



UNIVERSITÉ DE LORRAINE

MÉMOIRE DE MASTER

Libre Arbitre à l'ère des Sciences Contemporaines

Guillaume ROLLIN

Master MADELHIS

Années 2016 – 2018

Directrice de Mémoire :

Mme. Valeria GIARDINO

Chargée de Recherche (CRCN), CNRS

Septembre 2018

À Julie,

« Donc, rappelle-toi que si tu tiens pour libre ce qui est naturellement esclave et pour un bien propre ce qui t'est étranger, tu vivras contrarié, chagriné, tourmenté ; tu en voudras aux hommes comme aux dieux ; mais si tu ne juges rien que ce qui l'est vraiment – et tout le reste étranger –, jamais personne ne saura te contraindre ni te barrer la route ; tu ne t'en prendras à personne, n'accuseras personne, ne feras jamais rien contre ton gré, personne ne pourra te faire de mal et tu n'auras pas d'ennemi puisqu'on ne t'obligera jamais à rien qui soit mauvais pour toi. »

Le «Manuel» d'Épictète

Remerciements

En premier lieu je souhaite remercier ma Directrice de mémoire, Madame Valeria Giardino, pour avoir accepté de tutorer ce travail. Ses conseils m'ont été d'une précieuse aide. Je remercie également Manuel Rebuschi d'avoir accepté la relecture de ce travail.

L'histoire des réflexions qui ont mené au présent écrit a des origines qui remontent à plus loin que le début de mes études universitaires. A cette époque, je découvrais « L'essai sur le libre arbitre » d'Arthur Schopenhauer qui a bouleversé ma vision des choses, du monde et de ma vie. Je suis donc extrêmement reconnaissant à Madame Giardino, et plus généralement, à tous les intervenants du Master Madelhis de me permettre de matérialiser ma réflexion sur la question. Je ressors presque serein vis-à-vis de ce questionnement passionnant étant, pour moi, amplifié de conséquences éthiques, morales et surtout pratiques.

Je tiens également à remercier mon collègue de l'Institut UTINAM de Besançon, José Lages, de sa bienveillance qui m'a permis de concilier travail et études. Je remercie également Mahé Arexis pour son aide concernant la chronométrie mentale. Je lui souhaite le meilleur pour la soutenance de sa thèse.

Aussi, je souhaiterais remercier mon entourage, mes parents et mes amis qui ont dû supporter mes sempiternelles rengaines : « Je n'ai pas l'temps, j'ai le mémoire à finir ».

Mes derniers mots iront à ma compagne, Julie, à qui est dédié ce manuscrit. Merci à toi de m'accompagner dans mes réflexions, philosophiques ou non, et de partager des discussions qui peuvent être très longues.

Table des matières

Introduction	11
1 Le Libre Arbitre	13
1.1 Introduction	13
1.2 Quelques précisions	14
1.2.1 Le déterminisme	14
1.2.2 Liberté et Libre Arbitre	15
1.3 Les positions contemporaines	16
1.3.1 La position Libertarienne	17
1.3.2 La position Compatibiliste	18
1.3.3 Les positions du Déterminisme “dur” et du Sceptiscisme	20
1.4 La Physique contemporaine et le Libre Arbitre	21
1.4.1 Physique et déterminisme	21
1.4.2 La théorie du Chaos	22
1.4.3 La théorie Quantique	24
1.5 Conclusion	26
2 Action volontaire et expériences neuroscientifiques	29
2.1 Introduction	29
2.2 Mesurer les processus mentaux	30
2.2.1 La chronometrie mentale	30
2.2.2 Le potentiel de préparation moteur (readiness-potential)	32
2.3 L’origine de l’action volontaire : études scientifiques	32
2.3.1 L’expérience de LIBET <i>et al.</i> (1983)	33
2.3.2 L’expérience de FRIED <i>et al.</i> (1991)	35
2.3.3 L’expérience de WEGNER et WHEATLEY (1999)	37
2.3.4 L’expérience de SOON <i>et al.</i> (2008)	39
2.3.5 L’expérience de KHALIGHINEJAD <i>et al.</i> (2018)	42
2.4 Conclusion	44
3 L’action volontaire face aux Neurosciences	47
3.1 Introduction	47
3.2 Le dualisme et l’épiphénoménalisme improbables	48
3.2.1 Contre le dualisme	48

3.2.2	Contre l'épiphénoménalisme	51
3.3	Une action, deux types d'intentions	58
3.3.1	Intentions préalables et intentions en actions	58
3.3.2	Indices du caractère double de l'intention	60
3.3.3	Répercussions du double aspect de l'intention	61
3.4	Conclusion	64
	Conclusion	67

Introduction

La question du Libre Arbitre est un problème qui n'a de cesse d'alimenter les réflexions philosophiques depuis des millénaires. En ouverture de ce mémoire, on trouve une citation d'ÉPICTÈTE. De cette dernière nous voyons qu'il était déjà important pour ce penseur antique de déterminer ce qui dépend des êtres humains, de leur liberté, de ce qui n'en dépend pas. Par ailleurs, les principes inscrits dans cette citation servent de cadre à toute la philosophie Stoïcienne. Ce mémoire n'est, au final, qu'une réflexion dans la quête des meilleurs outils nous permettant de tracer cette distinction. La question qui servira de *leitmotiv* à notre réflexion sera la suivante : *Les sciences contemporaines peuvent-elles nous aider à éclaircir le débat sur le Libre Arbitre ?*

Qu'entend-t-on par sciences contemporaines ? Un travail exhaustif nécessiterait d'envisager cette question d'un point de vue pluridisciplinaire d'une colossale ampleur. Il faudrait peut-être traiter le sujet en prenant en considération la Biologie, la Sociologie, l'Histoire des idées, la Psychologie etc... Pour la présente étude, nous nous limiterons aux cas des Sciences Physiques et des Neurosciences. Nous choisissons ces deux exemples dans la mesure où ces sciences évoquent beaucoup la notion de Libre Arbitre. La première parce qu'une notion du déterminisme en est issue et que l'idée induite par un monde déterministe a un impact sur la production d'une action et sur sa prédictibilité. La seconde parce que ses avancées récentes ont permis aux neuroscientifiques d'avoir un aperçu sans précédent du fonctionnement cérébral qui est aujourd'hui reconnu comme le moteur des actions humaines. Pour effectuer notre analyse, nous procéderons en trois temps.

Dans un premier chapitre, les bases philosophiques sur lesquelles reposent la question du Libre Arbitre seront exposées. Nous verrons alors que le déterminisme causal joue un rôle majeur. En effet, ce dernier sera l'une des deux préoccupations des différents courants actuels concernant la question du Libre Arbitre. Nous constaterons alors que le déterminisme causal, utilisé par ces thèses, prend appui sur la notion de déterminisme utilisé en Physique. C'est ainsi que cela nous conduira à décrire deux grandes révolutions du domaine de la Physique : la Théorie du Chaos et la Mécanique Quantique. Ces révolutions ont profondément modifié notre vision du déterminisme. Les constatations que nous ferons nous mèneront à proposer que ces sciences ne sont pas les plus adéquates pour traiter de la question du Libre Arbitre. Nous en concluons que seules des sciences étudiant le moteur du choix, c'est-à-dire le cerveau et son fonctionnement, sont à même de donner des réponses assez précises à la question que l'on se pose. C'est ce qui nous fera nous tourner vers les Neurosciences pour le reste de ce mémoire.

Dans un second chapitre, nous étudierons la manière dont les Neurosciences s'emparent du débat concernant le Libre Arbitre. Nous resterons alors limités à l'étude de l'action volontaire pour trois raisons. Premièrement, cela permet de restreindre l'étude qui sans cela pourrait potentiellement être très vaste. Deuxièmement parce que le thème de l'action volontaire nous semble être un point

d'achoppement vis-à-vis de la question concernant le Libre Arbitre. En effet, comment comprendre ce concept sans que le processus de l'action ne soit évoqué ? Enfin, troisièmement, il se trouve que l'action volontaire est un domaine beaucoup étudié par les Neurosciences. Nous passerons donc en revue un certain nombre de protocoles expérimentaux qui nous renseigneront sur la manière dont ces sciences abordent la question philosophique de l'action.

Dans un troisième chapitre, nous nous efforcerons de tirer des conclusions philosophiques aux expériences neuroscientifiques que nous aurons vues. Ces conclusions mettront en question deux types de visions qui ont trait au rapport qu'entretiennent le corps et l'esprit. Il s'agira alors de parler du dualisme des substances et d'une forme de dualisme des propriétés s'appelant épiphénoménalisme. Nous verrons alors qu'un modèle philosophique particulier, précisant la notion d'intention, en faisant la distinction entre intention distale et intention proximale, nous sera secourable pour interpréter les résultats de nombreuses expériences.

Enfin, nous conclurons ce mémoire en résumant en quoi les sciences contemporaines nous semblent aujourd'hui primordiales pour juger de la manière dont un processus d'action se produit. Par ce biais, il nous sera possible de généraliser une partie de ces constatations à l'ensemble du Libre Arbitre. Ainsi, nous affirmerons que les sciences contemporaines peuvent être exhibées comme un outil pour aider à trancher la question.

Chapitre 1

Le Libre Arbitre

1.1 Introduction

Le Libre Arbitre pourrait être défini comme la capacité pour des individus rationnels à choisir entre plusieurs alternatives d'actions¹. Cependant, il est difficile d'extraire une définition générale qui serait susceptible de satisfaire toutes les propositions existantes de la notion de Libre Arbitre. C'est ce que nous allons voir dans ce chapitre.

Trois types de thèses philosophiques principales peuvent être donnés. Ces dernières pourraient être associées à des figures historiques de la philosophie. La première, la position libertarienne, pourrait être en partie représentée par Emmanuel KANT dans la mesure où il défend l'idée de l'existence de deux types de causalités dans le monde : l'une physique et l'autre agentielle (causalité de l'agent)². La seconde, la position compatibiliste, peut être représentée par John LOCKE ou David HUME dans la mesure où ils proposent que la volonté dépend elle-même de causes mais que cela est compatible avec l'idée de Libre Arbitre^{3, 4}. Enfin la troisième, regroupant déterminisme dur, pessimisme ou encore scepticisme vis-à-vis de la question du Libre Arbitre, pourrait être en partie représentée par la position de Baruch SPINOZA dans la mesure où pour lui l'idée de Libre Arbitre constitue surtout une illusion⁵. Dans la suite de ce chapitre nous allons donner les principales caractéristiques de ces positions contemporaines.

En premier lieu, il nous faut éclaircir la distinction entre la notion de Libre Arbitre et la notion plus générale de liberté. Il nous faut également définir ce que le concept de déterminisme recouvre. C'est ce que nous allons faire dans la prochaine section de ce chapitre. Ensuite, nous nous concentrerons sur les manières dont les diverses thèses philosophiques que nous avons évoquées articulent ces deux aspects.

1. O'Connor, Timothy, "Free Will", The Stanford Encyclopedia of Philosophy (Summer 2016 Edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/sum2016/entries/freewill/>

2. Emmanuel Kant, *Critique de la raison pure*, 1787

3. John Locke, *Essai sur l'entendement humain*, 1689

4. David Hume, *Traité sur la nature humaine*, 1739

5. Baruch Spinoza, *Lettre 58 à Schuller*, 1674

1.2 Quelques précisions

1.2.1 Le déterminisme

Il est important de distinguer les différentes définitions du déterminisme. Notons que ces dernières sont légions et que le débat sur cette question est loin de laisser émerger une définition claire et unique du concept [Pé12]. Ici, nous choisirons une définition du déterminisme qui menace le plus, et ce depuis des siècles, la question de l'existence du Libre Arbitre. Cette définition est celle du déterminisme qui est dit "causal". Précisons les termes de cette forme de déterminisme⁶ :

« [...] we shall define determinism as the metaphysical thesis that the facts of the past, in conjunction with the laws of nature, entail every truth about the future. »

Ainsi, par cette définition, nous spécifions que le déterminisme repose sur l'idée que tout événement est causalement impliqué par les événements du passé en conjonction avec les lois de la nature. Cette définition a des répercussions sur le problème de l'action volontaire. En effet si l'action est *in fine* gouvernée par les lois de la nature, sur lesquelles l'auteur n'a aucun contrôle, comment concevoir la notion de liberté de cette action ? La position du déterministe causal trouve un représentant historique en la personne de Pierre-Simon de LAPLACE, nous y reviendrons lorsque nous parlerons du lien entre Libre Arbitre et Science Physique.

Suite à cette définition, nous pouvons noter que certaines confusions peuvent apparaître. Comme le fait remarquer Krystèle APPOURCHAUX [App14], il est important de rectifier de possibles malentendus. Citons les principaux :

Le déterminisme causal n'implique pas le fatalisme : dans l'idée du fatalisme, il y a la proposition que quelque soit l'action menée, le futur restera identique. La définition du déterminisme, que nous avons donnée, ne contient rien de tel.

Le déterminisme causal n'implique pas la coercition : dans la définition de la coercition il y a l'idée que l'individu est mû, de façon contrainte, par des forces extérieures. Nous pouvons imaginer faire des choix libres, au sens où ils ne dépendent que de nous et de notre volonté, et que ces choix soient déterminés par des causes. En ce sens, au contraire justement d'une coercition, le déterminisme causal rendrait efficace nos états mentaux sur le monde physique.

Le déterminisme causal n'implique pas une forme de mécanisme et réciproquement : le déterminisme causal ne dit rien vis-à-vis de la réduction des états mentaux aux états physiques. De la même façon, si une explication de l'esprit implique ce type de réduction, rien n'indique qu'un déterminisme causal y est partie prenante.

Le concept de déterminisme causal est à présent défini. Nous avons évoqué en quoi il pouvait poser problème dans la définition que nous donnons au caractère libre d'une action. Nous avons également dissipé d'éventuelles associations, qui n'auraient pas lieu d'être, entre cette idée et d'autres.

6. McKenna, Michael and Coates, D. Justin, "Compatibilism", The Stanford Encyclopedia of Philosophy (Fall 2018 Edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/fall2018/entries/compatibilism/>

Il apparaît à présent que des phénomènes qui ne répondraient pas à la principale caractéristique de cette définition seraient qualifiés d'indéterministe. Une propriété indéterministe d'un événement serait, par exemple, que ce dernier advienne sans qu'une cause ne l'ait provoqué. Avant de passer à la suite, nous souhaiterions effectuer encore une distinction entre deux types d'indéterminismes :

L'indéterminisme ontologique : c'est une forme d'indéterminisme au sens le plus fort. Il relève d'une propriété fondamentale et révèle une impossibilité absolue de prédire un événement.

L'indéterminisme épistémologique : c'est une forme d'indéterminisme dans un sens plus faible que la précédente. Il relève d'une méconnaissance du système que l'on étudie et non d'une de ses propriétés fondamentales.

1.2.2 Liberté et Libre Arbitre

Pour parler du concept de Libre Arbitre, nous pouvons remarquer que ce dernier fait lui-même appel à la notion de liberté. La question de la liberté est un sujet de philosophie de premier plan. Pour déterminer les caractéristiques de cette idée, il convient d'effectuer un certain nombre de distinctions [Mic11]. Ici, nous allons emprunter celles listées par APPOURCHAUX. Il nous semble que cette liste constitue un ensemble de définitions complètes vis-à-vis des différentes facettes de la notion de liberté. L'idée principale est que la liberté n'est pas une, elle pourrait être vue sous cinq angles.

Type 1 : une liberté d'action ou d'autoréalisation. Ce type de liberté est souvent cité dans les ouvrages consacrés au Libre Arbitre [Mic11, App14]. En philosophie classique, il constitue la liberté qu'a un individu lorsqu'il n'est obligé par une contrainte extérieure. On qualifie parfois ce type de liberté, de liberté d'ordre politique. En philosophie contemporaine des contraintes internes à l'individu, telles que les compulsions, sont également ajoutées à ce type. En jouissant de cette liberté, l'individu est capable d'exécuter une action ou de s'en abstenir. Pour cela, il faut que l'individu ne soit soumis à aucune contrainte. Dans ce cadre, la volonté peut être déterminée mais aucune condition extérieure ne l'empêcherait de s'exercer.

Type 2 : une liberté de contrôle réflexif sur soi. Il s'agit de la capacité qu'a un individu à penser, par un retour sur soi, les raisons de ses actes. L'individu a ici la capacité d'évaluer le pour et le contre, de délibérer, d'une action qu'il va mener.

Type 3 : une liberté d'autoperfectionnement. Ici, il y a l'idée que le contrôle réflexif a besoin de "bonnes raisons" pour pouvoir juger. Il faut donc que l'individu ait la liberté de reconnaître ces bonnes raisons. ⁷

7. Ce point est particulièrement important vis-à-vis de la vision de la responsabilité qu'ont les compatibilistes.

Type 4 : une liberté d'autodétermination. C'est la capacité à agir selon des motifs qui ont été créés par l'individu sans intervention extérieure. Ainsi, dans ce cas, il n'est pas nécessaire que l'individu soit en tout temps capable d'agir autrement qu'il ne le fait. Il faut simplement qu'il puisse agir selon des motifs qui eux ont été créés de telle façon qu'ils auraient pu être construits autrement. C'est pour cela que cette liberté nécessite la liberté suivante, celle d'autoformation. C'est-à-dire, la capacité de construire librement ces motifs.

