines dimensions de la théorie prospective (section 2.3).

### 2.1) Approche holistique de la neuroéconomie

C'est par la fusion de trois disciplines : les neurosciences, la psychologie et l'économie, qu'a été rendue possible l'émergence de la neuroéconomie. Si l'origine et l'évolution des liens entre la psychologie et l'économie sont aujourd'hui établis et reconnus (voir par exemple Zwick, Erev et Budescu, 1999; Van Raij, 1999; Lewin, 1994; Lopes, 1994; Rabin, 1998), ceux concernant le croisement des neurosciences et de l'économie sont plus récents et *a priori* moins naturels. En fait, ce sont des chercheurs issus des neurosciences qui, les premiers, ont construit un pont entre les deux disciplines. Mais cette rencontre n'aurait pu survenir sans l'évacuation préalable, au sein des neurosciences cognitives, de la conception behavioriste pour laquelle le cerveau n'est qu'un organe stimuli-réponses. C'est l'abandon graduel de ce paradigme, que certains ont appelé *la théorie du réflexe* ou le *paradigme Descartes-Sherrington* (Glimcher, 2003), qui a permis cette fusion. En effet, Descartes (1637, cité dans Glimcher, *op. cit.*, p. 28), qui a eu une profonde influence sur la méthodologie des neurosciences, considérait que le comportement humain était constitué de deux volets: le simple et le complexe. Pour Descartes, seul le comportement simple, initié par un processus de stimuli-réponses était compréhensible et explicable. Le comportement complexe, ou volitif, était considéré par lui comme étant imprévisible et incompréhensible parce que né d'une âme humaine immatérielle et hors de portée du lien sensation action. C'est sous l'influence de Descartes que le physiologiste anglais Charles Scott Sherrington construisit par la suite les

bases biologiques du comportement, avec le réflexe comme ancrage à l'analyse, parce que cette approche, quoique incomplète, était néanmoins déterministe.

La *théorie du réflexe* servit de base méthodologique pour des générations de chercheurs issus des neurosciences; elle n'eut cependant guère d'emprise sur les recherches en matière de comportement animal. Dès 1966, Emlen (cité dans Glimcher, 2002), puis MacArthur et Pianka (*ibid.*, p. 324), suggèrent que la décision prise par un animal concernant son choix de nourriture, sa sélection de proie, implique un problème d'optimisation qui peut être quantifié en termes mathématiques. Charnov et Orrians (1973, *ibid.*, p.324), puis Stephens et Krebs (1986, *ibid.*, p. 324), ont développé un modèle de sélection optimale de la proie en des termes mathématiques. D'autres recherches, Herstein (1961, *ibid.*, p. 327), Harper (1982, *ibid.*, p. 329), Maynard Smith (1982, *ibid.*, p. 329) suggèrent qu'on peut non seulement associer le choix optimal de l'animal à ses chances de survie, mais également le prévoir par des modèles mathématiques. Pour l'écologie comportementale, l'évolution conduit les espèces à résoudre des problèmes d'optimisation, et à le faire efficacement. À partir de là, l'association entre les choix optimaux et leurs corollaires cérébraux devient plus facile à accepter et un premier pont entre les approches économique et écologique est établi.

En neurobiologie, c'est Randy Gallistel et ses collègues (1981, *ibid.*, p. 327) qui furent les premiers à utiliser une approche économique pour comprendre le cerveau. Pendant plus de vingt ans, ils étudièrent les décisions prises par des rats sous l'influence de stimuli électriques. À partir des années 1970, les limites du paradigme *Descartes-Sherrington* furent sérieusement remises en cause par le théoricien des neurosciences David Marr (1976, 1982, cité dans Glimcher, 2003, p. 133). Marr considérait le cerveau comme une machine, comme un dispositif de traitement d'information plutôt que comme un simple organe stimuli-réponses. Le cadre de *Descartes-Sherrington*, qui ne pouvait analyser le comportement que dans ses dimensions les plus simplificatrices, se

devait d'être élargi. Marr plaida pour un nouveau paradigme à l'intérieur duquel la compréhension et l'explication des fonctions du cerveau joueraient un rôle primordial :

« To understand the relationship between behavior and brain one has to begin by defining the function, or computational goal, of a complete behavior. Only then can a neuroscientist determine how the brain achieves that goal.»<sup>44</sup> [...] « An algorithm is likely to be understood more readily by understanding the nature of the problem being solved than by examining the mechanism (and the hardware) in which it is embodied.»<sup>45</sup>

Le tableau 2.1 illustre les trois niveaux, selon Marr, permettant de comprendre n'importe quel dispositif de traitement d'information. Les trois étages de cette figure peuvent illustrer les contributions respectives des trois disciplines que sont l'économie, la psychologie et les neurosciences. L'économique, de par ses descriptions d'objectifs et de choix formels, contribue à la théorie computationnelle<sup>46</sup>. La psychologie et les neurosciences correspondent respectivement aux deuxième et troisième étages du tableau 2.1. C'est donc afin de combler un vide théorique que des chercheurs issus des neurosciences ont commencé à trouver dans les théories économiques une complémentarité qui s'est, à certains égards, déjà avérée pour eux fructueuse :

« Until recently, the sensory-motor reflex was tacitly accepted by many physiologists as an appropriate model for describing the neural processes that underlie complex behavior. More recent findings indicate that, at least in some cases, the neural events that connect sensation and movement may involve processes other than classical reflexive mechanisms...»<sup>47</sup>.

---

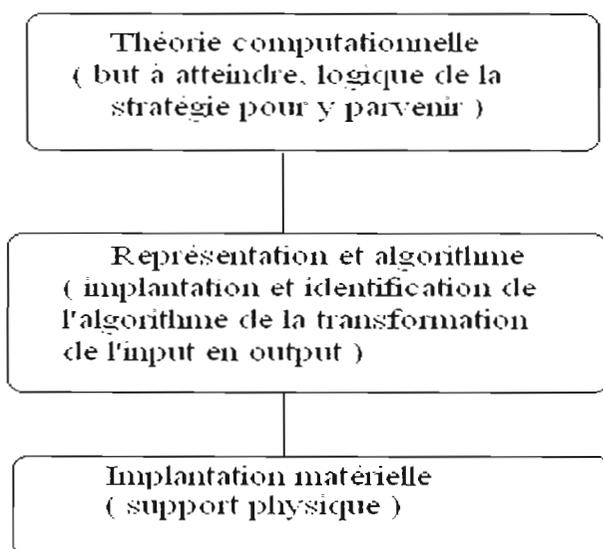
<sup>44</sup> Glimcher, 2003. *Ibid.*, p. 143

<sup>45</sup> Glimcher, 2003. *Ibid.*, p. 142

<sup>46</sup> De l'anglais *computational*. Nous utilisons ce terme pour l'ensemble du texte.

<sup>47</sup> Platt et Glimcher, 1999, p. 233.

Tableau 2.1: Les trois niveaux de compréhension d'un dispositif de traitement de l'information selon l'approche de Marr



Source : Glimcher<sup>48</sup>

Mais l'orientation future de cette complémentarité pourrait faire ressortir des tensions déjà existantes au sein de l'économie. Des chercheurs issus des neurosciences, Wise (1989, 2002), Shizgal (1998), Montague et Berns (2002), Glimcher (2002, 2003, 2004), Fiorillo, Tobler et Schultz (2003), Zalla *et al.* (2000), Gallistel *et al.*, (1981), Gallistel (1994), Platt et Glimcher (1999), ainsi que bien d'autres encore, ont déjà initié l'émergence d'une branche de la neuroéconomie, que les économistes approuvent ou non leur démarche. Dans certains cas, les outils théoriques utilisés pour la spécification de buts computationnels, outils qui furent parfois falsifiés par l'économie expérimentale (la rationalité de l'agent, l'espérance de l'utilité, les calculs bayésiens, l'équilibre de Nash), ont été par ailleurs partiellement corroborés par certaines de leurs recherches. Et l'utilisation de ces outils suscite

<sup>48</sup> *Ibid.* Adaptation et traduction par l'auteur du présent mémoire.

passablement d'enthousiasme. Par exemple, en référence aux modèles de décisions rationnelles de Von Neumann et Morgenstern (1944), Platt et Glimcher (1999) soulignent :

« These models which carry information about the environment, are extracted in advance of response selection, aid in the interpretation or processing of sensory data. »<sup>49</sup>

Au sein des neurosciences, l'utilisation d'une nouvelle théorie computationnelle, ayant pour base des formalisations issues de la théorie économique, suscite passablement d'enthousiasme pour l'avenir; des comportements qui n'auraient jamais pu être expliqués à l'intérieur de l'ancien paradigme pourraient l'être dans le futur: « [...] models of decision making rooted in economic theory should provide powerful insights into brain function over the next half century. »<sup>50</sup>

L'objet de ce mémoire n'étant pas de faire ressortir ces tensions ou de conjecturer sur l'avenir, nous nous contenterons de souligner ici que la neuroéconomie et l'économie comportementale peuvent parfois évoluer en complémentarité, parfois de façon parallèle. Quoi qu'il en soit, la complémentarité interdisciplinaire ne peut manquer d'être soulignée :

« What is striking about explanations of choice behavior by economists, psychologists, and neurobiologists is the different levels at which they operate. The economic approach attempts to describe globally all choice behavior with a single logically consistent formalism. The psychological approach examines the ways in which subjective and objective estimates of value differ and posits psychological modules that account for these observed behavioural preferences. The neurobiological explanation starts with the simplest possible neural circuitry that can account for the simplest measurable elements of behavior. »<sup>51</sup>

---

<sup>49</sup> *Ibid.*, p.233.

<sup>50</sup> Glimcher, 2002. *Ibid.*, p. 331.

<sup>51</sup> Glimcher et Rustichini, 2004.

Ce qui est intéressant, c'est que ces disciplines, qui ont depuis longtemps évolué en parallèle, cherchent à expliquer, à différents niveaux, certains aspects des processus décisionnels. Incorporer à l'analyse ces aspects pourrait élargir considérablement la compréhension fondamentale du comportement économique. Dans les sections qui suivent, nous chercherons à illustrer comment cette complémentarité permet d'enrichir la compréhension des fondements de la décision économique.

## 2.2) Les préférences et l'utilité subjective

Pour illustrer l'importance de tout modèle mental de décision, rappelons ici une fable, celle de l'âne de Buridan, qui avait faim et soif. Placé à égale distance d'un sceau d'eau et d'une botte de foin, il se devait de prendre une décision, mais laquelle : manger ou boire ? Incapable de choisir, la pauvre bête fut terrassée par l'indécision. Elle resta sur place pendant des jours, immobile, et finit par mourir affamée et assoiffée. Dans la nature, tout choix effectué par un animal a des répercussions sur ses chances de survie. L'objectif ultime et le sens pratique de l'âne de Buridan laissent perplexe; son cerveau semble étrangement « programmé » et c'est pourquoi il ne survivra pas. Damasio (1994) relate l'histoire d'un patient souffrant de dommages préfrontaux<sup>52</sup> qui pouvait passer toute une soirée à évoquer les avantages et les inconvénients à choisir un restaurant plutôt qu'un autre. Il était incapable d'arrêter sa décision. L'âne de Buridan et le patient du Docteur Damasio illustrent, bien malgré eux, à quel point le cerveau est à la base même de tout processus décisionnel. Pour expliquer la décision d'une façon générale, les économistes ont développé le concept de l'utilité. L'utilité représente une structure mentale de prise de décision qui permet de représenter le fondement des choix et des décisions encourues. L'économie expérimentale, la psychologie et les neurosciences peuvent, en validant ou en falsifiant le modèle

---

<sup>52</sup> Comme nous l'avons souligné à la section 1.4, les personnes souffrant de dommages préfrontaux ont des difficultés à traiter les informations émotionnelles. Elles sont en « déficits émotionnels ».

de l'utilité, permettre d'enrichir la compréhension du comportement économique.

Dans le modèle de l'utilité, un choix arrêté est supposé avoir une plus grande utilité que ses alternatives concurrentes. L'utilité peut donc être conçue comme un système mental où s'effectue une estimation subjective de différentes alternatives avec comme hypothèse la maximisation. Pour le philosophe anglais Jeremy Bentham (1748-1832), c'est la dimension hédonistique de l'utilité qui était au cœur de la prise de décision. Le choix se devait d'être orienté par la recherche du plaisir et l'évitement des déplaisirs. Plus récemment, c'est à partir du moment où fut acceptée la métaphore du cerveau comme dispositif de buts computationnels que de nouveaux champs d'études furent développés par des chercheurs issus des neurosciences et de la psychologie. Ces recherches, reliées à la mémoire, aux processus de décisions et de solutions de problèmes, faisaient appels à certaines conceptions propres à l'économie, telles que l'utilité, comme le soulignent Camerer et Loewenstein (2002) : « ...These new topics were more obviously relevant to the neoclassical conception of utility maximization than behaviorism had appeared to be... ».<sup>53</sup> L'utilisation du concept de l'utilité pour expliquer les choix et les comportements complexes a donc permis d'établir certains ponts entre les neurosciences et l'économie. Ces liens sont cependant récents :

« The central concept in modern economic theory is the notion of subjective utility : Preferences must be described as subjective properties of the chooser. Surprisingly, the notion that preferences are represented in the nervous system, that these preferences are subjective, and that they guide the production of action has only recently entered the neurobiological mainstream ».<sup>54</sup>

L'évaluation des substrats biologiques de la valeur, celle partant des concepts d'utilité, est donc un enchaînement naturel. Les sections 2.2.1 à 2.2.7 illustrent cet enchaînement.

<sup>53</sup> Camerer et Loewenstein, 2002, p. 5.

### 2.2.1) Évaluation et sélection d'objectifs : vers une computation neuronale du concept standard de l'utilité

Certains aspects du comportement humain ont été décrits et graduellement formalisés autour d'un ensemble de règles ou d'axiomes par la théorie économique. Des hypothèses découlant du modèle de l'utilité ont permis de nombreuses avancées dans d'autres disciplines : pour être falsifiées ou redéfinies par la psychologie (Kahneman et Tversky, 1979; Kahneman, 1994), afin de prévoir et décrire le comportement animal (Maynard Smith, 1982), ou encore pour combler un vide théorique en neurosciences (Platt et Glimcher, 1999; Shizgal, 1998). Nous verrons maintenant en quoi la neuroéconomie permet un certain retour de l'ascenseur, validant parfois certains axiomes de la microéconomie, et les falsifiant en d'autres occasions.

#### La fonction d'utilité

L'hypothèse standard d'une fonction d'utilité repose sur l'idée que l'utilité marginale est positive et décroissante, c'est-à-dire que  $U' > 0$  et  $U'' < 0$ . La figure 2.1 illustre une telle fonction d'utilité conventionnelle. Une expérimentation remarquable de Platt et Glimcher (1999) faite sur des singes rhésus permet d'illustrer que le cerveau pourrait réellement encoder ce genre de fonction. La figure 2.2 permet de dénoter une similarité entre ce concept de base de la science économique et les mesures enregistrées à partir du cerveau des singes; au fur et à mesure de l'arrivée de récompense en jus, l'activité neuronale augmente graduellement mais à un rythme décroissant jusqu'à l'atteinte d'un plateau. Bien que ce graphique ne prouve rien du tout, il permet néanmoins d'illustrer comment certains concepts économiques peuvent être combinés avec des recherches neurologiques.

---

<sup>54</sup> Glimcher et Rustichini, 2004, p. 448.

Figure 2.1 : Une fonction d'utilité conventionnelle

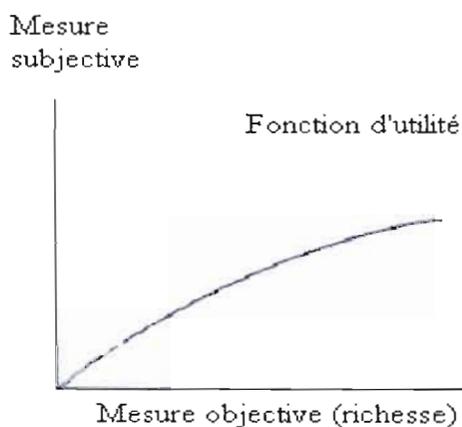
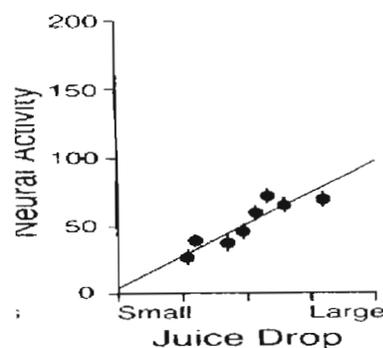


Figure 2.2 : Réminiscence d'une fonction d'utilité dans le cerveau



Source : Glimcher <sup>55</sup>

### L'espérance de l'utilité

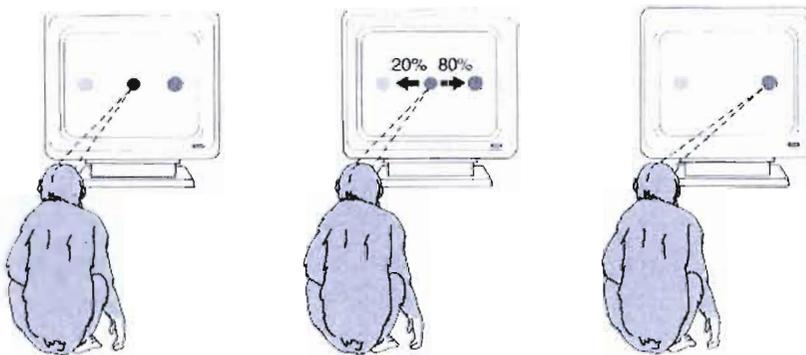
L'espérance de l'utilité est un autre concept économique qui a su capter l'attention de chercheurs qui veulent comprendre le fonctionnement du cerveau. Lorsque définie empiriquement, l'espérance de l'utilité peut être testée en laboratoire. Du point de vue des neurosciences, pour que le concept de l'espérance de l'utilité soit empiriquement démontrable, deux types d'évaluations subjectives doivent être comptabilisées dans le cerveau. L'une concerne la valeur attribuée à un choix, l'autre correspond aux probabilités lui étant associées. Une étude pionnière de Platt et Glimcher (1999)<sup>56</sup> a permis, en utilisant le concept de l'espérance de l'utilité et la technique de mesure d'un seul neurone, de démontrer que la région latérale intrapariétale du cerveau des singes rhésus pouvait à la fois encoder les probabilités et la valeur associées à une récompense. Platt et Glimcher enseignèrent à des singes à sélectionner et

<sup>55</sup> Glimcher, 2003. *Ibid.*, p. 262.

<sup>56</sup> Platt et Glimcher, 1999. Platt et Glimcher se sont en partie inspirés des recherches faites par Newsome *et al.*, 1989.

à choisir entre deux points lumineux. Lorsqu'ils effectuaient le bon choix, par un mouvement de tête, les singes obtenaient une récompense (du jus). Pour maximiser son utilité, un singe se devait de connaître les probabilités antérieures associées à son choix, mais aussi la valeur de la récompense. Ainsi, si le point lumineux ne se déplaçait que rarement dans une direction donnée, les anticipations du singe ne l'incitaient évidemment pas à regarder vers celle-ci. Et si la récompense associée à une direction était insignifiante, la motivation du singe à regarder vers celle-ci était proportionnelle. Au cours de l'expérimentation, le cerveau du singe devait donc procéder à une estimation s'apparentant à une forme d'espérance de l'utilité. La figure 2.3 illustre l'encodage des probabilités faite par un singe ainsi que son choix rationnel, maximisant ses chances d'obtenir récompense.

Figure 2.3 : L'encodage des probabilités de récompenses



Source : Glimcher (2003)

La tâche expérimentale s'effectua par blocs de 100 essais. Dans certains de ces blocs, les probabilités de renforcement d'un mouvement s'établissaient à 80 % vers la droite et à 20 % vers la gauche. Dans d'autres blocs, ces probabilités étaient inversées. Ce faisant, les probabilités antérieures et postérieures à chacun de ces blocs se devaient d'être continuellement estimées par le cerveau du singe et pouvaient donc ne plus concorder. L'objectif était de varier les probabilités de récompense tout en maintenant constants les stimuli

visuels et les mouvements moteurs. Ainsi, il serait alors possible de vérifier si l'activation des neurones de la région latérale intrapariétale était corrélée d'une façon ou d'une autre avec les probabilités antérieures et postérieures. Ce que Platt et Glimcher découvrirent est fascinant :

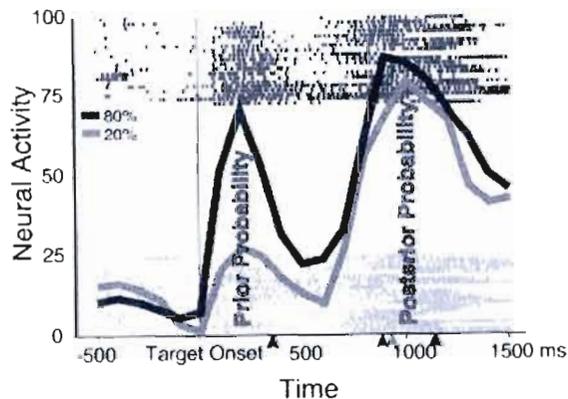
« Once we had selected from our data set only those trials on which the same stimulus was presented and the same movement was produced, what we found was very clear. When there was a high prior probability that an upward movement would be rewarded, LIP neurons associated with the upward movement responded very strongly as soon as the upper target was turned on. When there was a low prior probability that an upward movement would yield a reward, the same LIP neurons responded weakly to target onset... To us, this strongly suggested that prior probabilities were being encoded by LIP neurons.»<sup>57</sup>

En maintenant constants le stimulus et le mouvement des singes, Platt et Glimcher constatèrent que les probabilités de récompenses, et leurs changements, étaient corrélés avec l'activation neuronale (figure 2.4). Un signal non sensoriel et non moteur était donc encodé par le cerveau des singes. Platt et Glimcher interprétèrent ce signal comme étant une preuve que ceux-ci cherchaient d'abord et avant tout à maximiser leurs probabilités de gains. Cette interprétation n'aurait jamais pu être possible à l'intérieur du paradigme du réflexe (puisque'un même stimulus, c'est-à-dire une même quantité de jus et un même mouvement moteur, n'engendrait pas une même réponse neurologique).

---

<sup>57</sup> Glimcher, 2003. *Ibid.*, p. 259.

Figure 2.4 : Activité neuronale et probabilités de récompense



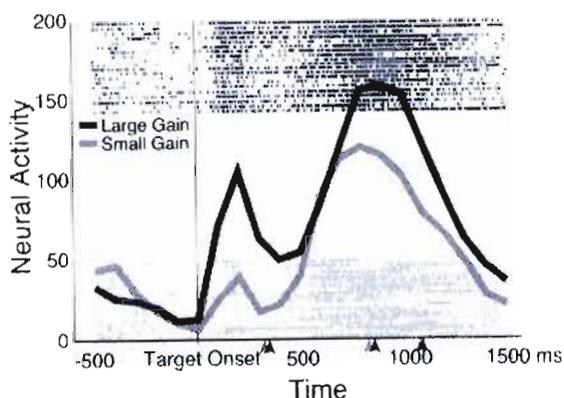
converger à la fin du bloc, ce qui montre que le cerveau des singes encode ce changement.

La figure 2.4 montre que pour un même mouvement moteur, l'activité neuronale est plus forte lorsque les probabilités antérieures de récompenses sont élevées (ligne noire) que lorsqu'elles sont faibles (ligne grise). La partie droite de la figure montre que lorsque les probabilités de récompenses sont modifiées pour atteindre 100 %, les probabilités antérieures et postérieures finissent par

Source : Glimcher (2003)

Après avoir montré que le cerveau des singes pouvait encoder les probabilités associées à une récompense, le deuxième volet de cette expérimentation consistait à vérifier si celui-ci était capable d'en faire autant avec la valeur de celle-ci. Platt et Glimcher maintinrent alors constants les mouvements, les stimuli et les probabilités de récompenses (qui furent fixées à 50 %). D'un bloc à l'autre, seules les quantités de récompense étaient changées. Dans certains des blocs, les quantités de jus obtenu étaient de 0,2 mL en regardant à gauche et de 0,1 mL en regardant à droite. Dans d'autres cas, les quantités étaient inversées. Les résultats (figure 2.5) illustrent que la valeur du gain fut aussi encodée par certains neurones. Alors que les éléments sensori-moteurs étaient maintenus constants, les neurones s'activèrent davantage lorsque l'espérance de gain était élevée.

Figure 2.5 : Activité neuronale et ampleur du gain



La figure 2.5 illustre que pour un même mouvement moteur du singe, l'activité d'un neurone n'est pas la même selon la valeur antérieure d'une récompense. La ligne noire montre que l'activité est plus élevée lorsque le gain passé fut plus grand et la ligne pâle montre une activité moindre lorsque les gains antérieurs furent plus faibles. En soi, ces résultats suggèrent que le cerveau encoderait la valeur associée à une récompense anticipée.

Source : Glimcher (2003)

Le fait que le cerveau des singes rhésus ait pu à la fois encoder l'ampleur des récompenses et leurs probabilités démontre deux choses. En premier lieu, que le concept de l'espérance de l'utilité<sup>58</sup> semble avoir trouvé un corollaire neurologique et qu'on ne peut donc le rejeter d'emblée en toute circonstance. En second lieu, que l'utilisation de certaines conceptualisations qui trouvent leur source dans la théorie économique peuvent élargir le cadre d'analyse des neurosciences ; puisque les mêmes stimuli-réponses n'ont pas ici les mêmes implications neurologiques, la théorie économique peut ici être utile aux neurosciences. L'économie et les neurosciences peuvent bénéficier l'une de l'autre, comme le soulignent Knutson et Peterson (2005), un retour d'ascenseur est aussi prévisible :

« As prescriptive approaches inform descriptive approaches, so can description inform prescription. As theory informs observation, observation informs theory. The dance of science continues. ».<sup>59</sup>

<sup>58</sup> Ou à tout le moins l'espérance de gain.

<sup>59</sup> Knutson et Peterson, 2005, p. 15.

Ces résultats, qui corroborent le modèle de l'espérance de l'utilité, montrent que dans certaines circonstances, le cerveau encode la probabilité et la valeur d'un gain comme le souligne Glimcher (2003):

« While these results can hardly be considered conclusive proof of anything, they do raise the possibility that even for a behavior as simple and deterministic as orienting toward a spot of light, economics may form the root of computational theory that will allow us to understand what the brain is trying to do when it makes a decision. Because we found evidence that neurons in area LIP encode both probability and value (or perhaps utility), it seems possible that neurons in area LIP encode something like classical expected utility.»<sup>60</sup>

#### Espérance de l'utilité et cerveau humain

Nous avons souligné, à la section 1.2.1, certaines similitudes entre le cerveau humain et celui des autres mammifères. Les résultats d'études faites sur ces derniers peuvent permettre de mieux comprendre le cerveau humain. Néanmoins, lorsque des études sont directement faites sur des êtres humains, cette inférence n'est évidemment plus nécessaire mais elle devient alors complémentaire. Ce qui est frappant, en regard des recherches de Platt et de Glimcher, c'est que des études subséquentes utilisant la technologie fMRI arrivent à des conclusions assez similaires chez l'humain. Par exemple, l'activation du cortex pariétal postérieur est corrélée avec l'ampleur du gain monétaire lors d'une loterie (Paulus, Hozack, Zauscher *et al.*, 2001) tandis que son anticipation serait positivement corrélée avec l'activation de la région du striatum ventral (Knutson et Peterson, 2005). Deux observations frappantes sont dénotées par Knutson et Peterson. Premièrement, le striatum ventral n'est activé que par l'anticipation d'un gain monétaire (figure 2.6) ; un gain, une perte, ou l'anticipation d'une perte monétaire n'ont apparemment pas d'effet sur cette région (figures 2.6 et 2.7). Le cadre d'interprétation de Knutson et Peterson diverge partiellement de celui de Platt et de Glimcher

---

<sup>60</sup> Glimcher, 2003. *Ibid.*, p. 263.

(1999) cependant. Knutson et Peterson interprètent leurs résultats dans l'optique de la théorie prospective de Kahneman et Tversky, c'est-à-dire que les perspectives de gains et de pertes sont supposées ne pas être traitées par les mêmes algorithmes et par les mêmes mécanismes neuronaux.<sup>61</sup>

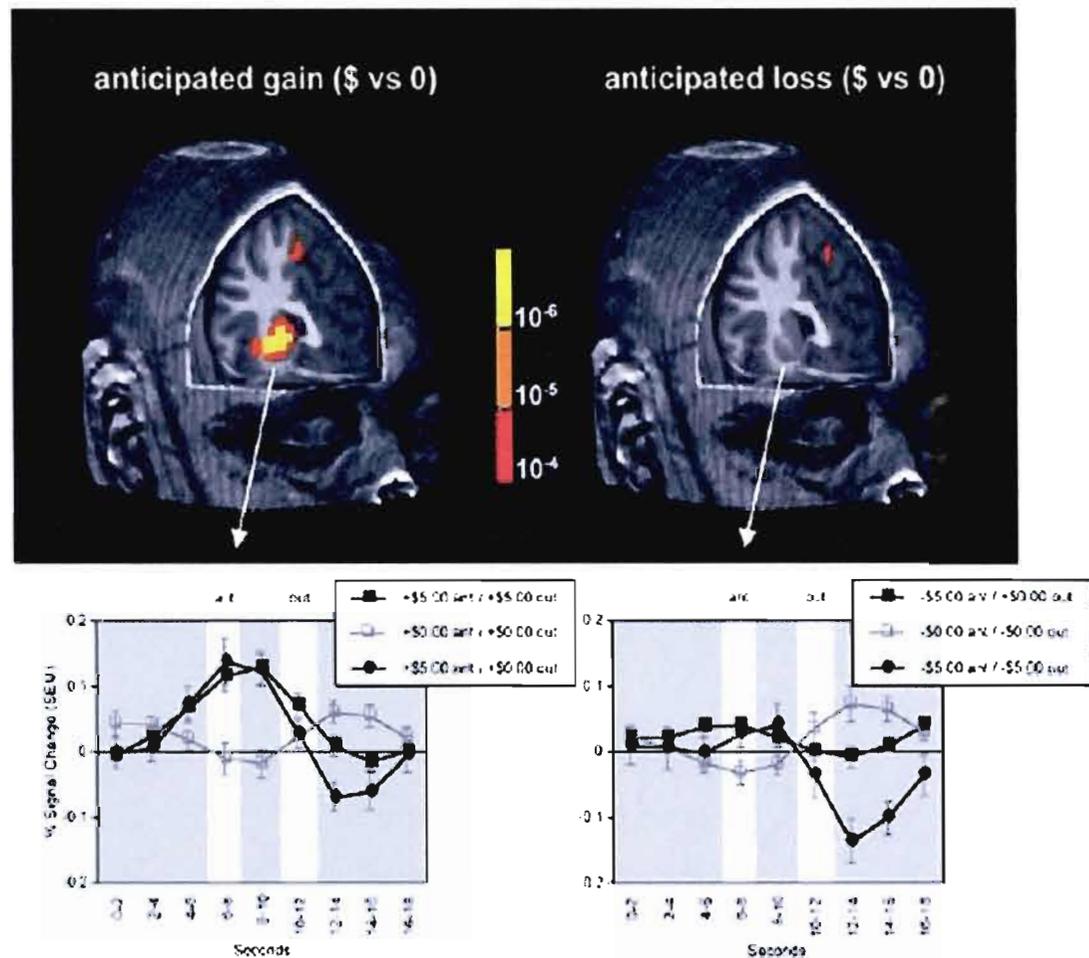
Une deuxième observation de Knutson et Peterson concerne la région médiale du cortex préfrontal. Selon eux, cette région enregistre les gains monétaires (figure 2.7). Et elle se désactive lorsque le gain monétaire est nul. Fait à noter, elle n'est activée ni par les anticipations (perte ou gain), ni par les pertes. Le cortex préfrontal n'est donc pas nécessaire pour l'anticipation d'une récompense préalablement apprise. Le cortex préfrontal peut cependant donner des *feedbacks* lorsque l'environnement change. Knutson *et al.* (2003)<sup>62</sup> font l'analogie suivante : ils comparent le noyau accumbens (dans lequel se trouve le striatum) à une pédale à gaz et la région médiale du cortex préfrontal à un volant qui peut orienter le choix vers la direction appropriée lorsque les paiements changent.

---

<sup>61</sup> La section 2.3.3 aborde les considérations neuroéconomiques de la théorie prospective. Pour plus d'explications concernant la théorie prospective, voir Kahneman et Tversky, 1979 ou encore Tversky et Kahneman, 1992.

<sup>62</sup> Knutson *et al.*, 2001, p. rc159.

Figure 2.6 : Anticipation d'un gain monétaire dans le striatum



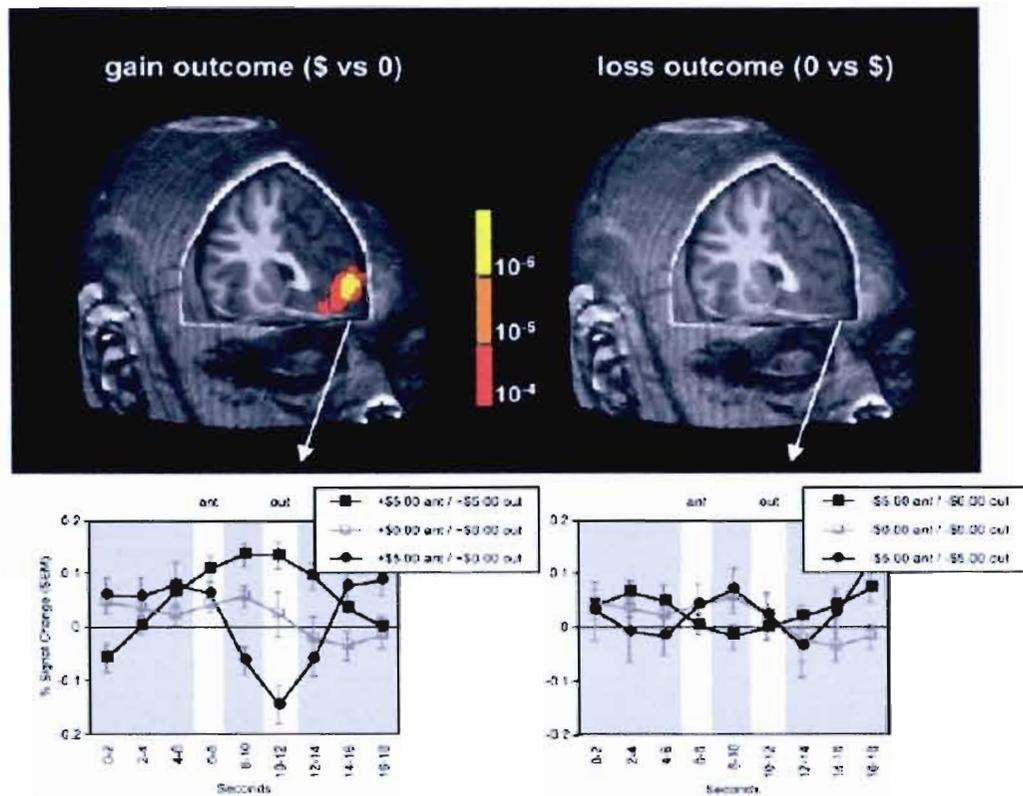
La partie gauche de la figure 2.6 montre la région du striatum (d'où part la flèche) activée par l'anticipation d'un gain monétaire mais non activée par l'anticipation d'une perte (partie droite de la figure).

Source : Knutson et Peterson (2005)

Ces résultats, en plus de valider l'existence de deux processus mentaux de prise de décision (les processus contrôlés du cortex préfrontal et les processus automatiques du striatum), corroborent le concept que Kahneman (1994) suggère pour l'utilité (voir section 2.2.4 pour plus de détails) ; les différentes formes d'utilité pourraient trouver leurs sources respectives dans différentes régions du cerveau. Knutson et Peterson suggèrent que le noyau accumbens

(striatum) pourrait avoir un rapport avec le concept d'utilité anticipée et que la région médiale du cortex préfrontal serait en lien avec le concept d'utilité instantanée; la région médiale du cortex préfrontal enregistrerait l'impact d'un gain monétaire. Leurs résultats sont également compatibles avec la distinction faite par Berridge et Robinson (1998) pour qui il existe deux processus neuronaux distincts pour ce qui est aimé (*liking*) et ce qui est voulu (*wanting*) en regard de la modulation dopaminergique du striatum ventral (section 2.2.4).

Figure 2.7 : Impact neurologique d'un gain ou d'une perte monétaire



La partie gauche de la figure 2.7 montre l'impact neurologique d'un gain monétaire dans la région du cortex médial préfrontal (la région activée d'où part la flèche). On constate que cette région n'est pas activée par une perte (partie droite de la figure 2.7).

Source : Knutson et Peterson (2005)

En somme, certaines interprétations et recherches permettent de corroborer certaines intuitions ou préceptes de l'approche néo-classique (fonction d'utilité, espérance de l'utilité comme le font Platt et Glimcher) mais aussi certaines de ses approches concurrentes ou complémentaires soulignées par Knutson et Peterson (théorie prospective). Nous n'y voyons pas nécessairement de contradiction. Il n'est pas impossible que sous certaines conditions, l'espérance de l'utilité puisse être une description valable du comportement économique. Néanmoins, dans les sections qui suivent, nous verrons comment la neuréconomie peut valider certaines approches concurrentes ou complémentaires au modèle standard de l'utilité.

### 2.2.2) L'utilité et la dopamine

« The full understanding of utility will come from biology and psychology by reduction to the elements of human behavior followed by a bottom-up synthesis, not from the social sciences by top-down inference and guesswork based on intuitive knowledge. It is in biology and psychology that economists and social scientists will find the premises needed to fashion more predictive models.»<sup>63</sup>

Mieux comprendre les mécanismes par lesquels le cerveau évalue et compare des alternatives et formes variées de récompenses peut nous permettre d'en apprendre sur les déterminants de nos choix et de nos comportements. Dans le cerveau, chacune des alternatives doit être incorporée et comparée de façon automatique ou délibérée, à partir de divers domaines d'information. Ces alternatives doivent donc être transformées par le cerveau en unités comparables. Il est possible d'établir certains parallèles entre le modèle de l'utilité et la dopamine. Nombre de recherches ont démontré le rôle primordial joué par la dopamine dans la sélection d'objectifs. Les neurones dopaminergiques, qui ne représentent qu'à peine 0,3 % de l'ensemble des cellules du cerveau, jouent pourtant un rôle crucial dans la prise de décision.

---

<sup>63</sup> Wilson, 1998 cité dans Glimcher et Rustichini, 2004.

Les premières évidences du rôle joué par la dopamine dans la prise de décision associaient celle-ci à la sensation de plaisir (Olds et Milner, 1954); des études avaient montré une augmentation de dopamine dans certaines régions du cerveau des rats quand ceux-ci étaient engagés dans des activités gratifiantes. Cela laissait entrevoir un lien entre la dopamine et le plaisir hédonistique. En fait, ces premières hypothèses établissaient un lien causal direct entre la sensation de plaisir et la dopamine. Cette interprétation est aujourd'hui remise en cause. De récentes découvertes en neurosciences ont plutôt démontré le lien entre la dopamine et l'apprentissage; avec l'expérience, la réponse dopaminergique serait transférée d'un stimulus inconditionnel (la récompense elle-même) à un stimulus conditionné (annonceur de la récompense). Les neurones dopaminergiques, qui seraient initialement activés par l'arrivée de la récompense,<sup>64</sup> ne le seraient donc par la suite que surtout par les stimuli conditionnés (Schultz, Dayan et Montague, 1997; Schultz 1998, 2000). La décision et les choix seraient alors orientés vers la satisfaction dopaminergique.

« It is the return to a reward previously experienced that is the essence of habit and addiction[...] The return to a previously experienced reward involves the return to reward-associated landmarks as much as it involves return to the reward itself [...] The sounds, sights, and smells associated with the food are clearly predictors of reward, and the efficiency of the animal increases with the identification of more and more distal predictors of reward, predictors that guide the foraging and that are important for the error signals that guide corrections to the foraging path». <sup>65</sup>

Pour assurer sa survie, toute espèce doit pouvoir combler des fonctions vitales aussi diversifiées que de se nourrir, de réagir à l'agression et de se reproduire. Le circuit cérébral de la récompense permet l'atteinte de ces objectifs. L'aire tegmentale ventrale (ATV), un groupe de neurones situés en plein centre du cerveau, est particulièrement importante dans le fonctionnement de ce circuit.

---

<sup>64</sup> Le terme « renforcement » que l'on doit à Pavlov (1903) est assez similaire mais le terme anglais « reward » est souvent utilisé en neurosciences.

<sup>65</sup> Wise, 2002, p. 233.

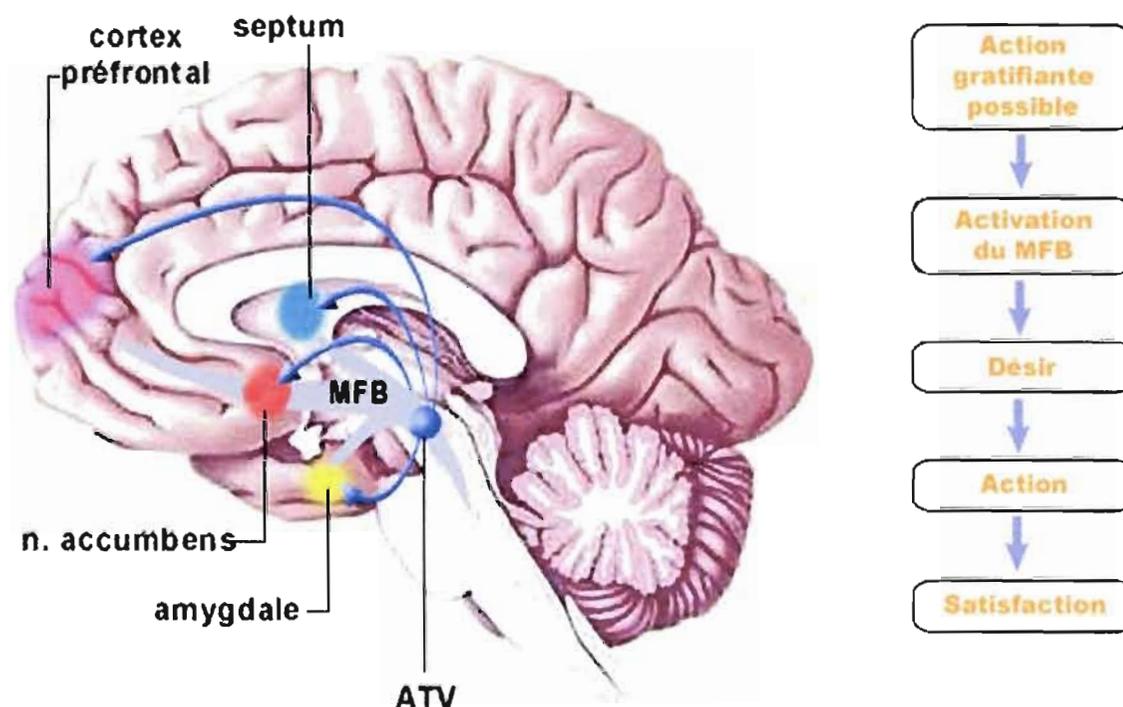
Elle reçoit des inputs de plusieurs autres régions qui l'informent du niveau de satisfaction des besoins fondamentaux (ou plus spécifiquement humains). À l'arrivée d'un signal annonçant une récompense, donc après traitement sensoriel par le cortex, l'activité de l'ATV se trouve augmentée. Cette région transmet alors cette information à l'aide d'un messenger chimique particulier, la dopamine, qui est alors libérée dans le noyau accumbens mais aussi dans le septum, l'amygdale et le cortex préfrontal (figure 2.8).<sup>66</sup> Le noyau accumbens intervient alors dans l'activation motrice et le cortex préfrontal dans la focalisation de l'attention. Ces régions sont reliées par ce que l'on appelle le faisceau de la récompense ou du plaisir. En terme neuro-anatomique, ce faisceau fait partie du *medial forebrain bundle* (MFB) dont l'activation mène à la répétition de l'action gratifiante pour en consolider les traces nerveuses. C'est donc cette sécrétion de dopamine qui aura alors un effet de renforcement sur des comportements permettant la satisfaction de besoins fondamentaux.

Par conséquent, on peut dire que le circuit de la récompense, ainsi que celui de la punition<sup>67</sup>, fournissent la motivation nécessaire à la plupart de nos comportements.

---

<sup>66</sup> Pour plus de détails, voir <http://www.lecerveau.mcgill.ca>

Figure 2.8 : Le circuit de la récompense



Source : Université McGill; <http://www.lecerveau.mcgill.ca>

### 2.2.3) Les *cues* environnementaux et la consommation

La section précédente nous permet maintenant d'introduire le phénomène de la tentation et de mieux comprendre son influence dans le comportement de consommation. Les *cues* ou signaux associés à une consommation passée peuvent engendrer une sécrétion dopaminergique qui orientera l'action. La force de l'impact de l'exposition à des *cues* environnementaux est donc tributaire des antécédents de consommation. Cette affirmation mérite quelques explications additionnelles. Le cerveau finit par encoder les *cues* associés à une récompense. Comme nous venons de le voir, la sécrétion de

<sup>67</sup> Pour plus de détails concernant le circuit de la punition, voir le site des neurosciences de l'Université McGill : <http://www.lecerveau.mcgill.ca> ou encore voir De Molina, 1962.

dopamine est initialement activée au moment de l'arrivée de la récompense. Mais au fur et à mesure que cette dernière agit comme agent de renforcement, ce sont plutôt les *cues* qui la précèdent qui finissent par provoquer la sécrétion de dopamine. Ceci n'est pas sans répercussion potentielle sur le choix du consommateur. Il s'agit d'une autre illustration de l'influence des processus automatiques. La libération de dopamine peut être déclenchée par l'environnement associé à la récompense, et non par la récompense elle-même. Par exemple, la simple vue d'un verre de scotch peut engendrer un impulsif désir (non conscient) de consommer une cigarette pour une personne ayant, par le passé, pris l'habitude de consommer ces deux substances conjointement. La dopamine serait alors responsable d'un ensemble de comportements destinés à atteindre diverses formes de récompenses. Le choix de consommation peut prendre naissance à partir d'une associativité initiée en dehors de la conscience. Cette idée a été incorporée dans les travaux de recherches d'économistes; par exemple, elle est présente dans l'élaboration des modèles économiques de Laibson<sup>68</sup> (2001) et de Bernheim et Rangel (2002).<sup>69</sup>

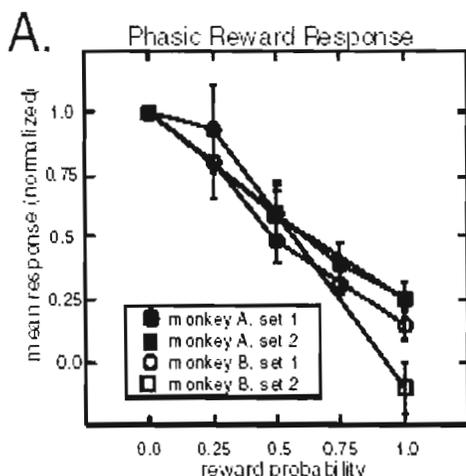
Fait intéressant dans la mesure où il montre que le cerveau encode bel et bien les anticipations, Fiorillo, Tobler, Schultz (2003, p. 1898) suggèrent que l'activation de la sécrétion de dopamine sur réception de récompense peut s'interpréter comme une erreur d'anticipation : « ...If dopamine neurons code reward prediction error, their activation after reward should decline monotonically as the probability of reward increases... ». La figure 2.9 illustre, au moment de l'arrivée d'un stimulus inconditionnel (une récompense en jus), la relation négative entre la probabilité de récompense et la sécrétion de dopamine.

---

<sup>68</sup> Laibson, 2001. David Laibson est un économiste de l'Université de Harvard.

<sup>69</sup> Bernheim et Rangel, 2002. Ces deux économistes enseignent l'économie à l'Université Stanford.

Figure 2.9 : L'encodage de l'erreur d'anticipation



Des probabilités certaines de récompenses (1) ou presque certaines (0,75) engendrent une sécrétion dopaminergique nulle ou faible au moment de la récompense. Lorsque les probabilités de récompenses sont nulles (0) ou faibles (0,25), et qu'une gratification arrive quand même, l'activation dopaminergique devient alors élevée. Fiorillo et ses collègues interprètent cette activité comme étant une erreur d'anticipation<sup>70</sup>. Sur le graphique, on voit que cette erreur est négativement corrélée aux probabilités de récompenses. L'apprentissage permettrait à la réponse dopaminergique d'être transférée du stimulus inconditionnel à un stimulus conditionné.

Source : Fiorillo *et al.*

Dans ce contexte, la capacité de résister ou non à la tentation initiée par un stimulus dépend à la fois de l'ampleur de la récompense anticipée (processus automatiques) et des capacités cognitives (processus contrôlés) à surpasser le désir.<sup>71</sup> Bernheim et Rangel, qui utilisent le terme *mécanisme de prévision hédonistique* (*Hedonic Forecasting Mechanism* ou HFM) en référence au système de la récompense, illustrent cette dualité :

« The HFM's main advantage is that it can produce rapid decisions with generally beneficial near-term outcomes provided the environment is stable. It cannot, however, anticipate sufficiently delayed consequences, and when the environment changes, it can neither ignore irrelevant past experiences nor adjust forecasts prior to acquiring further experience. The competing cognitive system addresses these shortcomings (albeit imperfectly), but is comparatively slow. Balanced competition between these two processes apparently emerged as evolution's best compromise...<sup>72</sup>

<sup>70</sup> D'autres interprétations sont possibles. Ce signal pourrait également être interprété comme une émotion dans le contexte de perceptions contrefactuelles (voir section 2.3.3).

<sup>71</sup> Berridge et Robinson (1998, 2003); Robinson et Berridge (1993, 2000, 2003) cités dans Bernheim et Rangel, 2004.

<sup>72</sup> Bernheim et Rangel, 2002. *Ibid.*, p. 9.

Le phénomène de l'*addiction* a auparavant été expliqué à l'intérieur du cadre rationnel où le choix résulte d'une délibération consciente et d'une estimation des prix anticipés (par exemple Becker et Murphy<sup>73</sup>, 1988 ; Hung, 2001 ; Gruber et Koszegi, 2001). Ce cadre est cependant limité et pourrait bénéficier de la neuroéconomie comme le soulignent Camerer *et al.* :

« It is relevant to rational models of addiction that every substance to which humans may become biologically addicted is also potentially addictive for rats. Addictive substances appear therefore to be “ hijacking ” primitive reward circuitry in the “ old ” part of the human brain. Although this fact does not disprove the rational model (since recently-evolved cortex may override rat-brain circuitry), it does show that rational intertemporal planning is not necessary to create the addictive phenomena of tolerance, craving, and withdrawal. It also highlights the need for economic models of the primitive reward circuitry, which would apply equally to man and rat. ».<sup>74</sup>

Pour Camerer *et al.* (2004), les modèles d'addiction basés sur la rationalité sont incomplets. Par exemple, l'élasticité-prix est un outil insuffisant qui ne tient pas compte, lors d'une hausse de prix, de la substitution faite par les fumeurs vers des marques contenant davantage de nicotine. En outre, le fait que certains drogués n'achètent qu'en petites doses, sans nécessairement profiter des fluctuations de prix, illustre chez ces consommateurs un dilemme intérieur difficilement réconciliable avec l'agent rationnel qui optimise de façon intertemporelle. Ce comportement cadre mieux avec deux processus mentaux antagonistes qui trouvent leur corollaire dans le cerveau: soit la gratification immédiate de prendre une dose ou les conséquences ultérieures des effets secondaires. Comme nous l'avons précédemment souligné, des régions spécifiques du cerveau sont sollicitées par chacune des composantes de ces alternatives. Certains autres comportements irrationnels comme l'achat compulsif, la boulimie, l'ergomanie,<sup>75</sup> le jeu compulsif<sup>76</sup>, pourraient impliquer

<sup>73</sup> Becker et Murphy, 1988. Une addiction « rationnelle » impliquerait des préférences stables où le choix délibéré ne ferait guère de place aux processus automatiques dans le processus décisionnel.

<sup>74</sup> Camerer, Loewenstein et Prelec, 2004.

<sup>75</sup> Le terme *workaholism* est souvent utilisé.

les mêmes circuits neuronaux. On sait qu'environ 1 % des joueurs seraient compulsivement pathologiques. Ceux-ci sont surtout des hommes ; ils consomment de l'alcool et d'autres drogues plus fréquemment que la moyenne. Certaines causes génétiques seraient à retenir (Comings, 1998 cité dans Camerer *et al.*, 2003)<sup>77</sup> et un traitement à la *naltrexone* réduirait le désir de jouer (*gambling*) (McElroy *et al.*, 1991 cités dans Camerer *et al.*, 2004).<sup>78</sup>

Toutes ces limitations justifient d'autres approches. Des économistes ont déjà commencé à intégrer certaines de ces notions dans leur modèle.<sup>79</sup> David Laibson de l'Université Harvard, a formalisé une théorie de la consommation (2001)<sup>80</sup> où l'équilibre atteint est tributaire du déclenchement ou non de *cues*, de signaux. Dans son modèle, l'utilité marginale de la consommation varie selon les *cues*. Laibson conserve à la base l'hypothèse de choix rationnel. Il intègre cependant une sensibilisation<sup>81</sup> endogène aux *cues* qui permet, entre autres choses, d'expliquer pourquoi les préférences peuvent parfois changer rapidement. Tandis que Douglas Bernheim et Antonio Rangel,<sup>82</sup> de l'Université Stanford, ont pour leur part développé une théorie économique de la dépendance aux drogues qui repose sur des bases empruntées aux neurosciences. Ils considèrent la consommation comme étant tributaire de deux systèmes, un état d'«impulsion chaude» et un autre de «délibération froide». Leur approche novatrice basée sur des processus biologiques leur permet de jeter un nouvel éclairage sur les interventions publiques appropriées :

---

<sup>76</sup> Une étude de Fiorillo *et al.*, 2003 montre un lien positif entre le degré d'incertitude d'une situation et l'activation des neurones dopaminergiques. L'incertitude pourrait donc en soi être gratifiante et ces découvertes suggèrent de nouvelles pistes dans la compréhension du phénomène du jeu compulsif.

<sup>77</sup> Un gène allèle (D<sub>2</sub>A1) serait présent chez la plupart des joueurs pathologiques.

<sup>78</sup> Cette drogue aurait également été utilisée pour soigner les acheteurs compulsifs.

<sup>79</sup> La finance comportementale également. Voir par exemple : Montier, 2005 ou encore Montier, 2002.

<sup>80</sup> *Ibid.*

<sup>81</sup> De l'anglais *sensitization* qui correspond à un état d'hyperexcitabilité du système dopaminergique.

<sup>82</sup> Bernheim et Rangel, 2004.

« Our theory also has potentially important implications for empirical studies of addiction. It suggests that users of addictive substances may respond very differently to changes in prices, with dramatically different implications for welfare, depending on whether decisions reflect “hot” impulses or “cold” deliberation. In contrast, existing studies treat data on consumption as if it were generated by a single process ». <sup>83</sup>

L'originalité des travaux de Bernheim et Rangel tient au fait qu'ils ont été capables de formaliser un modèle économique de la décision, non pas seulement axé sur les processus délibérés des quadrants I et II (voir chapitre 1), mais qui incorpore également les dimensions cérébrales des processus automatiques (quadrants III et IV, voir chapitre 1).

#### Choix intertemporel et activité cérébrale

La capacité d'associer le plaisir immédiat à ses conséquences futures, de s'en inquiéter même, est une particularité humaine (Camerer *et al.*, 2003). En soi, cela montre que le système affectif permet d'appréhender le futur, et de favoriser des décisions impliquant des conséquences bénéfiques à plus long terme. Un choix intertemporel implique que la décision de remettre à plus tard une satisfaction présente nécessite une collaboration entre les aspects affectif et cognitif des processus décisionnels. Un choix intertemporel nécessite donc une conscience cognitive des satisfactions futures mais cette conscience seule est insuffisante; les émotions jouent un rôle critique dans cette prise de décision. Pour Cottle et Klineberg (1974), les gens ne se soucient des conséquences futures de leurs actions que dans la mesure où ces conséquences ont un effet immédiat.

Le choix intertemporel est un arbitrage entre différents niveaux d'utilité dans le temps et cet arbitrage, supposé constant, s'exercerait par le biais d'un taux d'escompte subjectif, taux à partir duquel les gens escompteraient les utilités futures en fonction du moment où elles surviennent. Les gens préfèrent

---

<sup>83</sup> *Ibid*, p. 35.

souvent les gratifications immédiates. Une perspective standard de la théorie économique est de modéliser ce type de préférence à l'aide d'une exponentielle. Ce type de modélisation sous-tend implicitement une constance intertemporelle : les préférences d'une personne ayant à choisir entre différents paiements demeureront les mêmes peu importe le moment où on lui demande de choisir. De nombreuses observations empiriques ont cependant montré que les gens n'escomptent pas les utilités futures à un taux constant.<sup>84</sup> Prenons l'exemple d'un arbitrage entre deux choix futurs. Supposons qu'une personne doive choisir, pour la même récompense, entre le fait de travailler 10 heures dans 100 jours ou encore 11 heures dans 101 jours. Dans la mesure où l'accomplissement d'une telle tâche implique une désutilité, il est normal de croire que la plupart des personnes préféreront la première option à la seconde. Pourtant, si on leur propose le même choix juste avant que le délai ne soit expiré, par exemple dans 99 jours, ils pourraient changer d'avis. La non constance intertemporelle impliquerait que leur décision serait cette fois différente. La procrastination, le fait de remettre à plus tard cette désagréable besogne, quitte à travailler une heure de plus le surlendemain, impliquerait un changement des préférences. En général, lorsque nous devons faire un choix entre deux options futures, nous donnons, au fur et à mesure que le temps passe, de plus en plus d'importance à celle qui est la plus rapprochée dans le temps (Rabin, 1998). Ce genre d'observation empirique a récemment capté l'attention de chercheurs de la neuroéconomie. Utilisant la technologie fMRI, McClure *et al.* (2004) ont montré qu'un gain monétaire, selon qu'il soit attendu dans un court ou long avenir, ne sollicitera pas les mêmes régions du cerveau. Une récompense attendue dans un délai de plus long terme solliciterait plus particulièrement les régions fronto-pariétales alors que les gratifications plus immédiates stimuleraient davantage certaines régions dopaminergiques.

---

<sup>84</sup> Pour plus de détails, voir Rabin, 1998.

#### 2.2.4) Préférences révélées et révision du concept de l'utilité

Nous avons vu à la section 2.2.1 que le modèle classique de l'utilité (fonction d'utilité et espérance de l'utilité) avait fait l'objet de recherches au sein des neurosciences. Nous avons alors aussi souligné que certaines approches alternatives, telles que la théorie prospective, avaient également attiré l'attention de chercheurs intéressés aux processus décisionnels. Kahneman (1994) et Kahneman, Wakker et Sarin (1997) ont procédé à la révision du concept de l'utilité même. Leur travaux n'ont pas manqué d'attirer l'attention de chercheurs issus des neurosciences qui les ont incorporés à leurs propres recherches (Shizgal, 1997, 1998; Berridge, 1999). Nous présentons dans la présente section certaines des implications de tout cela.

Le principe des préférences révélées, sur lequel repose une importante littérature économique, suggère que les préférences, puisqu'elles sont considérées comme non observables, peuvent être inférées à partir des choix. Deux arguments fondamentaux défendent ce principe : le premier est que l'on ne peut observer l'expérience hédonistique d'une préférence parce qu'elle est subjective ; le second est que la décision implique nécessairement que l'agent rationnel optimisera son expérience hédonistique. Nombre d'évidences démontrent cependant les limites de ces présuppositions ; il est effectivement possible de mesurer la composante hédonistique de l'utilité et celle-ci peut diverger de l'utilité décisionnelle (Kahneman, Wakker et Sarin, 1997). De récentes recherches issues des neurosciences et de la psychologie démontrent que le choix et les préférences peuvent parfois diverger. Nous avons vu au premier chapitre que les processus automatiques, notamment parce qu'ils sont moins coûteux en terme neurologique, permettent la prise de décision rapide. Bien que ces processus nous aident très souvent à prendre les bonnes décisions, ils peuvent cependant engendrer certaines erreurs systématiques. Par exemple, le touriste américain en visite à Londres qui omet de regarder à droite avant de traverser la rue n'a pas des préférences masochistes. Il

commet une erreur qui est simplement due à un automatisme qui trouve son origine dans l'architecture du cerveau humain.<sup>85</sup>

Le choix n'est pas toujours le reflet d'une quête hédonistique et peut s'expliquer par d'autres facteurs. Par exemple, il a été démontré qu'avec l'usage répété d'une substance illicite<sup>86</sup>, les *cues* lui étant associés exagèrent grossièrement la récompense anticipée et provoquent une forte impulsivité à la consommer. On sait que toutes les substances psychoactives ont un effet sur les neurones du cerveau qui forment le circuit de la récompense. En fait, les drogues semblent augmenter directement ou indirectement la quantité de dopamine dans ce circuit. C'est ce qui décrit le phénomène de l'addiction dans une perspective neuroéconomique. Robinson et Berridge (1993) ont suggéré l'idée qu'un état d'hyperexcitabilité ou *sensitization* du système dopaminergique mésolimbique serait à l'origine de la sensation de *craving* des dépendances. Le système dopaminergique attribue une valeur ou *saliency* à des stimuli associés à l'activation du système, les rendant ainsi attirants et désirables. Le manque de contrôle du cortex préfrontal (manque de contrôle inhibiteur du quadrant I) peut aussi mener à l'impulsivité des personnes dépendantes. Ce qui remet en question la théorie des préférences révélées, c'est notamment que la valeur associée aux stimuli peut être en dissonance avec le réel plaisir de consommer.

Kahneman (1994) et Kahneman, Wakker et Sarin (1997) fragmentent le concept d'utilité. L'utilité mémorisée (*Remembered Utility*) correspond à l'évaluation d'une expérience vécue. C'est ce souvenir ou cette évaluation qui permettra par la suite de discriminer entre diverses alternatives et de procéder à un choix. L'utilité mémorisée est donc essentielle à l'utilité décisionnelle qui elle-même correspond au degré qu'un événement est voulu ou non. L'utilité décisionnelle initie toute forme d'action et c'est de cette forme d'utilité dont

---

<sup>85</sup> Cet exemple est tiré de Bernheim et Rangel, 2004.

<sup>86</sup> Mentionnons ici la cocaïne, les amphétamines et l'héroïne.

était tristement dépourvu l'âne de Buridan. Comme il est possible que le calcul de l'utilité décisionnelle soit influencé par d'autres facteurs que l'utilité mémorisée, Kahneman et ses collègues (1997) suggèrent également d'incorporer l'utilité anticipée (*expected, predicted*) qui correspond à l'utilité prévue qui est associée à un événement futur. L'utilité instantanée<sup>87</sup> (*Instant Experienced Utility*) correspond à la dimension hédonistique d'un événement, le plaisir ou déplaisir lui étant associé. C'est à partir de cet événement, de son évaluation subjective, que se formera l'utilité mémorisée.<sup>88</sup> Le tableau 2.2 illustre les différentes composantes de l'utilité suggérées par Kahneman.

Tableau 2.2 : Les différentes formes d'utilité suggérées par Kahneman

Expected/Remembered Utility	Decision Utility	Instant Experienced Utility
Beliefs about value	Wants	Likes

Source : Berridge<sup>89</sup>

Pour un esprit cohérent et rationnel, les différentes formes d'utilité devraient correspondre. Si un événement fut plaisant, on devrait s'en souvenir comme tel et prévoir qu'il soit agréable à expérimenter à l'avenir. On devrait donc vouloir le reproduire. Mais des expérimentations ont démontré que ce n'était

<sup>87</sup> Kahneman fait la distinction entre l'utilité vécue (*Experienced Utility*) et l'utilité décisionnelle (*Decision Utility*). Il associe la dimension hédonistique de l'utilité (au sens de Bentham) à l'utilité vécue (*Experienced Utility*).

<sup>88</sup> Bien qu'au cours d'une expérience vécue l'utilité instantanée puisse varier d'un moment à un autre, le souvenir que l'on en gardera correspondra à une estimation discrète et non continue. Kahneman suggère, basé sur ses observations empiriques, que l'utilité mémorisée puisse correspondre à la moyenne entre l'utilité instantanée maximale (*Peak*) et l'utilité observable à la fin de l'expérience (*End*), ce qu'il qualifie de *Peak-End Rule*. Il semble que la durée de l'expérience ne soit pas un facteur (*Duration Neglect*) dans la détermination de l'utilité mémorisée. Il est donc possible qu'entre une expérience «A» procurant en moyenne une utilité instantanée totale qui soit plus grande qu'une expérience «B», ce soit quand même cette dernière qui soit choisie parce que son utilité mémorisée fut préalablement plus grande. Cela implique donc que l'utilité décisionnelle ne maximise pas nécessairement toujours l'utilité totale. D'autres études montrent cependant une forte corrélation entre des mesures subjectives hédonistiques et le choix (voir Cabanac et Leblanc, 1983).

<sup>89</sup> Berridge, 1999.

pas toujours le cas. Les différentes formes d'utilité peuvent ne pas correspondre. Ainsi, la dissonance entre l'utilité anticipée et l'utilité instantanée a été démontrée par Kahneman et Snell (1992). Par ailleurs, il a été démontré que l'utilité mémorisée peut être découplée de l'utilité décisionnelle dans le cas où surviennent des distorsions de mémoire. Des dissonances entre ces deux composantes de l'utilité ont été démontrées par Kahneman *et al.* (1993); des gens peuvent choisir, en autant qu'elle soit accompagnée d'un décretement à la fin, la procédure qu'ils avaient pourtant préalablement classifiée comme étant la plus douloureuse (mettre sa main dans l'eau froide). Une telle dissonance a été aussi démontrée au cours d'une douloureuse procédure médicale (Redelmeier et Kahneman, 1996). Mais de toutes ces dissonances, c'est celle qui vient scinder l'utilité décisionnelle de l'utilité instantanée qui ébranle le plus la théorie des préférences révélées.