Type 5 : une liberté d'autoformation. C'est la capacité à forger des motifs d'actions qui ne sont pas déterminés par le passé de l'individu. Ces motifs, en conjonction avec la liberté d'autodétermination, sont importants pour une vision incompatibiliste du Libre Arbitre. En substance, pour rendre possible cette liberté, il faut que l'agent soit ultimement la source de la formation de ces motifs d'actions.

Maintenant que nous avons distingué entre ces différents types de libertés, et que nous avons convenablement défini une notion de déterminisme, voyons les relations que les positions philosophiques contemporaines sur le Libre Arbitre entretiennent avec le concept de déterminisme. Comme nous allons le voir, les **Types 1, 2 et 3** de libertés, restent accordables avec une thèse compatibiliste, les **Types 4 et 5** cadrent plutôt avec une vision incompatibiliste qui définit une version forte du Libre Arbitre.

1.3 Les positions contemporaines

		SOMMES-NOUS LIBRES D'AGIR	
		OUI	NON
LE LIBRE ARBITRE EST-IL COMPATIBLE AVEC LE DETERMINISME ?	OUI	COMPATIBILISME	
	NON	LIBERTARIANISME	DETERMINISME "DUR", PESSIMISME, SCEPTICISME

TABLE 1.1 – Tableau résumant les principales positions vis-à-vis de la question du Libre Arbitre et du déterminisme. Ces positions sont dépendantes de la réponse donnée à deux questions fondamentales : *sommes nous libre d'agir ?* et *le Libre Arbitre est-il compatible avec le déterminisme ?*

Dans notre introduction, nous avons vu que les positions existantes vis-à-vis de la question du Libre Arbitre pouvaient être distinguées selon trois grands courants de pensée. Ces courants peuvent être réunis dans un tableau selon les réponses qu'ils donnent aux deux questions suivantes : sommes nous libre d'agir ? Et, le Libre Arbitre est-il compatible avec le déterminisme ? Des réponses minimales à ces questions sont précisées dans le tableau Tab.1.1. Voyons à présent plus en détail en quoi consistent ces réponses.

1.3.1 La position Libertarienne

La position Libertarienne est une position radicalement incompatibiliste⁸. Pour elle envisager que le déterminisme causal, tel que nous l'avons défini, soit l'entière réalité du monde rend la notion de Libre Arbitre impossible. Ce type de pensée présuppose pourtant l'existence du Libre Arbitre et, pour le justifier, prétend que le déterminisme n'est pas l'unique vérité de production des événements du monde. Comme nous l'avons vu une des solutions que le Libertarianisme propose est celle de supposer qu'il y a, au moment de la délibération des individus, une part d'indéterminisme. Cette ouverture causale se loge dans le **Type 4** et le **Type 5** des libertés soit, respectivement, la liberté d'autodétermination et la liberté d'autoformation.

Il apparaît ainsi que, pour un libertarien, la volonté ne peut posséder de cause même si ces dernières sont probabilistes. Alors, dans exactement les mêmes circonstances (tous les états physiques étant les mêmes), il y a une possibilité que l'individu face un autre choix.

« Pour le libertarianisme, une personne, dans un contexte donné, aurait pu former une intention d'action différente de celle qu'elle a effectivement formée, cela étant donné exactement les mêmes circonstances et les mêmes désirs, et en ayant pondéré exactement de la même manière les mêmes raisons d'agir. » [Esf12]

Cette citation de Michael ESFELD a le mérite de clairement énoncer cette idée. Il y a donc, dans cette dernière, la possibilité de monde alternatif. La volonté d'un individu n'y est pas figée par le passé en conjonction avec les lois de la nature. C'est une vision forte du Libre Arbitre. Un indéterminisme de type ontologique est nécessaire (mais non suffisant) ici pour que la thèse soit soutenable. **Si l'on en croit les études de philosophie expérimentale, il semble que ce soit ce type de vision indéterministe que la plupart des gens possèdent intuitivement quand on leur demande de définir une action libre**⁹. Notons que, comme le précise APPOURCHAUX, plusieurs études ont été menées sur la question en philosophie expérimentale. Les conclusions furent partagées entre vision compatibiliste et libertarienne. **Il semble cependant que c'est bien la vision intuitive indéterministe qui finit par l'emporter [App14].**

Sur le plan classique, comme nous le stipulions dans notre introduction, on pourrait proposer qu'Emmanuel KANT soit l'un des acteurs de ce courant. En effet, il propose l'idée de deux types de causalité : une *causalité physique* et une *causalité de l'agent*. La causalité physique est celle que nous connaissons bien. Elle est responsable des relations de causalité que l'on constate dans le monde. La causalité de l'agent est de nature différente. Cette fois, l'individu aurait la capacité de démarrer de nouvelles chaînes causales non déterminées par celles qui précèdent. En quelque sorte, l'individu devient substance et un dualisme apparaît. On comprend ainsi que cette vision a une forte connotation métaphysique même si le tenant moderne de cette position tente de la réduire.

Quels sont les avantages, et les inconvénients, de ce type de position ? Dans la colonne avantage, nous pouvons noter que la nécessité d'un indéterminisme semble proposer une vision intuitive de ce que l'on présuppose du Libre Arbitre. Qui plus est, cette vision semble aussi conserver toute la force que l'on attribue à cette hypothétique faculté. Dans la colonne des inconvénients, nous pouvons, comme nous l'avons dit juste avant, arguer que cette position porte un bagage métaphysique fort

8. Nous parlons ici de la position Libertarienne vis-à-vis du Libre Arbitre et non de la position politique qui n'a rien à voir

9. Evoqué dans : Julien Ouellet, *Le scepticisme à propos du libre arbitre*, Université Laval, Mémoire de Master, 2016

flirtant avec un dualisme des substances. De plus, il apparaît que des arguments philosophiques forts contredisent cette version. L'un d'entre eux est *l'argument de la chance*, il suit le raisonnement suivant :

1. Les libertariens proposent que si les actions des agents ne dépendent pas, in fine, d'un processus indéterministe alors les agents ne peuvent être qualifiés de libre.
2. Si les actions des agents sont totalement indéterminées alors l'action ne peut résulter que de la chance.
3. La chance n'est pas contrôlée par l'agent.
4. En conséquence, si l'action résulte d'un processus indéterminé, l'agent n'est pas libre.
5. Les libertariens ne font finalement que proposer un modèle où aucune action libre n'est possible.

L'argument est très fort et il semble même fatal à un certain nombre de thèses libertariennes (au moins les plus basiques). Comme OUELLET le fait remarquer¹⁰, il existe des thèses libertariennes qui résistent à cet argument en y attaquant les prémisses. C'est le cas de la thèse de Laura EKSTROM. Cependant, cette dernière semble incapable de rendre compte de la liberté d'autoformation de l'individu.

Dans le modèle de EKSTROM, l'indéterminisme apparaît au moment de la délibération qui est partie prenante de la formation du caractère de l'agent et c'est le caractère qui détermine l'action. Or, forger son caractère ne peut se faire que sur la base de principes que nous noterons P_1 . Pour être libre selon le critère d'EKSTROM, l'agent doit aussi avoir choisi ces principes P_1 . Cela ne peut se faire que sur la base de principes que nous noterons P_2 . Si l'on suit ce raisonnement cela conduit soit à une régression à l'infini, soit à des principes de bases P_n que l'agent n'a pas pu choisir. Structurer son caractère semble donc impossible. En effet, ce qui est appelé *l'argument de la base* est difficilement digérable par ce type de thèse.

1.3.2 La position Compatibiliste

Comme son nom l'indique, la thèse Compatibiliste propose que le déterminisme soit vrai et que cela ne pose aucun problème à la question du Libre Arbitre. En effet, les compatibilistes proposent que la volonté elle-même est déterminée mais que cela ne nuit en rien à la question de la liberté. Pour eux la notion intuitive de liberté, telle qu'employée par les libertariens, est absolument inintelligible. Pour éviter cette écueil, les compatibilistes proposent une sorte de redéfinition de la notion de la liberté.

D'un point de vue classique, nous avons mentionné LOCKE et HUME comme des représentants de ce courant. La raison en est simple : dans la pensée de ces auteurs, la liberté ne relevait que des

10. *Ibid.*

deux premiers types. De plus, pour ces auteurs, parler de volonté libre n'avait pas de sens : seuls les individus pouvaient être libres et pas les volontés. Il y avait une distinction claire qui était faite entre la volonté et la liberté. On retiendra l'exemple célèbre de LOCKE où un agent, enfermé dans une cellule, décide de rester à sa place. Le prisonnier fait, en ce sens, preuve de volonté mais n'est pas libre parce que la porte est verrouillée. En effet, s'il avait voulu autrement, il n'aurait de toute façon pas pu sortir ¹¹.

Dans les types de libertés que nous avons listés, les compatibilistes contemporains ne vont conserver que les trois premiers (le troisième type vient s'ajouter aux deux premiers soutenus classiquement). Ainsi, pour eux, il ne s'agit pas de dire que dans les mêmes circonstances un choix pourrait être différent mais que si l'individu avait voulu différemment alors le choix aurait été différent. Il s'agit d'une définition qui relève de l'analyse conditionnelle de la liberté [Mic11]. Cependant, ce type d'analyse semble mener à une définition insatisfaisante et c'est pourquoi des auteurs contemporains ont proposé d'autres façons de définir la liberté.

Parmi ces auteurs, on retrouve un compatibiliste célèbre en la personne de Harry FANCKFURT. Son idée de la liberté repose sur la cohérence entre deux types de désirs :

Les désirs de premier ordre : ce sont les désirs de base que nous ressentons tous. Ces désirs sont du type : "je désire X"

Les désirs de second ordre : ce sont des désirs qui portent sur des désirs. Ils sont formés du désir que les désirs de premier ordre soit effectif. Ces désirs sont du type : "je désire désirer X".

Pour FANCKFURT, la définition de la liberté est que les deux types de désir soient cohérents. Ainsi, pour prendre un de ses exemples célèbres, il existerait des drogués réticents et des drogués consentants. Les deux drogués auraient le désir de premier ordre de consommer de la drogue, c'est leur désir de second ordre qui serait différent. Le drogué réticent aurait un désir de ne pas désirer de la drogue, il ne serait pas libre d'en consommer. Le drogué consentant aurait un désir de continuer à désirer de la drogue, il serait libre dans sa consommation ¹².

Quelle sont les avantages et les inconvénients de la position compatibiliste ? Dans la colonne des avantages nous pouvons noter que la redéfinition du concept de Libre Arbitre, afin de le rendre compatible avec la notion du déterminisme, offre un réel plus puisque le déterminisme semble bien ancré dans les lois naturelles découvertes par les sciences. De fait, on voit que l'inintelligibilité de la vision Libertarienne, qui a besoin d'un indéterminisme fondamental au sein des actions de délibérations, est elle aussi dépassée. Dans la colonne désavantages on peut noter au moins deux arguments principaux : l'*argument de la conséquence* et l'*argument de la manipulation* ^{13, 14}.

11. John Locke, *Essai sur l'entendement humain*, 1689. Exemple cité dans [Mic11].

12. Faisons remarquer que nous n'avons pas parlé du problème de la responsabilité, c'est un thème cher aux compatibilistes. En effet, concilier déterminisme et responsabilité n'est pas chose aisée. Dans le dernier exemple, le drogué réticent ne serait pas responsable s'il se mettait à consommer de la drogue car il ne serait pas libre de le faire. Le drogué consentant en revanche pourrait être considéré comme responsable de ses actes puisque, par la définition de FANCKFURT, il est libre.

13. McKenna, Michael and Coates, D. Justin, "Compatibilism", The Stanford Encyclopedia of Philosophy (Fall 2018 Edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/fall2018/entries/compatibilism/>

14. Les deux arguments sont aussi synthétisés dans : Julien Ouellet, *Le scepticisme à propos du libre arbitre*,

Le premier argument peut s'énoncer en ces termes :

1. Personne n'a de pouvoir sur le passé et sur les lois de la nature.
2. Personne n'a de contrôle sur le fait que les lois de la nature entraînent les événements futurs (déterminisme causal que nous avons défini).
3. En conclusion, personne n'a de contrôle sur les événements futurs.

En substance, cet argument montre notre incapacité à contrôler les événements présents si le déterminisme est vrai. Les prémisses de ce raisonnement sont l'impossibilité à contrôler les événements du passé ainsi que les lois de la nature. L'individu ne disposant d'aucun contrôle sur ce qui se passe, on en conclut que Libre Arbitre et déterminisme ne peuvent pas être conciliés.

Le second argument repose sur une continuité entre plusieurs expériences de pensées décrivant des situations de manipulations psychologiques d'un agent. Au départ de cette continuité, appelons cette expérience de pensée Ex_i , la manipulation est flagrante et le compatibiliste ne peut arriver qu'à la conclusion de la non liberté de l'agent ¹⁵. A la fin, la manipulation est plus subtile et le cas décrit ressemble à une situation réelle ¹⁶. Le compatibiliste ne peut plus alors conclure à la non liberté de l'agent ; appelons cette expérience Ex_f . Tout repose sur des cas d'expériences de pensées intermédiaires. Ces cas mènent de Ex_i à Ex_f sans changer fondamentalement la nature de la manipulation. Reconnaisant la non liberté dans la première expérience et la liberté dans la dernière, le compatibiliste se retrouve en contradiction et en bien mauvaise posture.

1.3.3 Les positions du Déterminisme “dur” et du Scepticisme

Il nous reste un type de position à voir, celui du déterminisme dur et du scepticisme que nous avons classé dans la même catégorie de notre tableau récapitulatif Tab. 1.1. Les déterministes durs croient le déterminisme vrai. En revanche, ils sont incompatibilistes du point de vue du Libre Arbitre. Ils en concluent naturellement que le Libre Arbitre est un concept impossible.

D'un point de vue classique, nous avons mentionné SPINOZA comme tenant de ce courant. Ce qui nous fait le dire c'est que la pensée de SPINOZA a pour conséquence d'affirmer que notre expérience de la liberté n'est qu'une pure illusion. Nous pouvons le constater dans une célèbre correspondance où l'auteur écrit :

« Concevez maintenant, si vous voulez bien, que la pierre, tandis qu'elle continue de se mouvoir, pense et sache qu'elle fait effort, autant qu'elle peut, pour se mouvoir. Cette pierre assurément, puisqu'elle a conscience de son effort seulement et qu'elle n'est en aucune façon indifférente, croira qu'elle est très libre et qu'elle ne persévère dans son mouvement que parce qu'elle le veut. »

Université Laval, Mémoire de Master, 2016

15. Si l'on part de la thèse de FRANCKFURT, dans ce cas, les désirs de second ordre ne proviennent pas de l'individu, ils sont forcés de l'extérieur via manipulation psychologique.

16. La “manipulation” n'est plus alors que la résultante de la vérité du déterminisme causal. L'individu a alors un maillage psychologique forgé de l'extérieur par les événements de la vie courante.

Telle est cette liberté humaine que tous se vantent de posséder et qui consiste en cela seul que les hommes ont conscience de leurs appétits et ignorent les causes qui les déterminent. » ¹⁷

En ce sens la pensée de SPINOZA sur la question du Libre Arbitre est une philosophie purement déterministe. Il est bien entendu délicat d'affirmer que la position de SPINOZA représente l'entièreté de la pensée du déterminisme dur et nous nous garderons bien de le faire. Ce que nous souhaitons simplement souligner ici c'est la position radicale de l'auteur qui affirme que la croyance dans le Libre Arbitre n'est qu'une illusion.

D'un point de vue contemporain, la pensée du déterminisme dur et la pensée sceptique, émergent d'abord et avant tout grâce à des arguments forts contre la position libertarienne (*argument d'intelligibilité* et *argument de la chance* par exemple) et de ceux contre la position compatibiliste (*arguments de la conséquence* et *argument de la manipulation* par exemple). Par suite, plusieurs auteurs ont aussi relevé une forme d'illusion de la croyance au concept du Libre Arbitre, ce qui mène à une forme de scepticisme vis-à-vis de cette question. Une différence entre les déterministes durs repose sur les conséquences qu'ils envisagent vis-à-vis de la prise de conscience dans le fait que l'idée du Libre Arbitre n'est qu'une illusion. Deux courants peuvent alors être mentionnés : ceux qui pensent qu'il y aurait un bienfait à la négation de l'idée de Libre Arbitre (Derk PEREBOOM) et ceux qui pensent, de manière pragmatique et normative, que la croyance dans le concept de Libre Arbitre est nécessaire à la société tout en sachant qu'il n'est qu'une illusion (Saul SMILANSKY).

1.4 La Physique contemporaine et le Libre Arbitre

À présent que nous avons énuméré les positions philosophiques principales concernant la question du Libre Arbitre, venons-en au sujet même de ce mémoire à savoir : quels aspects du Libre Arbitre les sciences contemporaines sont-elles susceptibles d'éclairer ? Dans la suite, nous nous concentrerons sur deux sciences particulières à savoir la Physique et les Neurosciences. Dès maintenant, nous allons parler de ce que la Physique apporte à la question. Nous constaterons que cet apport concerne d'abord la question du déterminisme. C'est ce constat qui nous fera accorder plus d'importance aux Neurosciences.