#### Dissonance entre l'utilité décisionnelle et l'utilité instantanée

A priori, l'utilité instantanée (le *liking*) et l'utilité décisionnelle (le *wanting*), semblent tellement intrinsèquement liées l'une à l'autre, qu'on pourrait vouloir remettre en question l'existence même d'une dissonance entre les deux. Après tout, si A et B sont disponibles, pourquoi irait-on choisir A si on préfère B ? Les premières expérimentations d'auto-stimulation électrique faites sur des rats associaient le choix des rats à une forme d'expérience hédonistique (Olds et Milner, 1954). Si le rat, plutôt que de boire ou de manger, préférait appuyer sur un levier lui donnant des stimulations électriques, n'était-ce pas parce que cela lui procurait un plus grand plaisir ? Comme nous l'avons préalablement souligné, ce genre d'inférence a depuis été remis en cause. De récentes études semblent illustrer des processus sous-jacents parfois séparables pour ces deux formes d'utilités (Berridge, 1999). Ainsi, certaines manipulations cérébrales peuvent modifier l'utilité décisionnelle sans pour autant affecter les composantes hédonistiques de l'action qui en découle. Ces manipulations, qui

modifient la sélection d'objectifs et de buts computationnels, permettent d'illustrer cette dissonance :

« This group of brain manipulations includes hypothalamic stimulation, drugs that act on dopamine receptors or on dopamine release, and 6-hydroxydopamine lesions that selectively destroy neurons that contain dopamine. All of these act on the massive projections of dopamine neurons that stretch from the midbrain to the forebrain, often called the mesolimbic or mesoaccumbens dopamine system. This neural system appears to be a neural common denominator or substrate shared in common by most manipulations that alter "wanting" without changing "liking."»<sup>90</sup>

On peut diminuer l'utilité décisionnelle tout en maintenant constante l'utilité instantanée. Il a été montré que certaines drogues<sup>91</sup> parviennent à bloquer les récepteurs de dopamine et ainsi à supprimer les incitatifs à quérir quelque forme de récompense. L'animal, à la manière de l'âne de Buridan, est alors prisonnier d'un déficit motivationnel et cesse de chercher à obtenir des récompenses. En d'autres termes, ces drogues réduisent ou éliminent l'utilité décisionnelle d'une récompense. Mais étonnamment, ces drogues laissent inchangée l'utilité instantanée ou le "liking" de la même récompense (Pecina, Berridge et Parker, 1997; Treit et Berridge, 1990). Un autre exemple est illustré par la dissociation la plus drastique qui soit, celle qui consiste en l'abolition partielle ou même complète du système dopaminergique de récompense. Des rats ayant subi des lésions dans ce système affichaient néanmoins des réactions hédonistiques aux goûts sucrés et des aversions aux goûts amers, tout comme les rats normaux (Berridge et Robinson 1998; Berridge, Venier et Robinson, 1989). D'autres types de manipulations cérébrales illustrent par ailleurs que l'on peut faire augmenter l'utilité décisionnelle sans pour autant modifier les composantes hédonistiques d'une récompense. L'usage de certaines drogues, comme la cocaïne, les

<sup>90</sup> Berridge, 1999, p. 540.

<sup>91</sup> Telles que le *haloperidol* et le *pimozide*.

amphétamines, l'héroïne et ses dérivés, de par leur *sensitization*<sup>92</sup> neuronale du système dopaminergique, vont faire augmenter l'utilité décisionnelle sans pour autant modifier l'utilité instantanée :

« For an addict, drugs of abuse have enormous decision utility. They are powerfully wanted and sought-often at great risk and great cost... Euphoric drugs are sought by addicts even when the available dose is so low or so poor in quality that it can't be expected to give much pleasure... ».<sup>93</sup>

Le fait que les différentes composantes de l'utilité puissent diverger devrait impliquer de nouvelles approches en économie du bien-être comme le soulignent Camerer, Loewenstein et Prelec (2003) :

« Economics proceeds on the assumption that satisfying people's wants is a good thing. While this is probably generally a safe assumption, if wanting and liking are two separate processes, then it cannot be assumed that satisfying someone's desires necessarily makes them better off. Welfare economics, then, would need to be augmented by an analysis of when and why wanting and liking do, sometimes, diverge. ».<sup>94</sup>

En somme, ces récentes découvertes du fonctionnement du cerveau ne permettent pas de rejeter l'intuition fondamentale du concept de l'utilité, mais bien de redéfinir celui-ci d'une façon plus réaliste et précise. Il existe différentes composantes à l'utilité, et celles-ci trouvent nécessairement un corollaire dans le cerveau, ce que des chercheurs tentent d'identifier. Chacune de ces composantes a une incidence, un rôle à jouer dans les processus décisionnels. La théorie des préférences révélées suggère que toutes ces composantes agissent conjointement de façon uniforme pour orienter la décision rationnelle. Mais des évidences prouvent que celles-ci peuvent parfois diverger. Comme nous l'avons vu à la section 1.4.1, la compétition est une caractéristique de l'interaction entre les différents processus cérébraux. C'est

<sup>92</sup> La *sensitization* implique une hyperactivité des neurones à un stimulus conditionnel. Cette hyperactivité est provoquée par l'usage répété de certaines drogues. Elle peut donc faire augmenter l'utilité décisionnelle.

<sup>93</sup> *Ibid.*, p. 543.

<sup>94</sup> Camerer, Loewenstein et Prelec, 2003, p. 21.

cela qui permet à la neuroéconomie de suggérer des approches plus réalistes et plus complètes que la simple théorie des préférences révélées.

### 2.2.5) Au-delà des préférences révélées

Les substrats associables à chacune des formes d'utilité peuvent parfois, puisqu'ils ne sont pas les mêmes, engendrer certaines dissonances. Ces découvertes pourraient permettre aux économistes d'améliorer leur modèle. Mais la neuroéconomie peut également permettre d'approfondir la compréhension des processus décisionnels, même lorsque les composantes de l'utilité coïncident et que les préférences sont, en apparence, révélées par le choix. En fait, l'approche bidimensionnelle du fonctionnement cérébral (proposée à la section 1.4) permet de dépaqueter le concept des préférences. Romer (2000) illustre le cas d'un homme, appelons le Al, et d'une femme, Naucia, qui refusent de manger des arachides. Al aime les arachides mais il pense en être allergique. Naucia déteste tout simplement les arachides depuis qu'elle a développé une aversion nauséuse<sup>95</sup> à cette nourriture. Tous les deux associent donc les arachides à quelque chose de désagréable et leur courbe d'indifférence a donc le même aspect (figure 2.10). Or tous deux se trompent; l'homme n'est pas allergique aux arachides mais bien à un préservatif que l'on retrouve sur certaines arachides. Et la femme n'a pas été malade à cause des arachides mais bien parce qu'elle avait contracté un virus au moment même de sa consommation antérieure. La théorie des préférences révélées suggérerait que ces personnes ont toutes deux une désutilité face aux arachides et là s'arrêterait l'analyse. Mais une préférence peut être délibérée ou implicite. Naucia n'est pas nécessairement consciente des motifs pour lesquels elle n'aime pas les arachides. Les rats possèdent un mécanisme similaire d'aversion nauséuse alors il n'y a pas de raison de croire que des processus délibérés soient nécessaires chez l'humain pour développer une telle aversion.

---

<sup>95</sup> Une aversion nauséuse est un mécanisme naturel qui amène à développer une aversion à une nourriture associée faussement ou non à une nausée. Voir Pelchat, Levin et Rozin, 1982.

Et même si Naucia mangeait à nouveau une arachide, même si elle comprenait les véritables motifs de son dédain pour celles-ci, les processus (automatiques) guidant sa préférence ne lui seraient pas plus accessibles qu'avant sa nausée, et elle continuerait de les trouver dégoûtantes.<sup>96</sup> Al, quant à lui, utilise des processus délibérés pour arriver à la conclusion qu'il est allergique. Si un docteur lui apprenait qu'il n'est pas allergique, il recommencerait à apprécier les arachides et à en consommer. Toutes ces considérations sont importantes; même si la décision de Al et de Naucia est la même, elle ne reflète pas la même réalité ; parce que les processus sous-jacents à leur décision sont distincts, l'aspect de leur courbe d'indifférence respective devrait différer l'une de l'autre advenant une modification de leur comportement. Par exemple, si Al et Naucia consommaient à nouveau des arachides, ne serait-ce que pour une seule fois, leur courbes d'indifférence ne seraient plus similaires (figure 2.11).

Cela s'explique par le fait que les régions du cerveau guidant la décision de Al et de Naucia ne sont pas les mêmes. Pour Al, ce sont les processus cognitifs du quadrant I qui guident sa décision; son cortex préfrontal lui permet de contrôler son envie de consommer des arachides, envie qui lui provient des processus automatiques du système dopaminergique. Il croit comprendre ce que seront les conséquences futures d'une consommation et il résiste alors à la tentation. Or, à partir du moment où il consommera des arachides, il constatera qu'il n'est pas allergique. Il recommencera alors à apprécier cette nourriture (figure 2.11) et ses préférences ainsi que sa courbe d'indifférence changeront. Pour Naucia, ce n'est pas le cas. Même si elle recommençait à consommer des arachides, elle ne pourrait plus apprécier cette nourriture parce que son dédain des arachides, son aversion nauséuse, lui provient de régions du cerveau qui sont inaccessibles à sa conscience. Même si elle savait que les arachides sont bonnes, elle ne pourrait contrôler son dégoût. Par conséquent, sa courbe d'indifférence demeurerait inchangée, même après avoir fait l'essai de celles-ci (figure 2.11). C'est ainsi que l'on peut « dépaqueter » le

---

<sup>96</sup> Voir Garcia et Koelling, 1966 ou encore Bernstein, 1977.

concept de préférence et séparer des phénomènes qu'on croyait initialement similaires.

Figure 2.10 : Courbes d'indifférence avec préférences non-dépaquetées

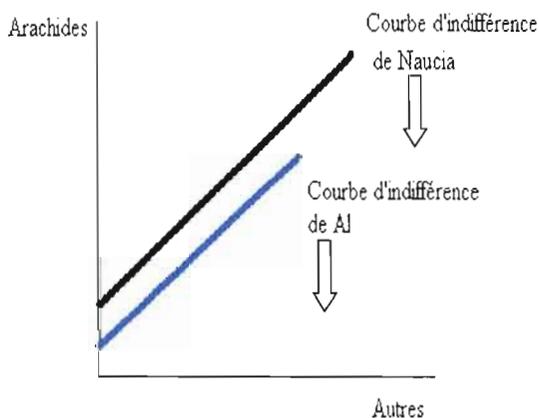
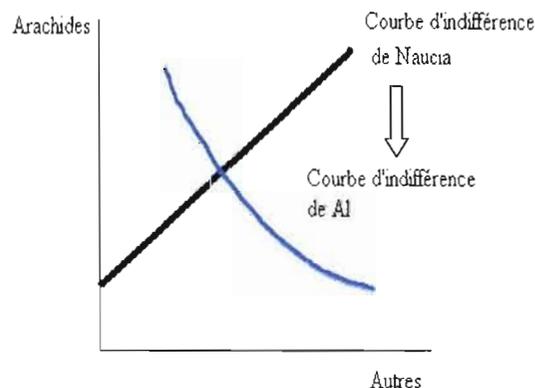


Figure 2.11 : Courbes d'indifférence avec préférences dépaquetées



La figure 2.10 illustre que les préférences de Al et Naucia semblent *a priori* similaires ; les arachides sont un mal tel que l'illustre la courbe d'indifférence à pente positive. Toutefois, après consommation (figure 2.11), les préférences de Al vont changer mais pas celle de Naucia

Cet exemple montre qu'en incorporant à l'analyse les substrats biologiques de deux préférences qu'on aurait initialement crues identiques, on peut enrichir le cadre de la théorie des préférences révélées.

#### 2.2.6) Homéostasie et *alliesthesia*

L'homéostasie est l'ensemble des processus organiques qui agissent pour maintenir l'état stationnaire de l'organisme, dans sa morphologie et dans ses conditions intérieures, en dépit de perturbations extérieures.<sup>97</sup> Les processus

<sup>97</sup> Définition de Wikipédia, l'encyclopédie libre : <http://fr.wikipedia.org/>

homéostatiques s'actionnent lorsque des *cues* internes et environnementaux viennent modifier un équilibre initial. Certains de ces mécanismes s'initient et s'effectuent en parallèle à la conscience. Ainsi, les frissons ou la sueur vont apparaître afin de maintenir constante la température du corps. D'autres de ces mécanismes sont volitifs et permettent une satisfaction bien consciente: par exemple, mettre son manteau quand il fait froid, et l'enlever quand il fait chaud afin de mieux se sentir. Les émotions peuvent donc être vues comme des processus homéostatiques parmi d'autres. En fait, l'idée même de l'homéostasie permet de resituer certaines dimensions économiques du comportement. Conventionnellement, les économistes considèrent les préférences comme étant stables et cohérentes (Rabin, 1998). Les préférences seraient le point de départ et le comportement, le point d'arrivée. La perspective des neurosciences est plus large; le comportement y est vu comme un des nombreux mécanismes permettant au cerveau de maintenir l'équilibre homéostatique. Dans cette perspective, les préférences sont transitoires et elles visent, à la base, à garantir la survie et la reproduction :

« The traditional economic account of behavior, which assumes that humans act so as to maximally satisfy their preferences, starts in the middle (or perhaps even towards the end) of the neuroscience account. Rather than viewing pleasure as the goal of human behavior, a more realistic account would view pleasure as a homeostatic cue – an information signal...»<sup>98</sup>

Et puisque l'évolution sélectionne les gènes qui survivent et se reproduisent, la dimension hédonistique des préférences n'est sans doute pas le but ultime de leur fondement. Une caractéristique fondamentale des processus homéostatiques est qu'ils sont foncièrement sensibles aux variations. Cette sensibilité neuronale aux variations pourrait expliquer pourquoi des indicateurs comme l'estimation du bonheur ou les suicides sont plus sensibles aux changements (sanitaires et monétaires) qu'à leur niveaux (Oswald, 1997 cité dans Camerer *et al.*, 2003 ). Après l'euphorie provoquée par un gain à la

<sup>98</sup> Camerer, Loewenstein et Prelec, 2003, p. 19.

loterie, les gagnants ne sont pas plus heureux que les personnes normales (Brickman, Coates et Janoff-Bulman, 1978, cités dans Rabin, 1998). Éventuellement, ce ne sont pas les niveaux absolus de consommation ou de revenus qui ont une incidence sur l'utilité mais bien leur déviation par rapport à un niveau de référence. Comme le soulignent Camerer *et al.* (2003) :

« Neural sensitivity to change is probably important in explaining why the evaluation of risky gambles depends on a reference point which encodes whether an outcome is a gain or a loss, why self-reported happiness (and behavioural indicators like suicide) depend on changes in income and wealth, rather than levels, and why violations of expectations trigger powerful emotional responses. ».<sup>99</sup>

Par ailleurs, maintes évidences ont démontré que les préférences ne sont pas stables et qu'elles peuvent être momentanément changées lors de l'activation des processus homéostatiques. Le changement de ces préférences provoquées par une modification physiologique est un phénomène qui est appelé « *alliesthesia* » (Cabanac, 1971).<sup>100</sup> Il n'est pas clair que certains processus homéostatiques puissent être agrégés pour former un macro-équilibre. Certaines observations semblent démontrer cette possibilité. Par exemple, le cycle menstruel de femelles qui cohabitent finit par converger (Camerer, Loewenstein et Prelec, 2004). Dans quelle mesure est-ce que les vagues d'euphorie et de panique sur les marchés trouveraient leur source dans ce que Keynes appelait les « esprits animaux » des investisseurs et jusqu'à quel point ce gréganisme peut s'expliquer en terme neurologique demeure une question ouverte.

---

<sup>99</sup> Camerer, Loewenstein et Prelec. 2003. *Ibid.*, p. 20. Il est tentant d'établir un lien entre ces propos et certaines considérations de la section 2.2.3 où nous avons souligné que l'activité dopaminergique, au moment de la récompense, finit par décroître lorsqu'elle est pleinement anticipée.

<sup>100</sup> Cabanac, 1971.

### 2.2.7) Utilité directe de la monnaie et de sa source

« La façon la plus sûre de tuer un homme, c'est de l'empêcher de travailler en lui donnant de l'argent ». <sup>101</sup>

À ce jour, nombre d'études ont cherché à mesurer l'impact hémodynamique de gains ou de pertes monétaires chez l'humain (Breiter *et al.*, 1996, 2001; Elliott *et al.*, 2000; Knutson *et al.*, 2000, 2001; Knutson et Peterson, 2005; Paulus *et al.*, 2001; Oderty *et al.*, 2001; Thut *et al.*, 1997). Il appert que l'obtention de sommes monétaires a l'effet de stimuler certains circuits neuronaux bien spécifiques, notamment les neurones dopaminergiques du mésencéphale. L'utilité de l'argent ne serait donc pas indirecte; elle serait directement gratifiante. Le fait que des ergomaniaques continuent de quérir l'argent bien après leur enrichissement personnel, alors que l'utilité marginale de leur consommation devrait normalement avoir depuis longtemps chuté, est une illustration de cet effet direct. Pour certaines zones spécifiques du cerveau, l'argent peut provoquer une activation hémodynamique supérieure à celles rencontrés lors de stimuli tactiles, gustatifs et même à certaines drogues aux propriétés euphorisantes. L'activation de régions communes à tous ces stimuli pourrait illustrer que des dysfonctions des mêmes mécanismes neuronaux pourraient engendrer des problèmes d'impulsivité comme ceux du jeu compulsif ou de la consommation de drogue (Breiter, Aharon, Kahneman, Dale et Shizgal, 2001). Par ailleurs, une expérimentation de Zink *et al.* (2004) a montré que le striatum était davantage activé par une récompense monétaire nécessitant un effort que par une même récompense ne nécessitant aucun effort. Si l'utilité monétaire est effectivement plus grande pour un gain nécessitant un effort, cette étude devrait pouvoir informer les théories du bien-être (Camerer, Loewenstein et Prelec, 2004).

---

<sup>101</sup> Félix Leclerc. 1972. Éd. Canthus. Extrait de la chanson « Les 100 000 façons de tuer un homme ».

### 2.3) Estimation des probabilités et révision du concept de l'espérance de l'utilité

Comme nous l'avons vu à la section 2.2.1, certaines des dimensions neurologiques de l'espérance de l'utilité<sup>102</sup> ont été corroborées, en partie du moins, par les travaux de Platt et Glimcher. Si le cerveau d'un singe rhésus est capable d'encoder les probabilités et la valeur associées à une décision, on peut raisonnablement inférer qu'il en est de même chez l'humain. La théorie de l'utilité, son hypothèse de rationalité et de maximisation, semble plausible en certaines circonstances. Lorsque questionné sur les raisons pour lesquelles le comportement rationnel était à ce point imbriqué dans la théorie économique, Daniel Kahneman n'a-t-il pas simplement répondu que cette hypothèse était tout à fait réaliste et valable dans plusieurs situations ?<sup>103</sup> Les travaux pionniers de Platt et de Glimcher illustrent d'une élégante façon que la fusion des neurosciences et de l'économie classique est, jusqu'à un certain point, possible. Les intuitions remarquables de Blaise Pascal, à savoir qu'une décision optimale est prise par l'identification de l'action qui maximise l'espérance du gain, et de Daniel Bernouilli, qui améliora la théorie de l'espérance du gain en y incorporant la notion de l'espérance de l'utilité<sup>104</sup>, sont donc partiellement corroborées par certaines recherches issues des neurosciences. Mais bien que les découvertes de Platt et de Glimcher soient remarquables, il n'en demeure pas moins qu'elles font fi de l'abondante littérature qui a, depuis les cinquante dernières années, remis en cause le principe même de l'espérance de l'utilité. Mais d'autres recherches sont en cours; dans un article paru dans la revue *Science* en 2004, Paul Glimcher et l'économiste Aldo Rustichini de l'Université du Minnesota soulignaient que la neuroéconomie, pour garantir une approche complète des processus décisionnels, et malgré le degré de difficulté que cela puisse représenter, aurait

---

<sup>102</sup> Ou à tout le moins de l'espérance du gain.

<sup>103</sup> Entrevue avec Daniel Kahneman. Voir Goode, 2002.

<sup>104</sup> Afin de résoudre ce que l'on a appelé le Paradoxe de Saint-Petersbourg en 1738.

à incorporer les avancées récentes de l'économie expérimentale et de la psychologie :

« But the complete reconstruction of the decision process, and hence of human behavior, is not going to be easy, because two of the cornerstones of economic analysis, subjective utility theory and Nash equilibrium, provide, even from the descriptive point of view, an incomplete picture... A more general notion is needed and, as we have seen, is beginning to be investigated and developed by psychologists and economists working together. The goal of the emerging neuroeconomic program will have to be a mechanistic, behavioural, and mathematical explanation of choice that transcends the explanations available to neuroscientists, psychologists, and economists working alone. ».<sup>105</sup>

Fait notable, les avancées théoriques rendues possibles par les approches alternatives à l'espérance de l'utilité trouvent également leurs corollaires neurologiques. Des chercheurs issus des neurosciences, de la psychologie et de l'économie ont déjà commencé à travailler ensemble pour avoir une vue plus complète du comportement humain. Nous chercherons maintenant à illustrer en quoi la neuroéconomie peut servir de support à certaines évidences formelles de falsifications de la théorie de l'espérance de l'utilité. Il sera question des paradoxes d'Allais et d'Ellsberg, ainsi que de certains éléments de la théorie prospective.

### 2.3.1) Le paradoxe d'Allais (1953)

Avec le recul, il est tentant de considérer les premières expérimentations économiques comme étant de naïves tentatives de validations du modèle de l'utilité. Thurstone (1931) tenta de déterminer les courbes d'indifférence individuelles de personnes ayant à choisir entre des chapeaux et des souliers; Rousseas et Hart (1951) poursuivirent dans cette voie en tentant de construire une courbe d'indifférence de diverses combinaisons de déjeuners; Mosteller et Nogee (1951) tentèrent de créer des courbes d'indifférence qui permettraient de

---

<sup>105</sup> Glimcher et Rustichini, 2004, p. 452.



Or, non seulement la réponse de L.J. Savage fut elle incompatible avec sa propre conception de la théorie de l'espérance de l'utilité, mais Allais démontra par la suite que près de 45 % des répondants violèrent également l'axiome d'indépendance (le principe des probabilités linéaires). Contrairement aux prescriptions suggérées par Morgenstern et Von Neumann, le paradoxe d'Allais démontre que l'estimation des probabilités faite par les gens n'est pas linéaire. Les espérances mathématiques attachées à C et D sont respectivement de 11 millions et de 50 millions. Si l'individu est neutre vis-à-vis du risque, il devrait préférer D à C. S'il choisit effectivement D, la cohérence implique que dans la première situation, l'individu aurait alors dû préférer B à A. Allais parvint à falsifier la théorie de l'espérance de l'utilité en montrant l'incohérence de ceux qui avaient préféré A à B dans le premier cas, mais qui avaient ultérieurement choisi D plutôt que C. Une hypothèse d'aversion pour le risque constante et la même courbure du champ d'indifférence aurait pourtant dû faire en sorte qu'ils préfèrent C à D. Mais les observations ont montré qu'à choisir entre un revenu certain et un revenu aléatoire, lorsque les montants au voisinage de la certitude sont très élevés, les individus préfèrent le revenu certain, même s'il y a une faible chance de ne rien percevoir. Et quand on est en présence de revenus aléatoires, avec des probabilités faibles mais proches l'une de l'autre, les individus choisissent le revenu le plus élevé. Ils révèlent alors une incohérence dans leurs choix, du moins en fonction de cet axiome de la théorie de l'espérance de l'utilité.

L'expérimentation de Allais a marqué le début d'une ère nouvelle en économie; après celle-ci, les approches normative et positive étaient plus clairement séparables l'une de l'autre. À partir de là, les expérimentations devinrent plus fréquentes et donnèrent naissance à des approches alternatives, notamment dans les années 1970 avec les travaux de recherche de Daniel Kahneman et d'Amos Tversky, pour ne nommer que ceux-là. Plus récemment, certains chercheurs ont tenté d'établir un lien entre les observations dénotées en laboratoire et leurs corollaires cérébraux respectifs. C'est ainsi que le paradoxe

d'Allais a, plus de cinquante années après la Conférence de Paris, récemment refait surface. Camerer, Loewenstein et Prelec (2004) soulignent que l'incohérence montrée par le paradoxe d'Allais a été observée chez d'autres espèces animales. Des mécanismes cérébraux, jusqu'à un certain point similaires, engendreraient ce type de choix chez différentes espèces animales; il est possible de conjecturer que ces similitudes pourraient trouver leur source dans les mécanismes évolutionnistes, notamment dans les exigences communes et antérieures de la survivance; les règles de l'évolution n'ayant pas toutes à être conformes aux axiomes de rationalité. Au niveau neurologique, la violation démontrée par Allais pourrait être attribuable au fait que nos choix sont guidés par plus d'un mécanisme d'évaluation. Ainsi, les processus cérébraux mis en branle quand un revenu certain est l'une des options ne sont pas les mêmes que lorsque les deux options sont incertaines (Dickhaut *et al.*, 2003 cités dans Glimcher et Rustichini, 2004).

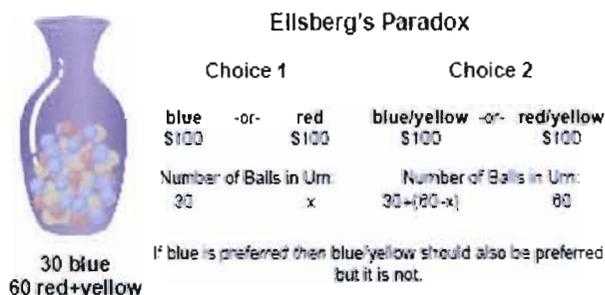
### 2.3.2) Le paradoxe d'Ellsberg (1961)

Une autre falsification de la théorie de l'espérance de l'utilité, possiblement plus déconcertante encore, fut illustrée par Daniel Ellsberg. Dans sa thèse de doctorat en économie de l'Université de Harvard, Daniel Ellsberg souligna dès 1959 que la théorie de l'espérance de l'utilité ne tenait pas compte de l'aversion qu'ont les gens face aux situations ambiguës. Dans un article paru par la suite dans le *Quarterly Journal of Economics*, Ellsberg (1961) illustra un type de préférence qui fut par la suite connu sous le nom de paradoxe d'Ellsberg.

Ce paradoxe (figure 2.12) s'explique de la façon suivante : dans une urne qui contient 90 boules, on retrouve 30 boules bleues. Les 60 boules restantes sont soit rouges soit jaunes dans une proportion qui est non spécifiée. Si les gens font les bons choix, ils obtiennent 100 \$. Entre un choix certain (l'option A de choisir la boule bleue qui donne 1/3 de chance de gagner) et un choix ambiguë

(l'option B de choisir la boule rouge qui donne entre 0 et 2/3 de chance de gagner), les gens vont typiquement préférer parier sur la boule bleue. Selon la théorie de l'espérance de l'utilité, si les gens font ce choix, c'est nécessairement parce qu'ils considèrent qu'il y a moins de 30 boules rouges dans l'urne ou de façon équivalente, qu'il y a plus de 30 boules jaunes. Le paradoxe est illustré par ce qui suit : avant même d'avoir effectué un autre tirage, et à partir du même boulier, on offre ensuite à ces participants de choisir entre l'une des deux options suivantes : soit une boule bleue ou jaune (l'option C), ou soit une boule rouge ou jaune (l'option D). Les gens devraient rationnellement opter pour l'option C (conformément à la théorie de l'espérance de l'utilité). Puisqu'ils avaient initialement choisi la boule bleue gagnante, c'est parce qu'ils avaient implicitement estimé qu'il y avait plus de 30 boules jaunes dans l'urne. Or, après avoir choisi l'option A, les gens choisissent typiquement l'option D. Pourquoi avoir choisi l'option D ? Puisqu'on parle de la même urne, on ne peut à la fois logiquement considérer qu'il y a moins de 30 boules jaunes en choisissant l'option D et plus de 30 boules jaunes en choisissant l'option A. Ce choix viole évidemment le principe de la rationalité suggéré par la théorie de l'espérance de l'utilité et illustre une aversion à l'ambiguïté (l'option C qui donne entre 1/3 et 1 chance de gagner et l'option D qui donne assurément 2/3 de chances de gagner).

FIGURE 2.12 : Le paradoxe d'Ellsberg



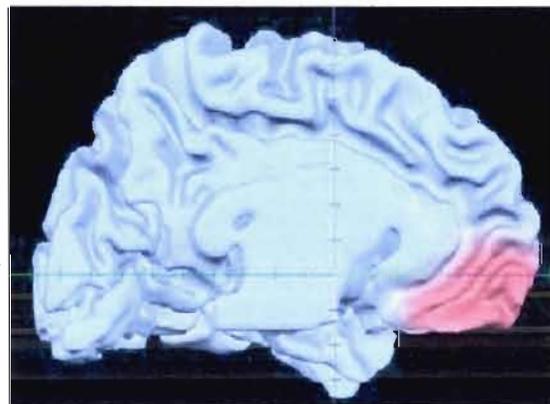
Source : Glimcher et Rustichini

Cette réfutation de la théorie de l'espérance de l'utilité a, près de quarante années plus tard, récemment de nouveau attiré l'attention de chercheurs intéressés par les théories de la décision. Certaines recherches en neurosciences permettraient aujourd'hui de décortiquer le paradoxe d'Ellsberg et d'identifier certains de ses substrats neuronaux. Hsu et Camerer (2004) ont montré qu'une région (*insula cortex*) n'est pas activée de la même façon par l'alternative d'un montant certain que par celle d'un montant incertain.

D'autres recherches faites par Bechara et ses collègues (1997) ont permis d'identifier un mécanisme émotif qui serait partiellement sensible aux situations d'ambiguïté et qui résiderait, en partie du moins, dans le cortex préfrontal ventromédial (CPFVM) (figure 2.13).

Figure 2.13 : Le cortex préfrontal ventromédial

Le cortex préfrontal ventromédial est représenté par la partie rouge de l'illustration. Le devant du cerveau est à droite. Il s'agit d'une vue médiale de l'hémisphère gauche du cerveau. Les personnes souffrant de dommages à cet endroit ont des déficits émotifs.



Source : Glimcher et Rustichini

Ce mécanisme serait nécessaire à certaines classes d'actions et ne serait sollicité qu'en certaines circonstances particulières. Comme nous l'avons préalablement souligné à la section 1.4, les patients ayant des dommages cérébraux à cet endroit spécifique auraient, à cause du non fonctionnement de certains de leurs mécanismes émotifs, des difficultés à prendre les bonnes décisions lorsque celles-ci nécessitent une évaluation correcte des

conséquences futures d'une action et ce, même si leur performance à des tests de quotient intellectuel est normale.

Lorsqu'une situation est risquée et ambiguë, Bechara et ses collègues montrèrent que ces patients, contrairement aux personnes normales, prennent les mauvaises décisions parce qu'ils manquent d'aversion aux situations ambiguës. Lorsqu'un choix a engendré de lourdes pertes, les patients sont beaucoup plus susceptibles de répéter l'expérience que les personnes normales. Leur manque d'input émotionnel leur est alors nuisible. Des observations faites à l'aide de fMRI semblent corroborer cette hypothèse. O'Doherty *et al.* (2001) ont montré que la région du CPFVM a un rapport avec l'apprentissage. Cette région est plus active lorsqu'il est nécessaire d'évaluer et d'apprendre la disponibilité de diverses récompenses à partir d'une situation ambiguë. Par contre, lorsque aucun apprentissage n'est possible et que les participants ne sont pas informés du résultat de la loterie après chacun des tirages, cette région ne montre aucune activation comme le soulignent Rustichini *et al.* (2002) : « So a possible explanation of the qualitative difference between our results [...] is that we study specifically choice, rather than learning and choice. ».<sup>107</sup> Tous ces résultats suggèrent également qu'il est possible de relier les recherches faites en économie et en neurosciences et que le rôle des émotions est donc primordial dans l'évaluation d'une situation ambiguë. Les circuits sous-jacents aux émotions seraient essentiels au processus d'apprentissage et à l'évaluation des choix risqués, tandis que la simple sélection d'alternatives ne les nécessiterait pas.

### 2.3.3) Éléments de la théorie prospective, heuristiques et effets de cadrage<sup>108</sup>

La théorie de l'espérance de l'utilité est parfois interprétée comme étant une *prescription* à la prise de décision rationnelle en situation d'incertitude. La

<sup>107</sup> Rustichini *et al.*, 2002, p. 25.

<sup>108</sup> Traduction de « Framing Effect ».

réfutation empirique de cette théorie a permis à Daniel Kahneman et Amos Tversky d'élaborer leurs travaux de recherche à partir du début des années 1970; pendant plus de vingt ans, ils ont, à partir d'une *approche descriptive*, cherché à comprendre les divers mécanismes menant à la prise de décision. L'un des moments culminants de leur collaboration scientifique fut l'élaboration de la théorie prospective (1979) qui, depuis lors, se veut une approche alternative à la théorie de l'espérance de l'utilité. Le but de ce mémoire n'étant pas de procéder à une élaboration exhaustive des travaux de Kahneman et de Tversky, nous ne soulignerons ici que les aspects qui, à notre connaissance, ont eu une incidence sur les avancées de la neuroéconomie. Essentiellement, la théorie prospective attaque l'espérance de l'utilité en trois points que nous allons maintenant couvrir.

1) L'existence d'une asymétrie entre l'impact d'un gain et l'impact d'une perte (aversion à la perte);

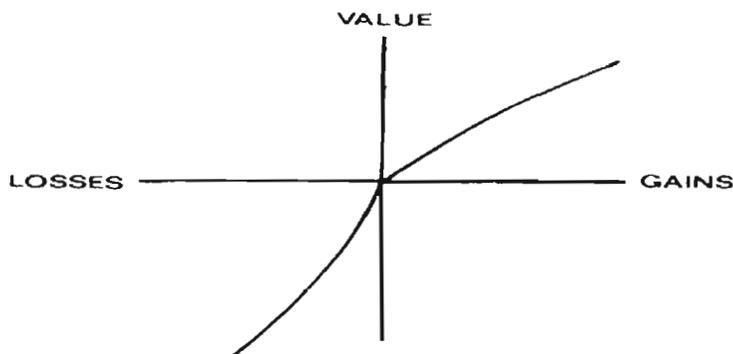
Kahneman et Tversky estiment que l'évaluation subjective que se font les gens de la fonction de valeur est qu'elle est concave pour les gains (impliquant une riscophobie) et convexe pour les pertes (impliquant une riscophilie), c'est-à-dire que :

$$| \text{désutilité de la perte } x | > | \text{utilité du gain } x |$$

Cette aversion à la perte peut impliquer des choix en contradiction avec la théorie de l'espérance de l'utilité. D'une façon systématique, les gens surestiment les probabilités d'événements catastrophiques; la forte dimension émotive rattachée à ces événements n'est pas étrangère à ce phénomène. Il est d'ailleurs intéressant de savoir qu'Adam Smith avait dénoté cette dimension émotive dès 1759 : « We suffer more...when we fall from a better to a worse

situation, than we ever enjoy when we rise from worse to a better... ».<sup>109</sup> Cette asymétrie peut parfois mener à des décisions sous-optimales, non conformes à la théorie de l'espérance de l'utilité. Par exemple, Shefrin et Statman (1985) dénotent le genre d'erreur engendrée par cette courbure de la fonction de valeur; parce que les investisseurs sont beaucoup plus sensibles aux pertes qu'aux gains, ils détiennent trop longtemps des titres qui se situent sous leur prix d'achat; pour ne pas ressentir l'effet d'une perte, ils refusent de vendre (ceci s'explique par la convexité de la fonction de valeur pour les pertes). Par ailleurs, lorsque le prix d'un titre se situe par-dessus sa valeur d'achat, les investisseurs ont tendance à le vendre trop rapidement (ceci s'explique par la concavité de la fonction de valeur pour les gains). Ce phénomène est connu sous le nom « d'effet de disposition ».<sup>110</sup> Chez les consommateurs, il a été montré que ceux-ci sont plus sensibles aux hausses qu'aux baisses de prix (Hardie *et al.*, 1993, cités dans Chen *et al.*, 2005). La figure 2.14 illustre l'asymétrie entre les gains et les pertes.

Figure 2.14 : La fonction de valeur de la théorie prospective



Source : Kahneman et Tversky (1979, p. 279)

Il est tentant de tracer un parallèle entre les processus homéostatiques décrits à la section 2.2.6 et le point d'inflexion de la fonction de valeur. Comme le soulignent Camerer *et al.*, 2003, *ibid.*, p. 19 : « Neural sensitivity to change is

<sup>109</sup> Cité dans Camerer et Loewenstein, 2002, p. 4.

probably important in explaining why the evaluation of risky gamble depends on a reference point which encodes whether an outcome is a gain or a loss». L'homéostasie implique pour la survie l'adaptation au changement; parallèlement nous avons préalablement souligné la plus grande sensibilité aux variations qu'aux niveaux absolus.

Kahneman et Tversky suggèrent que la fonction de valeur s'applique aux gains et aux pertes (notamment monétaires) et que leur évaluation se fait selon un point neutre; la perte est plus douloureuse que le bonheur engendré par un gain similaire. La courbure de la fonction de valeur pourrait très bien s'expliquer d'un point de vue évolutionniste; on la retrouverait chez d'autres espèces animales; pendant des millions d'années, les conséquences subies suite à une perte résultant d'un mauvais choix auraient été lourdes, voire fatales. Il est par conséquent tentant de suggérer que ce type de préférence aurait pu être essentiel à la survie. Différentes recherches ont récemment tenté d'établir des preuves biologiques et neurologiques à l'asymétrie entre les gains et les pertes. L'aversion à la perte serait innée et serait le fruit d'un processus évolutionniste vieux de plusieurs millions d'années<sup>111</sup> (Chen *et al.* 2005). Des études d'imagerie cérébrale ont montré que les gains et les pertes n'activent pas les mêmes régions du cerveau, et que les pertes provoquent, dans certaines régions spécifiques du cerveau, davantage d'activation que les gains (Camerer *et al.*, 1993; Smith *et al.*, 2002). D'autres études tentent d'identifier les régions du cerveau activées par diverses alternatives : Dickhaut *et al.* (2003) ont trouvé que le cortex orbitofrontal est davantage activé par l'idée d'un gain que par l'idée d'une perte. Knutson *et al.* (2000) ont trouvé que le cortex cingulaire antérieur et le thalamus étaient activés lors d'une perte. Knutson et Peterson (2005) croient que l'anticipation des gains et celle des pertes ne sont pas traitées par les mêmes mécanismes neuronaux.

---

<sup>110</sup> De l'anglais « *disposition effect* ».

2) Nombre d'heuristiques ne permettent pas la prise de décision délibérée et rationnelle;

Afin de mieux comprendre la décision en situation d'incertitude, Kahneman et Tversky proposent de remplacer les probabilités associées à un choix visant une gratification par des pondérations décisionnelles qui tiendraient davantage compte de leurs observations empiriques : les gens n'évaluent pas très bien les probabilités associées à un événement et ont des difficultés à appliquer les outils d'analyse statistique en dehors des applications purement mathématiques (voir par exemple l'erreur de la conjonction présentée à la section 1.3). En particulier, les gens ont tendance à surestimer la probabilité d'événements rares et à sous-estimer la probabilité d'événements fréquents. La surestimation des probabilités d'événements rares pourrait expliquer ce qui a priori semblait irréconciliable : le désir simultané de s'acheter une assurance et de faire du *gambling*. Cette mauvaise estimation des probabilités découlerait de ce que les auteurs appellent des heuristiques. Dans leur article de 1974, Tversky et Kahneman (1974) ne définissent pas un jugement heuristique mais le décrivent comme étant un ensemble de processus permettant de réduire le degré de difficulté associé à l'évaluation de la valeur et des probabilités. Trois types de jugements heuristiques (représentativité, disponibilité, ancrage) ainsi qu'une douzaine de biais furent présentés dans leur article. Nous présentons brièvement ces jugements heuristiques parce qu'ils auraient une influence prépondérante sur la décision.

La *représentativité* concerne le degré de similarité. Par exemple, quand A est hautement représentatif de B, la probabilité que A provienne de B est automatiquement jugée élevée. Le fait que B puisse être un événement statistiquement peu probable risque de ne pas suffisamment être dénoté. Les gens tiennent effectivement compte des probabilités lorsqu'il n'y a pas d'autres

---

<sup>111</sup> L'aversion à la perte a été observée chez les singes capucins. Nos ancêtres communs avec cette lignée auraient vécu il y a environ 40 millions d'années (Schneider *et al.*, 2001 cités dans Chen *et al.*, 2005).

informations disponibles mais ils sont insensibles à ces probabilités lorsqu'une description leur est donnée. Par exemple, si Steve est décrit comme serviable mais timide, renfermé et peu intéressé par le monde réel et qu'on demande ensuite à des gens de sélectionner la profession de Steve à partir d'une liste, ils ne tiendront pas suffisamment compte du pourcentage de la population exerçant cet emploi. Ils pourraient le classer comme étant bibliothécaire alors que la vaste majorité de la population exerce une autre profession.

La *disponibilité* correspond à un jugement heuristique où les gens qui font l'estimation de la probabilité d'un événement sont davantage influencés par la facilité avec laquelle certaines associations leur viennent à l'esprit. Au cours d'une expérience rapportée par Tversky et Kahneman (1974),<sup>112</sup> on demanda à des gens d'identifier le sexe majoritaire de divers groupes formés à partir de listes de noms énumérés séquentiellement. Dans chacune de ces listes, les noms des personnes étaient généralement inconnus mais il y avait aussi des célébrités. Il appert que le sexe des personnes célèbres eût trop d'impact sur l'estimation que faisaient les gens. Par exemple, lorsque la majorité des célébrités nommées étaient de sexe masculin, les gens estimèrent erronément que la plupart des personnes nommées dans la liste étaient de genre masculin. Et lorsque les célébrités étaient surtout de sexe féminin, les gens estimèrent que la liste contenait surtout des femmes.

L'*ancrage* implique un manque de distance entre l'évaluation subjective d'un nombre incertain et une quelconque valeur initiale (même arbitraire).<sup>113</sup> Par exemple, Tversky et Kahneman (1974) demandèrent à des personnes d'estimer le pourcentage de pays africains membres des Nations unies. Avant de donner leur réponse, les participants devaient faire tourner une roue de fortune affichant différents nombres de 0 à 100. Puis, ils devaient dire si leur estimation était plus ou moins élevée que le chiffre obtenu et ensuite déplacer

<sup>112</sup> Tversky et Kahneman, 1974, p. 1127.

<sup>113</sup> Ce phénomène a été observé par Slovic et Lichtenstein en 1971.

la roue dans la direction de ce nombre. Les résultats de cette expérience ont montré que la médiane des estimations des personnes ayant eu comme valeurs initiales les nombres 10 et 65 fut respectivement de 25 et de 45, illustrant par le fait même un effet d'ancrage.

Le parallèle entre les jugements heuristiques et les processus automatiques des quadrants III et IV du modèle de Camerer *et al.* (chapitre I) semble naturel : parce qu'ils ne nécessitent guère d'efforts et qu'ils sont faits rapidement, les jugements heuristiques sont souvent utiles mais ils peuvent néanmoins mener à des erreurs systématiques. Essentiellement, ces jugements heuristiques peuvent mener à de mauvaises évaluations des probabilités.

- 3) L'évaluation d'un prospect dépend des gains et des pertes par rapport à un point de référence (point d'ancrage) et celui-ci peut être modifié par un effet de cadrage;

L'idée de Bernouilli, à l'effet que la prise de décision soit basée sur l'utilité d'une position finale d'actifs, est remise en cause par les travaux de Kahneman et de Tversky<sup>114</sup> :

« Bernouilli's model is flawed because it is *reference-independant* (en italiques dans le texte)[...] This assumption flies against a basic principle of perception, where the effective stimulus is not the new level of stimulation, but the difference between it and the existing adaptation level. The analogy to perception suggests that the carriers of utility are likely to be gains and losses rather than states of wealth, and this suggestion is amply supported by the evidence of both experimental and observational studies of choice. »<sup>115</sup>

Kahneman et Tversky (1979) ont demandé à des participants de choisir entre différentes alternatives ( $x, p ; 0, 1-p$ ) où  $x$  est un prospect avec probabilité  $p$  comprise entre 0 et 1 et  $(x)$  est un prospect certain. Le choix proposé aux participants était le suivant :

<sup>114</sup> Kahneman a qualifié cette idée « d'erreur de Bernouilli ». Voir Kahneman, 2002.

<sup>115</sup> *Ibid.*, p. 460.

On vous a donné 1000. Vous devez maintenant choisir entre A ou B :  
 A (1000, 0.50) ou B ( 500 )

Le prospect B fut préféré par 84 % des participants alors que 16 % choisirent le prospect A. Ces choix illustrent la concavité de la fonction de valeur et l'aversion au risque pour les gains. La question fut ultérieurement reformulée de la façon suivante :

On vous a donné 2000. Vous devez maintenant choisir entre le prospect C ou D : C (-1000, 0.50) ou D (- 500 )

Ici, 31 % des participants ont préféré le prospect D tandis que 69 % ont opté pour le prospect C. Ces observations permettent de dénoter plusieurs anomalies en contradiction avec la théorie de l'espérance de l'utilité. En terme de position de l'actif final, les couples A et C ainsi que B et D sont pourtant exactement les mêmes choix :

$A = ( 2000, 0.50; 1000, 0.50 ) = C$  et  $B = ( 1500 ) = D$

Or, ce n'est pas la position de l'actif final qui est évaluée par les gens mais bien leurs gains ou pertes par rapport à un point neutre. En termes absolus, le choix révèle une incohérence; alors que seulement 16 % des participants avaient choisi A dans un contexte de gains, préférant ainsi l'option certaine, 69 % ont choisi l'option risquée C dans un contexte de pertes. Cet exemple est une illustration de l'effet de cadrage (*framing effect*). L'effet de cadrage survient lorsque la description d'une option en termes de gains (cadrage positif) engendre des choix systématiquement différents d'une description faite en termes de pertes (cadrage négatif); ainsi, de mêmes alternatives ne sont plus considérées sous le même angle. Dans l'exemple de Kahneman et Tversky, cette incohérence est expliquée par le fait que les prospects étaient présentés sous forme de gains en A et B, tandis qu'ils étaient présentés sous forme de

pertes en C et D. L'aversion à la perte a poussé les participants à devenir riscophiles en choisissant C dans le contexte de perte.<sup>116</sup>

De nombreux comportements économiques peuvent être expliqués à l'intérieur de ce cadre.<sup>117</sup> Par exemple, la courbe d'offre de travail négative des chauffeurs de taxis peut s'expliquer par la décision de continuer de travailler ou non selon l'atteinte d'un montant fixé à l'avance (point d'ancrage) (Camerer, Babcock, Loewenstein et Thaler, 1997). Barberis, Huang et Santos (2001) ont quant à eux élaboré un modèle basé sur la théorie prospective où les résultats passés des investisseurs modifient leur point d'ancrage (à partir duquel l'évaluation subjective d'un gain ou d'une perte s'établit) et ont donc une incidence sur leur aversion à la perte, leur perception et leur comportement face au risque. Benartzi et Thaler (1995) utilisent des éléments de la théorie prospective, notamment un coefficient d'aversion à la perte, pour expliquer l'énigme que représente le paradoxe de la prime de risque sur actions.<sup>118</sup>

Des chercheurs issus des neurosciences et de la neuroéconomie ont commencé à utiliser certains phénomènes reliés à la théorie prospective pour tenter d'établir leur corollaires neurologiques. Gonzalez *et al.* (2005) expliquent l'effet de cadrage à l'aide d'une théorie incorporant à la fois des considérations cognitives et affectives. Autant les processus cognitifs qu'affectifs sont coûteux en terme neurologique. Pour Gonzalez et ses collègues, toute décision prise s'analyse dans une perspective coût-bénéfice où l'individu cherche à obtenir la plus satisfaisante réponse (bénéfice) au coût neurologique le moins élevé possible. L'effet de cadrage peut s'expliquer par un arbitrage entre les coûts provoqués par les processus cognitifs et ceux des processus affectifs. Il a été

---

<sup>116</sup> En soi, cela pousse certains à remettre en cause la validité de quelques axiomes de la microéconomie. Puisque le point de référence et le contexte sont continuellement modifiés d'une situation à l'autre, il serait difficile de considérer que le choix révèle réellement les préférences. Pour plus de détails, voir le site <http://prospect-theory.behaviouralfinance.net/>

<sup>117</sup> Pour de nombreux et intéressants exemples, voir par exemple Camerer, 2000.

montré que le temps pris pour prendre une décision est plus élevé dans un contexte de perte que dans un contexte de gain (Payne, Bettman et Johnson, 1993, cités dans Gonzalez *et al.*, 2005), suggérant par le fait même une plus forte activité neurologique lorsque le choix s'effectue dans un cadrage négatif. Aidé de la technologie fMRI, Gonzalez et ses collègues ont ainsi observé que dans un contexte de gain, les efforts cognitifs requis sont moins élevés (moins d'activité neurologique dans certaines régions d'intérêt, notamment dans le cortex préfrontal et pariétal) pour l'option certaine que pour l'option risquée. Une personne a donc tendance à choisir l'option certaine (d'autant plus qu'elle est peu coûteuse affectivement). Lorsqu'il s'agit d'un contexte de perte par contre, les coûts neurologiques associées à l'option certaine sont élevés parce que les processus affectifs entrent en action (on se sent mal de choisir une perte) et c'est pourquoi l'option risquée est choisie (les coûts neurologiques engendrés par le calcul cognitif de l'espérance de gain de l'option risquée ne seraient donc pas plus élevés que ceux engendrés par les processus affectifs de l'option certaine) :

« ...the high emotional cost of the negative-certain alternative and the high computational cost of the negative-risky alternative can explain our findings here of similar levels of activation irrespective of participants' choices in response to negatively framed problems. Since participants chose the risky option more often than the certain option in response to such problems, our results suggest that people are more willing to accept a computational rather than an emotional cost, despite the fact that both kinds of costs are similarly demanding in terms of brain activity. »<sup>119</sup>

Une étude pionnière de Breiter, Aharon, Kahneman, Dale et Shizgal (2001) a permis, en utilisant la technologie fMRI, de vérifier certains des principes dérivés de la théorie prospective. L'objet de l'étude était de vérifier si :

- les perceptions pouvaient être « contrefactuelles<sup>120</sup> » ?

---

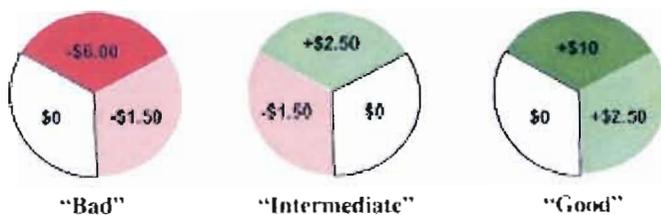
<sup>118</sup> Rappelons ici que le paradoxe de la prime sur les actions montre que les titres boursiers ne sont pas suffisamment risqués pour justifier l'écart de 6 % entre leur rendement et ceux des bons du trésor américains. Pour plus de détails, voir Mehra et Prescott, 1985.

- l'évaluation cérébrale d'une alternative risquée dépend moins de la position de l'actif en niveau que de sa variation (gains ou pertes) relatives à un point neutre ?

- l'impact cérébral d'une perte est plus élevé que l'impact d'un gain de même ampleur ?

Les participants furent informés qu'ils allaient participer à un jeu de chance. On leur octroya 50 \$ en leur disant que le jeu déterminerait s'ils allaient pouvoir conserver, augmenter ou perdre cette somme. Le jeu consistait en plusieurs tirages de deux phases. Lors de la première phase (phase d'anticipation), les participants voyaient apparaître sur un écran un des trois bouliers possibles (figure 2.15). Puis lors de la deuxième phase (phase du rendement), le boulier se mettait à tourner jusqu'à ce qu'il indique le montant obtenu. Les valeurs espérées des trois bouliers sont indiquées à la figure 2.15.

Figure 2.15 : Les trois bouliers



Source : Breiter *et al.*

Le rôle joué par les émotions dans la décision est de plus en plus sous investigation. Mellers et ses collègues (1997, 1999)<sup>121</sup> suggèrent que les émotions peuvent avoir quelque chose à voir avec les perceptions *contrefactuelles*. Une perception contrefactuelle implique qu'on se sent mieux

<sup>119</sup> Gonzalez *et al.*, 2005, p.16.

<sup>120</sup> De l'anglais *Counterfactual*. Nous utilisons l'expression contrefactuelle.

<sup>121</sup> Breiter *et al.*, *ibid.*, en référence à Mellers *et al.*, 1997 et à Mellers, Schwartz et Ritov. 1999.

de recevoir 0 \$ devant l'alternative d'une perte (- 6 \$) que devant l'alternative d'un gain (10 \$). Les probabilités associées à chacun des prospects peuvent alors rattacher une dimension émotive à la prise de décision. Breiter *et al.* établissent certains parallèles entre les perceptions contrefactuelles et le rôle joué par la dopamine. À la section 2.2.3, nous avons vu que l'arrivée de dopamine au moment de la récompense peut être interprétée comme une erreur d'anticipation du fait d'une récompense inattendue. Parallèlement, recevoir 10 \$ devant l'alternative de ne recevoir que 2,50 \$ pourrait très bien engendrer une émotion positive. Le contraire est aussi vrai comme le soulignent Breiter *et al.*:

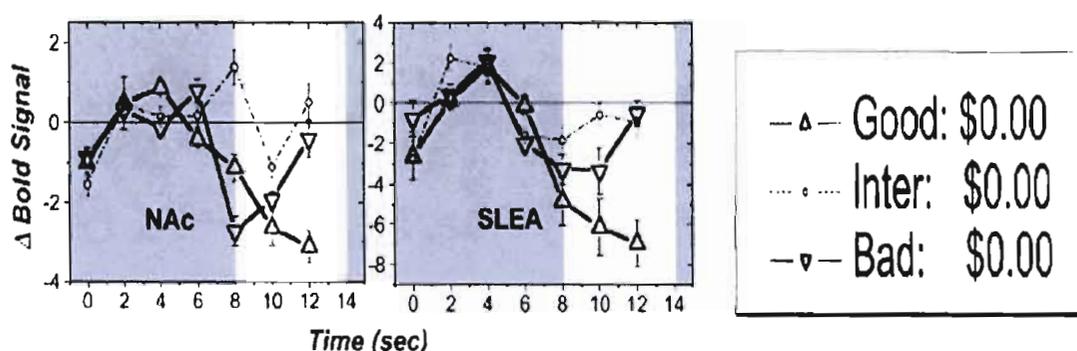
« During delivery and consumption of the reward, the activity of the dopamine neurons appears to reflect a comparison between expectation and outcome. When the expected reward is obtained, the dopaminergic neurons maintain their baseline activity, but when the reward is omitted, firing decreases. The inhibition due to reward omission can be construed as the result of a counterfactual contrast between the positive outcome that was expected and the null outcome that was obtained instead. Similarly, the activation of dopaminergic neurons by unexpected but not by expected rewards may bear some relation to the more positive emotional responses of humans to low-probability gains than to high-probability gains of the same magnitude. »<sup>122</sup>

L'étude de Breiter *et al.* corrobore l'hypothèse de comparaisons contrefactuelles de manière équivoque (figure 2.16). La zone blanche des graphiques illustre l'activation cérébrale de deux régions sous investigation au moment de l'arrivée du résultat. Le fait de « gagner » 0 \$ est stimulant dans le cas où les alternatives étaient pires (mauvais boulier) et « désactivant » dans le cas où les alternatives étaient meilleures (bon boulier). Les résultats enregistrés par le boulier intermédiaire ne corroborent cependant pas l'hypothèse de Mellers *et al.*; on aurait pu attendre de l'activation provoquée par le boulier intermédiaire qu'elle soit justement intermédiaire et comprise entre les valeurs enregistrées pour les deux autres bouliers.

---

<sup>122</sup> Breiter *et al.*, *ibid.*, p. 619.

Figure 2.16 : L'hypothèse de comparaisons contrefactuelles sous investigation



Source : Breiter *et al.*

Par ailleurs, les résultats hémodynamiques démontrèrent que certaines régions du cerveau enregistraient bel et bien la valeur espérée des trois bouliers et que l'activation neurologique engendrée par les résultats du meilleur boulier augmentait monotoniquement avec les montants en dollars. Enfin, lorsqu'on interroge les participants sur la position finale de leurs actifs, un seul d'entre eux est en mesure de répondre correctement et la plupart sous-estiment leur position finale. Tous ces résultats tendent à valider, en partie du moins, les hypothèses de la théorie prospective sous investigation. Ce genre de recherche est pionnière. Elle ne corrobore que partiellement certaines hypothèses de la théorie prospective. Néanmoins, elle soulève passablement d'enthousiasme. À plus long terme, l'enjeu des recherches en cours serait de développer une nouvelle théorie de la décision comme le soulignent Glimcher et Rustichini :

« If we succeed in understanding mechanistically how choices that violate expected utility are made at a neural level, then a new global theory of choice will be developed. ».<sup>123</sup>

<sup>123</sup> Glimcher et Rustichini, 2004, p. 451.

#### 2.4) Conclusion

L'évacuation graduelle, au sein des neurosciences, de la conception béhavioriste pour laquelle le cerveau n'est qu'un organe stimuli-réponses, permettrait à l'économie de servir d'outil conceptuel au sein des neurosciences et de la neuroéconomie. À l'origine, toute décision provient nécessairement du cerveau et du système nerveux; une meilleure compréhension de ce que le cerveau cherche à faire permettrait d'informer toute science du comportement. La fusion des neurosciences, de l'économie, et de la psychologie, est donc une séquence naturelle; chacune de ces disciplines étudiant le comportement à différents niveaux. Des approches normatives et descriptives peuvent être combinées pour l'analyse des fondements de la décision économique. Ainsi, les concepts de l'utilité et des préférences ont, depuis peu, fait l'objet d'études au sein des neurosciences. Ces outils économiques standards ont été utilisés par des chercheurs issus des neurosciences pour leurs propres recherches; et ces dernières permettent un retour d'ascenseur qui peut à son tour informer la théorie économique. Les intuitions fondamentales se cachant derrière la fonction d'utilité et le concept de l'espérance de l'utilité seraient validées par l'observation du fonctionnement cérébral de certains primates. Il a notamment été démontré que le cerveau des singes rhésus pouvait, dans certaines circonstances, à la fois encoder les probabilités et les valeurs associées à une décision de nature économique. Chez l'humain, des études subséquentes arrivent à des résultats assez similaires. Cependant, le fait que les perspectives de gains et de pertes ne soient pas traitées par les mêmes mécanismes neuronaux permettrait de nuancer ces résultats en y incorporant un autre cadre d'analyse, celui de la théorie prospective. Il est par ailleurs possible d'établir certains parallèles entre le concept de l'utilité et le rôle joué par la dopamine dans la prise de décision; les neurones dopaminergiques fourniraient la motivation nécessaire à la plupart de nos comportements. De par l'activation dopaminergique qu'elle initie, la rétribution monétaire aurait une utilité directe et celle-ci serait même plus grande dans la région du

striatum lorsque la gratification nécessiterait un effort. Mais l'activation dopaminergique serait également initiée par d'autres facteurs, notamment par les *cues* environnementaux ; ceux-ci pourraient enclencher un désir automatique de consommer, que seul le cortex préfrontal serait capable d'inhiber. Les modèles d'addiction rationnelle basés sur une planification intertemporelle délibérée sont donc remis en cause par les récentes découvertes des neurosciences. Des économistes ont d'ailleurs commencé à intégrer ces considérations biologiques dans leurs travaux de recherches, illustrant par le fait même que la neuroéconomie a déjà commencé à influencer les théories économiques.

Le concept de l'utilité, qui a par ailleurs été redéfini par Kahneman et ses collègues, implique des dissonances potentielles entre ses différentes composantes. En ayant pu identifier leurs composantes cérébrales respectives, des recherches issues des neurosciences viennent fondamentalement remettre en cause le concept des préférences révélées, et les théories du bien-être pourraient bénéficier d'une analyse incorporant certaines de ces dissonances. Puisqu'il est finalement possible de décomposer l'utilité et d'observer chacune de ses composantes, toutes ces recherches suggèrent des approches plus réalistes et plus complètes que la simple théorie des préférences révélées. Lorsque le choix est rationnel, qu'aucune dissonance n'est observée et que les préférences sont révélées par le choix, une approche neuroéconomique permet quand même de dépaqueter le concept de préférence et de séparer des phénomènes qu'on croyait similaires ; ceci puisque les préférences peuvent être parfois délibérées, parfois implicites.

Conventionnellement, les économistes considèrent que les préférences sont stables et cohérentes. Cette vision est difficilement réconciliable avec les observations empiriques issues de la psychologie et de l'économie expérimentale. En outre, si on y rajoute les dimensions biologiques de l'homéostasie et du phénomène de *alliesthesia*, et si on prend en

considération la sensibilité neuronale aux variations par rapport à un point de référence (plutôt qu'au niveau), d'autres approches deviennent encore plus acceptables.

Enfin, les falsifications de l'espérance de l'utilité ont été étudiées à l'aide des neurosciences. C'est ainsi que les incohérences démontrées par le paradoxe d'Allais peuvent s'expliquer par le fait que les processus cérébraux ne sont pas les mêmes quand un prospect certain est l'une des options que lorsque les deux options sont incertaines. Le paradoxe d'Ellsberg pourrait également s'expliquer par le fait que la région du cortex insulaire n'est pas activée de la même façon par l'alternative d'un montant certain que par celle d'un montant incertain. En outre, le cortex préfrontal ventromédial jouerait un rôle dans l'aversion émotive aux situations ambiguës; les personnes ayant des dommages dans cette région du cerveau n'afficheraient pas ce genre d'aversion, ce qui en soi sous-tend qu'il serait normal de vouloir éviter les situations ambiguës. Enfin, passablement d'aspects de la théorie prospective ont également été étudiés et partiellement corroborés par les recherches de la neuroéconomie : les gains et les pertes, ainsi que leur anticipation, ne seraient pas traités par les mêmes mécanismes neuronaux ; et les pertes engendreraient davantage d'activation neurologique que les gains. Par ailleurs, il est aisé d'établir un parallèle entre les jugements heuristiques et les processus automatiques décrits au chapitre 1. L'effet de cadrage, les perceptions contrefactuelles, l'évaluation cérébrale d'une alternative risquée et de sa position par rapport à un point neutre, les préférences intertemporelles, ont tous fait l'objet de recherches de la neuroéconomie. Bien que la neuroéconomie n'en soit qu'à ses premiers balbutiements, il est maintenant clairement établi que des chercheurs issus des neurosciences, de la psychologie et de l'économie peuvent travailler ensemble pour les avancées de cette nouvelle discipline.

## CHAPITRE III

### THÉORIE DES JEUX ET NEUROÉCONOMIE

#### 3.0) Introduction

Dans ce chapitre, nous présentons certaines utilisations de la théorie des jeux comme outils conceptuels des neurosciences. Ces liens récents permettent de nouvelles façons d'étudier le cerveau; nous croyons qu'ils peuvent tout à la fois informer les neurosciences que la théorie économique. À la section 3.1, nous présentons des résultats empiriques qui valident une conception standard de la théorie des jeux (notamment le comportement rationnel et l'atteinte d'un équilibre de Nash) et montrons en quoi l'utilisation de ces outils fournit à des chercheurs des neurosciences un nouveau cadre théorique leur permettant de jeter un autre regard sur le cerveau et ce qu'il cherche à accomplir.

Après avoir montré que certaines recherches de la neuroéconomie peuvent parfois corroborer les prédictions de comportements rationnels suggérées par la théorie des jeux, nous présentons à la section 3.2 des résultats appuyant davantage l'économie comportementale;<sup>124</sup> celle-ci, comme nous l'avons préalablement souligné à la section 1.1, remplace les hypothèses de rationalité substantive par des alternatives cognitives plus plausibles, dans la mesure où ces dernières sont concrètement observables en laboratoire. Certaines recherches suggèrent que le comportement observé en laboratoire, qui est parfois soumis à d'autres règles que celles des dimensions rationnelles de la

---

<sup>124</sup> Plus précisément la théorie comportementale des jeux.

théorie des jeux, pourrait être associé à certains substrats cérébraux; il est suggéré que ceci pourrait confirmer leur fondement. En ce sens, la neuroéconomie permet de valider les observations préalablement faites par les économistes comportementaux et cet apport est non négligeable.

Comme l'étude du cerveau incorpore nécessairement des dimensions évolutionnistes, nous soulignons au passage le rapport entre certaines hypothèses de la psychologie évolutionniste et les avancées récentes de la neuroéconomie (section 3.3). Il est suggéré qu'une meilleure compréhension de ces dimensions pourrait potentiellement aider les économistes à mieux comprendre les fondements de certaines interactions économiques.

### 3.1) Théorie des jeux, rationalité et neurosciences

À la section 2.2.1, nous avons souligné que des chercheurs avaient pu identifier, dans la région latérale intrapariétale, certains neurones capables d'encoder quelque chose s'apparentant aux conceptions standards de l'utilité; les probabilités et la valeur associées à une action gratifiante sont encodées par le cerveau, du moins dans un environnement statique. Cette capacité du cerveau permettrait ainsi l'orientation du comportement dans la direction maximisant les récompenses. Mais dans un monde complexe où la rentabilité d'une décision est également influencée par les actions des autres agents, cette approche est clairement insuffisante. C'est justement afin d'ériger une théorie économique plus complète qui incorporerait les actions stratégiques d'opposants intelligents, que John Von Neumann et Oskar Morgenstern (1944) créèrent la théorie des jeux. Pour les neurosciences, la théorie des jeux, tout comme certaines conceptions classiques de l'utilité, permet l'ajout d'outils computationnels. Des recherches récentes ont aidé certains chercheurs des neurosciences à réduire l'écart entre leur compréhension des causes proximales et ultimes du comportement. Comme le souligne Glimcher

(2002)<sup>125</sup>: « It seems imperative that these economic approaches to behavior be employed as tools to bridge the gap between ultimate and proximal causes of behavior ».