1.4.1 Physique et déterminisme

La définition du déterminisme que nous avons adoptée (celle du déterminisme causal) est liée à une vision en vogue à partir du XIX^{ème} siècle. Cette position est dite Laplacienne dans la mesure où Pierre-Simon de LAPLACE, le mathématicien, astronome et physicien français, s'exprime en ces termes :

« Une intelligence qui pour un instant donné connaîtrait toutes les forces dont la nature est animée et la situation respective des êtres qui la composent, si d'ailleurs elle était assez vaste pour soumettre ces données à l'analyse, embrasserait dans la même formule les mouvements des plus grands corps de l'univers et ceux du plus léger atome : rien ne serait incertain pour elle, et l'avenir, comme le passé, serait présent à ses yeux. » ¹⁸

17. Baruch Spinoza, *Lettre 58 à Schuller*, 1674

18. Pierre-Simon de Laplace, *Essai philosophique sur les probabilités*, 1814

Il serait alors aisé de proposer que LAPLACE soit le père de cette vision déterministe. Comme l'article [Pé12] en témoigne, cela est sans doute plus complexe. Quoi qu'il en soit, nous voyons ici que la Physique, via la mécanique des corps (en particulier célestes), fait une proposition qui ne peut qu'impacter la vision du Libre Arbitre.

Avant l'avènement de la Physique de Newton, l'omniscience divine était ce qui posait problème vis-à-vis de la question du Libre Arbitre. En effet, comment comprendre l'existence d'actions libres si Dieu est capable de connaître à l'avance cette action ? Dans les propos de LAPLACE, entre les lignes, nous pouvons lire ici que sa proposition remplace l'omniscience divine par la vérité des équations et de leurs prédictions mathématiques. Ainsi, la calculabilité des propriétés d'un corps rendrait le futur aussi bien que le passé de ce corps connu. On en arrive bien à la conclusion que tous les événements seraient la résultante des événements du passé en conjonction avec un certain nombre de lois fondamentales de la nature. Par exemple la dynamique, résultante de la loi de la gravitation de NEWTON, cadre parfaitement avec cette idée. Le futur de tous les objets soumis à la gravitation, dans la mesure où l'on connaît leurs états passés et les lois de la nature en vigueur, est connu par le calcul. Nous pouvons reconnaître quatre points caractéristiques à la vision Laplacienne [Bis02].

1. **Dynamique différentielle (DD)** : il existe un algorithme non-probabiliste qui lie l'état d'un système à l'instant t à l'état d'un système à l'instant t' .
2. **Evolution Unique (EU)** : un système dans un état E suit toujours la même évolution.
3. **Détermination (D)** : tout état E d'un système peut être décrit avec une erreur arbitrairement petite.
4. **Précision (P)** : grâce à l'algorithme, les états E d'un système peuvent être prédits avec une erreur arbitrairement petite.

Si nous en restons à ces propositions, propositions que nous qualifierons de classiques, nous voyons que la Physique a tendance à donner raison à la thèse déterministe. Si le déterminisme était vrai, la thèse libertarienne se retrouverait affaiblie et il ne resterait plus que les thèses compatibilistes et celles du déterminisme "dur". Pourtant trois révolutions scientifiques, dans le milieu de la Physique, sont apparues durant le siècle dernier. Ces révolutions sont la théorie de la Relativité, la théorie du Chaos et la théorie de la Mécanique Quantique. La théorie de la Relativité ne faisant qu'apporter de l'eau au moulin du déterminisme, celle-ci restant classique de ce point de vue, nous nous concentrerons sur les deux dernières qui ont entâché les certitudes apportées par la vision Laplacienne. C'est de ces révolutions dont nous allons à présent parler.

1.4.2 La théorie du Chaos

La théorie du Chaos vient de l'étude de la dynamique d'un système classique. Dans ce cas, un ensemble d'équations différentielles mathématiques décrivent la position d'un point dans l'espace des états dont la dimension correspondra aux nombres d'équations dynamiques en présence. La

connaissance des propriétés de ce point, ainsi que la connaissance d'une loi de la nature, suffisent pour donner l'évolution du système considéré, aussi bien son passé que son futur. Afin de donner un exemple simple, si l'on considère la dynamique d'un corps céleste, les informations nécessaires à la reconstruction de toute sa dynamique sont données par les coordonnées de ses positions dans l'espace réel à trois dimensions (x, y, z) et les coordonnées de sa vitesse (v_x, v_y, v_z) suivant ces trois axes. Avec la connaissance d'une loi dynamique naturelle, on peut reconstruire l'évolution passée et future de toute la trajectoire dans cet espace à 6 dimensions.

Le problème du chaos naît de notre incapacité à connaître parfaitement les conditions initiales d'un système physique. C'est sur le point (D) que la théorie du chaos joue. En effet, le manque de connaissance des conditions initiales d'un système entraîne une incertitude vis-à-vis de la prédiction d'un état ultérieur. Le fait qu'un état initial E_i , très peu différent, entraîne un état final E_f différent n'a rien de surprenant mais c'est dans les proportions des différences qu'intervient la vraie surprise. En effet, deux trajectoires très proches au départ peuvent diverger extrêmement rapidement et de manière très importante durant la dynamique.

Notons t_i le moment initial de la dynamique d'un système, E_i l'état initial de ce système, t_f le moment final de la dynamique et E_f l'état final du système. Si nous avons une méconnaissance minimale de l'état E_i alors les résultats de l'algorithme portant sur l'état E_f sont potentiellement incertains si la dynamique du système est chaotique. Il est important de noter que dans le cas classique, on parle de chaos déterministe. Les lois qui gouvernent ce type de système restent déterministes au sens où la connaissance exacte des conditions initiales mène à une connaissance exacte de l'état final. L'indéterminisme ne vient que d'une inévitable erreur portant sur les conditions initiales. Ce type de considération est particulièrement représentée dans la culture populaire sous le terme d'*effet papillon*. Ce terme fait suite aux travaux d'Edward LORENZ [Lor72] sur un modèle météorologique très simplifié où une différence minimale dans les conditions initiales d'un système entraîne des dynamiques très différentes¹⁹. Le chaos apparaît aussi, par exemple, dans le cas du problème à trois corps (par exemple la dynamique d'une comète soumise à l'attraction conjointe du Soleil et de Jupiter). La complexité des trajectoires engendrées avait déjà été supposée par Henri POINCARÉ [Poi90] dès la fin du XIX^{ème} siècle. Plus généralement, ce problème apparaît aussi beaucoup dans les systèmes dynamiques composés d'équations différentielles non-linéaires.

La conclusion philosophique que l'on peut tirer du constat imposé par cette théorie c'est que l'idéal déterministe Laplacien est loin d'être effectif. Ainsi, la distinction entre déterminisme ontologique et épistémologique trouve ici tout son sens. Cela représente un aspect important pour la question du Libre Arbitre parce qu'il complexifie l'idée que l'on peut se faire du déterminisme causal. Il montre que le simple fait de connaître de façon partielle les conditions initiales d'un système peut entraîner une méconnaissance radicale de son futur. La connaissance approximative des conditions initiales ne sera donc plus de mise pour pouvoir prédire un état final, même approximativement, si le système est chaotique. Il apparaît ainsi que l'impossibilité à connaître de manière infiniment précise l'état d'un système a des conséquences néfastes sur la prédictibilité de sa dynamique passée et future. Bien que cet aspect crée une faille dans la prédictibilité absolue supposée par la thèse déterministe, l'indéterminisme qu'il produit ne peut suffire aux thèses Libertariennes. Ces dernières ont en effet besoin d'un indéterminisme de type ontologique. C'est cela

19. L'image du papillon fait référence au titre de la conférence d'Edward LORENZ de 1972 : *Predictability : Does the Flap of a Butterfly's Wings in Brazil Set off a Tornado in Texas ?*

que va apporter la thèse suivante que nous allons aborder.

1.4.3 La théorie Quantique

La révolution apportée par la mécanique quantique est qu'elle remet en question le point (EU) de la vision Laplacienne. En effet, l'unicité des résultats n'a plus le même sens que celui que nous donnions dans le cas de la dynamique classique. A présent, l'état E d'un système ne sera plus donné de façon absolue mais sous forme de probabilité. La mécanique quantique a aussi la particularité d'apporter des caractéristiques tout à fait nouvelles vis-à-vis de notre appréhension du monde. Ces caractéristiques peuvent être résumées en 3 points [Bis02] :

1. Indéterminisme
2. Non-localité
3. Intervention de l'observateur

Pour le premier point, nous l'avons vu précédemment, dans le cas classique, un état est représenté par un point dans un espace à N dimensions où N correspond au nombre de variables dynamiques du système. Dans le cas quantique, on parle de vecteur d'état. Traditionnellement, on le note $|\Psi\rangle$ et c'est un vecteur représenté dans un espace dit de Hilbert [CTL07]. Ce vecteur est composé de vecteurs propres $|\Psi_i\rangle$ associé à des valeurs d'états propres que l'on peut noter K_i . Ces valeurs propres sont les valeurs mesurables d'un système. Là où il y a un réel changement par rapport à la mécanique classique c'est que les états propres $|\Psi_i\rangle$ d'un système sont dans un état superposé avant la mesure. On parle de principe de superposition quantique. Ainsi, l'état $|\Psi\rangle$ du système est dans une superposition d'états $|\Psi_i\rangle$ ce qui est un comportement qui n'a pas d'équivalent classique lorsque l'on parle de l'énergie, ou de la position d'un objet, par exemple. C'est aussi tout à fait différent d'un mélange statistique où le système est bien dans un unique état mais où cet état n'est connu que statistiquement. Ici, l'état superposé est solution de l'équation fondamentale de la mécanique quantique, ou équation de Schrödinger. C'est l'opération de mesure qui réduira l'état superposé en l'un des états $|\Psi_i\rangle$ de valeur propre K_i . Au sein de l'interprétation orthodoxe de la mécanique quantique, il y a donc une incertitude fondamentale quant au résultat après mesure qui ne découle nullement d'une méconnaissance du système. Ici, on a une théorie qui possède un indéterminisme de type ontologique (au moins pour son interprétation orthodoxe).

Le second point, la non localité, fait référence au fait que la théorie autorise qu'un état $|\Psi\rangle$ puisse être lié à un autre état que l'on notera $|\Phi\rangle$, on parlera alors d'intrication. Le fait qui a troublé les physiciens des années 30 est que cette intrication puisse exister entre deux états infiniment éloignés (au sens de la Physique) l'un de l'autre. En effet, dans le cas d'une intrication, si une mesure est effectuée sur $|\Psi\rangle$ alors instantanément une réduction des possibilités de mesure s'opère sur $|\Phi\rangle$. Or, si les deux états sont infiniment éloignés (au sens de la Physique), comment comprendre que l'un communique à l'autre son état ? Il faudrait que l'information voyage plus vite que la lumière ce que la théorie de la relativité interdit. On parle alors, dans les années 30, de paradoxe EPR pour EINSTEIN-PODOLSKY-ROSEN. Il est aujourd'hui accepté qu'aucune information ne voyage entre les

deux états. Il apparaît plutôt que des états quantiques intriqués se comportent comme un système quantique même s'ils sont très éloignés. Cela rend, par conséquent, la théorie quantique non-locale.

Enfin, le troisième point repose sur l'idée que l'expérimentateur jouerait un rôle dans la sélection des états superposés au moment de la mesure. Cette idée remonte à la célèbre expérience de pensée de Erwin SCHRÖDINGER connue sous le nom de *chat de Schrödinger*. Dans cette expérience de pensée, un chat vivant est contenu dans une boîte où une fiole de poison peut se rompre ou non en fonction de l'état quantique de désintégration d'une particule. La particule est donc dans un état superposé de type $|\text{désintégrée}\rangle + |\text{non désintégrée}\rangle$; par continuité, la fiole de poison est dans un état $|\text{rompue}\rangle + |\text{non rompue}\rangle$ et, encore par continuité, le chat serait dans un état $|\text{vivant}\rangle + |\text{mort}\rangle$. Ce n'est que la mesure, c'est-à-dire le moment d'ouverture de la boîte, qui déterminera l'état du chat. Dans cette expérience de pensée l'idée était de montrer comment les possibilités quantiques étaient susceptibles de se propager à un état macroscopique. Une des interprétation fût de dire que c'est la mesure, dans ce cas la conscience de l'expérimentateur, qui pourrait agir sur la matière en provoquant l'état final. Cependant, cette position semble aujourd'hui marginale. Ce problème, dans toute sa généralité, est qualifié de problème de la mesure.

Voilà qui liste les apports de la mécanique quantique vis-à-vis de la Physique classique et donc aussi de la vision classique que l'on a du déterminisme causal. En quoi, philosophiquement parlant, cela pourrait avoir un impact sur la question du Libre Arbitre ? Nous l'avons vu, il semble que la question du déterminisme causal soit importante aux différentes thèses sur la question. La mécanique quantique, en minant l'universalité de la notion même de déterminisme causal des lois de la nature, donne des arguments pour les tenants d'une thèse indéterministe. Nous reconnaissons que quelques auteurs comme Roger PENROSE ou Henry STAPP proposent des idées intéressantes vis-à-vis de l'importance de la mécanique quantique sur la question de la conscience [Hod02]. Ces auteurs proposent en effet que les processus quantiques soient parties prenantes du fonctionnement cérébral. Cependant, nous avons choisi de ne pas prêter attention à ces propositions dans ce mémoire. Comme de nombreux autres auteurs [Hod02], nous jugeons la théorie quantique comme non pertinente vis-à-vis de questions relevant du problème de Libre Arbitre. En effet, il nous semble que ce qui est vraiment important pour cette question est l'étude du cerveau car c'est le cerveau qui constitue le moteur du choix d'un agent. De plus, il apparaît que les processus physiques présents dans le cerveau, par leurs caractéristiques, en particulier leurs tailles et leurs températures, relèvent de la mécanique classique et que des phénomènes quantiques effectifs et mesurables y semblent absents ou non pertinents vis-à-vis de l'explication du mécanisme cérébral. A vrai dire cela n'a rien d'étonnant car nous nous rallions, dans ce mémoire, à la thèse dite de la *décohérence quantique* qui en Physique explique une bonne partie (mais pas tous les problèmes) du passage du monde quantique au monde classique. Ce phénomène rend les spécificités des propriétés quantiques extrêmement fragiles vis-à-vis de leur environnement (en particulier de la température). Cela pose de nombreux problèmes aussi bien aux physiciens, lorsqu'ils souhaitent utiliser les propriétés quantiques (voir les problèmes de l'informatique quantique), qu'aux philosophes, par exemple libertariens, lorsqu'ils souhaitent y voir l'origine de l'indéterminisme nécessaire à leur position. En effet, pour donner raison à leur thèse, ces derniers doivent montrer comment des phénomènes quantiques dans le cerveau sont susceptibles de produire des effets indéterministes macroscopiques. La tâche est loin d'être aisée. Certains proposent, par exemple, d'allier la mécanique quantique à la théorie du chaos pour exporter des effets indéterministes à ces plus grandes échelles (voir [Bis02, Hod02]).

L'idée, bien qu'intéressante, reste cependant hypothétique et, d'un point de vue pragmatique, il nous semble que les conséquences de ces hypothétiques caractéristiques devraient être constatées par les sciences du cerveau. C'est donc tout l'intérêt de tourner notre attention vers ces dernières qui sont en mesure de décrire précisément le moteur du choix.

1.5 Conclusion

Dans ce premier chapitre, nous avons d'abord appris à distinguer entre différents types de liberté. Suite à cela, nous nous sommes placés dans le contexte du déterminisme causal qui constitue la définition actuellement reconnue et utilisée du déterminisme. Ensuite, il a été possible de donner les principales positions philosophiques articulant la question du Libre Arbitre et du déterminisme. Nous avons alors constaté que certaines positions s'accommodaient d'un déterminisme total des lois de la nature (la position Compatibiliste), alors que d'autres le refusaient absolument (les positions Libertariennes et Déterministe Dur). Il a donc paru important de creuser cette notion de déterminisme causal.

Nous avons montré que le déterminisme causal était intrinsèquement lié à la vision physicienne que l'on se faisait du monde. Il est alors devenu nécessaire de passer en revue deux grandes révolutions de la Physique du siècle dernier, la théorie du Chaos et la théorie Quantique. Ces grandes révolutions remettaient en cause la vision du déterminisme absolu qui avait cours depuis le XIX^{ème} siècle. Nous avons volontairement laissé de côté la théorie de la Relativité dans la mesure où cette dernière restait purement déterministe au même titre que la mécanique classique. Nous ne nions cependant pas que cette théorie a aussi eu un impact vis-à-vis de la question du Libre Arbitre (voir [BA02]).