Michael Dorris et Paul Glimcher (2003, 2004) ont utilisé le jeu « Work or Shirk » d'une façon inédite pour leurs recherches en neurosciences : ils ont tenté de trouver les corollaires neurologiques de personnes interagissant dans un équilibre de Nash. Dans ce jeu, on cherche à prévoir la stratégie optimale et l'interaction entre deux joueurs qui s'affrontent : un employé et un employeur. La rémunération du travailleur est  $W$  mais travailler lui coûte un effort  $E$  qu'il préférerait ne pas avoir à donner. Les profits de l'employeur sont mesurés par la production du travailleur  $P$  moins ce qu'il lui en coûte en salaire  $W$  et en coût de vérification ou d'inspection  $I$ . L'employeur préférerait ne pas avoir à vérifier la production de l'employé car cela lui occasionne des coûts additionnels. La matrice des paiements relatifs est présentée au tableau 3.1. Les paiements du travailleur sont situés dans chacune des cellules, à la gauche de la virgule, et les paiements de l'employeur se trouvent à la droite de la virgule. La décision de l'employeur consiste à inspecter ou à ne pas inspecter tandis que le choix du travailleur se pose entre le fait de travailler ou de se la couler douce (resquiller).

Tableau 3.1 : La matrice des paiements du jeu de l'inspection et du resquillage

	Inspecter	Ne pas inspecter
Travailler	W-E, P-W-I	W-E, P-W
Resquiller	0, -I	W, -W

Source : Glimcher (2003)

Dans ce jeu, un équilibre de Nash sera atteint lorsque le travailleur et l'employeur se trouveront tous deux indifférents entre le choix de l'une ou de

<sup>125</sup> Glimcher, 2002, p. 331.

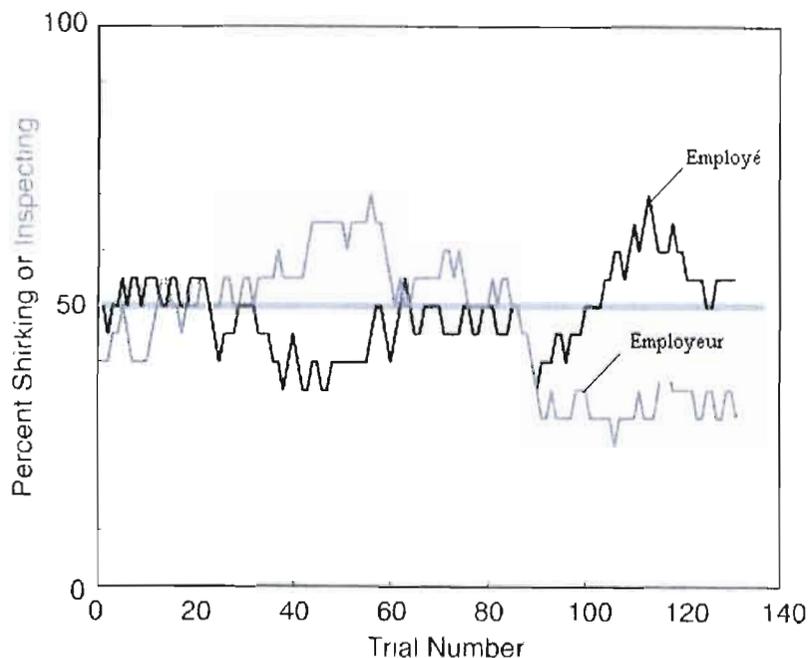
l'autre de leurs options respectives. La solution à ce jeu est que le travailleur devrait resquiller dans une proportion de  $I/W$  alors que l'employeur devrait inspecter dans une proportion de  $E/W$ .<sup>126</sup>

Dorris et Glimcher demandèrent à des humains de jouer à ce jeu à une centaine de reprises pour des périodes de 1 à 2 heures, toujours avec le même adversaire. Afin de les garder motivés, les joueurs étaient rémunérés et pouvaient gagner jusqu'à 30 \$ par heure selon leur performance. Les deux joueurs étaient séparés, chacun se trouvant dans une pièce distincte et le résultat de leur décision était transmis par un terminal après chacun de leurs choix. Il est intéressant de dénoter que les participants, malgré qu'ils ne furent jamais informés des paiements de la matrice, et bien qu'ils n'avaient aucune idée de ce que pouvait être une stratégie mixte ou un équilibre de Nash, parvinrent à se maintenir très proche de la solution d'équilibre. Par exemple, au cours d'un jeu où la solution optimale était d'inspecter et de resquiller dans une proportion de 50 %, les deux chercheurs obtinrent des résultats assez fidèles à la solution optimale de l'équilibre de Nash (figure 3.1).

---

<sup>126</sup> Pour le travailleur, l'équilibre de Nash survient lorsque  $EU(W) = EU(S)$ . En supposant que les probabilités d'inspection sont données par  $P(I)$  et que la fonction d'utilité est linéaire, on a que :  $P(I)(W-E) + (1-P(I))(W-E) = P(I)0 + (1-P(I))W$ , donc  $(W-E) = (1-P(I))W$  et en isolant  $P(I)$ , on trouve  $P(I) = E/W$ . Pour l'employeur, la condition d'équilibre est la même, c'est-à-dire que l'espérance de l'utilité de ses deux options doivent s'équivaloir.

Figure 3.1 : Résultats du jeu de l'inspection et du resquillage



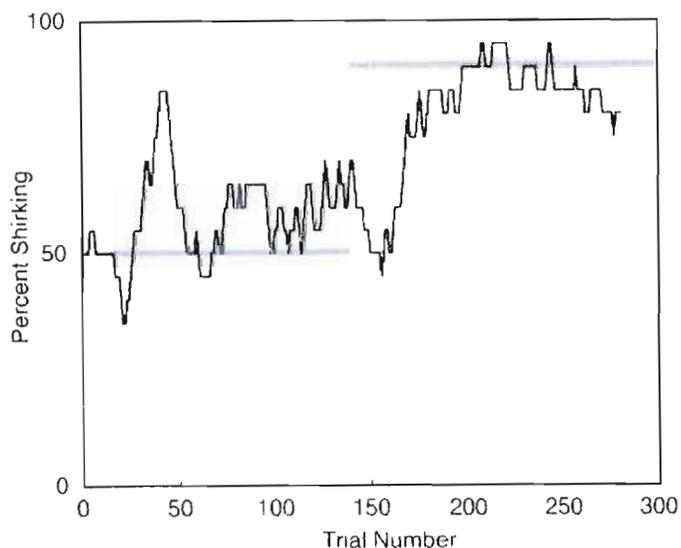
Explications : Le trait horizontal indique le pourcentage de la stratégie optimale (inspecter et resquiller la moitié du temps). Le comportement de l'employé est identifié par le trait foncé et celui de l'employeur correspond au trait pâle. Le graphique révèle que les joueurs, malgré qu'ils n'avaient jamais été informés des paiements de ce jeu, ne s'éloignent pas vraiment de l'équilibre de Nash. À la lumière du graphique, on constate que des inspections moins fréquentes engendraient davantage de resquillage, et vice-versa.

Source : Glimcher (2003)

Dorris et Glimcher utilisèrent ensuite un programme d'ordinateur pour jouer le rôle de l'employeur. Les participants ne furent cependant pas informés qu'ils jouaient contre un logiciel. Le programme de cet ordinateur avait la particularité d'être capable de détecter des comportements non aléatoires chez les joueurs placés dans le rôle du travailleur. Si un joueur ne se comportait pas réellement de façon aléatoire, ses paiements allaient en être affectés (un comportement prévisible diminue les paiements dans ce jeu). Après avoir fait

jouer les participants, Dorris et Glimcher constatèrent que ceux-ci agissaient, à peu de choses près, encore une fois en conformité avec un équilibre de Nash (figure 3.2). Puis, ils montrèrent à des singes (qui n'étaient pas rémunérés en argent mais avec du jus<sup>127</sup>) à jouer à ce jeu contre le même ordinateur et les bêtes obtinrent des résultats très similaires à ceux des être humains (figure 3.3.). Avant chacune des décisions prises par les travailleurs, et ce autant pour les humains que pour les singes, le programme de l'ordinateur fut incapable de détecter un comportement prévisible<sup>128</sup>.

Figure 3.2 : Comportement humain dans le jeu de l'inspection et du resquillage



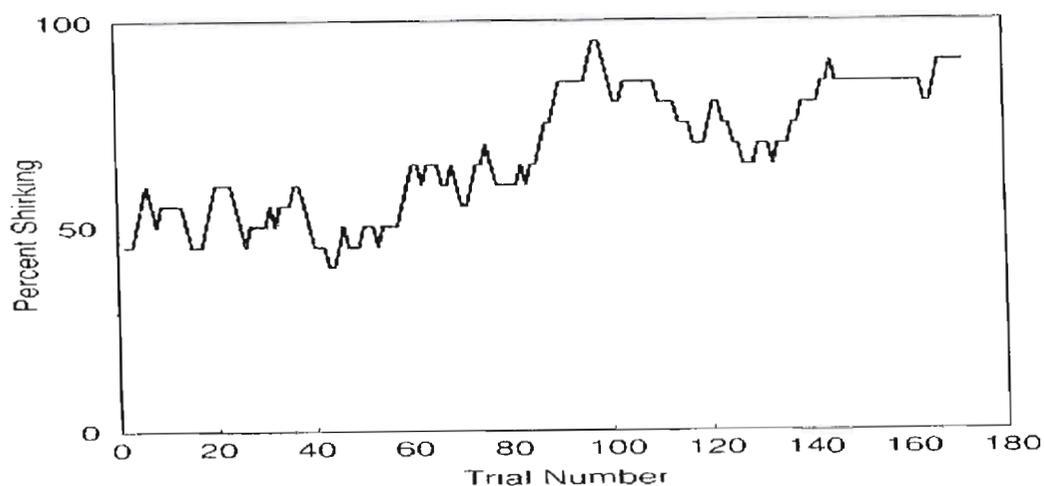
La figure 3.2 illustre un comportement tendant vers un équilibre de Nash. Ici, le travailleur est joué par un volontaire humain alors que l'ordinateur joue le rôle de l'employeur. Lorsque la stratégie optimale passe d'un resquillage de 50 % à 90 %, le travailleur se met à resquiller sensiblement selon les mêmes proportions.

Source : Glimcher (2003)

<sup>127</sup> Dans Dorris et Glimcher (2004), les auteurs disent que la rétribution fut en eau. Dans Glimcher (2003), l'auteur dit qu'elle fut donnée en jus.

<sup>128</sup> Dorris et Glimcher testèrent mathématiquement cette hypothèse à l'aide d'un test statistique (*Ljung-Box Q-Statistic*). Ils tentèrent de trouver chez les participants un quelconque *pattern* à l'aide d'un logiciel spécialisé mais ils ne furent pas capable d'en détecter un. D'une séquence à l'autre, le comportement était donc indéterminé dans ce jeu.

Figure 3.3 : Comportement du singe dans le jeu de l'inspection et du resquillage



La figure 3.3 illustre que le singe se comporte aussi conformément à un équilibre de Nash lorsqu'il est confronté à un programme d'ordinateur jouant le rôle de l'employeur. Lorsque la stratégie optimale passe d'un resquillage de 50 % à 90 %, le singe se met à resquiller proportionnellement. À noter la similitude entre le comportement du singe et celui de l'humain de la figure 3.2.

Source : Glimcher (2003)

### Corollaire cérébral

Il est clair que la solution à ce jeu implique une stratégie mixte chez les deux joueurs<sup>129</sup>. La stratégie mixte requiert d'un individu qu'il soit capable de produire des comportements incertains aux yeux de ses opposants afin de pouvoir maximiser ses gains. Par exemple, si l'employé se mettait à resquiller un jour sur deux par alternance, l'employeur aurait tôt fait de le remarquer, et finirait par choisir de n'inspecter précisément que ces jours-là. Il en va de la

<sup>129</sup> Pour de récentes recherches sur les liens entre le cerveau et les stratégies mixtes, voir Wilson *et al.* (2001) qui suggèrent que les stratégies dominantes et mixtes ne sont pas traitées par les mêmes mécanismes cérébraux ou encore Barraclough *et al.* (2004) qui cherchent à comprendre le rôle du cortex préfrontal dans l'élaboration d'une stratégie mixte.

même logique pour l'employé; s'il parvenait à détecter une séquence particulière chez son employeur, il choisirait de ne se la couler douce que les jours où celui-ci n'inspecte pas.

Mais en quoi cela peut-il être relié au cerveau ? En fait, l'idée de Nash, à l'effet qu'une stratégie mixte puisse engendrer un équilibre, vient frapper le paradigme du dualisme cartésien de plein fouet. Rappelons qu'un équilibre de Nash requiert que les valeurs espérées de chaque choix ou ressources soient équivalentes, de telle sorte qu'un individu soit indifférent entre une ou l'autre de ses options. À la section 2.1, nous avons souligné que pour Descartes, seul le comportement simple et déterminé, initié par un processus de stimuli-réponses, pouvait être compréhensible et explicable. Or, l'existence d'un équilibre mixte requiert à la fois un comportement prédéterminé (celui d'agir selon les proportions de la stratégie optimale) mais également un comportement indéterminé (le fait d'agir de façon imprévisible à chacun des tirages). Avant d'arrêter son choix, le joueur doit adopter un comportement imprévisible aux yeux de ses opposants. Dans la nature, certaines classes d'actions nécessitent un comportement imprévisible. Quand on analyse ces actions dans le cadre de la théorie des jeux, elles deviennent explicables. Alors que le comportement d'un joueur est, d'une séquence à une autre, imprévisible, le choix de la stratégie optimale est, quant à lui, mathématiquement très prévisible. Pour comprendre le fonctionnement du cerveau, Glimcher postule qu'il faille accepter l'idée qu'il est possible d'étudier certains types de comportements imprévisibles en autant que l'on dispose des outils appropriés. Cette imprévisibilité ou cet indéterminisme est de deux ordres : le premier se réfère aux limitations épistémologiques auxquelles fait face tout animal; avant de prendre une décision, on fait nécessairement face à une information incomplète. Pendant des millions d'années, le cerveau a évolué dans un environnement incertain. Le second est un non déterminisme nécessaire pour faire face à la compétition des autres organismes. Tenter de comprendre ce que le cerveau cherche à accomplir sans tenir compte de ces

notions est selon lui impossible. En référence aux limites d'une approche déterministe<sup>130</sup>, Glimcher (2003, *ibid.*, p. 341) dit :

« The great challenge faced by classical monism is that it cannot explain non determinate behavior. The dualist postulate of free will solves this problem by attributing indeterminacy to a process that cannot, in principle, be studied. Game theory, however, provides the tools for understanding both determinate and indeterminate behavior without resort to an ill-defined non physical process ».

La théorie des jeux peut être considérée comme un outil des neurosciences pour certaines de ses considérations stratégiques et dynamiques. Elle vient accroître, de par les interactions avec d'autres opposants intelligents, le degré d'analyse que ne permettait pas le concept d'utilité statique. À la section 2.2.1, nous avons souligné que les travaux de Platt et Glimcher avaient montré que les neurones du cortex pariétal étaient fortement corrélés avec la valeur d'une rétribution ainsi qu'avec ses probabilités. Mais l'environnement entourant ces expérimentations était statique. Afin de vérifier s'il était possible d'analyser les corollaires neurologiques de choix impliquant des interactions stratégiques, Dorris et Glimcher enregistrèrent l'activité d'un seul neurone du cerveau des singes dans la région du cortex latéral pariétal pendant que ceux-ci participaient au jeu de l'inspection et du resquillage. L'objectif était de vérifier si certains types de comportements complexes pouvaient s'analyser à l'aide de la théorie des jeux :

---

<sup>130</sup> Glimcher fait référence au déterminisme limitatif de Descartes, pour qui le comportement complexe doit être exclu de toute analyse parce qu'il est né d'une âme immatérielle et ne peut donc faire l'objet d'étude. Il réfère également à une approche physiologique moniste où tous les comportements ne sont que le produit de réflexes déterministes. Cette approche ne peut non plus, selon lui, parvenir à expliquer le comportement complexe parce qu'elle fait fi des mathématiques probabilistes.

« The theoretical work of Von Neumann and Morgenstern argued that all kinds of behavior, both simple and complex, could be viewed as points along a single mathematical continuum, a continuum that ranges from determinate behavior to probabilistic behavior. This theoretical claim raised what Michael Dorris and I considered a crucial question : Did neurons in area LIP (*Lateral Intraparietal*<sup>131</sup>) also see this as a single continuum, a single continuum governed by the same rules and computations ? Or, alternatively, did neurons in area LIP see simple and complex behaviors as distinct in the way Descartes had originally proposed ? ». <sup>132</sup>

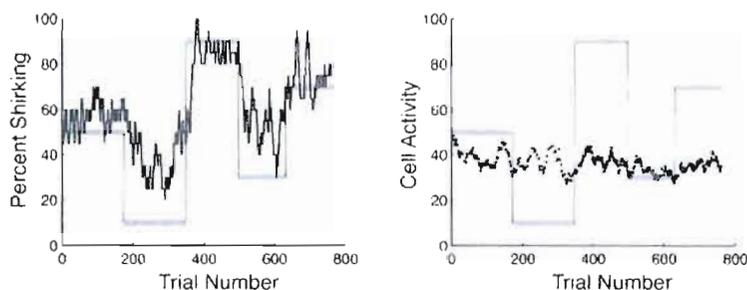
Nous avons préalablement mentionné qu'au cours de l'expérimentation de Platt et Glimcher (1999), des neurones de la région latérale intrapariétale s'activaient beaucoup, de blocs en blocs, lorsque les montants en rétribution de jus étaient augmentés. Dorris et Glimcher (2004) modifièrent aussi les valeurs des rétributions, de blocs en blocs, mais dans leur expérimentation les singes se trouvaient continuellement dans, ou très près d'un équilibre de Nash. Nous savons qu'en étant dans un tel équilibre, les valeurs relatives de l'espérance de l'utilité sont exactement les mêmes quelle que soit l'alternative; c'est ce qui définit un équilibre de Nash. Au niveau de l'expérimentation, cela implique donc que même si les montants de jus furent modifiés d'un bloc à un autre (par le biais d'une nouvelle solution optimale), les valeurs relatives de l'utilité des deux options se devaient de demeurer les mêmes. Si les singes maintinrent un équilibre de Nash, l'espérance de l'utilité relative des deux options se devaient d'être équivalentes. En observant l'activité de certains neurones du cortex latéral pariétal, c'est exactement ce que Dorris et Glimcher observèrent (figure 3.4).

---

<sup>131</sup> Ce qui est entre parenthèses et en italique à été ajouté par l'auteur du présent mémoire.

<sup>132</sup> Glimcher, 2003, p. 314.

Figure 3.4 : Comportement du singe et activité neuronale



La partie gauche de la figure 3.4 montre que le comportement du singe demeure à peu près optimal d'un équilibre de Nash à un autre. Simultanément, la partie droite de la figure 3.4 mesure l'activité d'un neurone de la région latérale intrapariétale (LIP). Le neurone ne capte pas les différentes valeurs des équilibres de Nash mais possiblement l'espérance de l'utilité relative des deux options qui, par définition, demeure la même d'un bloc à l'autre.

Source : Glimcher (2003), p. 316

Ce que Dorris et Glimcher ont cherché à faire, c'est du « monisme indéterminé ». En reliant les causes proximales et ultimes du comportement, ils ont montré qu'un choix incertain (avant chacun des tirages) pouvait néanmoins être déterministe (chance de survie associée à la maximisation des gains). L'avantage de cette approche neuroéconomique, c'est qu'elle leur permet un plus grand pouvoir explicatif. L'activation neurologique observée dans le cortex latéral intrapariétal n'aurait jamais pu être expliquée à l'intérieur du paradigme stimuli-réponses.

### 3.2) Théorie comportementale des jeux et neurosciences

Au cours des dernières années, plusieurs falsifications de la théorie des jeux rationnels ont été rendues possibles par l'essor de l'économie expérimentale. Ces falsifications ont initié une nouvelle branche de la théorie des jeux

(Camerer, Ho et Chong, 2001) que l'on a appelée la théorie comportementale des jeux (*Behavioral Game Theory*).

Pour donner un avant-goût du type de recherches rendues possibles par la fusion des neurosciences et de la théorie comportementale des jeux, nous présentons ici trois types d'interactions stratégiques et abordons certaines de leurs dimensions neurobiologiques. Dans la présente section, nous verrons les jeux de la confiance, du dilemme du prisonnier et de l'ultimatum.

### 3.2.1) Stratégie, échange et confiance

L'échange, qui implique nécessairement la coopération et la confiance mutuelle, est à la base de presque toutes les activités économiques humaines. On sait que diverses formes de coopération ont été observées chez d'autres espèces animales (par exemple la chasse, les soins parentaux ou le partage de la nourriture), mais jamais elles n'atteignent l'ampleur des interactions humaines. C'est par le biais de l'échange et de certaines règles de conduite qu'ont été rendues possibles les imposantes dimensions des économies humaines. Le fait que des milliards de transactions économiques volontaires ont quotidiennement lieu sur cette planète n'est pas un phénomène banal. La main invisible, rendant possibles toutes ces transactions, implique que chaque individu qui entrevoit l'échange croit non seulement pouvoir y trouver son compte, mais également la confiance de ne pas être floué dans la transaction. Un échange économique volontaire ne peut avoir lieu sans confiance mutuelle. Cette caractéristique ne date pas d'hier : des instincts génétiquement hérités servaient à renforcer la coopération des membres vivant en bandes. L'individualisme primitif est un mythe. Le sauvage n'est pas solitaire, et ses instincts sont collectivistes (Hayek, *ibid.*, p. 20). Pour nos ancêtres hominidés, le partage de nourriture était essentiel à la survie de l'espèce (Lee et Devore, 1968 cités dans Rilling *et al.*, 2002). L'hypothèse de l'intelligence sociale, à l'effet que le cerveau ait évolué pour faciliter l'interaction avec les autres, suggère des habiletés « à se mettre dans la peau de l'autre » (Robson, 2001).

Une telle capacité implique une « théorie de l'esprit »<sup>133</sup>. Des expériences ont montré qu'un enfant de 4 ans est capable de comprendre qu'une autre personne a une fausse croyance, ce que ne peut pas encore faire un enfant de trois ans. Comme nous l'avons souligné à la section 1.3.3, les autistes ont des déficits en matière d'inférence sur les intentions des autres. Certaines recherches suggèrent donc qu'un module spécifique de la théorie de l'esprit aurait été identifié dans notre cerveau (Lieberman *et al.* 2002; Robson, 2001; Rilling *et al.*, 2004). L'évolution aurait favorisé une telle capacité d'inférence car elle aurait permis des réponses flexibles et adaptables lors de certaines sortes d'interactions stratégiques (Robson, 2001). Le développement de l'intelligence sociale a facilité les échanges. Il aurait permis la renonciation à des gratifications immédiates lors de certaines interactions stratégiques. Ce faisant, il aurait rendu possibles les gratifications futures accompagnant la coopération ou la réciprocité.

Maintes observations rendues possibles par l'essor de l'économie expérimentale ont permis de développer des modèles d'utilité sociale où les paiements des autres sont également pris en compte (par exemple, voir Rabin, 1993). Essentiellement, ces observations empiriques ont démontré que les gens peuvent faire preuve d'une réciprocité forte, c'est-à-dire qu'ils vont répondre à une action perçue comme étant positive par une action positive, et à une action perçue comme étant hostile par une action hostile (même si cela leur est coûteux de le faire).<sup>134</sup> Il est plausible de croire que cette réciprocité forte aurait pu jouer un rôle essentiel d'un point de vue évolutionniste. La réciprocité forte, qui implique une propension à fournir des punitions et

---

<sup>133</sup> L'expression provient de Premack et Woodruff, 1978, cités dans Robson, 2001. Un individu capable d'attribuer à un autre un état d'esprit possède une telle théorie. Pour une revue des corollaires cérébraux associés à la théorie de l'esprit, voir Rilling *et al.*, 2004.

<sup>134</sup> Pour une explication sur les différentes natures des préférences sociales, voir Fehr et Fischbacher, 2001.

récompenses altruistes, seraient particulièrement importante et jouerait un rôle fondamental dans la coopération humaine.

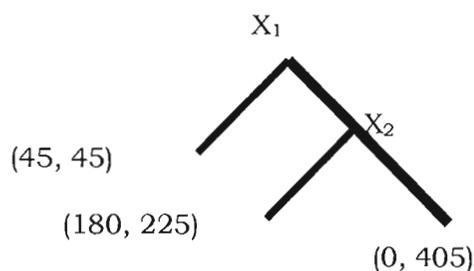
De récentes études montrent que la réciprocité forte est cruciale pour l'établissement de comportements coopératifs dans des groupes où l'on retrouve un certain nombre d'individus égoïstes (Fehr et Rockenbach, 2004). Des punitions altruistes, même si elles sont coûteuses pour ceux qui les appliquent, encouragent par la suite les individus égoïstes à adopter des stratégies plus coopératives, stratégies qui seront alors bénéfiques à l'ensemble du groupe. Si cette sorte d'interaction avait pu avoir des conséquences évolutionnistes avantageuses, c'est notamment parce qu'elle aurait favorisé la coopération à l'intérieur des petits groupes; pendant des millénaires, les sociétés humaines étaient dépourvues d'institution juridiques capables de punir les personnes trichant lors de transactions économiques et les sanctions privées auraient permis de renforcer les normes sociales (De Quervain *et al.*, 2004).

De récents développements en neurosciences permettent aujourd'hui de tester certaines hypothèses concernant les fondements biologiques des préférences sociales. Les techniques d'imagerie cérébrale ont ainsi permis d'intéressantes avancées. Par exemple, il a été montré que la coopération mutuelle et la punition de tricheurs activaient certaines zones du circuit de la récompense (voir figure 2.8), suggérant par le fait même que l'évolution aurait doté l'humain de mécanismes proximaux gratifiant le comportement altruiste. Les trois jeux que nous allons maintenant présenter permettent d'illustrer en quoi une combinaison de l'économie expérimentale et des neurosciences ouvrent de nouveaux champs de recherches. Ces dernières pourraient nous éclairer sur les bases des comportements d'interactions sociales que l'on retrouve dans les sociétés humaines.

### 3.2.2) Le jeu de la confiance

On doit à Kevin McCabe et à ses collègues (2001) la première expérience (initiée par des économistes) combinant les neurosciences et la théorie des jeux. McCabe et ses collègues demandèrent à douze personnes de participer à une version du jeu de la confiance (figure 3.5) afin de tester si un module précis favorisant le comportement coopératif pouvait être localisé dans le cerveau. Dans ce jeu, le joueur  $X_1$  peut prendre la décision de ne pas faire confiance à l'autre, auquel cas le jeu se termine et les deux joueurs obtiennent 45 sous. Si au contraire le joueur  $X_1$  fait confiance au joueur  $X_2$ , ce dernier peut alors décider de garder toute la mise (405 sous) en ne donnant rien au joueur  $X_1$ . L'autre possibilité s'offrant au joueur  $X_2$ , c'est d'opter pour la réciprocité et de partager la mise (paiements de 180 sous pour le joueur  $X_1$  et de 225 sous pour le joueur  $X_2$ ).

Figure 3.5 : Jeu de la confiance et de la réciprocité



Source : McCabe *et al.*

McCabe et ses collègues rapportent qu'en d'autres expériences de laboratoire, on observe que le joueur  $X_1$  fait confiance au joueur  $X_2$  la moitié du temps et que trois-quarts des joueurs  $X_2$  font preuve de réciprocité. Pour la présente étude, McCabe et ses collègues firent jouer les participants entre eux, dans un rôle ou un dans un autre, ainsi qu'avec un ordinateur. Avant de jouer, les

participants apercevaient, par le biais d'une photographie apparaissant sur un écran, le partenaire avec qui ils seraient jumelés pour la transaction. On leur expliqua également que l'ordinateur jouant le rôle du joueur  $X_1$  transférerait à tout coup l'argent au joueur  $X_2$ . Quand l'ordinateur se trouverait à jouer le rôle de  $X_2$ , il opterait de coopérer 75 % des fois. Utilisant la technologie fMRI, McCabe et ses collègues observèrent que les choix les plus coopératifs engendraient une plus grande activation cérébrale dans certaines régions préfrontales.<sup>135</sup> Cette activation était perceptible lorsque les participants jouaient avec un partenaire humain alors qu'elle ne l'était pas lorsqu'ils affrontaient l'ordinateur. Chez les joueurs les moins coopératifs, cette différence d'activation cérébrale ne fut pas observée. Pour interpréter ces résultats, McCabe *et al.* ont suggéré l'existence d'une toile corticale commune qui permettrait la coopération humaine. Cela a fait dire à Kevin McCabe, du Centre d'étude en neuroéconomie de l'Université George Mason : « We are biologically endowed to engage in personal exchange ».<sup>136</sup> Et cette toile ne s'activerait que lors d'un choix coopératif avec un humain. L'attention portée aux gains mutuels, l'inhibition nécessaire de la gratification immédiate qui rend possible les futures coopérations, ne seraient nécessaires que lorsqu'un humain est confronté à un autre humain. Si les régions identifiées par les auteurs sont effectivement nécessaires aux choix coopératifs, le fait qu'elles ne soient pas activées chez les participants choisissant de ne pas coopérer pourrait être une indication que d'autres processus guident leur décision, et que celle-ci est d'abord vue sous un angle de maximisation d'utilité.

Une autre version de ce jeu (figure 3.6) a récemment été utilisée par Kosfeld *et al.* (2005) pour chercher à comprendre le rôle joué par l'ocytocine dans le comportement de confiance. L'ocytocine est une petite peptide composée de neuf acides aminés que produit notamment l'hypothalamus. Dans le règne

---

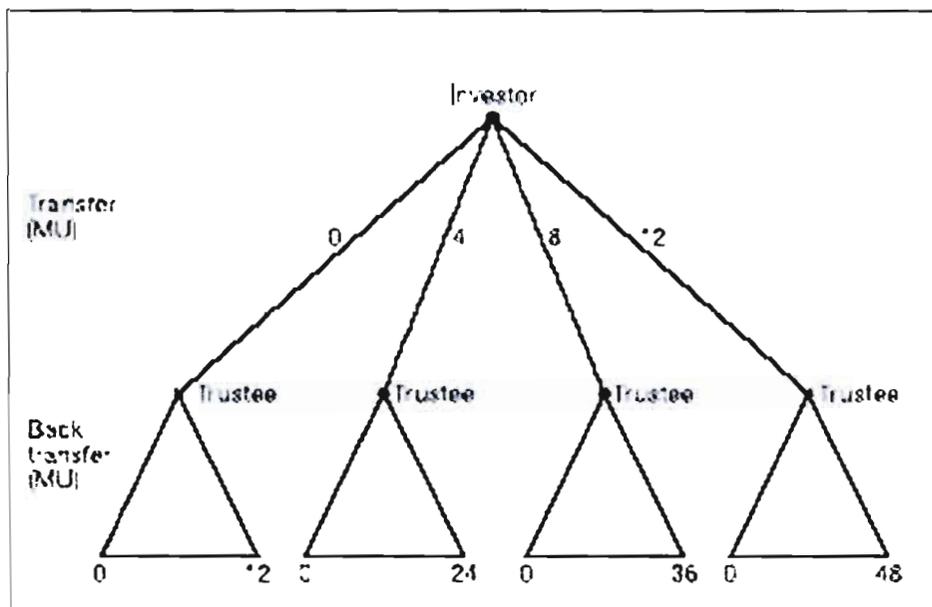
<sup>135</sup> Notamment dans certaines régions médiales du cortex préfrontal. Les régions Broadman 8 et 10 étaient activées lorsque le joueur  $X_1$  décidait de faire confiance au joueur  $X_2$ .

<sup>136</sup> Powell, 2003.

animal, l'ocytocine jouerait un rôle dans certains comportements sociaux; elle contribuerait notamment aux liaisons après l'accouplement, aux soins maternels et au comportement d'approche (Kosfeld *et al.* 2005; Damasio, 2005). Voulant vérifier si elle pouvait avoir quelque impact chez l'humain, Kosfeld et ses collègues utilisèrent le procédé du double-aveugle pour vérifier si l'administration d'une dose d'ocytocine par voie nasale pourrait avoir un impact dans le choix de l'investisseur. Dans cette version du jeu, les deux joueurs commencent avec 12 unités monétaires et l'investisseur peut conserver ou transférer sa dotation en tout ou en partie. Plus l'investisseur envoie d'argent au second joueur, plus les montants disponibles pour le partage sont augmentés (ils sont triplés dans chacun des cas). Une plus grande confiance du joueur 1 permet donc la possibilité d'obtenir davantage d'unités monétaires pour les deux joueurs. Cependant, un plus grand transfert peut également impliquer une plus grande perte pour l'investisseur, si la confiance de celui-ci n'est pas honorée par le second joueur.

Pour vérifier si l'ocytocine pouvait augmenter le degré de confiance entre humains, Kosfeld *et al.* invitèrent des étudiants à jouer à ce jeu en utilisant des rétributions monétaires. Les résultats de leur étude sont probants : l'ocytocine aurait favorisé un plus haut degré de confiance chez les investisseurs au cours de cette expérience; des 29 sujets investisseurs dans le groupe ayant reniflé de l'ocytocine, 13 (45 %) ont choisi d'investir le montant maximal contre seulement 6 (21 %) dans le groupe ayant reniflé un placebo.

Figure 3.6 : Le jeu de la confiance



L'investisseur peut choisir de transférer différentes sommes monétaires (0, 4, 8, 12) au second joueur. Ce dernier peut alors choisir de transférer à son tour un montant compris entre les deux extrémités du triangle.