La conclusion que nous pouvons tirer de ces considérations, c'est que la Physique contemporaine apporte un certain nombre d'arguments concernant la question du déterminisme des lois de la nature :

La théorie du Chaos, porte en son sein un déterminisme de type épistémologique. Elle ne peut donc vraiment trancher un débat concernant le déterminisme ou l'indéterminisme de la nature. Il n'en reste pas moins qu'elle remet en question la pureté de la vision déterministe classique. Nous proposons également ici l'idée que le chaos serait susceptible d'entretenir une certaine forme d'illusionnisme vis-à-vis de la question du Libre Arbitre. Dans cette idée, le monde serait bien déterminé mais ce sont les incertitudes sur les conditions initiales observées qui nous donneraient l'illusion que deux systèmes identiques choisissent. Ce processus donnerait l'impression que, dans deux situations identiques, le choix non déterminé est possible alors que les deux systèmes ne seraient simplement pas dans un même état initial mais dans deux états très proches. Cette illusion ne ferait qu'encourager une vision commune, plutôt incompatibiliste, du Libre Arbitre. Il nous semble que cette question mériterait d'être approfondie, cela pourrait être le sujet d'un futur travail.

La théorie Quantique, quant à elle, bien que supportant un indéterminisme ontologique ne semble pas adéquate à décrire les propriétés du cerveau. Ce dernier est pourtant le producteur de la capacité que nous aurions à être libre (que cela s'entende par le prisme libertarien

ou compatibiliste). Les propriétés quantiques se verraient par conséquent inadéquates vis-à-vis de la question que nous nous posons ici. Dans cette idée, nous rejoignons la parole de nombreux auteurs. Nous pouvons d'ailleurs noter que le philosophe David CHALMERS semble se moquer de cette idée dans la mesure où il avance que vouloir lier la mécanique quantique au processus de la conscience est un acte de *réduction des mystères* [Bis02]. Le raisonnement serait le suivant : la mécanique quantique est mystérieuse $\Rightarrow$ la conscience est mystérieuse $\Rightarrow$ le mystère doit être le même.

Une approche de la Physique, vis-à-vis du problème du Libre Arbitre, nous semble plus intéressante et plus solide que celle que nous avons vu jusqu'ici. Cette approche tire parti de la possible émergence de certaines propriétés physiques dans des systèmes hors équilibre comportant un certain degré de complexité. Les propriétés d'un tel système peuvent être résumées en quelques points [Bis02] :

1. Un grand nombre de particules
2. Haut degré de structuration et d'ordre
3. Comportement collectif
4. Irréversibilité des équations
5. Propriétés émergentes

Des propriétés émergentes seraient donc susceptibles d'apparaître d'un système possédant ces caractéristiques. Typiquement, le cerveau pourrait répondre à tous ces critères et, nous l'avons évoqué précédemment, nous pensons que l'étude du cerveau est ce qu'il y a de plus adéquate pour éclairer scientifiquement la question du Libre Arbitre. Il en découle qu'il nous apparaît que c'est par cette voie que la Physique pourrait apporter de réels arguments aux diverses thèses concernant la question. Il nous semble cependant que les conséquences de ces propriétés, si elles existent, devraient être observables par les Neurosciences. C'est en cela que nous pensons que quelque soit la réponse proposée, existence d'un processus émergent ou non, c'est le moteur du choix qui doit être étudié et ce dans son ensemble. Pour cette raison, ce sont les Neurosciences qui seront, à partir d'ici, le sujet de notre étude et ce jusqu'à la fin de ce mémoire.

Chapitre 2

Action volontaire et expériences neuroscientifiques

2.1 Introduction

Le problème du Libre Arbitre recouvre beaucoup d'aspects et les problématiques qui lui sont associées sont nombreuses. Par exemple citons le thème de la responsabilité, celui du problème du sentiment d'agentivité, ou encore, comme nous en avons parlé au premier chapitre, le déterminisme ontologique des lois de la nature. Dans ce mémoire, nous souhaiterions nous pencher sur un autre de ces aspects. Il s'agit du problème de l'action volontaire libre. Il nous semble qu'il s'agit d'un questionnement fondamental pour toutes les autres problématiques relatives au Libre Arbitre. En effet, que peut-être notre conception des autres caractéristiques de ce dernier, si nous n'avons pas une idée claire quant au rapport qu'entretiennent action, volonté et liberté ? Le problème de l'action volontaire libre est donc aussi important dans la mesure où, étant étudié par les sciences, à travers lui, il nous sera possible d'analyser le rapport que les sciences entretiennent avec la question philosophique plus générale du Libre Arbitre.

Justement, dans le chapitre précédent, nous avons émis l'idée que les sciences contemporaines pourraient être utiles à cadrer ce problème. Il arrive donc une question : en plus de la philosophie, quel domaine des connaissances serait le plus adéquat pour donner des informations concernant la problématique relative à l'action volontaire libre ? Nous avons déjà montré que les apports de la Physique se concentrent sur la question du déterminisme des lois de la nature. Si nous voulons aborder le problème du rapport entre action, volonté et liberté, il va donc falloir traiter la question par un autre axe que celui de la Physique. Même si dans ce cas, il ne parle que de la thèse Libertarienne, les propos de Michaël ESFELD confirment cette idée. Il écrit :

« Il est sans importance ici de savoir si la causalité physique est déterministe ou probabiliste. Si l'explication de la volonté libre que propose le libertarianisme est correcte, les actions libres sont également incompatibles avec les lois probabilistes physiques, et pas uniquement avec les lois déterministes. En effet, les probabilités objectives que ces lois déterminent ne peuvent pas être influencées par un sujet agissant [...] » [Esf12]

Nous voyons donc que sur ce terrain au moins, les conclusions physiques ne nous avanceront guère concernant notre problématique. C'est le point de départ que nous prendrons et, nous le verrons, il

permet de tirer des conclusions ayant des répercussions importantes concernant la conception que l'on se fait de l'action volontaire.

Il semble que l'action volontaire d'un agent, pour être qualifiée de libre, doit répondre à au moins une condition : il faut qu'un état mental conscient, endogènement construit, soit la source de l'action impactant le monde physique. Citons quelques contres exemples : l'épiphénoménalisme (dont nous allons longuement parler par la suite), les actions automatiques, les actions réflexes qui ne peuvent être qualifiées d'actions volontaires. Les Neurosciences semblent parfaitement aptes à fournir des pistes de réflexions intéressantes à ce sujet. Ces dernières font une dichotomie pour distinguer entre actions auto-initiées, et actions déclenchées par un stimulus extérieur [Hag08]. Le premier type serait donc l'exemple d'une action produite par un individu et lui seul de façon consciente et endogène (via une délibération). Le second type serait un exemple d'une action déclenchée de façon exogène. Cette distinction est essentiellement la définition d'une action volontaire pour les Neurosciences. Cette définition a le mérite d'être simple et aussi d'être fructueuse du point de vue de la compréhension que l'on se fait du phénomène. En effet il nous semble que les données empiriques que les Neurosciences peuvent produire, concernant les processus cérébraux en jeu dans ces deux types d'actions, peut éclairer (voir contraindre) les réflexions philosophiques concernant l'action volontaire libre.

Dans un premier temps, nous allons remonter aux éléments de bases qui ont permis aux Neurosciences expérimentales de mesurer certaines caractéristiques de l'action volontaire libre définie comme précédemment. Ces éléments sont constitués d'une grandeur à mesurer ; *le potentiel de préparation motrice*, et d'une méthode de mesure ; *la chronométrie mentale*. Dans un deuxième temps, nous nous intéresserons à l'origine de l'action. Nous illustrerons notre propos par la description de protocoles de plusieurs types d'expériences contemporaines. Une fois ces protocoles décrits, et leurs conclusions exposées, nous serons mieux armés pour discuter de ce qui nous semble être un apport capital des Neurosciences pour dessiner les contours envisageables des positions philosophiques concernant l'action volontaire. Nous le ferons dans le troisième et dernier chapitre de ce mémoire.

2.2 Mesurer les processus mentaux

Le but de cette section est simplement d'introduire l'importance prise par la mesure scientifique dans l'étude des états mentaux. Rappelons que notre souhait est de traiter de l'action volontaire et, selon la définition que nous avons admise, cette action se doit d'être consciente et endogène à l'individu. En étudiant les propriétés du cerveau nous remonterons donc à la racine de ce qui produit ce type d'action.

Dans le cadre qui nous intéresse dans ce mémoire, une grandeur peut-être clairement mesurée. Cette grandeur, obtenue de manière non invasive par les sciences cognitives, c'est *le potentiel de préparation moteur* (readiness-potential). Mais, avant d'établir l'origine de ce potentiel, nous allons décrire une technique fondamentale en psychologie cognitive : *la technique de la chronométrie mentale*.

2.2.1 La chronometrie mentale

Historiquement, l'idée de ce type de chronométrie vient en premier lieu de la prise de conscience, établie au XIX^{ème} siècle, que les signaux circulant dans les nerfs ne se propagent pas instantané-

ment. C'est par les travaux de Hermann Ludwig Ferdinand von HELMOLTZ, où ce dernier montre qu'il est possible de mesurer une vitesse de conduction nerveuse [Hel50, Sch02], que cette prise de conscience a été initiée. L'existence de cette vitesse de propagation laisse immédiatement présager d'une chose : quel que soit le processus menant à une action, ce dernier va mettre du temps à s'accomplir. Il est donc naturel de penser qu'une mesure de durée peut contribuer à la compréhension des processus en oeuvre dans une action.

En 1868, le néerlandais Francissus Cornelis DONDERS effectue une première mesure des opérations mentales [Don69]¹. Suite à ses prédécesseurs il propose qu'un processus mental, enclenché par un stimulus sensitif, a lui aussi un temps d'exécution. Il va soustraire le temps nécessaire pour une réaction à un stimulus unique, au temps nécessaire pour la même réaction mais avec deux stimuli distincts. Se faisant, il est en mesure de donner le temps pris par l'esprit pour discriminer entre deux stimuli et l'action à effectuer. Par mesure du temps, il mesure donc une information relative à un processus mental précédant une action.

La chronométrie mentale est massivement utilisée. Dans un article, daté de 2005, le psychologue américain Michael I. POSNER écrit :

« Many cognitive and emotional tasks studied with neuroimaging have implicated a small number of brain areas that are consistently active. Mental chronometry plays a role in suggesting the cognitive operations that each of these areas performs and how they are organized in real time. » [Pos05]

Comme l'avait montré DONDERS au XIX^{ème} siècle, les processus mentaux, qui mènent à des actions, suivent des délais de propagation. Etudier des temps de latence mis en jeu dans un processus mental doit donc être en mesure de nous renseigner sur son fonctionnement. Dans la conclusion de son article POSNER écrit :

« In describing the links between brain and behavior, mental chronometry is still a cornerstone that binds psychology to the techniques of neuroscience. » [Pos05]

Nous comprenons donc bien en quoi cette méthode est centrale dans l'étude de l'action volontaire. Cette dernière, par simple mesure de temps, va permettre de mettre en évidence certaines caractéristiques des processus mentaux associés à une action. En particulier, ces mesures vont montrer que ces processus mentaux sont souvent dotés d'une part fortement inconsciente avant la réalisation d'une action. Ce type de résultat est à la base de la production d'un doute concernant la conception, pourtant souvent admise de tous, qu'une action volontaire est produite par une forme de conscience des individus.

Dans la suite, nous allons passer en revue quelques expériences de Neurosciences relatives à l'idée d'action volontaire. La publication de ces expériences s'étale sur une durée comprise entre 1983 et 2018 avec un pas de temps de plus ou moins une dizaine d'années. Nous verrons alors ce qui est susceptible (ou non) de déclencher l'action volontaire. Mais, avant cela, décrivons ce que les neuroscientifiques ont l'habitude de mesurer dans leurs études concernant ce type d'action.

1. Article en langue originale : [Don69]

2.2.2 Le potentiel de préparation moteur (readiness-potential)

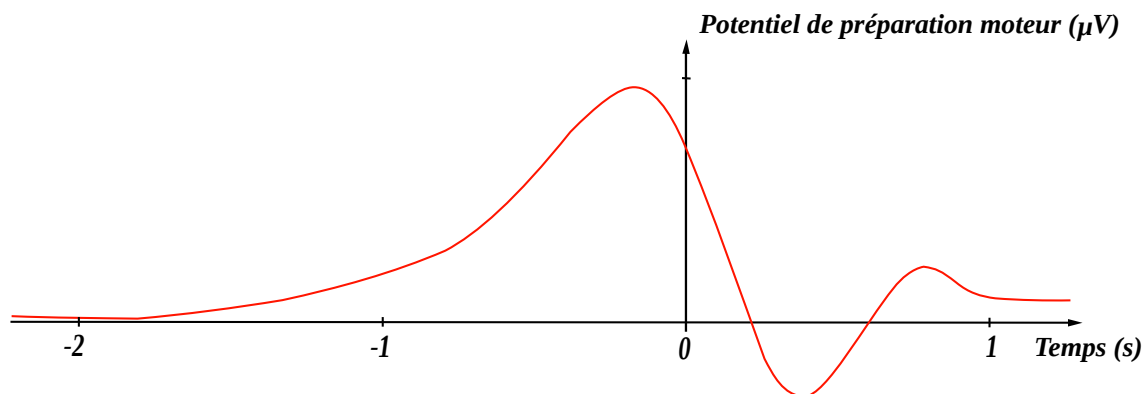


FIGURE 2.1 – Schéma, très idéalisé, d'un potentiel de préparation moteur.

Commençons par une rétro-ingénierie de l'action². D'où vient l'action ? Elle est la résultante d'une contraction musculaire stimulée par des nerfs dont l'activité est elle-même déclenchée par le cerveau. On parle alors d'activité prémoteur. Cette activité a lieu principalement dans une zone qui, dans la classification de Brodmann, correspond à l'aire 6 du cerveau³. Le potentiel électrique du cerveau est alors conduit en dernière instance vers la zone du cortex moteur primaire ($M1$ dans l'aire 4 de Brodmann). Partant de $M1$ le signal conduira, via la colonne vertébrale, à la contraction des muscles induisant l'action [Hag08].

Le potentiel de préparation moteur est un potentiel lié à l'activité cérébrale en phase de préparation d'une action. Ce potentiel a été découvert dans les années 60 par Hans Helmut KORNHUBER et Lüder DEECKE [KD65, SH06]. Cette activité cérébrale est susceptible d'être mesurée et enregistrée via un procédé d'encéphalographie (EEG). Des capteurs sont placés sur le crâne d'un sujet et ceux-ci mesurent l'activité électrique du cerveau alors que ce dernier exécute une tâche particulière.

Dans le cas d'une action physique volontaire, comme appuyer sur un bouton ou fléchir un doigt quand le sujet le souhaite, un potentiel va donc apparaître. Ce potentiel est mesuré en microvolt μV dans le temps. L'axe des abscisses de la représentation de ce signal est le temps. Par convention l'axe des ordonnées montre les valeurs négatives du potentiel en les orientant vers le haut. Nous obtenons alors un signal qui peut être schématisé à la manière de la Fig. 2.1.

2.3 L'origine de l'action volontaire : études scientifiques

Dans cette section nous allons illustrer, par quelques exemples, de quelle façon les Neurosciences s'emparent de la question de l'action volontaire. Nous parlerons du protocole de ces expériences et de leurs résultats. La première expérience que nous évoquerons n'est autre que l'expérience princeps de l'étude scientifique de l'action volontaire. Il s'agit d'un travail réalisé par Benjamin LIBET et ses collaborateurs [LGWP83].

2. Propos de Patrick HAGGARD, lors de la remise du Prix Jean-Nicod le mardi 17 mai 2016 : <http://www.institutnicod.org/seminaires-colloques/conferences-et-prix-jean-nicod/archives/article/prix-et-conferences-jean-nicod-1378>

3. *Ibid.*

2.3.1 L'expérience de LIBET *et al.* (1983)

Travaux publiés en 1983, l'expérience de Benjamin LIBET *et al.* constituent le socle de l'approche scientifique de la question de l'action volontaire. Cette expérience, au premier abord, donne lieu à des conclusions surprenantes.

Protocole et résultats

L'idée du travail de LIBET est de traiter expérimentalement la question de l'action volontaire. En effet, dans cette expérience, le but est de mesurer l'activité neuronale relative à ce type d'action. En particulier, l'auteur et son équipe s'intéressent à l'importance de la volonté consciente d'effectuer une action dans ce processus. Les mots importants dans cet article sont la formulation “urge to move”⁴. C'est le moment de conscience de “l'impériosité de la volonté de mouvement” que l'on va demander aux participants de noter. Quand ont-ils conscience de cette volonté ? Comme notre expérience quotidienne nous le ferait penser, la conscience de l'envie de bouger précède-t-elle vraiment le départ de l'activité neuronale menant au mouvement ? En est-elle la cause directe ? Quand, par rapport à l'activité cérébrale, cette conscience apparaît-elle ?

Nous l'avons vu, du fait de la conduction nerveuse, une différence de temps entre le désir de mouvement, activité cérébrale, et mouvement réel apparaît. Notre idée première, issue de l'expérience quotidienne, mais aussi selon nous d'une forme de pensée dualiste, nous incline à penser que le processus de l'action volontaire suit le schéma suivant : un état mental de désir de mouvement est produit par l'agent (de manière *ex nihilo* ?) $\Rightarrow$ il provoque une mise en fonctionnement du moteur cérébral $\Rightarrow$ ce dernier transmet un signal aux muscles pour exécuter l'action. Nous comprenons que tout ce processus ne peut pas être instantané, cependant, le laps de temps révélé par LIBET *et al.* va être complètement contre intuitif. Il va révéler un écart important entre le temps du début de l'activité cérébrale et le temps de conscience de l'envie de bouger elle même. On a donc un renversement des deux premiers termes de la chaîne des événements. Comment LIBET va-t-il expérimenter cela ? Grâce à la quantité physique et à la technique présentées dans la section précédente, respectivement, le potentiel de préparation moteur et la chronométrie mentale.

Le dispositif expérimental était composé d'une horloge et l'expérience commençait quand le participant se disait prêt. L'instruction donnée au sujet était simplement de fléchir un doigt quand il le souhaitait. Il devait alors relever, en mémorisant la position de l'aiguille de l'horloge, le moment où est apparu son désir de bouger ; le “urge to move”. Trois types de mesures furent réalisées lors de session de test différentes :

Type (M) : Il est demandé aux participants de mémoriser le temps où ils ont conscience de bouger réellement.

Type (S) : Il est demandé aux participants de mémoriser le temps où ils ressentent un stimulus. Ce dernier était aléatoirement produit par l'expérience, à l'arrière de la main. Ce type sert de référence pour corriger d'éventuels biais.

4. A beaucoup d'endroits dans ce chapitre nous rencontrerons de nouveau cet anglicisme couramment utilisé par les neuroscientifiques

Type (W) : Il est demandé aux participants de mémoriser le temps où ils ont conscience de leur envie impérieuse d'effectuer un mouvement ("urge to move").

Les temps, mémorisés par les participants, étaient rapportés par ces derniers quelques secondes après l'expérience. Deux types de retour de temps aux expérimentateurs furent utilisés lors de sessions différentes.

Type (A) : Il est demandé aux participants de donner le temps absolu qu'ils ont mémorisé à l'horloge. Dans ce cas, après l'expérience, l'horloge continuait à tourner sur un temps aléatoire pour éviter de fournir un indice.

Type (O) : Il est demandé aux participants de donner l'ordre de présence de l'aiguille par rapport à leur temps de conscience mémorisé. Dans ce cas, après l'expérience, l'aiguille de l'horloge se figeait à un endroit aléatoire. Le participant devait alors dire s'il estimait que l'aiguille était présente avant, après ou au même endroit que ce qu'il avait mémorisé.

Ces deux différentes méthodes permettaient d'avoir une meilleure confiance concernant les retours de temps que les participants donnaient.

Les résultats qui vont plus particulièrement nous intéresser sont ceux qui rapportent les temps de type (W). En effet, ce sont eux qui donnent une référence quant au temps de conscience de la volonté de produire une action spontanée. Ces temps sont, dans l'esprit commun, le moment de la prise de décision de bouger. Ils sont donc un marqueur important concernant l'action volontaire.

Durant l'expérience, deux types de mesures objectives étaient effectuées : un EEG et un électromyogramme EMG. L'encéphalogramme mesurait l'activité cérébrale, l'électromyogramme l'activité du muscle mit à contribution dans le mouvement. L'origine des temps qui fût choisie était le mouvement lui-même. Les séries d'expériences ont montré que les temps des expériences de type (S) diffèrent d'un individu à l'autre. Cet effet était d'ailleurs assez important pour qu'ils révèlent un biais relatif à chaque individu. Le temps des séries (W), quant à lui, correspondait en moyenne à un temps de 200 *ms* avant l'action. En corrigeant le biais révélé par les expériences de type (S), le temps moyen de conscience du déclenchement de l'action tombe à 150 *ms* avant sa réalisation. La conclusion est donc que les agents avaient, en moyenne, conscience de leur désir de mouvement ("urge to move") 150 *ms* avant le début de l'action.

LIBET et son équipe ont alors mis en rapport trois choses : le temps rapporté par les sujets, les mesures de l'encéphalogramme et les mesures de l'électromyogramme. Ce que les résultats semblent montrer, c'est que les temps de début du potentiel de préparation moteur précèdent de façon importante les temps de conscience des individus. La Fig. 2.2 récapitule, de manière schématique, les conclusions de ce travail expérimental. D'après cette étude il semblerait que l'activité cérébrale, source de l'action, démarre bien avant que le sujet ait conscience de son envie impérieuse d'exécuter cette action. Il est évident que ce type de résultats a une répercussion majeure sur la question philosophique de l'action volontaire.

Dans ce travail de recherche, la sensation "urge to move" jouait un rôle prépondérant. Voyons à présent une expérience réalisée 8 années plus tard qui va nous en dire un peu plus sur les zones semblant impliquées dans cette sensation.

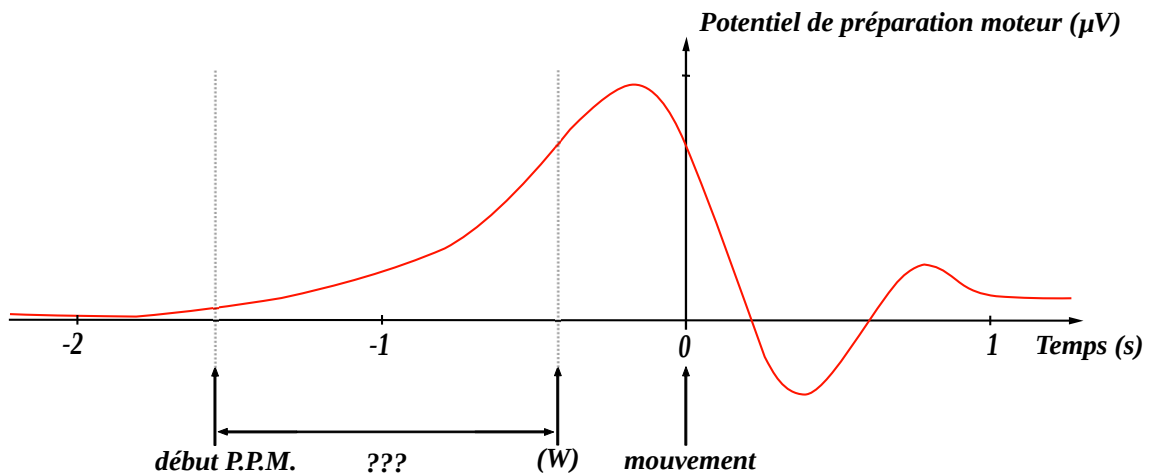


FIGURE 2.2 – Schéma montrant l'avance du potentiel de préparation moteur avant la conscience qu'une action va être débutée. Ici P.P.M. signifie potentiel de préparation moteur et (W) est le temps de conscience du sujet.

2.3.2 L'expérience de FRIED *et al.* (1991)

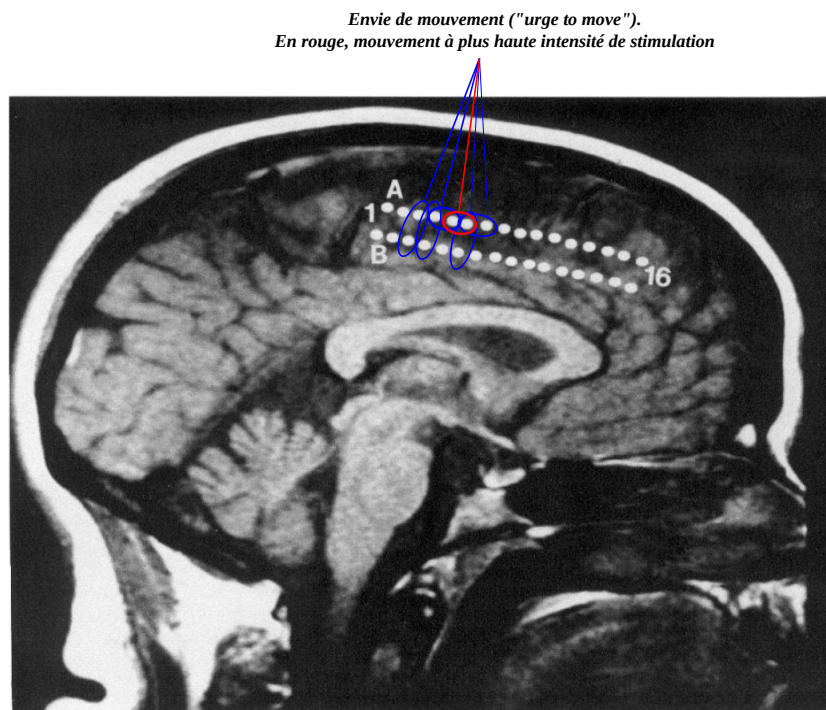


FIGURE 2.3 – Exemple d'électrodes placées au contact du cerveau d'un patient dans l'expérience de FRIED *et al.* Les marques bleues représentent les sites qui ont provoqués une sensation de "urge to move", la marque rouge montrant une zone où la stimulation a provoqué une sensation "urge to move" et à plus haute intensité, une réaction du corps à un endroit très proche de ce que prédisait la sensation.

Les travaux de Itzhak FRIED *et al.* [FKM⁺91] n'ont pas pour but premier l'étude de l'action volontaire. Cependant ses conclusions sont importantes vis-à-vis de cette problématique. Dans cette étude ce sont les fonctions d'une partie du cerveau, le SMA (Supplementary Motor Area) vis-à-vis du contrôle moteur du corps, qui sont étudiées par stimulations électriques. Les réactions des patients, possédant des électrodes implantées à la surface du cerveau, sont étudiées lorsque l'on fait varier la position et l'intensité de la stimulation.

Protocole et résultats

Dans cette expérience, les neuroscientifiques vont étudier la réaction moteur et sensitive de 13 patients soumis à des stimulations d'une partie de leurs cortex. Chaque sujet étudié portait une série d'électrodes directement en contact avec leur cerveau et, plus particulièrement, dans une zone comprenant le SMA. Les électrodes avaient été implantées chirurgicalement dans le cadre du traitement de l'épilepsie dont souffraient ces patients.

Comme le rappellent les auteurs dans leur article, un long historique existe quant à la stimulation électrique de cette zone, en particulier sur des sujets animaux non humains. Ce que ces précédentes expériences suggèrent, c'est que le SMA est une zone du cerveau très importante dans la préparation de l'action moteur. Le but des auteurs est donc d'essayer de décrypter l'organisation fonctionnelle, vis-à-vis des réactions du corps, de cette zone. Avec la méthode employée, qui consiste à stimuler le cerveau via des sortes de chapelets d'électrodes (voir la Fig. 2.3), les auteurs sont en mesure de cartographier l'organisation somatotopique du SMA.

La procédure était alors la suivante : on augmentait l'intensité du courant en parcourant chacune des électrodes ; le courant était contraint à subir des sauts réguliers de 0.5 ou 1 mA. Une fois une réponse comportementale obtenue, les auteurs arrêtaient d'augmenter le courant. Une limite était fixée à l'augmentation de ce courant à 10 mA. Il est important de noter que dans ces expériences, on s'intéressait aussi au ressenti des patients. En effet, ils étaient encouragés à témoigner des événements dont ils avaient la sensation au moment de la stimulation électrique. Lors de ces sessions, plusieurs types de réponses ont alors été notés :

Les réponses moteurs : il s'agissait de mouvement plus ou moins complexes que les auteurs ont classé en plusieurs catégories (voir [FKM⁺91]).

Les réponses de sensations subjectives : ces réponses allaient du simple picotement d'une partie du corps, à une sensation de mouvement d'un membre alors que ce dernier ne bougeait pas. Le point important pour notre étude et qu'il y eut aussi des sensations de "urge to move" similaires à celles que l'on demandait de noter aux participants de l'expérience de LIBET.

Cette sensation de "urge to move", ressentie quelque fois par les patients, est ce qui est le plus intéressant vis-à-vis de la question que l'on se pose dans ce mémoire. Ces événements étaient suivis d'un mouvement lorsque l'intensité du courant était augmentée. Ce mouvement ne se produisait pas toujours dans la zone concernée par l'envie impérieuse de mouvement rapporté par le patient. Il s'agissait cependant presque toujours d'une zone proche. Il est alors important de préciser que,

d'un point de vue expérimental, bien que les expériences de l'époque produisaient des résultats inédits, il semblait délicat d'être très précis quand aux zones stimulées.

Au final cette étude donne la preuve, comme dans les études passées faites sur les animaux non humains, de l'importance de l'activité de la zone SMA dans l'action. Ces recherches suggèrent donc que le SMA est partie prenante dans les volontés relatives à l'activité moteur. Il apparaît que cette étude est capitale à notre investigation de l'analyse neuroscientifique de l'action volontaire. Elle l'est a au moins deux titres : il semble qu'une zone soit ici clairement identifiée comme étant importante dans le déclenchement (ou le contrôle) d'une action, et également, il semble que cette zone soit aussi impliquée dans cette sensation de "urge to move" dont le timing a été étudié dans l'expérience de Benjamin LIBET.

Dans l'expérience suivante nous allons décrire des résultats qui ont provoqué, et provoquent encore, beaucoup de remous et de débats dans la communauté philosophique s'intéressant à l'action volontaire et à ses potentiels précurseurs. Cette expérience conduit ses auteurs à des conclusions philosophiques très fortes vis-à-vis d'une forme d'épiphénoménalisme des états mentaux sur les états physiques.

2.3.3 L'expérience de WEGNER et WHEATLEY (1999)

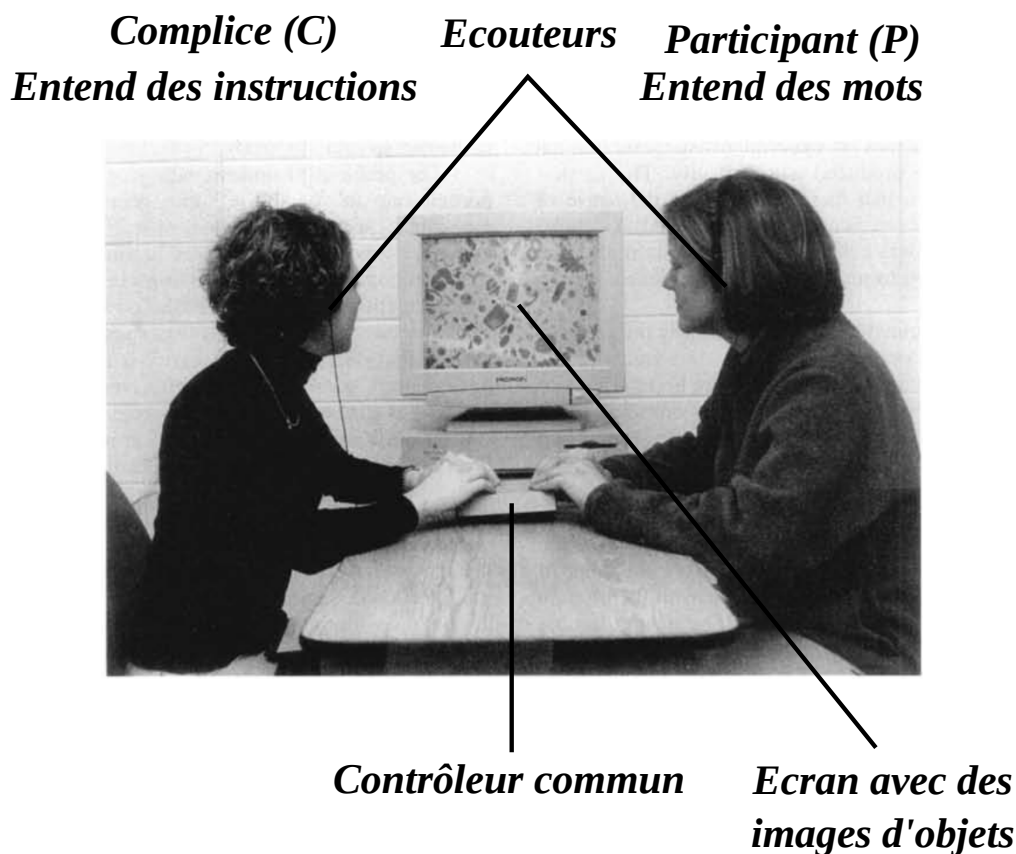


FIGURE 2.4 – Photographie présentant l'expérience réalisée par WEGNER et WHEATLEY en 1999.

Avec l'expérience de LIBET *et al.* nous avons vu qu'une idée, celle que la conscience soit une source de l'action, pouvait être mise en doute. Un autre protocole expérimental, réalisé par

WEGNER et WHEATLEY [MWW99], se propose de mettre à l'épreuve de l'expérience un modèle théorique. Ce modèle postule une relation de causalité fautive, induite par l'esprit, entre pensée et action dont nous verrons les détails au chapitre suivant.

Protocole et résultats

L'idée que la conscience implique l'action est très tenace dans nos esprits. WEGNER et WHEATLEY commencent leur article par cette constatation. Cette expérience, réalisée en 1999, montre qu'il semble possible d'induire l'idée à un agent que c'est son intention qui a causé une action alors que ce n'est de toute évidence pas le cas. Ici, nous employons pour la première fois le terme "intention". Il est à prendre au sens commun, en effet, nous verrons dans le chapitre suivant que cette notion a une définition plus précise. Dans cette expérience, il ne sera pas question de mesures neuroscientifiques à l'aide d'encéphalogrammes par exemple. Les auteurs ont étudié un comportement et le sentiment relatif à ce comportement. Nous discuterons, dans le chapitre suivant, en quoi cette expérience étaye un modèle bien particulier concernant l'origine de l'action volontaire.