Source : Kosfeld *et al.*, 2005

En moyenne, les transferts de l'investisseur sont 17 % plus élevés dans le groupe ayant consommé de l'ocytocine (test unilatéral Mann-Whitney :  $z = -1,897$ ,  $p = 0,029$ ). Afin de vérifier que l'ocytocine avait une incidence sur le sentiment de confiance interpersonnel, et non pas plutôt spécifiquement sur l'aversion au risque, Kosfeld et ses collègues demandèrent à des participants placés dans le rôle de l'investisseur de jouer contre un programme d'ordinateur affichant le même degré de risque que celui observé chez les participants humains. Leurs résultats (tableau 3.2) suggèrent que l'effet de l'ocytocine est spécifiquement relié aux interactions humaines (les transferts moyens de 7,5 unités monétaires sont les mêmes dans les deux groupes jouant contre un ordinateur).

Fait à noter, le second joueur n'est pas influencé par sa consommation d'ocytocine. Kosfeld *et al.* écartent donc l'hypothèse que l'ocytocine puisse favoriser une inclination pro-sociale plus générale; seule la décision d'investir en semble tributaire. Les principes de réciprocité et de confiance n'engendreraient pas les mêmes mécanismes psychiques. L'ocytocine jouerait sur le comportement d'approche et de confiance, et impliquerait donc davantage la décision de l'investisseur que celle de l'autre joueur (pour qui le comportement est davantage relié à un principe de réciprocité qu'à un comportement d'approche).

Tableau 3.2 : L'impact de l'ocytocine dans la décision d'investissement

Table 1 | Median and average transfer behaviour of investors

	Trust experiment		Risk experiment	
	Oxytocin group	Placebo group	Oxytocin group	Placebo group
Mean average transfer (MU)	9.6	8.1	7.5	7.5
Median average transfer (MU)	10	8	8	8
Standard deviation of transfers (MU)	2.8	3.1	3.3	3.4
Number of observations	29	29	31	30

Source : Kosfeld *et al.*

Presque toutes les interactions sociales chez l'humain impliquent, à différents degrés, la confiance vis-à-vis d'autrui. Mais encore peu d'études avaient abordé les dimensions neurobiologiques de la confiance. Les auteurs concluent que leur découverte pourrait, en plus d'informer les théories de la décision, permettre de soigner certains problèmes mentaux comme les phobies sociales. Enfin, les auteurs soulignent également le danger que pourrait représenter l'usage de cette substance à des fins plus stratégiques (techniques marketing, etc.).

Un autre groupe de recherche (King-Casas *et al.*, 2005) a récemment poursuivi les travaux sur les fondements biologiques de la confiance en combinant les

techniques d'imagerie cérébrale<sup>137</sup> à un jeu pour localiser et analyser les mécanismes cérébraux activés par ce type de comportement. Dans leur version du jeu de la confiance, l'investisseur peut transférer 20 dollars ou moins au second joueur. Une fois le transfert effectué, le montant est triplé et le second joueur peut alors décider de faire preuve de réciprocité en retournant une partie de cette somme à l'investisseur. Afin de tenir compte du fait que les échanges sont souvent répétés et non isolés, et que les algorithmes cérébraux sont supposés refléter cet état de la nature, les participants furent jumelés avec la même personne (sans savoir laquelle) et jouèrent ce jeu en conservant le même rôle pendant dix reprises. King-Casas et ses collègues notèrent, d'une ronde à une autre, une forte corrélation entre les transferts marginaux des deux joueurs.<sup>138</sup> Utilisant ensuite la technologie fMRI pour observer l'activation neurologique des deux joueurs simultanément, ils notèrent que « l'intention de faire confiance » (d'envoyer plus d'argent) était corrélée avec l'activité neurologique du noyau caudé. Fait notable, il semble que la réputation des joueurs finissait par être encodée avec l'expérience. Par exemple, alors que dans les premières rondes le noyau caudé du second joueur était activé brièvement après la décision de l'investisseur; cette activation avait lieu, dans les dernières rondes, avant même que ne soit révélée la décision de l'investisseur. Ceci n'est pas sans rappeler l'encodage des erreurs d'anticipation décrit à la section 2.2.3. Or, la région du noyau caudé est en lien avec le système dopaminergique (King-Casas *et al.*, *ibid.*, p. 82). Cette région permettrait donc de formuler un modèle d'anticipation des actions des autres.

<sup>137</sup> Plus précisément la technologie *hyperscanning*.

<sup>138</sup> C'est leur définition empirique de la réciprocité. Plus spécifiquement, King-Casas *et al.* définissent la réciprocité comme étant le transfert monétaire marginal en réponse au transfert marginal de l'autre joueur. Par exemple, la réciprocité de l'investisseur à la ronde  $j$  se mesure par  $\Delta I_j - \Delta R_{j-1}$  où  $\Delta I$  est le changement dans les quantités investies entre les rondes  $j-1$  et  $j$  et  $\Delta R_{j-1}$  correspond à  $R_{j-1} - R_{j-2}$ .

### 3.2.3) Le dilemme du prisonnier

Le dilemme du prisonnier est un autre jeu permettant à l'économie expérimentale de vérifier si les équilibres prévus par les hypothèses de maximisation sont plausibles. Dans ce jeu, deux personnes doivent choisir entre la coopération ou la défection. Dans une version du jeu à une seule ronde, l'équilibre de Nash ne mène pas à un optimum de Pareto puisque les incitatifs à ne pas coopérer sont trop élevés. Cependant, lorsque la version du jeu est présentée de façon répétée, la coopération mutuelle est une solution plausible (Varian, 1990). C'est d'ailleurs une telle version du jeu qui a permis à Rilling *et al.* (2002) d'identifier certaines régions du cerveau activées par une coopération mutuelle. Leurs recherches suggèrent que le fait d'adopter un comportement coopératif, et de recevoir la réciprocité, engendrerait une activation du circuit de la récompense (décrit à la section 2.2.2)<sup>139</sup> et impliquerait donc une forme de gratification. Adoptant une approche de biologie évolutionniste, les auteurs suggèrent que cette gratification permettrait de réduire le désir plus primaire ou immédiat de la défection (au bénéfice d'une satisfaction ultérieure plus grande amenée par un macro-équilibre où la coopération est plus fréquemment observée). L'activation de ce circuit, en permettant le renforcement positif d'actes de réciprocité altruiste, aiderait donc à repousser la tentation d'agir « égoïstement ».

Suggérant qu'il est primordial d'être doté de mécanismes neuronaux permettant d'établir qui est, ou n'est pas digne de confiance, Rilling et ses collègues (2004) ont utilisé une version du dilemme du prisonnier à une seule ronde (figure 3.7) pour vérifier si le cerveau de dix-neuf participants (11 femmes et 8 hommes) pouvait encoder de tels mécanismes. Les chercheurs observèrent l'activité cérébrale des participants alors que ceux-ci jouaient soit

---

<sup>139</sup> Selon cette étude, la coopération mutuelle avec un partenaire humain activerait plus spécifiquement certaines parties du striatum et du cortex cingulaire antérieur rostral. D'autres régions sont aussi activées. Pour plus de détails, voir Rilling *et al.*, 2002.

contre des partenaires humains, soit contre un programme d'ordinateur. En laboratoire, Rilling et ses collègues notèrent que les participants coopéraient significativement plus souvent contre des partenaires humains (81 % des essais) que contre le programme d'ordinateur (66 % des essais). Les chercheurs identifièrent ensuite les régions affectées par les actes de coopération et de défection, afin d'analyser leur impact dopaminergique.

Figure 3.7 : Dilemme du prisonnier

(a)

		Player A (scanned)	
		Cooperate	Defect
Player B	Cooperate	\$5 (\$5)	\$6 (\$0)
	Defect	\$0 (\$6)	\$1 (\$1)

La figure 3.7 montre les deux alternatives des joueurs A et B (coopérer, faire défection). Les paiements du joueur sous investigation neurologique, le joueur A, correspondent aux montants situés à gauche alors que les paiements du joueur B sont entre parenthèses.

Source : Rilling *et al.* (2004)

Nous avons préalablement souligné (section 2.2.3) que l'activité dopaminergique survenant après l'arrivée d'une récompense<sup>140</sup> peut être interprétée comme une erreur d'anticipation. L'erreur d'anticipation est donc l'écart entre la récompense anticipée et la récompense obtenue.

<sup>140</sup> Dans le cas présent, il s'agit de la coopération ou de la défection de l'autre joueur. La défection peut être vue comme « une récompense négative ».

Une activation *BOLD*<sup>141</sup> négative implique qu'un niveau de récompense attendu ne s'est pas matérialisé; tandis qu'une activation *BOLD* positive peut être interprétée comme une « agréable surprise ». Après avoir décidé de coopérer, le joueur se fera une estimation idiosyncrasique des probabilités de réciprocité de son partenaire. Considérant que le comportement humain n'est que rarement parfaitement prévisible, cette estimation a peu de chance d'être de 0 ou de 100 %. Par conséquent, les résultats du jeu engendreront toujours une incertitude et donc une erreur d'anticipation plus ou moins grande. Rilling *et al.* s'intéressèrent plus particulièrement à l'activation dopaminergique du cerveau d'un joueur coopératif réagissant à la décision de l'autre. Les chercheurs portèrent leur attention sur deux régions du cerveau (du joueur A) qui étaient sensibles aux décisions du joueur B. L'activation *BOLD* d'une partie du striatum et du cortex préfrontal ventromédial (VMPFC) fut ainsi mesurée selon différents contextes (tableau 3.3).

---

<sup>141</sup> Comme nous l'avons préalablement souligné au chapitre 1, *BOLD* signifie *Blood Oxygen Level Dependent Signal*.

Tableau 3.3 : Dilemme du prisonnier et activation dopaminergique dans certaines régions du cerveau

**Table 2.** Regression coefficients for GLM of striatal and ventromedial frontal ROIs from Figure 2.  $n=19$ ,  $p < 0.005$ .

<b>Striatum</b>				
Regressor	Value	Se	t	p
CC human	<b>0.30</b>	<b>0.10</b>	<b>2.96</b>	<b>0.003</b>
CD human	<b>-0.21</b>	<b>0.10</b>	<b>-2.03</b>	<b>0.04</b>
\$0 control	<b>-0.22</b>	<b>0.10</b>	<b>-2.15</b>	<b>0.03</b>
\$5 control	-0.02	0.10	-0.17	0.87
CC computer	-0.01	0.10	-0.11	0.92
CD computer	<b>-0.29</b>	<b>0.13</b>	<b>-2.16</b>	<b>0.03</b>
<b>VMPFC</b>				
Regressor	Value	Se	t	p
CC human	<b>0.34</b>	<b>0.10</b>	<b>3.37</b>	<b>0.0008</b>
CD human	<b>-0.36</b>	<b>0.10</b>	<b>-3.47</b>	<b>0.0005</b>
\$0 control	-0.11	0.10	-1.07	0.28
\$5 control	-0	0.10	-0.02	0.98
CC compute	0.03	0.10	0.32	0.75
CD computer	-0.14	0.13	-1.08	0.27

Betas for regressors in bold type differed significantly from zero at  $p < 0.05$ .

Le tableau 3.3 (Table 2) indique les différents contextes au cours desquels fut mesurée l'activation neurologique du cerveau des joueurs A (région du striatum et du cortex préfrontal ventromédial). Les différents contextes sont : CC (coopération des deux joueurs), CD (coopération du joueur A et défection du joueur B), 0 \$ et 5 \$ (contrôles expérimentaux). Les joueurs firent face à des partenaires humains ainsi qu'à un ordinateur. Les valeurs  $p < 0,05$  impliquent des coefficients statistiquement significatifs.

Source : Rilling *et al.*, 2004

Le tableau 3.3 montre que l'activation dopaminergique est initiée dans les deux régions d'intérêt, autant par une coopération que par une défection du joueur B. Cette activation n'est cependant pas significative lorsque le joueur A affronte un ordinateur (sauf dans le striatum lors d'une défection). La nature de l'interaction a donc une incidence sur l'activation *BOLD*. Ainsi, l'activation positive ne survient que lorsqu'un partenaire humain, et non un ordinateur,

fait preuve de réciprocité. Rilling et ses collègues suggèrent qu'il y aurait donc quelque chose de particulièrement gratifiant à une interaction sociale positive avec les autres. Ces observations neurologiques complètent assez bien les observations empiriques dénotées précédemment : les gens ne coopèrent pas autant lorsque confrontés à un ordinateur. Ce serait moins gratifiant de le faire. Comme nous l'avions souligné à la section 2.2.2, les décisions et les choix sont orientés vers la satisfaction dopaminergique (Wise, 2002). Les comportements les plus gratifiants sont ceux qui sont répétés le plus souvent. Rilling et ses collègues suggèrent donc que certaines régions associées au système dopaminergique nous permettraient d'apprendre, par le biais des erreurs d'anticipation, qui sont les personnes dignes de confiance et qui ne le sont pas; cela nous inciterait à chercher de nouvelles interactions avec les premières, et d'éviter les transactions avec les secondes.

D'autres recherches très récentes poursuivent dans la même veine et illustrent la rapidité des nouvelles découvertes de la neuroéconomie. Par exemple, Singer *et al.* (2004) utilisent à leur tour le dilemme du prisonnier pour montrer que la simple vue du visage d'une personne ayant préalablement coopéré dans un dilemme du prisonnier à un seul tour active certaines zones du circuit de la récompense. En soi, ceci permet aussi de comprendre en quoi les valeurs objectives de la matrice des paiements pourraient différer de leur attribution subjective. Par ailleurs, une autre étude de Rilling *et al.* (2004)<sup>142</sup> utilise le dilemme du prisonnier et le jeu de l'ultimatum pour tenter de nouvelles avancées dans l'identification des composantes neurobiologiques de la théorie de l'esprit.

L'approche neuroéconomique du jeu de la confiance et du dilemme du prisonnier permet à des chercheurs d'identifier des liens entre les hypothèses découlant de l'économie expérimentale et leurs fondements physiologiques. À

---

<sup>142</sup> Rilling *et al.*, 2004.

la lumière des deux jeux que nous venons de présenter, il semblerait que la réciprocité positive ait des composantes réellement physiologiques qui faciliteraient l'échange. Mais la réciprocité peut également être négative. Dans certaines situations, cela peut réduire les possibilités d'échanges. Aidé du jeu de l'ultimatum et d'une approche neuroéconomique, nous verrons maintenant pourquoi.

#### 3.2.4) Le jeu de l'ultimatum

L'une des expérimentations les plus marquantes de la neuroéconomie a été effectuée par Sanfey *et al.* (2003) qui utilisèrent le jeu de l'ultimatum pour évaluer les bases neurologiques de la décision économique. Au cours de ce jeu, on offre un montant d'argent à une personne et on lui demande de diviser la somme entre elle et une deuxième personne. Si cette dernière accepte l'offre, les sommes sont ainsi partagées et le jeu s'arrête là. Par contre, si elle refuse l'offre, les deux repartent les mains vides. La solution standard au jeu de l'ultimatum serait, dans le cas du premier joueur, d'offrir le plus petit montant possible; dans le cas du second joueur, la solution serait d'accepter tout montant (sur la base raisonnable que toute somme monétaire est préférable à zéro) (Rubinstein, 1985). Mais les gens ne sont pas si rationnels. Contrairement à la prédiction des modèles standards, les gens offrent souvent de 40 % à 50 % du montant disponible et le contexte dans lequel le jeu est présenté a une forte incidence sur les résultats (Chorvat et McCabe, 2004). En fait, des études faites dans une vingtaine de pays ont montré que la vaste majorité des personnes offrent entre le tiers et la moitié de la mise, tandis que les offres inférieures à 20 % sont rejetées la moitié du temps (Camerer et Loewenstein, 2002). Le rejet d'une offre est souvent accompagné d'un sentiment de colère (Pillutla et Murnighan, 1996, cités dans Sanfey *et al.*, 2003). Ce sentiment aurait son importance puisque le rejet d'une offre permettrait le maintien d'une bonne réputation; accepter des offres basses nuirait à la réputation et augmenterait les probabilités ultérieures de recevoir de faibles offres (Nowak *et al.*, 2000, cités dans Sanfey *et al.*, 2003). En fait,

depuis que Güth *et al.* (1982) ont introduit le jeu de l'ultimatum en laboratoire, les économistes expérimentaux ont constaté à maintes reprises que le comportement humain différait fondamentalement des simples prévisions standards. Les quelques cas d'exceptions où les gens offrent peu et acceptent peu, comme par exemple chez les adultes autistes, montrent ironiquement que le comportement rationnel est surtout observable chez les personnes socialement mal adaptées ou ayant des déficits en matière de théorie de l'esprit (Camerer, 2003).

Aidé de la technologie fMRI, Sanfey et ses collègues étudièrent le cerveau de dix-neuf participants jouant au jeu de l'ultimatum. Les chercheurs portèrent leur attention sur l'activation cérébrale des participants assignés au rôle d'un répondant recevant des offres inéquitables (20 % et moins de la mise). Partant de l'hypothèse qu'une telle offre engendrerait un conflit cognitif entre, d'une part, le désir émotionnel de ne pas l'accepter, et, d'autre part, la volonté d'accumuler le plus d'argent possible, Sanfey et ses collègues identifièrent des zones du cerveau pouvant potentiellement être en lien avec de tels processus mentaux. Les auteurs portèrent une attention plus particulière à trois zones d'intérêt : le cortex cingulaire antérieur (ACC), la région bilatérale antérieure (insula droite et insula gauche) et le cortex préfrontal dorsolatéral (DLPFC). Ils remarquèrent que ces trois régions étaient comparativement plus activées par des offres inéquitables que par des offres équitables.

Nous avons préalablement souligné le rôle du cortex cingulaire antérieur à la section 1.3.2. Mentionnons ici que des études ont montré que cette région était activée lorsqu'une personne est confrontée à un conflit cognitif (MacDonald *et al.*, 2000; Botvinick *et al.*, 1999, cités dans Sanfey *et al.*, 2003), ou quand un automatisme est une réponse inappropriée qui entre en conflit avec l'objectif en cours (Bush *et al.*, 2000, cités dans Eisenberger *et al.*, 2003). D'autres recherches ont montré que cette région est activée par la sensation de détresse accompagnant la douleur physique (Rainville *et al.*, 1997). Une

recherche de Eisenberger *et al.* (2003) montre que le rejet social entraîne une activation du cortex cingulaire antérieur<sup>143</sup>. Quoiqu'il en soit, pour leur étude, Sanfey et ses collègues interprètent l'activité de la région ACC comme étant le reflet de la compétition qu'une offre inéquitable provoque entre les processus cognitifs (maximiser les gains) et affectifs (vouloir rejeter l'offre). Par ailleurs, d'autres études ont montré que la région DLPFC a été associée avec la maintenance d'un objectif en cours (Miller et Cohen, 2001, cités dans Sanfey *et al.*, 2003) et c'est pourquoi Sanfey *et al.* interprètent l'activation de cette région comme reflétant la volonté de maintenir l'objectif en cours, c'est-à-dire de faire le plus d'argent possible dans le jeu de l'ultimatum. Enfin, la région bilatérale antérieure (insula droite et gauche) a notamment été associée à des émotions négatives telles que la colère et le dégoût (Phillips *et al.*, 1997; Damasio *et al.*, 2000; cités dans Sanfey *et al.*, 2003).

Au cours de leur expérimentation, Sanfey *et al.* demandèrent à leurs participants de compléter 30 rondes du jeu de l'ultimatum (dix avec différentes personnes, dix avec un ordinateur et dix contrôles)<sup>144</sup>. Chacune des rondes durait 36 secondes. Avant que l'offre du partenaire ne soit révélée, une photo apparaissait à un écran pour montrer avec qui le participant était jumelé (soit une personne, un ordinateur ou une roulette) (voir figure 3.8). Sanfey et ses collègues observèrent des résultats comportementaux assez similaires à ce qui fut noté dans d'autres études : le taux d'acceptation était négativement corrélé avec le degré d'iniquité des offres. Il est intéressant de noter que le comportement des joueurs différait selon qu'ils recevaient une offre inéquitable

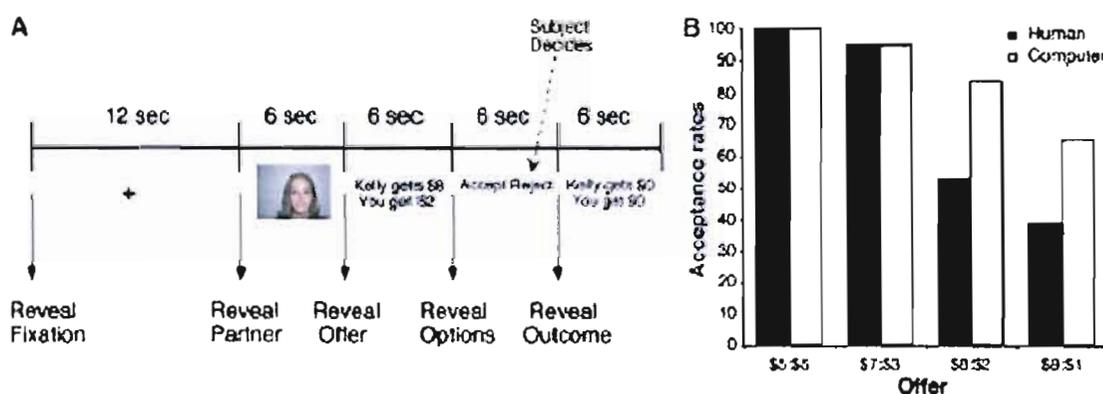
---

<sup>143</sup> La question se pose : est-ce que le fait de recevoir une offre inéquitable serait perçu comme une forme de rejet social ?

<sup>144</sup> Les offres effectuées par les humains obéissaient dans les faits à un algorithme prédéterminé qui était similaire pour l'ordinateur. Les contrôles permettaient de voir la réaction cérébrale provoquée par un gain monétaire sans interaction sociale.

de la part d'un autre joueur ou de l'ordinateur. Le rejet était significativement plus élevé lorsque l'offre provenait d'un être humain (figure 3.8).<sup>145</sup>

Figure 3.8 : Ronde typique du jeu de l'ultimatum et résultats comportementaux



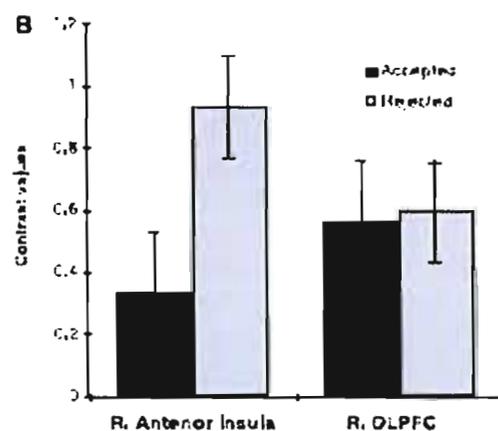
Source : Sanfey *et al.*

L'activation des trois régions d'intérêt, suite à une offre inéquitable, indique que celles-ci sont non seulement sensibles au montant d'argent proposé, mais également au contexte (recevoir une offre inéquitable de la part d'un partenaire humain). La région de l'insula était significativement plus activée par une offre inéquitable en provenance d'un humain qu'en provenance de l'ordinateur (Insula gauche :  $t = 2,52$  et  $p < 0,02$  et insula droite :  $t = 2,2$  et  $p < 0,03$ ). Comme l'activité de la région de l'insula a préalablement été associée avec des états émotionnels négatifs, son activation suite à une offre inéquitable est interprétée comme étant l'impact d'une émotion par Sanfey et ses collègues. Les auteurs suggèrent que la forte activation de cette région permet même d'annoncer à l'avance le rejet de l'offre et illustre par le fait même l'importance

<sup>145</sup> Le pourcentage élevé de rejets d'offres « inéquitables » en provenance d'une machine était assez élevé comme l'illustre la figure 3.8. Cet anthropomorphisme s'explique peut-être en partie par le contexte artificiel de l'étude.

du rôle joué par les émotions dans la décision économique. Par ailleurs, il ressort de cette expérimentation que l'activation d'une autre région d'intérêt, le DLPFC, qui est supposée être en lien avec la volonté de maximiser les gains monétaires, n'est pas significativement corrélée avec le taux d'acceptation des offres inéquitables. Si cette région à elle seule ne peut pas permettre de prévoir le comportement, les auteurs suggèrent cependant qu'elle puisse néanmoins être en compétition avec les autres régions pour le contrôle de la décision. C'est pourquoi ils ont comparé l'activation de cette région avec celle de l'insula. Leurs résultats indiquent que si l'activité dans l'insula est plus grande que celle observée dans le DLPFC droit, l'offre est rejetée. Et quand c'est l'activité dans le DLPFC droit qui est relativement plus élevée, les offres sont acceptées (figure 3.9).

Figure 3.9 : Compétition entre raison et émotions pour le contrôle du comportement dans le jeu de l'ultimatum

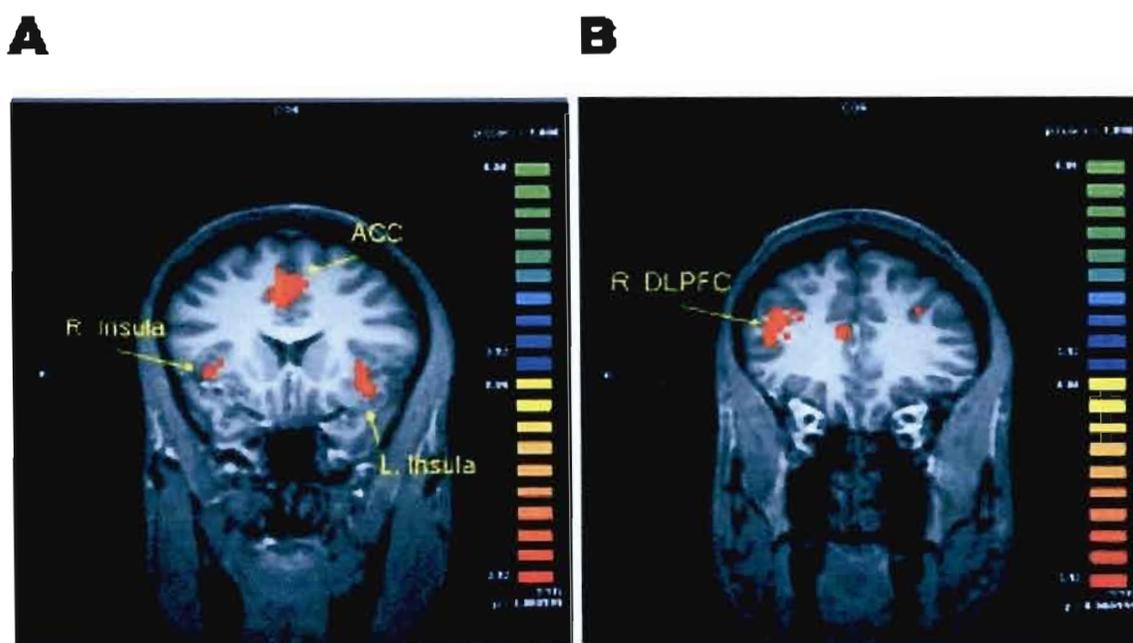


Source : Sanfey *et al.*

Si leur interprétation est la bonne, le pessimisme souligné par Jevons (voir section 1.1) n'aurait pas lieu d'être, et une approche neuroéconomique permettrait d'incorporer les émotions dans l'analyse du comportement

économique. Les techniques d'imagerie cérébrale pourraient même permettre d'observer le poids d'une émotion dans la décision ainsi que la compétition entre les différents processus mentaux (figure 3.10).

Figure 3.10 : Régions activées par une offre inéquitable au jeu de l'ultimatum



Les régions activées par une offre inéquitable dans le jeu de l'ultimatum. La forte activité observée dans l'insula droite et gauche suggère que l'offre sera rejetée (image A) mais l'activation encore plus forte de la région du DLPFC droit (image B) suggère au contraire que l'offre sera finalement acceptée.

Source : Sanfey *et al.*

### 3.3) Réciprocité forte et considérations évolutionnistes

Utilisant les technologies fMRI pour cerner les différentes facettes du comportement économique, l'étude de Sanfey *et al.* suggère que l'on peut identifier et démontrer le rôle joué par les émotions dans une décision

économique. Selon eux, les émotions et leurs dimensions biologiques doivent être incorporées à l'analyse de la décision économique :

« Therefore, not only do our results provide direct empirical support for economic models that acknowledge the influence of emotional factors on decision-making behavior, but they also provide the first step toward the development of quantitative measures that may be useful in constraining the social utility function in economic models. Models of decision-making cannot afford to ignore emotion as a vital and dynamic component of our decisions and choices in the real world. »<sup>146</sup>

Les fonctions d'utilité sociale utilisées comme outils d'analyse par les économistes comportementaux pourraient donc être enrichies par les recherches de la neuroéconomie. Les jeux de la confiance et du dilemme du prisonnier ont permis de relier le cerveau à la notion de réciprocité positive; recevoir la réciprocité de quelqu'un en qui on a fait confiance sollicite des régions du cerveau associées au circuit de la récompense. C'est cette gratification qui pousse ensuite les gens à répéter l'expérience de la coopération. Mais les résultats du jeu de l'ultimatum, eux, nous permettent d'incorporer une autre dimension aux préférences sociales, celle des « punitions altruistes »<sup>147</sup>. Nous avons préalablement souligné que la réciprocité négative peut limiter les échanges. Placés dans certains contextes, beaucoup de gens vont rejeter des sommes monétaires, même si cela leur est nuisible (Loewenstein, 1996). Le jeu de l'ultimatum et nombre de comportements économiques illustrent très bien cette idée. De récentes recherches pourraient permettre de comprendre pourquoi les gens ont de telles motivations à agir ainsi. De Quervain *et al.* (2004) ont utilisé le jeu de la confiance avec possibilité de punir l'autre suite à une défection. Ils ont pu

<sup>146</sup> Sanfey *et al.*, 2003, p. 1758.

<sup>147</sup> Nous empruntons l'interprétation de De Quervain *et al.*, 2004. La définition biologique de la punition altruiste implique qu'un acte est coûteux pour la personne qui punit mais bénéfique pour l'ensemble du groupe. L'altruisme est ici défini en termes de ses conséquences, et non selon une quelconque volonté à faire le bien. La définition psychologique ou philosophique suggère que les motivations de punitions altruistes ne soient pas orientées par la satisfaction hédonistique. Les résultats de recherches de De Quervain *et al.* suggèrent que la définition d'une punition pourrait être altruiste au sens biologique mais non au sens psychologique.

observer que l'action de punir activait certains circuits neuronaux reliés au circuit de la récompense. Punir le tricheur, se venger de lui serait, en terme neurologique, une activité gratifiante. Dans leur expérimentation, De Quervain *et al.* soulignent l'activation du striatum (plus précisément du noyau caudé) lors d'un acte punitif. Cette région a préalablement été associée à l'anticipation d'une gratification. Dans ce contexte, la punition de l'autre serait une action gratifiante qui provoquerait donc une satisfaction anticipée. Il est intéressant de noter que ces résultats relient aux causes proximales du comportement les récents modèles de préférences sociales. (Rabin, 1993; Fehr et Schmidt, 1999; Camerer, 2003).

Lorsque les causes proximales du comportement sont associées à des décisions de nature économique, il devient naturel d'intégrer à l'analyse des considérations évolutionnistes. Les liens entre la biologie, l'évolution et l'économie sont grandissants mais ils ne sont pas si récents. Il a été reconnu que Malthus et Adam Smith auraient exercé sur l'œuvre de Charles Darwin une réelle influence (Robson, 2001). Les sciences cognitives et la psychologie évolutionniste reconnaissent généralement comme ancestral un mécanisme qui explique à la fois les comportements humains et celui d'autres primates (Chen *et al.*, 2005)<sup>148</sup>. Il est intéressant de souligner que l'aversion à l'inégalité a récemment été observée chez des singes capucins (Brosnan et De Wall, 2003). Au cours de leur expérimentation, Brosnan et De Wall ont montré que lorsqu'un singe (*Cebus apella*) obtenait moins qu'un autre après avoir travaillé autant pour l'obtention d'une récompense, ou encore pire lorsqu'il obtenait moins alors que l'autre singe n'avait fait aucun effort pour mériter une meilleure récompense, sa réaction devenait très émotionnelle et il pouvait

---

<sup>148</sup> Au cours d'une fascinante série d'expérimentations, Chen *et al.* (2005) ont pu montrer que les singes capucins possédaient les mêmes biais que ceux observés chez les humains. Les choix des singes reflétaient notamment une sensibilité à l'effet de cadrage et à l'aversion à la perte. Selon les auteurs, ces préférences seraient donc innées.

refuser un échange qu'il avait pourtant préalablement accepté lorsqu'il était seul. En d'autres termes, il préférerait ne pas consommer du tout lorsqu'il percevait avoir été injustement traité. Contrairement à d'autres primates caractérisés par des hiérarchies despotiques, les singes capucins sont reconnus pour leur coopération et leur partage de nourriture; ils semblent, tout comme les humains, mesurer leurs récompenses en des termes relatifs ; ils comparent leur niveau d'effort et leurs gains en fonction de ceux des autres.