Dans cette expérience, pour chaque session, deux personnes sont choisies. L'une d'elle sera le véritable participant, appelons la *P*, l'autre, *C*, sera un complice des expérimentateurs. *P* et *C* sont placés devant un écran d'ordinateur où figurent cinquantes petites images d'objets (l'expérience prendra le nom de "I Spy experiment" en référence à une suite de livres pour enfants comprenant des pages avec des collections de photos d'objets). Via un contrôleur commun (une souris d'ordinateur), *P* et *C* peuvent faire bouger un pointeur sur l'écran. Toutes les trente secondes environ, il est demandé de stopper le pointeur sur une des images de l'écran. Après cet arrêt il est attendu que chacun, individuellement, note sur une règle graduée personnelle le degré d'intention placé dans l'action : "arrêt du pointeur". La graduation allait de "j'ai laissé décider de l'arrêt" à "j'ai initié l'arrêt". Les résultats sont ensuite convertis en pourcentage d'intention. Dans ce protocole on dit aux deux agents de l'expérience que, dans les écouteurs qu'ils portent, deux types de son seront entendus.

Type 1 : Une musique de 10 secondes signalant qu'un arrêt devait être décidé.

Type 2 : Des mots correspondants (ou non) à des images figurant sur l'écran. Par exemple : voiture, dinosaure, singe...

En réalité, *C* n'entend ni musique ni mots. Il va entendre dans ses écouteurs les instructions de mouvements donnés par les expérimentateurs. Une photographie illustre le moment du test de cette expérience sur la Fig. 2.4.

Toutes les trente secondes, un arrêt est donc effectué. La plupart du temps, il est demandé à *C* de laisser stopper *P*. Cela permettra d'effectuer un contrôle sur la corrélation entre les mots entendus par *P* et sa position d'arrêt. L'idée étant de constater si *P* a plus tendance à stopper sur des images en cohérence avec les mots entendus ou non. De temps en temps, il est demandé à *C* de forcer l'arrêt sur une image. C'est le coeur de l'expérience. En effet, quand l'arrêt est forcé par *C*, trois laps de temps sont décidés avant de faire entendre un mot à *P*. Ce mot était entendu quelques secondes avant ou après que *C* ne stoppe sur l'image évoquée par les expérimentateurs et

P entendait alors le mot correspondant à cette image. Les timing étaient les suivants : 30 s avant, 5 s avant, 1 s avant ou 1 s après. Le dessein derrière ces temps est d'induire, à différent moment, à *P* une pensée qui reflète l'action en train de se produire. Ce timing vise donc à étudier à partir de quel moment l'esprit associe cette pensée par un lien de causalité à l'action en train de se produire.

Comme pour les expériences (*S*) de LIBET, le fait de laisser *P* s'arrêter à sa guise la plupart du temps jouait le rôle de contrôle. Il est évident que s'il y avait une corrélation majeure entre le mot entendu et l'image d'arrêt, alors cet effet jouerait aussi un rôle dans les cas des arrêts forcés. En quelque sorte, dans ce cas, *P* contribuerait à l'arrêt forcé par *C*. WEGNER et WEATLEY montrent qu'il n'en est aucunement le cas. Il n'y a aucune corrélation entre le fait que *P* entende un mot et l'objet sur lequel il va stopper le curseur. On peut ainsi en conclure que l'expérience est sauve de ce type de biais.

C'est alors qu'intervient le résultat capital de l'étude. Lorsque l'on dépouille les données, dans les cas d'arrêts forcés, on voit clairement que *P* tend à estimer une intention de s'arrêter sur l'objet. Pourtant, nous avons dit qu'il n'avait aucun contrôle sur l'événement se produisant devant ses yeux. Par l'émission d'un mot en cohérence avec l'objet qui allait être atteint par le pointeur, il semblerait que l'on ait simplement incité son esprit à établir un lien de causalité entre l'action et sa pensée. Un autre résultat important est le timing de ce processus. Les données montrent clairement que ce sentiment particulier est accru si le mot est prononcé à un temps suffisamment proche du temps d'arrêt. Si le mot est prononcé 30 secondes avant l'arrêt ou 1 seconde après, le sentiment d'intention semble perdu. On a ainsi une idée de l'ordre de grandeur du temps où une pensée est susceptible d'être associée causalement à une action par l'esprit. Ce temps est de l'ordre de la seconde avant l'action. C'est d'ailleurs cet ordre de grandeur que les auteurs estimaient déjà par d'autres considérations.

L'expérience présentée ici semble rendre compte d'une position forte. Son idée est qu'il existe une forme d'épiphénoménisme des pensées sur les actions. Ainsi, le monde psychique conscient n'aurait pas d'influence sur le monde physique. Via son dispositif, l'expérience évoquée tend à mettre à l'épreuve cette idée. Ses résultats sont concluants dans la mesure où ils montrent clairement qu'une personne, dont on a initié l'idée, a tendance à associer cette idée à l'action qu'il voit devant ses yeux. C'est alors un lien de causalité, entre idée et action, qui est créé de toute pièce par l'esprit *a posteriori*. On pourrait schématiser ce lien par la phrase : "si mon idée concorde en temps, en cohérence, et en exclusivité à l'action physique que je vois, c'est qu'elle en est la cause première, c'est que j'en suis l'auteur volontaire". Les trois concepts relatifs au timing, à la cohérence et à l'exclusivité, sont les trois piliers du modèle développé par WEGNER et WEATLEY. Nous en discuterons plus tard.

L'expérience que nous avons à présent choisi d'évoquer a eu lieu en 2008 et se base sur une autre technique, celle de l'analyse de données d'IRM fonctionnelle. Cette analyse mettrait en évidence une activité du cerveau prédisant l'action non pas quelques centaines de millisecondes avant l'action mais plusieurs secondes avant celle-ci.

2.3.4 L'expérience de SOON *et al.* (2008)

Ce travail est une sorte d'amélioration de l'expérience de LIBET qui va s'affranchir de certaines des critiques qui lui ont été portées. Les résultats de cette étude, réalisée par SOON *et al.* [SBHH08], publiés en 2008, vont suggérer que les informations encodées dans l'activité électrique du cerveau

permettent de prédire le choix d'un sujet bien avant que celui-ci n'en ait conscience.

Protocole et résultats

Pour commencer, les auteurs notent trois problèmes dans les études précédentes, surtout celles qui se sont servies du potentiel de préparation moteur pour évoquer la possibilité que ce dernier apparaisse avant la conscience d'avoir envie de produire une action.

Problème 1 : le p.p.m. vient du SMA, il arrive très tard dans le processus moteur.

Problème 2 : en lien avec le problème précédent, la mesure du p.p.m. arrive assez tôt avant le temps de conscience rapporté par les agents. Il pourrait y avoir des erreurs de mesures mettant en doute l'ordre : début p.p.m. $\Rightarrow$ (W).

Problème 3 : le p.p.m, mesuré par LIBET, ne mesure *a priori* pas une action spécifique. Il faudrait alors mieux étudier une situation de choix.

Pour faire face à ces problèmes, les auteurs ont choisi d'utiliser l'IRM fonctionnelle qui leur permet de scanner toute l'activité du cerveau avant un choix. Couplé, après coup, à un algorithme de reconnaissance de schémas neuronaux ils ont ainsi tenté de prédire les choix que les agents allaient faire.

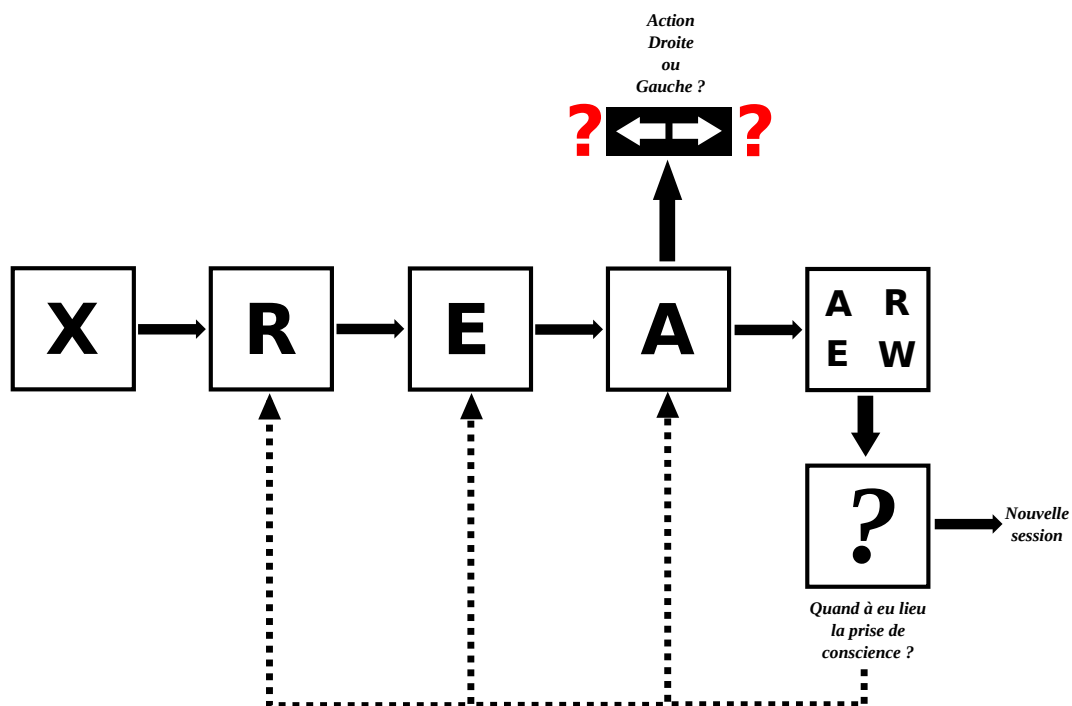


FIGURE 2.5 – Schéma représentant l'expérience réalisée par SOON *et al.* Les agents devaient réaliser une tâche moteur pendant qu'ils subissaient une IRM fonctionnelle.

Le schéma Fig. 2.5 résume le protocole utilisé. Des agents, placés dans une IRM fonctionnelle voyaient une succession de lettres sur un écran. Ces lettres changeaient tous les 500 ms. Les agents devaient appuyer sur un bouton à droite ou à gauche quand ils le souhaitaient. Il était aussi demandé de se rappeler la lettre au moment où leur décision vis-à-vis du choix apparaissait. Ici, le défilement des lettres jouait le rôle de l'horloge présente dans l'expérience de LIBET. Après avoir appuyé, l'écran affichait des lettres et l'agent devait choisir celle qu'il avait retenu au moment de la prise de conscience de sa décision. Après cela, le processus recommençait pour un autre essai. L'expérience ainsi faite a 3 avantages :

1. Elle permet de regarder les déterminants de l'action à long terme.
2. Elle permet d'étudier toutes les régions du cerveau en activité et non plus seulement l'activité du SMA.
3. Elle permet de voir une activité spécifique (ici il y a un choix) plutôt que de refléter un processus de préparation potentiellement non spécifique.

Les résultats de l'expérience montrent qu'il y a deux régions qui semblent majoritairement impliquées dans la réalisation d'un choix simple :

1. Le cortex fronto-polaire. Il montre une activité jusqu'à 10 secondes avant la décision.
2. Le cortex pariétal.

L'algorithme utilisé par les auteurs montre que 10 secondes avant l'action, avec 60% de précision, il est possible de prédire le choix qui allait être fait. Bien que faible, la prédiction est meilleure que le hasard, elle semble donc significative. Qui plus est, il a été possible de prédire le départ de l'action 5 secondes avant la décision motrice.

Suite à ces résultats les auteurs font la suggestion que, quand la décision du sujet atteint la conscience, celle-ci a été influencée par une activité cérébrale inconsciente pendant près de 10 secondes. L'étude va donc plus loin que les travaux précédents. Elle montre que les premières informations prédictives d'une action volontaire simple sont codées dans des régions spécifiques du cortex frontal et pariétal et non pas uniquement dans le SMA. Les informations recueillies par l'expérience suggèrent aussi que les décisions d'actions ont des premiers précurseurs inconscients dans le cortex fronto-polaire. Ces précurseurs influenceraient l'activité dans le cortex pariétal et, en dernière instance, l'activité du SMA.

La prochaine expérience que nous allons détailler sera la dernière. Cette expérience a le mérite de grandement améliorer la prise en considération de la différence entre actions auto-initiées et actions stimulées par un processus extérieur.

2.3.5 L'expérience de KHALIGHINEJAD *et al.* (2018)

Les résultats d'une expérience, réalisée par KHALIGHINEJAD *et al.*, ont été publiés en janvier 2018 [KSD⁺18]. Il s'agit du protocole le plus récent dont nous parlerons, ce dernier reprend les éléments d'un dispositif qui avait déjà été mis à l'épreuve sur des animaux non humains [GS07]. Il s'agit alors de comparer l'activité cérébrale entre une action auto-initiée par un agent et le même type d'action mais cette fois déclenché par un stimulus extérieur.

Protocole et résultats

Depuis le début de ce chapitre, deux types d'actions furent dessinés. Ces deux types d'actions sont utiles à la définition scientifique de l'action volontaire.

Les actions auto-initiées : elles sont le résultat d'un processus purement endogène.

Les actions stimulées : elles sont une réponse à un stimulus et ont donc une origine exogène.

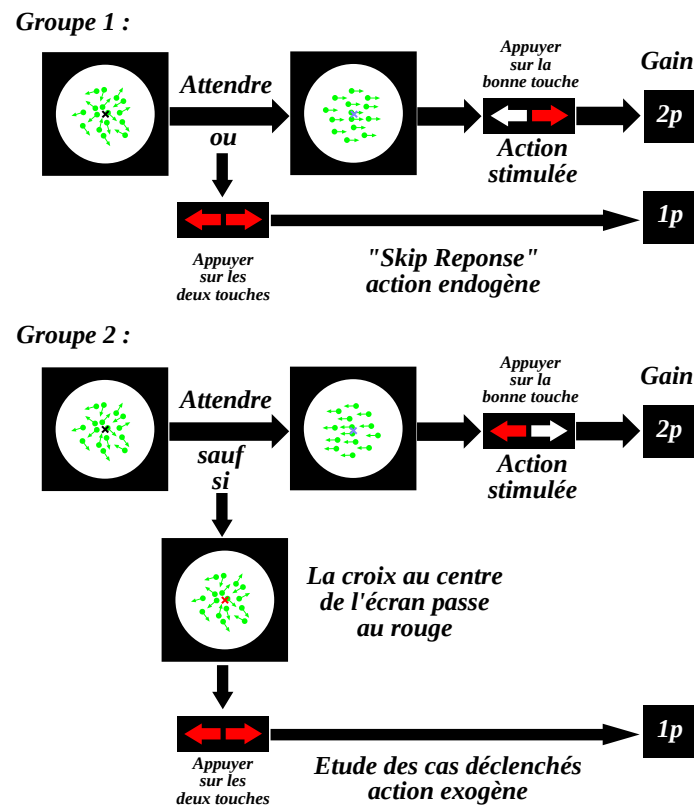


FIGURE 2.6 – Illustration du protocole utilisé par Khalighinejad *et al.* Les deux types de session effectués par les deux groupes y sont représentés. Le groupe 1 correspond à l'étude des cas d'action auto-initiées. Le groupe 2, placé exactement dans les mêmes conditions d'expérience, à l'exception que pour ceux-ci la "skip response" était remplacée par une réponse déclenchée par un stimuli.

Dans le cadre de notre étude, c'est évidemment le premier type d'action qui nous intéresse. C'est aussi ce type d'action qui étaient censé être étudié dans l'expérience de LIBET. Ces actions font naître un potentiel de préparation moteur qui, dans les conclusion de LIBET, semblait être la source inconsciente d'une action endogène. Vraiment endogène? A cause du protocole de LIBET, qui implique une instruction, c'est toute la question et nous y reviendrons au chapitre suivant. Ici, le travail que nous allons décrire ne va pas mettre en doute ce point.

Dans cette expérience, on ne donne plus d'instruction du type "appuyez quand vous le voulez" aux participants. A la place, on leur laisse le "choix du saut" à un essai suivant. C'est la méthode dite de la "skip response". Un schéma, la Fig. 2.6, résume le protocole de cette expérience. Les participants portaient un casque EEG afin de mesurer leur activité cérébrale et ils faisaient face à un écran sur lequel figurait une croix qui changeait de couleur aléatoirement. Sur l'écran, était aussi visible un nuage de points qui se déplaçait de manière aléatoire (tout en restant concentré dans un cercle). A un instant, tous les points se mettaient à avoir un mouvement cohérent à droite ou à gauche. Le participant devait alors appuyer sur les touches d'un clavier. Soit il appuyait sur la flèche droite soit sur la flèche gauche en fonction de ce qu'il voyait. Le participant gagnait 2 pence si la réponse était correcte et perdait 1 penny si la réponse était mauvaise (ou s'il répondait trop tôt ou trop tard).

Lors de cette étude, 2 groupes furent formés. Il était question d'étudier 2 cas : le cas des actions auto-initiées et le cas des actions stimulées. Pour ces deux groupes, 10 essais étaient effectués et un temps limite d'une heure à l'expérience était fixé. On remarquera que, pour les deux groupes, les expériences furent réalisées exactement dans les mêmes conditions. La seule différence entre les deux groupes était alors la suivante :

Groupe 1 (les actions auto-initiées) : les participants peuvent, s'ils trouvent que le temps d'attente est trop long avant de voir un mouvement cohérent des points, appuyer sur les touches droite et gauche en même temps pour passer à l'essai suivant. Il est important de noter que dans ce cas ils ne gagnent que 1 penny. Cette action est alors un pur produit interne. Les participants devaient peser le pour et le contre entre un long temps d'attente ou un plus faible gain.

Groupe 2 (les actions stimulées) : les participants devaient appuyer sur les deux touches en même temps uniquement si la croix au centre de l'écran passait au rouge. Cette action est alors déclenchée par un stimulus externe.

Le potentiel de préparation moteur est ce qui va être étudié en priorité et, plus particulièrement, la variabilité d'un grand nombre de mesures de p.p.m va être passée au crible. Depuis déjà plusieurs années avant cette expérience, il avait été montré que la forme typique du potentiel de préparation moteur pouvait être obtenue par simulation via un modèle mathématique stochastique. Ces modèles portent le nom de "modèle d'accumulateur" [OF02, SSD12, Sch18] et semblent pouvoir simuler, et même prédire, les temps d'actions des sujets. Sur la même base KHALIGHINEJAD *et al.* vont eux aussi proposer un modèle mathématique capable de simuler le comportement électrique capté par leur EEG. Dans cette idée, le potentiel d'action n'est pas un signal fixe du cerveau mais plutôt la résultante d'une moyenne du bruit neuronal qui est un signal électrique aléatoire de

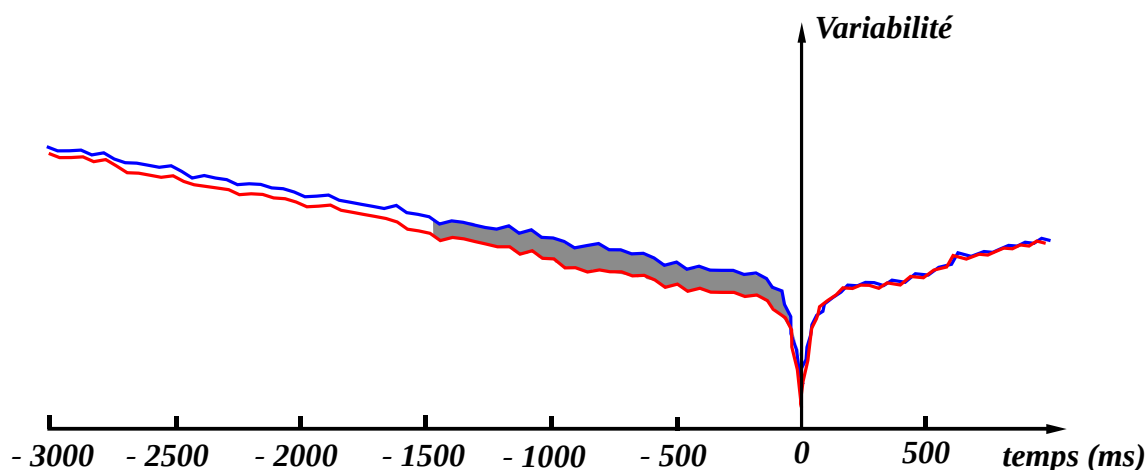


FIGURE 2.7 – Représentation schématique de la différence de variabilité des signaux EEG dans le cas d’une action déclenchée de l’extérieur (en bleu) et EEG dans le cas d’une action auto-initiée (en rouge). Ici, c’est la variabilité du potentiel de préparation moteur qui est étudiée. La zone grise marque la différence dont est responsable un signal spécifique aux actions auto-initiées.

l’activité cérébrale. Dans le modèle mathématique, les auteurs ont légèrement modifié le modèle de base pour avoir un paramètre de contrôle sur la valeur du bruit. Leur objectif est de rendre compte, par la simulation mathématique, d’une différence qu’ils ont remarquée dans leurs données expérimentales. Cette différence apparaît entre les EEG résultant d’une action auto-initiée et ceux résultant d’une action stimulée. Comme le montre la Fig. 2.7 il est clair que l’EEG relatif au potentiel de préparation moteur, issue d’une action auto-initiée, semble voir sa variabilité diminuer plus rapidement que la variabilité issue d’une action déclenchée de l’extérieur. L’hypothèse des auteurs est que cette diminution plus rapide dans le premier cas serait liée à une diminution du bruit neuronal, c’est-à-dire une forme de régularisation de l’activité électrique du cerveau. Le signal réduisant le bruit neuronal serait alors susceptible d’être le déclencheur de l’action volontaire. Cette idée est aussi confirmée par le modèle mathématique.

Bien que les auteurs discutent de nombreux autres points techniques, en particulier, de la cohérence de leur idée et de la fiabilité de celle-ci, il serait inutile de détailler ces points ici. Le résultat qui est vraiment le pivot central de ce travail est la différence de variabilité, certes faible mais non négligeable, entre les EEG de p.p.m issus d’une action auto-initiée et les EEG de p.p.m issus d’une action déclenchée. Ce que montre le modèle mathématique c’est que cette différence de variabilité pourrait bien être induite par une diminution du bruit neuronal. Un signal caractéristique de la volonté d’action semble ici se manifester. Ce signal est appelé par Patrick HAGGARD *volitional signal*.

2.4 Conclusion

L’objectif de ce chapitre était de fournir les bases nécessaires pour comprendre l’implication des Neurosciences vis-à-vis de la question du Libre Arbitre. Nous nous sommes alors centrés sur un aspect très étudié par les Neurosciences qui est celui de l’action volontaire. Dans ce mémoire,

nous souhaitons argumenter qu'il nous semble que ces sciences contemporaines sont aujourd'hui incontournables pour traiter de cette question philosophique. Il n'en reste pas moins qu'il faut être prudent et en éviter des conclusions trop hâtives.

Dans ce chapitre, nous avons rappelé que là où la Physique ne fournissait que des arguments en faveur (ou en défaveur) du déterminisme ontologique des lois de la nature, les Neurosciences qui travaillaient au problème de l'action volontaire, étaient plus aptes à donner des arguments en liaison directe avec la question d'un Libre Arbitre effectif. L'action volontaire est une pierre d'achoppement de toutes les problématiques relatives à cette question. Deux points ont alors été détaillés, l'un était relatif à une grandeur physique mesurée : le potentiel d'action moteur, l'autre était centré sur une technique de mesure : la chronométrie mentale. Ces deux aspects ont permis un bond majeur dans la compréhension des actions volontaires.

Un des exemples les plus marquant, qui met en rapport les deux points précédents, n'est autre que l'expérience de LIBET et al. Cette expérience arrive à la conclusion que le potentiel électrique qui déclenche une action démarre bien plus tôt que la prise de conscience de vouloir faire cette action. Les certitudes concernant l'action volontaire du sens commun, qui se résument souvent à la proposition "ma pensée implique mon action", sont donc clairement ébranlées.

Dans le même ordre d'idées, nous avons vu que l'expérience de WEGNER et WHEATLEY et celle de SOON *et al.* enfoncent le clou. Dans le premier cas, les auteurs affirment qu'il n'y a rien qui puisse confirmer avec certitude qu'un état mental soit la source d'une action physique. Leurs résultats d'expérience semblent bien aller dans ce sens puisqu'ils arrivent à créer ce sentiment de causalité de toute pièce. Dans le second cas, les capacités prédictives à long terme du choix d'une action ont évidemment des répercussions également très fortes sur la question. Cette expérience est, de fait, une sorte d'expérience de LIBET améliorée.

À la fin de ce chapitre, via l'utilisation de la procédure de "skip response" par KHALIGHINEJAD *et al.*, nous avons décrit l'étude d'un type d'action volontaire qui était clairement la résultante d'un processus délibératif entièrement endogène à l'individu. Il a également été montré que l'outil mathématique prenait une place importante dans les travaux neuroscientifiques. Dans cette expérience, il a été possible d'avancer qu'en plus de donner des résultats satisfaisants, la modélisation mathématique permettait de clarifier l'interprétation du signal électrique du potentiel de préparation moteur. Suivant la ligne des modèles d'accumulateurs, les auteurs montrent qu'il se trouve que le potentiel d'action pourrait bien ne pas être le précurseur de l'action volontaire, comme le pensait LIBET, mais l'aperçu d'une moyenne de signaux aléatoires neuronaux. Pour autant, un nouveau précurseur voit le jour dans cette hypothèse puisque le modèle montre l'importance d'un signal réduisant le bruit neuronal avant une action auto-initiée. Cette importance étant même confirmée par l'apport d'autres études dans d'autres domaines (voir conclusion de [KSD+18]).

Pour conclure, les exemples que nous avons développés avaient pour but de montrer en quoi les Neurosciences apportaient des éléments majeurs au questionnement philosophique concernant la question de l'action volontaire. Les matériaux apportés à cette réflexion ne sont pas eux même exempts de questionnements. C'est ce que nous allons présenter, et que nous allons analyser, dans le prochain chapitre de ce mémoire.

Chapitre 3

L'action volontaire face aux Neurosciences

3.1 Introduction

Dans le chapitre précédent, un certain nombre d'expériences neuroscientifiques ont été présentées. De nombreuses autres expériences concernant l'action volontaire auraient pu être évoquées, cependant, nous allons voir que sur la base de ces cinq travaux il est possible de proposer un discours susceptible d'éclairer les positions philosophiques concernant l'action volontaire. Nous allons constater que ces expériences se parlent entre elles et que les points forts de l'une peuvent répondre aux lacunes d'une autre. Notre proposition est d'affirmer que les Neurosciences sont aujourd'hui indispensables à l'étude de l'action volontaire et, par conséquent, à l'étude de la question du Libre Arbitre. Nous allons donc tenter de décrire des propositions philosophiques, en cohérence avec les exemples caractéristiques des expériences que nous avons vus.

Dans [Wal02], sous la plume du neuroscientifique Henrik WALTER, nous pouvons lire trois raisons de soutenir l'importance des résultats neuroscientifiques dans le débat concernant le Libre Arbitre. En effet, ces résultats peuvent être :

1. Une source d'inspiration pour les thèses philosophiques.
2. Une source de modération pour les thèses philosophiques.
3. Une source de nouvelles problématiques, par exemple éthiques, via les neurotechniques.

Concernant notre étude, restreinte à l'action volontaire, il nous semble que les deux premières raisons restent légitimes. C'est ce que nous tenterons de défendre dans la suite. Dans cette idée, l'aspect inspirant est susceptible de produire une modification des théories sur l'action. Il peut aussi faire apparaître de nouvelles idées. L'aspect empirique servira plutôt à distinguer les théories philosophiques restant plausibles avec l'expérience et celles ne l'étant plus. Bien que le lien entre expérience et théorie soit complexe, nous nous contentons ici d'une définition baconnienne, celle

de l'*instance de la croix* ou de l'expérience cruciale. Il semble que cette définition soit suffisante dans ce cas et des expériences pourront montrer que certaines thèses philosophiques ne sont plus tenables.

Les expériences que nous avons décrites sont traversées de deux grands thèmes de la philosophie de l'esprit :

Le dualisme des substances. Nous verrons des éléments empiriques contre la possibilité qu'une action volontaire soit guidée, *in fine*, par un esprit fonctionnant sur un mode dualiste. Nous constaterons alors ici l'aspect modérateur de l'expérimentation neuroscientifique.

L' épiphénoménalisme. Nous verrons des éléments empiriques qui alimentent une thèse épiphénoménaliste que l'on pourrait qualifier de forte. Nous constaterons alors ici l'aspect inspirant de l'expérience.

Dans la suite, nous allons revenir sur ces deux points. Nous présenterons en quoi des éléments importants sont amenés par les Neurosciences. Ces éléments permettront d'éclairer le débat sur l'action volontaire ainsi que de lui poser certaines limites. Sur la base des expériences vues précédemment nous argumenterons donc en quoi, pour nous, une thèse dualiste ou une proposition épiphénoménaliste n'est pas tenable. Ce sera le premier temps de ce chapitre.

Dans un second temps, nous présenterons un modèle théorique faisant la distinction entre deux types d'intention. Nous confronterons ensuite ce modèle aux expériences neuroscientifiques dont nous avons parlé. Nous verrons alors que certains résultats trouveront un éclairage nouveau.

Enfin, dans un troisième temps, nous conclurons ce chapitre et nous cheminerons vers la conclusion générale de ce mémoire.

3.2 Le dualisme et l'épiphénoménalisme improbables

3.2.1 Contre le dualisme

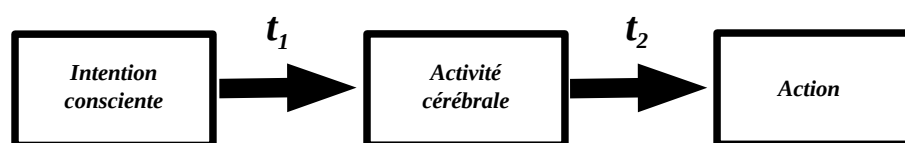


FIGURE 3.1 – Représentation schématique d'une vision dualiste de l'action volontaire. Les flèches noires représentent les temps t_1 et t_2 qui s'écoulent avant l'action.

Dans son introduction à la philosophie de l'esprit, Michaël ESFELD écrit que « *Le libertarianisme entraîne ainsi le dualisme interactionniste.* » [Esf12]. Cette déclaration semble pouvoir prêter à polémique si nous ne faisons pas attention aux termes de la phrase. Pourquoi ESFELD pose-t-il cela ?

Le dualisme interactionniste est l'une des modalités de réponse au problème de la philosophie de l'esprit. Ce problème, si on exclut la possibilité de la surdétermination, consiste à articuler trois

grands principes : (1) Le principe de distinction entre états mentaux et états physiques, (2) le principe de causalité mentale, (3) le principe de complétude causale, nomologique et explicative du domaine physique. Pour le dire rapidement, le dualisme choisit de répondre au problème en affirmant : (1) et (2) et en rejetant (3). Il nous semble que c'est par ces trois affirmations que le libertarianisme (au moins sa forme la plus simple, celle dont nous avons parlé au premier chapitre) entraîne également une forme de dualisme.

L'idée du dualisme, introduite par René DESCARTES, est d'abord celle d'un dualisme des substances. Les états mentaux et les états physiques sont alors de natures différentes et le premier type de substance interagit sur le second type (via la glande pinéale chez DESCARTES.). Il s'ensuit que les états physiques du corps, dont les actions, sont engendrés par un esprit étant d'une autre nature. Les états physiques ne se suffisent donc plus à eux-mêmes pour expliquer toutes les causalités présentes dans le monde. Le principe de complétude causale est alors mis au rebut. Sur la Fig. 3.1, nous avons représenté un schéma de la conception dualiste de l'action.

Il est aujourd'hui admis, de quasiment tout le monde (scientifiques et philosophes), que ce type de pensée dualiste n'est pas tenable face aux explications données par les sciences contemporaines. Par ailleurs, ce n'est pas un seul de nos exemples qui valide cet état de fait mais l'ensemble des cinq travaux que nous avons décrits. En effet, si une position dualiste de ce type était tenable, il semble que les Neurosciences dans leur quête d'explication se verraient confrontées à une sorte de mur explicatif. Il serait pour elles, à un moment, impossible de remonter à une certaine cause puisque cette dernière sortirait complètement du champ scientifique. Cette position porte en effet un bagage métaphysique extrêmement lourd.

Bien qu'il ne soit pas exclu que les Neurosciences se retrouvent confrontées un jour à ce type de mur explicatif, il semble que plus ces dernières avancent, moins cette possibilité apparaît crédible. Ces sciences arrivent à remonter aux causes des phénomènes d'origine mentale sans faire appel à des processus inintelligibles ; toutes ces causes sont du domaine physique. Nous avons ici l'exemple de l'aspect empirique des Neurosciences qui vient fermer la porte de ce que certains ont appelé le problème du *fantôme dans la machine*. Nous pouvons constater que toutes les expériences modernes tendent à mettre un point final à la plausibilité de cette hypothèse [San13].

Les Neurosciences donnent des éléments contre le dualisme des substances. En guise d'exemple, analysons les expériences précédemment décrites. En quoi l'expérience de LIBET s'oppose-t-elle à cette forme de dualisme mais, dans le même temps, semble la réintroduire ? En ce qui concerne le dualisme des substances, cela paraît évident, l'expérience menée ne relève la nécessité d'aucune substance mystérieuse pour expliquer l'apparition de l'action. En effet, c'est même le contraire, la découverte de l'existence d'une activité électrique particulière (le potentiel de préparation moteur) précédant toute action volontaire vient argumenter en faveur d'une cause purement physique aux mouvements volontaires. Pourtant, il est important de noter que l'expérience de LIBET souffre de certains problèmes. Parmi ces derniers notons :

PB1 : l'activité inconsciente, mesurée par LIBET, est-elle la cause de l'action ? N'est-elle pas simplement un processus de préparation du cerveau tout à fait normal avant une action ? La décision effective de l'agent ayant lieu plus tard, dans le laps de temps conscient.