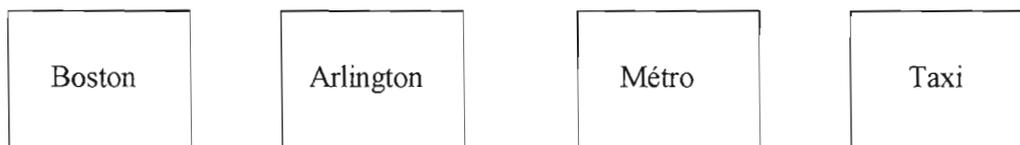
### 3.3.1) Psychologie évolutionniste et modularité

Les profondes structures du comportement humain relèvent de trois ordres distinguables : l'ordre interne de l'esprit, l'ordre externe des interrelations sociales, et l'ordre étendu des marchés (Smith, 2004). Chacun de ces ordres forme un système complexe qui reflète l'évolution biologique et culturelle. L'ordre interne de l'esprit relève des structures neurobiologiques de l'être humain. Ce sont ces structures qui font l'objet d'investigations des neurosciences. L'ordre externe des interrelations sociales a fait l'objet de recherches au sein de l'économie expérimentale, particulièrement en laboratoire lors de transactions anonymes. Il ressort de ces études que cet ordre est gouverné par la réciprocité (positive et négative). Comme ces observations ont été dénotées dans plusieurs pays et cultures distinctes, elles permettent l'élaboration d'hypothèses pouvant orienter les recherches de la neuroéconomie. L'évolution humaine s'est bâtie sur l'évolution déjà accomplie par les primates, en particulier par les hominidés (Eccles, 1992). La langue et la socialisation sont présumées avoir évolué au cours des deux ou trois millions d'années au cours desquelles les hominidés ont évolué comme chasseurs-cueilleurs. C'est pendant cette période que nos ancêtres auraient développé des mécanismes d'échanges sociaux dans lesquels on retrouvait l'assistance, les faveurs et le partage de nourriture. Cette évolution n'aurait pas été sans conséquences sur l'ordre interne de l'esprit. Des instincts de solidarité et d'altruisme, génétiquement hérités, auraient servi à renforcer la coopération des membres du groupe (Hayek, 1988) et à éventuellement

favoriser l'émergence spontanée des marchés (Cosmides et Tooby, 1992). Il est plausible de croire que les fortes émotions rattachées à l'inégalité feraient partie de ce bagage préhistorique. Le partage des fruits de la chasse implique un échange intertemporel hasardeux; plus la variance des gains de la chasse est élevée, plus le partage de nourriture devient avantageux. Mais dans un monde sans monnaie, une faveur pourrait ne pas être honorée. Développer un jugement sur qui doit, ou ne doit pas être digne de confiance devient alors un mécanisme de survie. Certaines recherches suggèrent l'existence d'un module de « détection du tricheur » (Cosmides et Tooby, 1997). Ce module aurait accompagné le développement de l'intelligence sociale et les recherches présentées à la section 3.2 complètent bien cette hypothèse.

Pour tester cette hypothèse en laboratoire, Cosmides et Tooby ont utilisé un problème de logique connu sous le nom de *Wason selection task*. Pour résoudre ce problème, le sujet doit trouver une violation à une règle conditionnelle de la forme « Si P alors Q ». Au tableau 3.4, on montre quatre cartes sensées représenter quatre personnes distinctes. La règle est la suivante : « Si une personne va à Boston, alors cette personne prend le métro ». On demande aux participants de ne retourner que les cartes nécessaires permettant de vérifier que la règle est violée.

Tableau 3.4 : Le problème de Wason



Source : Cosmides et Tooby (1997)<sup>149</sup>

<sup>149</sup> Traduction par l'auteur du présent mémoire.

D'un point de vue logique, la règle est violée dès qu'une personne va à Boston sans prendre le métro. C'est pourquoi la bonne réponse est de ne retourner que la première et la quatrième cartes. Mais en général, moins de 25 % des personnes parviennent à résoudre ce problème. Cosmides et Tooby ont utilisé ce problème pour tester l'hypothèse d'un module spécialisé dans la détection de tricherie. Utilisant la même structure logique que le problème de Wason, ils ont modifié le contexte; leurs résultats sont frappants : des personnes qui ne parvenaient pas à solutionner ce problème dans sa forme normale arrivaient facilement à le faire dans un contexte où la tâche était d'identifier la violation d'un contrat social<sup>150</sup> (tableau 3.5).

Tableau 3.5 : Détection d'une tricherie

*Quelle(s) carte(s) devez-vous retourner pour tester la violation de la règle suivante : « Si vous prenez le bénéfice, vous payez le coût » ?*



Source : Cosmides et Tooby (1997)<sup>151</sup>

Tel que souligné à la section précédente, il est intéressant de voir que l'hypothèse du module de détection du tricheur a fait l'objet de récentes recherches des neurosciences. Dans une interaction stratégique, la réputation des joueurs serait encodée avec l'expérience (King-Casas *et al.*, 2005) et certaines régions associées au système dopaminergique permettraient d'apprendre, par le biais des erreurs d'anticipation, qui est digne de confiance et qui ne l'est pas (Rilling *et al.*, 2004).

<sup>150</sup> 65 % à 80 % trouvent la bonne réponse, choisir la première et la dernière carte.

<sup>151</sup> Traduction par l'auteur du présent mémoire.

Enfin, la troisième catégorie de l'organisation des structures du comportement humain est l'ordre étendu des marchés. Cet ordre n'implique pas le même genre de coopération que le précédent. Il est plus impersonnel. Il relève des règles de conduite humaine, tout spécialement celles concernant la propriété plurielle, l'honnêteté, le contrat, l'échange, le commerce, la compétition, le profit et la vie privée (Hayek, 1988). Il n'est pas le résultat d'un esprit ordonnateur et aurait été rendu possible par la tradition, l'enseignement et l'imitation. Hayek (1988, p. 28) souligne les tensions entre les différents ordres : « Nous ne nous sommes pas défaits de notre héritage de bande en état de confrontation, et nos instincts n'ont été ni pleinement « adaptés » à un ordre étendu au fond récent, ni été réduits par lui à l'innocuité ». Il poursuit en faisant état des dangers de ces tensions :

Si, comme nos penchants instinctifs et sentimentaux nous incitent souvent à le désirer, nous devons appliquer les règles non modifiées, non limitées, du microcosme (c'est-à-dire de la petite bande ou troupe ou, disons, de nos familles) au macrocosme (notre civilisation élargie), nous détruirions celui-ci. Si à l'inverse nous devons appliquer constamment les règles de l'ordre étendu aux groupes plus restreints et plus intimes en lesquels nous évoluons, nous les *broierions*.<sup>152</sup>

La conception de Hayek est similaire à certains développements récents de la psychologie évolutionniste (Smith, 2001). L'économie expérimentale a montré que les gens se comportent plutôt rationnellement lorsque confrontés à un ordinateur. Lorsqu'ils interagissent avec d'autres humains par contre, les gens utilisent d'autres règles, d'autres algorithmes psychiques. Il a été montré que le cerveau n'était pas activé de la même manière par ces deux alternatives. Cette situation pourrait-elle être le reflet de quelque chose de plus profond ? Dans un cadre impersonnel, comme celui de l'ordre étendu, le comportement deviendrait plus rationnel. Et face à un autre humain, la collaboration et la réciprocité deviendraient le comportement adapté à l'ordre externe des relations sociales.

---

<sup>152</sup> *Ibid.*, p. 28.

### 3.4) Conclusion

La théorie des jeux peut être considérée comme un outil des neurosciences permettant d'analyser le comportement humain lors de certaines considérations stratégiques et dynamiques. En prédisant les interactions entre individus intelligents, la théorie des jeux permet d'étudier le cerveau en dehors d'un cadre statique, ce qui correspond très souvent à l'état de la nature. L'utilisation de cet outil conceptuel permet aux chercheurs des neurosciences d'étudier des comportements qu'ils n'auraient jamais pu analyser dans un paradigme stimuli-réponses. Les observations rendues possibles par l'économie expérimentale, plus précisément par la théorie comportementale des jeux, et qui ont notamment remis en cause les hypothèses de maximisation, d'égoïsme et de rationalité, ont aussi fait l'objet d'études des neurosciences. Il ressort de ces études que les comportements altruistes observés en laboratoire auraient une base biologique. Les propriétés d'un organisme comme le cerveau sont sensées refléter les tâches qu'il eut à accomplir pendant son évolution. Comme le cerveau a évolué pendant des millions d'années, il est suggéré d'incorporer à l'analyse du comportement économique ses justes parts de considérations évolutionnistes. Ce faisant, il serait plus aisé de comprendre pourquoi certains comportements sont parfois irrationnels et mal adaptés aux exigences contemporaines de nos économies. La réciprocité, qu'elle soit positive ou négative, serait l'héritage d'un ordre ancien. Alors que l'évolution biologique est très ancienne, les exigences contemporaines de l'ordre étendu des marchés sont par contre très récentes. En soi, cela peut créer des confusions et des tensions.

De par ses récentes découvertes, la neuroéconomie corrobore la théorie comportementale des jeux, et conserve, pour objectif ultime, la compréhension du comportement économique.

## CONCLUSION

Depuis le début des années 2000, la rencontre entre chercheurs issus des neurosciences, de la psychologie et de l'économie, a permis d'initier l'émergence d'une nouvelle discipline : la neuroéconomie. Les avancées récentes des neurosciences et de la neuroéconomie ont démontré que les approches économiques qui sont uniquement basées sur des processus de délibérations conscientes et rationnelles sont incomplètes. Pour tenir compte de ces lacunes, des économistes comportementaux et des neuroéconomistes suggèrent d'inclure aux théories économiques les dimensions automatiques et affectives des processus décisionnels (Camerer *et al.*, 2003, 2004, 2005; Camerer et Lowenstein, 2002 ; Sanfey *et al.*, 2003). Des recherches ont récemment entrepris d'établir les circonstances dans lesquels l'un ou l'autre de ces processus entre en action, afin d'établir leur corollaires cérébraux. Aidé d'outils des neurosciences, il est maintenant possible de mieux comprendre ce qui initie l'action ou la décision économique. On peut ainsi regrouper des phénomènes qu'on croyait différents et en séparer d'autres que l'on croyait similaires, notamment parce que des préférences peuvent être délibérées ou implicites.

Cette approche neuroéconomique n'aurait pu être possible sans l'évacuation préalable, au sein des neurosciences, d'un paradigme béhavioriste qui ne permettait d'analyser la décision que dans un cadre stimuli-réponses (Glimcher, 2003, 2004). Pour mieux comprendre le cerveau, il faut chercher à savoir ce qu'il cherche à accomplir et c'est pourquoi la théorie microéconomique intéresse les chercheurs des neurosciences. Des outils analytiques comme l'utilité subjective, l'espérance de l'utilité, la théorie des jeux et l'équilibre de Nash, ont attiré l'attention de ces chercheurs qui les ont incorporés avec enthousiasme à leur méthodologie ; et le fait de travailler avec ces nouvelles bases conceptuelles leur a d'ailleurs déjà permis de mieux

comprendre le fonctionnement du cerveau (Platt et Glimcher, 1999 ; Glimcher, 2003). Que des chercheurs issus des neurosciences puissent s'intéresser aux théories économiques ne peut assurément pas être considéré comme étant un phénomène banal. Que ces chercheurs parviennent même à corroborer biologiquement, en partie du moins, ces outils de la théorie économique devrait, il nous semble, intéresser tous les économistes. Mais en fait, l'enjeu ultime de la neuroéconomie est bien plus grand : il pourrait permettre de développer une nouvelle théorie de la décision qui permettrait de réduire l'écart entre les approches normatives de la théorie économique, et les approches descriptives de l'économie comportementale. Il est vrai que les intuitions se cachant derrière l'utilité, l'espérance de l'utilité ou l'équilibre de Nash ont été en partie validées par les neurosciences (ce qui est en soi non négligeable), mais ce qui importe davantage, c'est que d'autres travaux, issus de l'économie comportementale, ont commencé à corroborer des approches alternatives ou complémentaires. Ainsi, les comportements auparavant observés par Maurice Allais et Daniel Ellsberg trouveraient leur origine dans l'architecture même du cerveau humain: les processus cérébraux activés quand un revenu certain est l'une des options ne sont pas les mêmes que lorsque les deux options sont incertaines (Dickhaut *et al.*, 2003) et l'aversion à l'ambiguïté aurait ses corollaires cérébraux (Hsu et Camerer, 2004; Bechara *et al.*, 1997; O'Doherty *et al.*, 2001). Mais ce n'est pas tout, d'autres recherches (Breiter *et al.*, 2001; Chen *et al.*, 2005; Dickhaut *et al.*, 2003 ; Knutson et Peterson, 2005; Knutson *et al.*, 2000;) ont permis d'établir des corollaires cérébraux à la théorie prospective. Et la liste des apports de la neuroéconomie ne s'arrête pas là : les préférences révélées (Berridge, 1999), les préférences intertemporelles (McClure *et al.*, 2004) et la théorie des jeux (rationnelle ou comportementale) (Dorris et Glimcher, 2003, 2004; Glimcher, 2002; Wilson *et al.*, 2001; Barraclough *et al.*, 2004; McCabe *et al.*, 2001, Kosfeld *et al.*, 2005; King-Casas *et al.*, 2005; Rilling *et al.*, 2002, 2004; Sanfey *et al.*, 2003) peuvent toutes être étudiées à l'aide d'une approche neuroéconomique. Il est d'ailleurs particulièrement intéressant de voir que de récents développements permettent aujourd'hui de

tester certaines hypothèses concernant les fondements biologiques des préférences sociales. Ainsi, la réciprocité et la confiance auraient des fondements physiologiques qui auraient favorisé, au travers des âges, les échanges et l'émergence des économies humaines. Les économistes pourront-ils se permettre d'ignorer ces découvertes ? En fait, certains économistes ont déjà commencé à intégrer à leurs propres recherches certaines avancées des neurosciences; ce faisant, ils ont ainsi sans doute augmenté le degré de réalisme de leur modélisation économique (Bernheim et Rangel, 2002 ; Laibson, 2001). La neuroéconomie a donc déjà commencé à influencer les théoriciens de l'économie. En somme, la neuroéconomie, en reliant les causes proximales et ultimes du comportement, et en réduisant ainsi l'écart entre les approches normatives et descriptives, cherche au bout du compte à jeter les bases d'une nouvelle théorie de la décision économique. On sait que l'histoire de l'économie a été marquée par toutes sortes de tensions entre l'approche normative de maximisation rationnelle de l'utilité, et l'approche descriptive d'observations faites en laboratoire. La neuroéconomie pourrait-elle finalement permettre de réduire ces tensions ?

La neuroéconomie est une science très jeune qui évolue très rapidement et qui suscite passablement d'enthousiasme. Pour les chercheurs des neurosciences, elle offre de nouveaux outils conceptuels leur permettant d'étudier le cerveau à l'intérieur d'un autre paradigme; pour les psychologues et les économistes comportementaux, elle est une suite logique de leurs observations empiriques, puisque le comportement trouve nécessairement son origine à partir du cerveau et non à partir d'abstractions mathématiques.

Enfin, puisque la neuroéconomie utilise le cerveau comme outil d'analyse, elle incorpore inévitablement des considérations évolutionnistes qui pourraient permettre de mieux comprendre pourquoi le comportement s'éloigne si souvent de la rationalité substantive, et pourquoi l'être humain semble si souvent

tirillé entre l'ordre interne de l'esprit, l'ordre externe des interrelations sociales et l'ordre étendu des marchés.

## RÉFÉRENCES

- Allais, Maurice. 1953. « Le comportement de l'homme rationnel devant le risqué: critique des postulats et axiomes de l'école américaine ». *Econometrica*, vol. 21, p. 503-546.
- Barberis, Nicholas, Ming Huang et Tano Santos. 2001. « Prospect Theory and Asset Prices ». *Quarterly Journal of Economics*, vol. 116, no 1, p.1-53.
- Bargh, John A. 1996. « Automaticity in Social Psychology ». In E.T. Higgins et A.W. Kruglanski (éd.), *Social Psychology : Handbook of Basic Principles*, p. 169-183. New York : Guilford.
- Bargh, John A. 1997. « The Automaticity of Everyday Life ». In R.S. Wyer (éd.), *Advances in Social Cognition*, vol. 10, p. 1-61. Mahwah (New Jersey) : Erlbaum.
- Bargh, John A. et Tanya L. Chartrand. 1999. « The Unbearable Automaticity of Being ». *American Psychologist*, vol. 54, no 7, p. 462-479.
- Baron-Cohen, Simon. 1995. *Mindblindness : An Essay on Autism and Theory of Mind*. Cambridge (Massachusetts) : MIT Press.
- Barracough, Dominic J., Michelle L. Conroy et Daeyeol Lee. 2004. « Prefrontal Cortex and Decision Making in a Mixed-Strategy Game ». *Nature Neuroscience*, vol. 7, no 4, p. 404-410.
- Baumeister. R.F., E. Bratslavsky, M. Muraven et D.M. Tice. 1998. « Ego Depletion : Is the Active Self a Limited Resource ? ». *Journal of Personality and Social Psychology*, vol. 74, p. 1252-1265.
- Baylor College of Medicine. 2005. « What is fMRI ? ». *Site internet* : <http://www.bcm.edu/news/packages/trust-fmri.cfmt>, Houston.
- Bechara, Antoine. 2004. «The Role of Emotion in Decision-Making : Evidence from Neurological Patients with Orbitofrontal Damage ». *Brain and Cognition*, vol. 55, p. 30-40.
- Bechara, Antoine, Hanna Damasio, Daniel Tranel et Antonio R. Damasio. 1997. « Deciding Advantageously before Knowing the Advantageous Strategy ». *Science*, vol. 12, no 5, p. 1293-1295.
- Bechara, Antoine, Daniel Tranel et Hanna Damasio. 2000. « Characterization of the Decision-Making Impairment of Patients with Bilateral Lesions of the Ventromedial Prefrontal Cortex ». *Brain*, vol. 123, p. 2189-2202.
- Bechara, Antoine, Daniel Tranel et Antonio R. Damasio. 2002. « The Somatic Marker Hypothesis and Decision-Making ». In F. Boller et J. Grafman (Éd.) *Handbook of neuropsychology : Frontal Lobes* , vol. 7, Deuxième édition, p. 117-143, Amsterdam : Elsevier.

- Bechara, Antoine et Hanna Damasio. 2002. « Decision-Making and Addiction (part I) : Impaired Activation of Somatic States in Substance Dependant Individuals when Pondering Decisions with Negative Future Consequences ». *Neuropsychologia*, vol. 40, p. 1675-1689.
- Bechara, Antoine et Hanna Damasio. 2002. « Decision-Making and Addiction (part II) : Myopia for the Future or Hypersensitivity to Reward ». *Neuropsychologia*, vol. 40, p. 1690-1705.
- Bechara, Antoine et Antonio R. Damasio. 2004. « The Somatic marker Hypothesis : A Neural Theory of Economic Decision ». *Sous presse*, p. 1-37.
- Becker, Gary et Kevin Murphy. 1988. « A Theory of Rational Addiction ». *Journal of Political Economy*, vol. 96, no 4, p. 675-700.
- Bentham, J. 1789. 1948. *An Introduction to the Principle of Morals and Legislations*. Oxford : Blackwell.
- Bernheim, Douglas B. et Antonio Rangel. 2002. «Addiction and Cue-Conditioned Cognitive Processes ». *NBER Working Paper*, no 9329.
- Berridge, Kent C. 1999. « Pleasure, Pain, Desire, and Dread: Hidden Core Processes of Emotions ». In D.Kahneman, E. Diener *et al.* (éd.) *Well-Being: The Foundations of Hedonic Psychology*, p. 525-557, New York, Russell Sage Foundation.
- Botvinick, M., L.E. Nystrom, K. Fissell, C.S. Carter et J.D. Cohen. 1999. « Conflict Monitoring Versus Selection-For-Action in Anterior Cingulate Cortex ». *Nature*, vol. 402, p. 179-181.
- Botvinick, M., T.S. Braver, D.M. Barch, C.S. Carter et J.D. Cohen. 2000. « Conflict Monitoring and Cognitive Control ». *Psychological Review*, vol. 108, p. 624-652.
- Breiter, Hans C., J. Berke, W. Kennedy, B. Rosen et S. Hyman. 1996. « Activation of Striatum and Amygdala During Reward Conditioning : An FMRI Study. » *Neuroimage*, vol. 3, p. S220.
- Breiter, Hans C., Itzhak Aharon, Daniel Kahneman, Anders Dale et Peter Shizgal. 2001. « Functional Imaging of Neural Responses to Expectancy and Experience of Monetary Gains and Losses ». *Neuron*, vol. 30, p. 619-639.
- Brickman, Philip, Dan Coates et Ronnie Janoff-Bulman. 1978. « Lottery Winners and Accident Victims : Is Happiness Relative ? ». *Journal of Personality and Social Psychology*, vol. 36, no 8, p. 917-927.
- Brooks, King-Casas, Damon Tomlin, Cedric Anen, Colin F. Camerer, Steven R. Quartz et P. Read Montague. 2005. « Getting to Know You : Reputation and trust in a Two-Person Economic Exchange ». *Science*, vol. 308, p. 78-83.
- Brosman, S. et F. de Waal . 2003. « Monkeys Reject Unequal Pay ». *Nature*, vol. 425, p. 297-299.

- Buck, Ross. 1999. « The Biological Affects : A Typology ». *Psychological Review*, 106 : vol. 2, p. 301-336.
- Bush, G., P. Luu, et M.I. Posner. 2000. « Cognitive and Emotional Influences in Anterior Cingulate Cortex ». *Trends in Cognitive Sciences*, vol. 4, p. 215-222.
- Cabanac, M. 1971. « Physiological Role of Pleasure. » *Science*, vol. 173, no 2, p. 1103-1107.
- Cabanac, M. et J. Leblanc. 1983. « Physiological Conflict in Humans : Fatigue versus Cold Discomfort ». *American Journal of Physiology*, vol. 244, p. R621-R628.
- Caggiula, Anthony R. et Bartley G. Hoebel. 1966. « Copulation-Reward Site in the Posterior Hypothalamus » *Science*, vol. 153 : 3741, p. 1284-1285.
- Camerer, Colin F. 1987. « Do Biases in Probability Judgment Matter in Markets? Experimental Evidence ». *American Economic Review*, vol. 77, p. 981-997.
- Camerer, Colin F. 1989. « An Experimental Test of Several Generalized Utility Theories ». *Journal of Risk and Uncertainty*, vol. 2, p. 61-104.
- Camerer, Colin F. 1989. « Does the Basketball Market Believe in the 'Hot Hand' ? ». *American Economic Review*, vol. 79, p. 1257-1261.
- Camerer, Colin F. 1992. « Recent Tests of Generalized Utility Theories ». In W. Edwards (éd.), *Utility Theories: Measurement and Applications*, p. 207-251, Cambridge: Cambridge University Press.
- Camerer, Colin F. 2000. « Prospect Theory in the Wild: Evidence from the Field » In D. Kahneman et A. Tversky (éd.), *Choices, Values, and Frames*, p. 288-300, Cambridge: Cambridge University Press.
- Camerer, Colin F. 2003. *Behavioral Game Theory: Experiments on Strategic Interaction*. Princeton: Princeton University Press.
- Camerer, Colin F. 2003. « Strategizing in the Brain ». *Science*, vol. 300, p. 1673-1675.
- Camerer, Colin F. 2004. « What is neuroeconomics ? ». [http://www.hss.caltech.edu/~camerer/web\\_material/n.html](http://www.hss.caltech.edu/~camerer/web_material/n.html)
- Camerer, Colin F., G. Loewenstein et M.W. Weber. 1989. « The Curse of Knowledge in Economic Settings : An Experimental Analysis ». *Journal of Political Economy*, vol. 97, no 5, p. 1232-1254.
- Camerer, Colin F. et M.W. Weber. 1992. « Recent Developments in Modelling Preferences: Uncertainty and Ambiguity ». *Journal of Risk and Uncertainty*, vol. 5, p. 325-370.

- Camerer, Colin F., Eric Johnson, Talia Rymon, et Sen, Sanker. 1993. « Cognition and Framing in Sequential Bargaining for Gains and Losses ». In K. Binmore, A. Kirman, and P. Tani (éd.), *Frontiers of Game Theory*, p. 27-48, Boston, MIT Press.
- Camerer, Colin. F. et R. Thaler. 1995. « Anomalies: Dictators, Ultimatums, and Manners ». *Journal of Economic Perspectives*, vol. 9, p. 209-219.
- Camerer, Colin F., L. Babcock, G. Loewenstein, et R. Thaler, R. 1997. « Labor Supply of New York City Cab Drivers: One Day at a Time ». *Quarterly Journal of Economics*, vol. 111, p. 408-441.
- Camerer, Colin. F. et T. Ho. 1999. « Experience Weighted Attraction Learning in Normal-Form Games ». *Econometrica*, vol. 67, p. 827-874.
- Camerer, Colin F. et R.M. Hogarth. 1999. « The Effects of Financial Incentives in Economics Experiments: A Review and Capital-Labor-Production Framework ». *Journal of Risk and Uncertainty*, vol. 19, p. 7-42.
- Camerer, Colin., T. Ho et K. Chong. 2001. « Behavioral Game Theory: Thinking, Learning and Teaching ». *Mimeo*, Caltech, Pasadena : Californie.
- Camerer, Colin F. et George Loewenstein. 2002. « Behavioral Economics : Past, Present, Future ». *Mimeo*, Division of Humanities and Social Science, 228-77, Californie : Caltech, Pasadena.
- Camerer, Colin. F., T. Ho et K. Chong. 2002. « Functional EWA: a One-Parameter Theory of Learning in Games. » *Mimeo*, Caltech, Pasadena : Californie.
- Camerer, Colin F., George Loewenstein et Drazen Prelec. 2003. « Neuroeconomics : How Neuroscience Can Inform Economics », *mimeo*, Pasadena : California Institute of Technology.
- Camerer, Colin, G. Loewenstein et D. Prelec. 2004. « Neuroeconomics: Why Economics Needs Brains ». *Scandinavian Journal of Economics*, vol. 106, p. 555-579.
- Camerer, Colin F., George Loewenstein et Drazen Prelec. 2005. « Neuroeconomics : How Neuroscience Can Inform Economics », *Journal of Economic Literature*, vol. XLIII, mars 2005, p. 9-64.
- Carter, Rita. 1999. *Mapping the Mind*. Berkeley, University of California Press : Californie.
- Charnov, E.L. 1973. « *Optimal Foraging : Some Theoretical Exploration* ». Thèse de doctorat, Université de Washington : Washington.
- Chen, Keith, Venkat Lakshminarayanan et Laurie Santos. 2005. « The Evolution of Our Preferences: Evidence From Capuchin-Monkey trading Behavior ». *Mimeo*, Yale School of Management and Cowles Foundation, p. 1-26.

- Chorvat, Terrence et Kevin McCabe. 2004. « The Brain and the Law ». *Philosophical Transactions of the Royal Society of London*, vol. 359, p. 1727-1736.
- Christoff, Kalina. 2000. « Brain and Behavior ». *Compte rendu de lecture de Psychologie 112*, 22 juin 2000, Département de psychologie, Université Stanford : Californie.
- Cohen, Jonathan D. et Kenneth I. Blum. 2002. « Reward and Decision ». *Neuron*, vol. 36, p. 193-198.
- Comings, D.E., R.J. Rosenthal, H.R. Lesieur, L. Rugle, D. Muhleman, C. Chiu, G. Dietz, et R. Gade. 1996. « A Study of the Dopamine D2 Receptor Gene in Pathological Gambling ». *Pharmacogenetics*, vol. 6, p. 223-234.
- Conover, K.L., B. Woodside et Peter Shizgal. 1994. « Effects of Sodium Depletion on Competition and Summation Between Rewarding Effects of Salt and Lateral Hypothalamic Stimulation in the Rat ». *Behavioral Neuroscience*, vol. 108, no 3, p. 549-558.
- Conover, K.L. et Peter Shizgal. 1994. « Differential Effects of Postingestive Feedback on the Reward Value of Sucrose and Lateral Hypothalamic Stimulation in the Rat ». *Behavioral Neuroscience*, vol. 108, no 3, p. 559-572.
- Conover, K. L. et Peter Shizgal. 2002. *Employing Labor Supply Theory to Measure the Reward Value of Electrical Brain Stimulation : Actes du premier colloque sur l'économie et les neurosciences* (Minnesota, 25-27 octobre 2002). Université du Minnesota : Minnesota.
- Cosmides, Leda et John Tooby. 1994. « Better than Rational : Evolutionary Psychology and the Invisible Hand ». *The American Economic Review*, vol. 84, no 2, p.327-332.
- Cosmides, Leda et John Tooby. 1997. « Evolutionary Psychology : A Primer ». *Mimeo*, p. 1-26, Center for Evolutionary Psychology, Université Santa Barbara : Californie.
- Cottle, Thomas J., et Stephen L. Klineberg. 1974. *The Present of Things Future : Explorations of Time in Human Experience*. New York : Free Press.
- Damasio, Antonio R. 1994. *Descartes' Error : Emotion, Reason, and the Human Brain*. New York : G.P. Putnam.
- Damasio, Antonio. R. 2005. « Brain Trust ». *Nature*, vol. 435, p. 571-572.
- Damasio, Antonio. R., T.J. Grabowski, A. Bechara, H. Damasio, L.L.L. Ponto, J. Parvizi et R.D. Hichwa. 2000. « Subcortical and Cortical Brain Activity During the Feeling of Self Generated Emotions ». *Nature Neuroscience*, vol. 3, p. 1049-1056.
- D'Antonio, Michael. 2004. « How We Think : Brain Researchers are Using MRIs to Predict Our Decisions Before They Are Made. The Results Are Intriguing, And a Little Disturbing ». *Los Angeles Times* (Los Angeles), 2 mai 2004, latimes.com..

Dawes, R.M et R. Hastie. 2000. *Rational Choice in an Uncertain World :The Psychology of Judgement and Decision Making*, San Diego (Californie) : Sage press.

Dehaene, S. 1992. « Varieties of Numerical Abilities ». *Cognition*, vol. 44, p.1-42.

Dehaene, S., G. Dehaene-Lambertz et L. Cohen. 1998. « Abstract Representations of Numbers in the Animal and Human Brain ». *Trends in Neuroscience*, vol. 21, no 8, p. 355-361.

Dehaene, S., E. Spelke, P. Pinel, R. Stanescu et S. Tsivkin. 1999. « Source of Mathematical Thinking : Behavioral and Brain-Imaging Evidence ». *Science*, vol. 284, p. 970-974.

Dehaene, S., Nicolas Molko, Laurent Cohen et Anna J. Wilson. 2004. « Arithmetic and the Brain ». *Current Opinion in Neurobiology*, vol. 14, p. 218-224.

Descartes, René. 1664;1972. *Treatise on Man*. Cambridge (Massachusetts) : Harvard University Press.

Devetag, M. Giovanna. 2000. « From Utilities to Mental Models : A Critical Survey on Decision Rules and Cognition in Consumer Choice ». *Industrial and Corporate Change*, vol. 8, no 2, p. 289-351.

De Quervain, Dominique J.-F., Urs Fischbacher, Valerie Treyer, Melanie Schellhammer, Ulrich Schnyder, Alfred Buck et Ernst Fehr. 2004. « The Neural Basis of Altruistic Punishment ». *Science*, vol. 305, p. 1254-1258.

Dickhaut, John, Kevin McCabe, Jennifer C. Nagode, Aldo Rustichini, et Jose V.Pardo. 2003. « The Impact of the Certainty Context on the Process of Choice ». *Proceedings of the National Academy of Sciences*, vol. 100, p. 3536-3541.

Dorris, Michael C. et Paul W. Glimcher. 2004. « Activity in Posterior Parietal Cortex is Correlated with the Relative Subjective Desirability of Action ». *Neuron*, vol. 44, p. 365-378.

Eccles, John C. 1992. *Évolution du cerveau et création de la conscience*. Trad. de l'anglais par Jean-Mathieu Luccioni. Paris : Flammarion.

Eisenberger, Naomi I., Matthew D. Lieberman et Kipling D. Williams. 2003. « Does Rejection Hurt ? An fMRI Study of Social Exclusion ». *Science*, vol. 302, p. 290-292.

Elliot, R., K.J. Friston et R.J. Dolan. 2000. « Dissociable Neural Responses in Human Reward System ». *Journal of Neuroscience*, vol. 20, p. 6159-6165.

Ellsberg, Daniel. 1961. « Risk, Ambiguity and the Savage Axioms ». *Quarterly Journal of Economics*, vol. 75, p. 643-669.

Emlen, J.M. 1966. « The Role of Time and Energy in Food Preference ». *American Naturalist*, vol. 100, p. 611-617.

- Fehr, Ernst et Bettina Rockenbach. 2004. « Human Altruism : Economic, Neural, and Evolutionary Perspectives ». *Current Opinion in Neurobiology*, vol. 14, p. 784-790.
- Fehr, Ernst et Urs Fischbacher. 2001. « Why Social Preference Matter : The Impact of Non-Selfish Motives on Competition, Cooperation and Incentives ». Actes de colloque : “Nobel Symposium on Behavioral and Experimental Economics”, Stockholm, p. 1-29.
- Fehr, Ernst et K.M. Schmidt. 1999. « A Theory of Fairness, Competition and Cooperation ». *Quarterly Journal of Economics*, vol. 114, p. 817-868.
- Fiorillo, Christopher D., Phillippe N. Tobler, et Wolfram Schultz. 2003. « Discrete Coding of Reward Probability and Uncertainty by Dopamine Neurons ». *Science*, vol. 299, p. 1898-1902.
- Fischman, M.W. et R.W. Foltin. 1992. « Self-administration of Cocaine by Humans : A Laboratory Perspective. In G.R. Bock et J. Whelan (Éd.) *Cocaine : Scientific and Social Dimensions*. Vol. 166, p. 165-180, Chichester, Eng : Wiley.
- Frith, Uta. 2001. « Mind Blindness and the Brain in Autism ». *Neuron*, vol. 32, no 6, p. 969-979.
- Fuster, J. 1999. *Memory in the Cerebral Cortex*. Cambridge (Massachusetts): MIT Press.
- Gallistel, C.R. 1994. « Foraging for Brain Stimulation : Toward a Neurobiology of Computation ». *Cognition*, vol. 50, p. 151-170.
- Gallistel, C.R., Peter Shizgal et J. Yeomans. 1981. « A Portrait of the Substrate for Self-Stimulation ». *Psychological Reviews*, vol. 108, p. 228-273.
- Gazzaniga Michael S. 1998. « The Split Brain Revisited ». *Scientific American*, July, p. 50-55.
- Gazzaniga, Michael S. et J.E. Ledoux. 1978. *The Integrated Mind*. New York : Plenum.
- Gazzaniga, Michael S., Richard Ivry et George R. Mangun. 1998. *Fundamentals of Cognitive Neuroscience*. New York : Norton.
- Gibbons, Ann. 1998. « Comparative Genetics : Which of Our Genes Make Us Human ? ». *Science*, vol. 281, no. 4, p. 1432-1434.
- Gilbert, Daniel et Michael T. Gill. 2000. « The Momentary Realist ». *Psychological Science*, vol. 11, p. 394-398.
- Glimcher, Paul W. 2002. « Decisions, Decisions, Decisions : Choosing a Biological Science of Choice ». *Neuron*, vol. 36, p. 323-332.
- Glimcher, Paul W. 2003. *Decisions, Uncertainty, and the Brain : The Science of Neuroeconomics*. Cambridge (Massachusetts) : MIT Press.

- Glimcher, Paul W. et Aldo Rustichini. 2004. « Neuroeconomics : The Consilience of Brain and Decision ». *Science*, vol. 306, p. 447-452.
- Gneezy, Uri et Jan Potters. 1997. « An Experiment on Risk Taking and Evaluation Periods ». *The Quarterly Journal of Economics*, vol. 112, p. 631-645.
- Gneezy, Uri et Aldo Rustichini. 2000. « A Fine is A Price ». *Journal of Legal Studies*, vol. 29, p. 1-17.
- Gobet, Fernand et Herbert A. Simon. 1996. « Recall of Random and Distorted Chess Positions : Implications for the Theory of Expertise ». *Memory and Cognition*, vol. 24, no 4, p. 493-503.
- Gonzalez, Cleotilde, Jason Dana, Hideya Koshino et Marcel Just. 2005. « The Framing Effect and Risky Decisions : Examining Cognitive Functions with fMRI ». *Journal of Economic Psychology*, vol. 26, p. 1-20.
- Goode, Erica. 5 novembre 2002. « On Profit, Loss, and the Mysteries of the Mind ». Entrevue avec Daniel Kahneman, *New York Times* (New York), section F, page 1.
- Green, Leonard et Howard Rachlin. 1991. « Economic Substitutability of Electrical Brain Stimulation, Food, and Water. ». *Journal of the Experimental Analysis of Behavior*, vol. 55, p. 133-143.
- Greene, Joshua D., R. Brian Sommerville, Leigh E. Nystrom, John M. Darley et Jonathan D. Cohen. 2001. « An fMRI Investigation of Emotional Engagement in Moral Judgment ». *Science*, vol. 293, p. 2105-2108.
- Güth, W., R. Schmittberger, B. Schwarze. 1982. « An Experimental Analysis of Ultimatum Bargaining ». *Journal of Economic Behavior and Organization*, vol. 3, p. 367-388.
- Haggard, P., S. Clark et J. Kalogeras. 2002. « Voluntary Action and Conscious Awareness ». *Nature Neuroscience*, vol. 5, no 4, p. 382-385.
- Hardie, Bruce G. S., Eric J. Johnson et Peter S. Fader. 1993. « Modeling Loss Aversion and Reference Dependence Effects on Brand Choice ». *Marketing Science*, vol. 12, no 4, p. 378-394.
- Harper, D.G.C. 1982. « Competitive Foraging in Mallards : Ideal Free Ducks ». *Animal Behavior*, vol. 30, p. 575-584.
- Hayek, F. A. 1952. *The Sensory Order*. Chicago : University of Chicago Press.
- Hayek, F. A. 1988. « Entre l'instinct et la raison ». *In La présomption fatale : les erreurs du socialisme*, p. 19-41. Trad. de l'anglais par Raoul Audoin et révisé par Guy Millière. Paris. Les Presses universitaires de France.

- Herrnstein, R. J. 1961. « Relative and Absolute Strength of Response as a Function of Frequency of Reinforcement ». *Journal of the Experimental Analysis of Behavior*, vol. 4, p. 267-272.
- Herrnstein, R. J. 1997. *The Matching Law*. Éditeurs H. Rachlin et D.I. Laibson, Cambridge (Massachusetts) : Harvard University Press.
- Hsee, C.K., G. Loewenstein, S. Blount et M. Bazerman. 1999. « Preference Reversals Between Joint and Separate Evaluations of Options: a Theoretical Analysis ». *Psychological Bulletin*, vol. 125, no 5, p. 576-590.
- Hsu, Ming et Colin F. Camerer. 2004. « Ambiguity-Aversion in the Brain ». *Mimeo*. Caltech.
- Jallais, Sophie et Pierre-Charles Pradier. 2003. « The Allais Paradox and its Immediate Consequences for Expected Utility » In Fontaine, P. et R. Leonard (éd.), *The 'Experiment' in the History of Economics*. Cambridge (Royaume-Uni) ; Cambridge University Press.
- Jevons, William Stanley. 1871. *The Theory of Political Economy*. MacMillan and Co : Londres et New York.
- Kagel, John et Alvin Roth. 1995. *The Handbook of Experimental Economics*. Princeton : Princeton University Press.
- Kahneman, Daniel. 1994. « New Challenges To The Rationality Assumptions ». *Journal of Institutional and Theoretical Economics*, vol. 150, p. 18-36.
- Kahneman, Daniel. 2000. « Experienced Utility and Objective Happiness: A Moment Based Approach ». In D. Kahneman et A. Tversky (éd.), *Choices, Values, and Frames*, p. 673-692, Cambridge (Royaume-Uni) : Cambridge University Press.
- Kahneman, Daniel. 2003. « Maps of Bounded Rationality: Psychology for Behavioral Economics ». *American Economic Review*, vol. 93, no 5, p. 1449-1475.
- Kahneman, Daniel et Amos Tversky. 1979. « Prospect Theory : An Analysis of Decision Under Risk ». *Econometrica*, vol. 47, p. 263-291.
- Kahneman, Daniel et Amos Tversky. 1984. « Choices, Values, and Frames ». *American Psychologist*, vol. 39, p. 341-350.
- Kahneman, D., B.L. Fredrickson, C.A. Schreiber et D.A. Redelmeier. 1993. « When More Pain is Preferred to Less : Adding a Better End ». *Psychological Science*, vol. 4, p. 401-405.
- Kahneman, Daniel, P.P. Wakker et R. Sarin. 1997. « Back to Bentham ? Explorations of Experienced Utility ». *Quarterly Journal of Economics*, vol. 112, no 2, p. 375-405.
- King-Casas, Brooks, Damon Tomlin, Cedric Anen, Colin F. Camerer, Steven R. Quartz et P. Read Montague. 2005. « Getting to Know You : Reputation and Trust in a Two-Person Economic Exchange ». *Science*, vol. 308, p. 78-83.

- King, Marie-Claire et Allan Wilson. 1975. « Evolution at Two Levels In Human and Chimpanzees ». *Science*, vol. 189, p. 446-447.
- Knutson, Brian, A. Westdorp, E. Kaiser et D. Hommer. 2000. « FMRI Visualisation of Brain Activity During a Monetary Incentive Delay Task ». *Neuroimage*, vol. 12, p.20-27.
- Knutson, Brian, C.M. Adams, G.W. Fong et D. Hommer. 2001. « Anticipation of Increasing Monetary Reward Selectively Recruits Nucleus Accumbens ». *Journal of Neuroscience*, vol. 21, p. RC159.
- Knutson, Brian, G. Fong, C. Adams, J. Varner et D. Hommer. 2001. « Dissociation of Reward anticipation and Outcome with Event-Related fMRI ». *Brain Imaging*, vol. 12, no 17, p. 3683-3687.
- Knutson, Brian et Richard Peterson. 2005. « Neurally Reconstructing Expected Utility ». *Mimeo*, Stanford University, Stanford : Californie.
- Körding, Konrad P., Izumi Fukunaga, Ian S. Howard, James N. Ingram, Daniel M. Wolpert. 2004. « A Neuroeconomics Approach to Inferring Utility Functions in Sensorimotor Control ». *Plos Biology*, vol. 2, no 10, p. E330.
- Kosfeld, Michael, Markus Heinrichs, Paul J. Zak, Urs Fischbacher et Ernst Fehr. 2005. « Oxytocin Increases Trust in Humans ». *Nature*, vol. 435, p. 673-676.
- Laibson, David. 2001. « A Cue Theory of Consumption ». *Quarterly Journal of Economics*, vol. 66, no 1, p. 81-102.
- Layard, Richard. 2003. *Happiness : Has Social Science A Clue ? : Actes de colloque Lionel Robbins Memorial Lectures 2003* (London, 3-5 mars 2003). London : London School of Economics.
- Leboeuf, Robyn Aimee. 2002. « Alternating Selves and Conflicting Choices : Identity Saliency and Preference Inconsistency ». *Dissertation Abstract International*, vol. 63, no 2B, p. 1088.
- LeDoux, Joseph E. 2000. « Emotion Circuits in the Brain ». *Annual Review of Neuroscience*, vol. 23, p. 155-184.
- Lee, R.B. et I. Devore. 1968. *Man the Hunter*. Chicago: Aldine.
- Leonard, Robert J. 1994. « Laboratory Strife : Higgling as Experimental Science in Economics and Social Psychology ». In *Higgling : Transactors and Their Markets in the History of Economics*, Annual Supplement to vol. 26, coll. History of Political Economy, Éd. Neil De Marchi et Mary S. Morgan, Duke University Press, Durham and : London.
- Lerner, Jennifer et Dacher Keltner. 2001. « Fear, Anger, and Risk ». *Journal of Personality and Social Psychology*, vol. 81, p. 146-159.

- Lewin, Shira B. 1996. « Economics and Psychology : Lessons for Our Own Day from the Early Twentieth Century ». *Journal of Economic Literature*, vol. 34, p. 1293-1323.
- Lieberman, M. D. 2000. « Intuition: A Social Cognitive Neuroscience Approach ». *Psychological Bulletin*, vol. 126, p. 109-137.
- Lieberman, M. D., R. Gaunt, D.T. Gilbert et Y. Trope. 2002. « Reflection and Reflexion: A Social Cognitive Neuroscience Approach to Attributional Inference ». *Advances in Experimental Social Psychology*, vol. 34, p. 199-249.
- Lieberman, M. D. 2003. « Reflective and Reflexive Judgment Processes: A Social Cognitive Neuroscience Approach ». In *Social judgments: Implicit and explicit processes*. J. P. Forgas, K. R. Williams, & W. von Hippel (éd.), p. 44-67, New York: Cambridge University Press.
- Lieberman, M. (sous presse). « The X and C systems : The Neural Basis of Reflexive and Reflective Social Cognition ». In : Harmon-Jones, E. et P. Winkelman (éd). (À paraître). *Fundamentals of Social Neuroscience*, New York : Guilford.
- Locke, John. 1690, 2004. *Essai sur l'entendement humain*. Genève : Ellipses.
- Lo, Andrew W. et Dimitry V. Repin. 2001. « The Psychophysiology of Real-Time Financial Risk Processing ». *Working Paper no 4223-01*, MIT Sloan School of Management; *Working Paper no LFE-1039-01*, MIT Laboratory for Financial Engineering.
- Lopes, L.L.. 1994. « Psychology and Economics : Perspective on Risk, Cooperation, and the Market Place ». *Annual Review of Psychology*, vol. 45, p. 197-227.
- Loewenstein, George. 1996. « Out of Control : Visceral Influences on Behavior ». *Organization Behavior and Human Decision Processes*, vol. 65, p. 272-292.
- Loewenstein, G. Frederick et T. O'Donoghue. 2002. « Time Discounting and Time Preference : A Critical Review ». *Journal of Economic Literature*, vol. 40, no 2, p. 351-401.
- Loewenstein, G. et T. O'Donoghue. 2004 « Animal Spirits : Affective and Deliberative Processes in Economic Behavior ». *Mimeo*, Pittsburg : Carnegie Mellon University.
- Loewenstein, Roger. 1996. *Buffet : The Making of an American Capitalist*. New York : Doubleday.
- MacArthur, R.H. et E.R. Pianka. 1966. « On optimal Use of a Patchy Environment ». *Am.Nat.*, vol. 100, p. 603-610.
- MacDonald, A.W., J.D. Cohen, V.A. Stenger et C.S. Carter. 2000. « Dissociating the Role of the Dorsolateral Prefrontal and Anterior Cingulate Cortex in Cognitive Control ». *Science*, vol. 288, p. 1835-1838.

- MacLean, Paul D. 1970. « The Triune Brain : Emotion and Scientific Basis. » In F.O. Schitt (éd.) *The Neurosciences : Second Study Program*. New York : RockFeller University Press.
- Macmillan, Malcolm. 2000. *An Odd Kind of Fame : Stories of Phineas Gage*. Cambridge (Massachussets) : MIT Press.
- Marr, David. 1982. *Vision*. San Francisco : W.H. Freeman.
- Maynard Smith, John. 1982. *Evolution and the Theory of Games*. Cambridge : Cambridge University Press.
- McCabe, Kevin. 2003. « Neuroeconomics ». In *Encyclopedia of Cognitive Science*. New York : Macmillan Publishing, p. 294-298.
- McCabe, Kevin, Daniel Houser, Lee Ryan, Vernon Smith et Theodore Trouard. 2001. « A functional Imaging Study of Cooperation in Two-Person Reciprocal Exchange ». *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, p. 11832-11835.
- McCabe, Kevin et Vernon Smith. 2001. « Goodwill Accounting and the Process of Exchange ». In *Bounded Rationality the Adaptive Toolbox*. MIT Press, p.319-340, Cambridge : Massachussets.
- McClure, S.M., D.I. Laibson, G. Loewenstein et J.D. Cohen. 2004. « Separate Neural Systems Value Immediate and Delayed Monetary Rewards ». *Science*, vol. 304, p. 503-507.
- Medin, Douglas L. et Max H. Bazerman. 1999. « Broadening Behavioral Decision Research : Multiple Levels of Cognitive Processes ». *Psychonomic Bulletin and Review*, vol. 6, no 4, p. 533-547.
- Mehra, R. et E. Prescott. 1985. « The Equity Premium : A Puzzle ». *Journal of Monetary Economics*, vol. 15, p. 145-161.
- Mellers, B.A., A. Schwartz, K. Ho et I. Ritov. 1997. « Decision Affect Theory : Emotional Reactions of the Outcomes of Risky Options ». *Psychol. Sci.*, vol . 8, p. 423-429.
- Mendelson, Joseph. 1967. « Lateral Hypothalamic Stimulation In Satiated Rats : The Rewarding Effects of Self-Induced Drinking ». *Science*, vol . 157:3792, p. 1077-1079.
- Miller, E.K. et J.D. Cohen. 2001. « An Integrative Theory of Prefrontal Cortex Function ». *Annu. Rev. Neurosci*, vol. 24, p. 167-202.
- Montague, P. Read et Gregory S. Berns. 2002. « Neural Economics and the Biological Substrates of Valuation ». *Neuron*, vol. 36, p. 265-284.
- Montague, P. Read, Gregory S. Berns, J.D. Cohen, S.M. McClure, G. Pagnoni, M. Dhamala, M.C. Wiest, I. Karpov, R.D. King, N. Apple et R.E. Fisher. 2002. « Hyperscanning : Simultaneous fMRI during Linked Social Interactions ». *Neuroimage*, vol. 16, no 4, p. 1159-1164.

- Montier, James. 2005. « Emotion, Neuroscience, and Investing : Investors as Dopamine Addicts ». *GRI Equity*, www.griequity.com, 20 janvier 2005.
- Mosteller, F. et P. Noguee. 1951. « An Experimental Measurement of Utility ». *Journal of Political Economy*, vol. 59, p. 371-404.
- Newsome, W.T., K.H. Britten et J.A. Movshon. 1989. «Neuronal Correlates of a Perceptual Decision ». *Nature*, vol. 341, p. 52-54.
- Nowak, M.A., K.M. Page et K. Sigmund. 2000. « Fairness Versus Reason in the Ultimatum Game ». *Science*, vol. 289, p. 1773.
- Ochsner, K.N. et Matthew D. Lieberman. 2001. «The Emergence of Social Cognitive Neuroscience ». *American Psychologist*, vol. 56, p. 717-734.
- O'Doherty, J., M.L. Kringelbach, E.T. Rolls, J. Hornack et C. Andrews. 2001. « Abstract Reward and Punishment Representations in The Human Orbitofrontal Cortex ». *Nat. Neuroscience*, vol. 4, p. 95-102.
- Olds, James et Peter Milner. 1954. « Positive Reinforcement Produced By Electrical Stimulation of Septal Area and Other Regions of Rat Brain ». *Journal of Comparative and Physiological Psychology*, vol. 47, p. 419-427.
- Öhman, Arne. 2002. « Automaticity and The Amygdala : Nonconscious Responses to Emotional Faces ». *Current Directions in Psychological Science*, vol. 11, no 2, p. 62-66.
- Paulus, M.P., N. Hozack, B. Zauscher, J.E., McDowell, L. Frank, G.G. Brown *et al.* 2001. « Prefrontal, Parietal and Temporal Cortex Networks Underlie Decision-Making in the Presence of Uncertainty ». *Neuroimage*, vol. 13, p. 91.
- Payne, J.W., J.R. Bettman et E.J. Johnson. 1993. *The Adaptive Decision Makers*. New York : Cambridge University Press.
- Pecina, S., B. Cagniard, K.C. Berridge, J.W. Aldridge et X. Zhuang. 2003. « Hyperdopaminergic Mutant Mice Have Higher “Wanting” But Not “Liking” for Sweet Rewards ». *Journal of Neuroscience*, vol. 23, p. 9395-9402.
- Phillips, M.L., A.W. Young, C. Senior, M. Brammer, C. Andrew, A.J. Calder, E.T. Bullmore, D.I. Perret, D. Rowland, S.C.R. Williams, J.A. Grey, et A.S. David. 1997. « A Specific Neural Substrate for Perceiving Facial Expressions of Disgust ». *Nature*, vol. 389, p. 495-498.
- Pillutla, M.M. et J. K. Murnighan. 1996. « Unfairness, Anger, And Spite : Emotional Rejections of Ultimatum Offers ». *Organizational Behavior and Human Decision Processes*, vol. 68, p. 208-224.
- Platt, M.L. et Paul W. Glimcher. 1999. « Neural Correlates of Decision Variables in Parietal Cortex ». *Nature*, vol. 400, p. 233-238.

- Popelard, Marie-Dominique et Denis Vernant. 1997. *Les grands courants de la philosophie des sciences*. Paris : Seuil.
- Powell, Kendall. 2003. « Economy of the Mind ». *Biology*, vol. 1, no 3, p. 312-315.
- Premack, David et Guy Woodruff. 1978. « Does the Chimpanzee Have a Theory of Mind ? ». *Behavior Brain Science*, vol. 1, p. 515-526.
- Rabin, M. 1993. « Incorporating Fairness into Game Theory and Economics ». *American Economic Review*, vol. 85, no 5, p. 1281-1302.
- Rabin, M. 1998. « Psychology and Economics ». *Journal of Economic Literature*, vol. 36, no 1, p. 11-46.
- Rainville, P., G. H. Duncan, D.D. Price, B. Carrier, M.C. Bushnell. « Pain Affect Encoded in Human Anterior Cingulate but not Somatosensory Cortex ». *Science*, vol. 277, p. 968-971.
- Reinders, A.A.T.S., J.A. DenBoer, J. Gläscher, J.R. de Jong, A.T.M. Williemsen et C. Büchel. 2003. « Detecting Fearful and Neutral Faces : BOLD Latency Differences in Amygdala-Hippocampal Junction ». *Methods to Meaning In Functional Neuroimaging*, p.123-139.
- Rilling, James K., David A. Gutman, Thorsten R. Zeh, Giuseppe Pagnoni, Gregory S. Berns et Clinton D. Kitts. 2002. « A Neural Basis for Social Cooperation ». *Neuron*, vol. 35, p. 395-405.
- Rilling, James K., Alan G. Sanfey, Jessica A. Aronson, Leigh E. Nystrom et Jonathan D. Cohen. 2002. « Opposing Bold Responses to Reciprocated And Unreciprocated Altruism in Putative Reward Pathways ». *Brain Imaging*, vol. 15, no 16, p. 2539-2543.
- Rilling, James K., Alan G. Sanfey, Jessica A. Aronson, Leigh E. Nystrom et Jonathan D. Cohen. 2004. « The Neural Correlates of Theory of Mind Within Interpersonal Interactions ». *Neuroimage*, vol. 22, p. 1694-1703.
- Rilling, James K., Alan G. Sanfey, Jessica A. Aronson, Leigh E. Nystrom et Jonathan D. Cohen. 2004. « Opposing BOLD responses to Reciprocated and Unreciprocated Altruism in Putative Reward Pathways ». *Brain Imaging*, vol. 15, p. 2539-2543.
- Robson, Arthur J. 2001. « The Biological Basis of Economic Behavior ». *Journal of Economic Literature*, vol. 39, no 1, p. 11-33.
- Romer, Paul M. 2000. « Thinking and Feeling ». *American Economic Review*, vol. 90, no 2, p.439-443.
- Ross, Lee, Mark R. Lepper et Michael Hubbard. 1975. « Perseverance in Self-Perception and Social Perception . Biased Attributional Processes in the Debriefing Paradigm ». *Journal of Personality and Social Psychology*, vol. 32, no 5, p. 880-892.

- Rousseas, S.W. et A.G. Hart. 1951. « Experimental Verification of a Composite Indifference Map ». *Journal of Political Economy*, vol. 59, no 4, p. 288-318.
- Rubinstein, Ariel. 1985. « A Bargaining Model With Incomplete Information About Time Preference ». *Econometrica*, vol. 53, p. 1151-1172.
- Rubinstein, Ariel. 1991. « Comments on the Interpretation of Game Theory ». *Econometrica*, vol. 59, no 4, p. 909-924.
- Rustichini, Aldo, John Dickhaut, Paolo Ghirardato, Kip Smith et Jose V. Pardo. 2002. « A Brain Imaging Study of the Choice Procedure ». *Mimeo*, Computable and Experimental Laboratory, no 17-02, University of Minnesota : Minnesota.
- Sanfey, Alan G., James K. Rilling, Jessica A. Aronson, Leigh E. Nystrom et Jonathan D. Cohen. 2003. « The Neural Basis of Economic Decision-Making in the Ultimatum Game ». *Science*, vol. 300, p. 1755-1758.
- Schneider, Horacio, Flavio C. Canavez, Iracilda Sampaio, Miguel Angelo M. Moreira, Claudia H. Tagliaro, Hector N. Seuanez. « Can Molecular Data Place Each Neotropical Monkey in its Own Branch ? ». *Chromosoma*, vol. 109, no 8, p. 515-523.
- Schneider, Walter et Richard M. Shiffrin. 1977. « Controlled and Automatic Human Information Processing : Perceptual Learning, Automatic Attending and a General Theory. *Psychological Review*, vol. 84, no 2, p. 127-190.
- Schultz, W., P. Dayan et P. Read Montague. 1997. « A Neural Substrate of Prediction and Reward ». *Science*, vol. 275, p. 1593-1599.
- Schultz, Wolfram. 1998. « Predictive Reward Signal of Dopamine Neurons ». *Journal of Neurophysiology*, vol. 80, p. 1-27.
- Schultz, Wolfram. 2002. « Getting Formal with Dopamine and Reward ». *Neuron*, vol. 36, p. 241-263.
- Schultz, Wolfram. 2003. « Certainly Nondeterministic ». *Science*, vol. 300, p. 1662-1663.
- Schultz, Wolfram. 2004. « Neural Coding of Basic Reward Terms of Animal Learning Theory, Game Theory, Microeconomics and Behavioral Ecology ». *Current Opinion in Neurobiology*, vol. 14, no 2, p. 139-147.
- Seegerstrale, Ullica. 2000. *Defenders of the Truth : The Battle for Science in the Sociobiology*. New York : Oxford University Press.
- Shallice, T et P. Burgess. 1998. « The Domain of Supervisory Processes and the Temporal Organizations of Behaviour ». In : Robert, A.C., T.W. Robbins et L. Weiskrahtzl (éd.) *The Prefrontal Cortex*, p. 22-35 Oxford University Press, Oxford.

Shefrin, Hersh et Meier Statman. 1985. « The Disposition to Sell Winners too Early and Ride Losers too Long ». *Journal of Finance*, vol. 40, p. 777-790.

Shiv, Baba, George Loewenstein, Antoine Bechara, Hanna Damasio et Antoine R. Damasio. 2005. « Investment Behavior and the Dark Side of Emotion ». *Psychological Science*, vol. 16, no 6, p. 435-439.

Shizgal, Peter. 1997. « Neural Basis of Utility Estimation ». *Current Opinion in Neurobiology*, vol. 7, no 2, p. 198-208.

Shizgal, Peter. 1998. « On the Neural Computation of Utility : Implications From Studies of Brain Stimulation Reward ». Mimeo, *Center for Studies in Behavioral Neurobiology*. Montréal : Université Concordia.

Shizgal, Peter et Andreas Arvanitogiannis. 2003. « Gambling on Dopamine ». *Science*, vol. 299, no 5614, p. 1856-1858.

Simon, Herbert A. 1997. *Models of Bounded Rationality : Empirical Grounded Economic Reason*. Cambridge (Massachusetts) : MIT Press.

Slooman, Steven A. 1996. « The Empirical Case for Two Systems of Reasoning ». *Psychological Bulletin*, vol. 119, no 1, p. 3-22.

Slovic, Paul et Sarah Lichtenstein. 1971. « Comparison of Bayesian and Regression Approaches to the Study of Information Processing in Judgment ». *Organizational Behavior and Human Performance*, vol. 6, no 6, p. 649-744.

Smith, Kip et John Dickhaut. 2002. « Auction Structure, Heart Rate, and Neuroeconomic Behavior ». *Management Science*, vol. 48, p. 182-189.

Smith, Kip, John Dickhaut, Kevin McCabe et José V. Pardo. 2002. « Neuronal Substrates for Choices Under Ambiguity, Risk, Gains, and Losses ». *Management Science*, vol. 48, no 6, p.711-718.

Smith, Vernon L. 1994. « Economics in the Laboratory ». *Journal of Economic Perspective*, vol. 8, no 1, p. 113-131.

Smith, Vernon L. 2001. « Mind, Reciprocity and Markets in the Laboratory ». Mimeo, Interdisciplinary Center for Economic Science and Mercatus Center at George Mason University.

Smith, Vernon L. 2002. « The Experimental Economist ». Interviewé par Mike Lynch et Nick Gillespie, *Reasononline*, 9 octobre 2002.

Smith, Vernon L. 2003. « What is Experimental Economics ? ». In *Encyclopedia of Cognitive Science*. Interdisciplinary Center for Economic Science at George Mason University.

- Smith, Vernon L. 2004. « Experimental Methods in (Neuro)Economics ». *In Encyclopedia of Cognitive Science*. Interdisciplinary Center for Economic Science at George Mason University.
- Squire, Larry R., Craig E.L. Stark et Robert E. Clark. 2004. « The Medial Temporal Lobe ». *Annual Review of Neuroscience*, vol. 27, p. 279-306.
- Stephens, D.W. et J. R. Krebs. 1986. *Foraging Theory*. Princeton (New Jersey) : Princeton University Press.
- Thurstone, L.L. 1931. « The Indifference Function ». *Journal of Social Psychology*, vol. 2, p. 139-167.
- Thut, G., W. Schultz, U. Roelcke, M. Nienhusmeier, J. Missimer, R.P. Maguire et K.L. Leenders. 1997. « Activation of the Human Brain by Monetary Reward ». *Neuroreport*, vol. 8, p. 1225-1228.
- Tooby, John et Leda Cosmides. 1990. « The Past Explains the Present : Emotional Adaptations and the Structure of Ancestral Environments ». *Ethology and Sociobiology*, vol. 11, p. 374-424.
- Tversky, Amos et Daniel Kahneman. 1974. « Judgment under Uncertainty : Heuristics and Biases ». *Science*, vol. 185, p. 1124-1131.
- Tversky, Amos et Daniel Kahneman. 1981. « The Framing of Decision and The Psychology of Choice ». *Science*, vol. 211, p. 453-458.
- Tversky, Amos et Daniel Kahneman. 1983. « Extensional Versus Intuitive Reasoning : The Conjunction Fallacy in Probability Judgment ». *Psychological Review*, vol. 90, no 4, p. 293-315.
- Tversky, Amos et Daniel Kahneman. 1992. « Advances in Prospect Theory : Cumulative Representation of Uncertainty ». *Journal of Risk and Uncertainty*, octobre 1992, p. 297-323.
- Varian, Hal R. 1990. *Intermediate Microeconomics : A Modern Approach*. New York : Norton and Company.
- Van Raij, W.F. 1999. « History of Economic Psychology ». In Earl, P. et S. Kemp (éd.). *The Elgar Companion to Consumer Research and Economic Psychology*, p. 289-296, E. Elgar : Cheltenham.
- Veblen, Thorstein. 1898. « Why Economics Not an Evolutionary Science ? ». *Quarterly Journal of Economics*, vol. 12, p. 373-397.
- Viner, J. 1925. « The Utility Concept in Value Theory and Its Critics ». *Journal of political Economy*, vol. 33, p. 369-387.
- Von Neumann, J. V. et O. Morgenstern. 1944. *Theory of Games and Economic Behavior*. Princeton (New Jersey) : Princeton University Press.

- Wilson, E. O. 1975. *Sociobiology*. Cambridge (Massachusetts) : Harvard University Press.
- Wilson, T.D., D.J. Lisle, J.W. Schooler, S.D. Hodges, K.J. Klaaren et S.J. Lafleur. 1993. « Introspecting About Reasons Can Reduce Post-Choice Satisfaction ». *Psychological Bulletin*, vol. 19, p. 331-339.
- Wilson, Rick, K., Randolph Stevenson et Geoffrey Potts. 2001. « Brain Activity in the Play of Dominant Strategy and Mixed Strategy Games ». Actes de colloques présentés lors de "Annual Meeting of the Midwest Political Science Association, p. 1-36.
- Winkielman, Piotr, Kent C. Berridge et Julia L. Wilbarger. 2005. « Unconscious Affective Reactions to Masked Happy Versus Angry Faces Influence Consumption Behavior and Judgments of Value ». *Personality and Social Psychology Bulletin*, vol. 31, no 1, p. 121-135.
- Wise, Roy A. 1989. « Brain Dopamine and Reward ». *Annual Review Psychology*. vol. 40, p. 191-225.
- Wise, Roy A. 2002. « Brain Reward Circuitry : Insights from Unsensed Incentives ». *Neuron*, vol. 36, p. 229-240.
- Wolford, George, Miller, Michael B., Gazzaniga, Michael. 2000. « The Left Hemisphere's Role in Hypothesis Formation ». *The Journal of Neuroscience*, vol. 20, p. RC64.
- Wyvell, Cindy L. et Kent C. Berridge. 2000. « Intra-Accumbens Amphetamine Increases the Conditioned Incentive Saliency of Sucrose Reward : Enhancement of Reward "Wanting" without Enhanced "Liking" or Responses Reinforcement ». *The Journal of Neuroscience*, vol. 20, p. 8122-8130.
- Zalla, T., E. Koehler, P. Pietrini, G. Basso, P. Aquino, A. Sirigu et J. Grafman. 2000. « Differential Amygdala Responses to Winning and Losing : A Functional Magnetic Resonance Imaging Study in Humans ». *European Journal of Neuroscience*, vol. 12, p. 1764-1770. *The Journal of Neuroscience*, September 3, 2003, 23(22): 8092-8097.
- Zink, Caroline F., Giuseppe Pagnoni, Megan E. Martin, Mukeshwar Dhamala et Gregory S. Berns. 2003. « Human Striatal Response to Salient Nonrewarding Stimuli ». *The Journal of Neuroscience*, vol. 23, no 22, p. 8092-8097.
- Zink, Caroline. F., G. Pagnoni, G., Megan E. Martin-Skurski, J.C. Chappelow et Gregory S Berns. 2004. « Human Striatal Response to Monetary Reward Depends on Saliency ». *Neuron*, vol. 42, p.509-517.
- Zwick, Rami, Erev, Ido et David Budescu. 1999. « The Psychological and Economical Perspectives on Human Decisions in Social and Interactive Contexts ». In Budescu, Erev et Zwick (éd.). *Games and Human Behavior : Essays in Honor of Amnon Rapoport*. Laurence Erlbaum Associates : New Jersey.