PB2 : ne pourrait-on pas, dans le cas du point précédent (**PB1**) , stipuler que l'action est

démarrée inconsciemment mais que le choix de la mener à son terme appartient à l'individu ? (La question plus général du Libre Arbitre serait alors transformée en une liberté de veto et le terme anglais de “Free won’t” serait employé plutôt que celui de “Free will” [OH04]).

PB3 : le protocole pourrait être remis en cause dans la mesure où LIBET prétend mesurer une action endogène alors qu'il donne une instruction aux individus : “fléchissez le doigt quand vous le voulez”. Peut-on encore considérer que l'action vient de l'individu et de lui seul ? Que mesure-t-on vraiment dans ce cas ? A quel moment se situerait la véritable décision consciente ? Au moment, de fléchir le doigt ? Au moment de l'acceptation de l'instruction ?

PB4 : Les temps mesurés avant l'action sont très courts. On pourrait faire remarquer que les temps de conscience pourraient bien se dérouler avant l'action mais que les mesures contiennent des erreurs à cause de ce laps de temps très ténu. Cette argument viendrait remettre en cause l'ordre causal début potentiel de préparation moteur $\Rightarrow (W)$ impliquant qu'un dualisme des substances resterait pensable.

Une idée est contenue dans le **PB1** et le **PB2**. Cette idée est que, au fond, le potentiel de préparation moteur ne ferait que donner la possibilité d'une action. Ce potentiel révélerait une préparation du cerveau à l'action mais la décision de mener l'action à son terme appartiendrait à l'agent. On aurait alors ici un modèle de décision tardive (late decision). Une forme de dualisme des substances pourrait réapparaître. C'est d'ailleurs une des positions qu'envisage LIBET, dès son expérience originale, en proposant la capacité du libre veto de l'agent. Son idée serait alors que le potentiel de préparation moteur précéderait une action de façon inconsciente mais que l'agent aurait la possibilité, dans le laps de temps conscient, de choisir de faire avorter l'action. A propos de cette notion de veto, MELE écrit dans [Mel02] que LIBET semble faire deux suggestions :

S1 : Même si les décisions libres de veto ont des processus inconscients parmi leurs causes, ces décisions ne sont pas déterminées de manière déterministe.

S2 : Les décisions libres de veto ne sont pas causalement dépendantes de processus inconscients qui les précèdent.

Ce que MELE souligne ici, dans **S1** et **S2**, c'est le caractère peu compréhensible de la vision proposée. Il semble bien que LIBET évoque la possibilité d'un processus mystérieux de veto, non déterministe, qui pourrait être vu comme une forme de réintroduction d'un dualisme des substances. Nous faisons alors remarquer que cette thèse souffre d'un défaut majeur : comment rendre compréhensible le fait que l'agent pourrait avorter son action mais pas la déclencher ? Le fait de stopper sa propre action, n'est-ce pas produire une action ? En conséquence il semble bien que cette proposition n'a pas plus de solidité que celle qui propose que l'agent déclenche consciemment une action, de façon *ex-nihilo* physiquement parlant, via une hypothétique capacité qu'il posséderait.

Le **PB3** est relatif au moment où se produit l'action. Nous reviendrons à ce point dans la prochaine section. Le modèle que nous souhaiterions proposer se basera sur ce problème pour

avancer l'idée que l'intention dans ce type d'actions, et probablement dans beaucoup d'autres, relève d'une intention précédant de loin l'action (première distinction concernant l'intention) alors que LIBET étudie des intentions se produisant juste avant l'action (seconde distinction concernant l'intention). Notre proposition sera alors que l'action pourrait démarrer de façon inconsciente, mais sur la base de délibérations conscientes, placées bien en amont de l'action à proprement parler.

Le **PB4** est intéressant parce qu'il remet en cause la précision de la prise de mesure. De nombreux facteurs peuvent être évoqués pour douter de la validité des temps obtenus dans l'expérience de LIBET. On peut par exemple postuler des problèmes relatifs à des difficultés de synchronisations des agents [Mel02]. En effet, les personnes testées devaient être conscientes de deux choses dans le même temps : la position de l'aiguille et leur conscience. Ne pourrait-il pas y avoir une forme d'anticipation, liée à la prédiction de la position de l'horloge à l'instant suivant ? Selon nous cet argument est balayé par une autre expérience que nous avons décrite : celle de Soon *et al.* Bien entendu, comme nous l'avons vu, le pouvoir prédictif de cette expérience n'est que de 60% ce qui n'est pas très important mais cela reste mieux que le hasard. Il apparaît donc que des processus contenant déjà l'information de l'action à venir apparaissent 10 secondes avant que l'action ne se produise. Ici, l'erreur de mesure est totalement exclue. L'interprétation de LIBET semble donc bien correcte, simplement, l'activité qu'il a mesuré ne correspondrait qu'à la dernière étape du long processus de prise de décision inconscient de certains modules du cerveau.

Les autres expériences que nous avons décrites, en remontant toujours plus loin la chaîne causale des processus menant à l'action, semblent clairement rendre caduque toutes les formes de dualisme concernant l'action volontaire. Toutes ? Non, car la proposition que nous allons à présent aborder, l'épiphiénoménalisme, peut être classée dans la catégorie du dualisme des propriétés. L'illustration de cette position est majoritairement représentée par l'expérience de WEGNER et WHEATLEY. Nous verrons que selon toute vraisemblance, bien que séduisante à la vue des expériences neuroscientifiques, cette position ne semble pas tenable dans toute sa généralité.

3.2.2 Contre l'épiphiénoménalisme

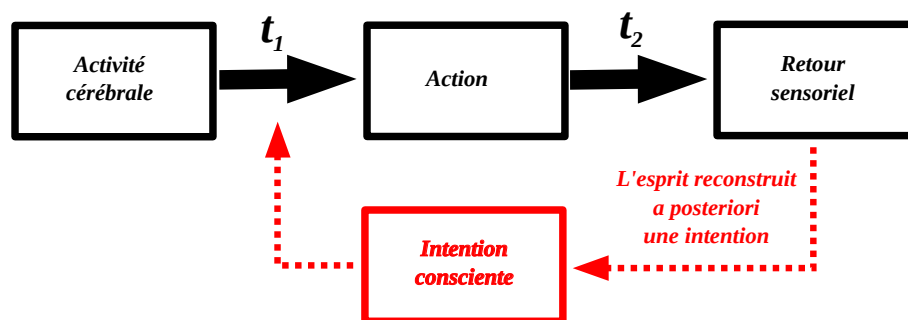


FIGURE 3.2 – Représentation schématique d'une vision épiphénoménaliste de l'action. Les flèches noires représentent les temps t_1 et t_2 qui s'écoulent avant l'action.

Une autre possibilité, impliquant une forme de dualisme, serait d'envisager un monisme des substances mais un dualisme concernant les propriétés. Les propriétés, associées à un type d'action qualifiée de libre, émergeraient de la structure complexe de la matière. Un peu comme les propriétés du vivant qui semblent émerger de la matière inerte. Dans le cas qui nous intéresse ici, la capacité

à effectuer des actions libres procéderait d'une propriété rendue possible par la complexité du cerveau. Pour échapper au même type d'argument que celui que nous avons évoqué précédemment, c'est-à-dire l'idée que les Neurosciences ne pourraient pas expliquer entièrement à leur niveau le phénomène de l'action, il nous semble qu'une caractéristique majeure se doit d'être respectée : il faut que cette propriété émergente soit observable.

Michael S. GAZZANIGA, dans son ouvrage *Le Libre Arbitre et la science du cerveau*, semble plutôt pencher pour une hypothèse émergentiste. Cependant, au détour d'un chapitre où il évoque les propos du physicien Robert LAUGHLIN, il écrit :

« *« La raison pour laquelle nous y croyons, comme pour la plupart des choses émergentes, est parce que nous les observons. » Bien que le physicien Robert Laughlin ait parlé des transitions de phase telles que celle de la glace en eau, il aurait pu dire la même chose sur nos sentiments de responsabilité et de liberté. »* [Gaz13]

Il apparaît donc que le caractère observable du phénomène émergent soit important. Mais, bien qu'observable et expérimentable par tous, qu'en est-il de cette liberté que nous sommes si sûrs de posséder ? Est-elle une propriété réellement émergente ? En étudiant le processus de l'action volontaire via les Neurosciences comme nous l'avons vu, et en reprenant encore l'image de Patrick HAGGARD, en effectuant une forme de rétro-ingénierie (reverse engineering) de ce processus, les Neurosciences nous informent sur la manière dont ce type d'action est produit. Il semble donc que ces sciences arrivent à expliquer de façon convaincante, à leur niveau, la production de l'action volontaire (au moins dans sa version minimale). Elles ne se heurtent alors à aucune propriété émergente qui serait de nature inexplicable pour elles. Définir la liberté d'action comme la somme d'une prise de décision (non forcément consciente), et le fait de pouvoir suivre de façon adéquate tous les processus en jeu dans le cerveau avant que celle-ci ne se produise, ne semble pas être mis en difficulté (si on exclut les difficultés techniques). Ici on opère bien à une forme de réduction. Par conséquent l'idée d'une propriété émergente, concernant l'action volontaire, qui ne serait plus explicable au seul niveau des Neurosciences ne semble pas d'actualité. C'est au moins le cas pour des processus simples et immédiats. Ici, nous précisons bien que nous parlons d'un processus juste avant l'action. La distinction entre "juste avant" une action et "bien avant" cette action nous paraît, en effet, essentielle et nous le verrons plus loin.

Il reste cependant une option qui peut être considérée. Elle consisterait à penser que notre conscience de la liberté d'action que nous ressentons est bien émergente, au sens où elle serait irréductible à des processus physiques, mais que cette propriété n'agit pas sur le monde qui nous entoure. La thèse de l'épiphiénoménalisme est classée dans la catégorie du dualisme des propriétés, il en s'agit même de la thèse la plus courante. Son idée est justement celle que nous venons de poser. En effet, cette position prétend que la conscience de la volonté de produire une action n'est qu'un pur produit émergent de notre cerveau n'ayant aucun impact causal sur les objets du monde physique. La figure Fig. 3.2 représente un type de cette position qui est appelé *confabulationisme*. L'une des expériences que nous avons présentée amène particulièrement des arguments en faveur de cette thèse. Il s'agit de l'expérience menée par WEGNER et WHEATLEY en 1999. Cette expérience montrait alors que, dans certaines conditions, les agents pensaient avoir déclenché une action volontairement mais que cette action leur était imposée de l'extérieur. WEGNER proposera que toutes les actions procèdent de la même façon et que les intentions que l'on associe à ces actions ne sont que des illusions. En effet, suite à une action et à son retour sensible (on se voit faire

l'action), l'esprit associerait une intention d'action au mouvement du corps. Cette intention n'aurait, en ce sens, aucun rôle causal dans l'action. La suite de notre propos peut être divisé en deux temps. Premièrement, nous montrerons en quoi l'expérience de WEGNER et WHEATLEY inspire la thèse épiphénoménaliste. Deuxièmement, nous évoquerons en quoi cette thèse est critiquable et ne parvient pas à convaincre totalement.

L'aspect inspirant de l'expérience WEGNER et WHEATLEY

Dans leur travail, les propos de WEGNER et WHEATLEY sont assez détaillés et ils ne se basent d'ailleurs pas que sur leur expérience. Pour étayer leur proposition, les auteurs parsèment leur écrit de multiples exemples. Certains relèvent de pathologies mentales, d'autres d'illusions paranormales. Ces exemples montrent les dysfonctionnements que la mise en relation causale peut subir. Il est par exemple question de la radiesthésie, des sourciers ou des phénomènes de troubles mentaux associés à la schizophrénie. Dans le cas des sourciers (ou des radiesthésistes), alors que le mouvement du bâton (ou du pendule) qu'ils utilisent est bien produit par l'agent, la mise en relation par causalité n'est pas effectuée par l'esprit. Il en résulte que le mouvement apparaît faussement produit par une cause mystique. Dans le cas de la schizophrénie, entendre des voix pourrait provenir du fait que les pensées des agents ne sont pas en cohérence avec les buts conscients au même instant. Il en résulterait une "voix intérieure" qui est sans rapport avec la pensée en cours. Ainsi des événements, qualifiés de paranormaux ou de pathologiques, peuvent tout à fait s'expliquer par un défaut de mise en relation pensée-corps qui dans ce modèle devrait se produire *a posteriori* de l'action. Le modèle que WEGNER et WHEATLEY proposent est donc basé sur la causalité apparente entre pensée et action. Détaillons un peu plus celui-ci.

WEGNER et WHEATLEY évoquent les travaux d'Albert MICHOTTE. Ce dernier avait déjà proposé que deux facultés rendaient la propriété de causalité des idées sur le monde physique possible à l'esprit : la capacité de prévoyance du résultat d'une action et la capacité de ressentir l'action en train de se dérouler. En ce sens, le sentiment que notre conscience est l'origine de notre action résulterait d'une mauvaise interprétation du flot des événements. Sans connaissance des mécanismes profonds mis en jeu dans l'action, il est fort probable que l'esprit ait tendance à attribuer une causalité apparente entre pensée consciente et événement du monde. Il y a ainsi une illusion qui se crée qu'un 'je' a pensé et que le corps exécute. Ce modèle expliquerait pourquoi les actions réfléchies sont ressenties comme provoquées par nos pensées conscientes alors que les actions automatiques ne le sont pas spécialement. Au début de leur article, WEGNER et WHEATLEY proposent une vision humienne¹ de l'expérience de la volonté. Il s'agirait d'une perception qui vient de la mise en conjonction constante des événements se produisant. Aussi, ils écrivent :

« The person experiencing will, in this view, is in the same position as someone perceiving causation as one billiard ball strikes another. »

L'exemple du billard sera explicité par la suite comme nous le verrons.

Le modèle développé ici est schématisé sur la Fig.3.3. Une manière de se représenter cette idée est de prendre l'angle logique. Imaginons deux objets *A* et *B*. Supposons ces deux objets comme ayant des existences qui se succèdent. Bien que cela soit la position la plus évidente, rien n'indique de $A \rightarrow B$. En effet, cela peut être vrai mais au même titre qu'un troisième objet *C* peut impliquer les deux : $C \rightarrow A$ et $C \rightarrow B$. Les auteurs précisent que ce problème, impliquant une troisième

1. Les auteurs citent un passage de : David Hume, *Traité sur la nature humaine*, 1739

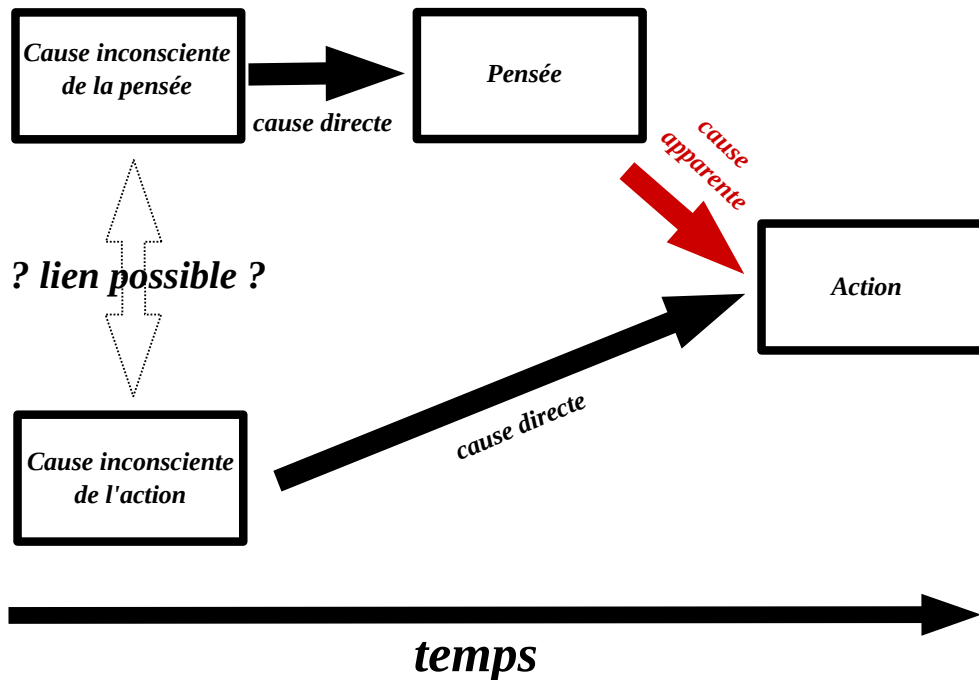


FIGURE 3.3 – Représentation schématique du modèle proposé par WEGNER et WHEATLEY.

cause, est exactement similaire au problème de la pensée impliquant l'action. Il est tout à fait admissible qu'une autre origine, par exemple inconsciente, soit la cause à la fois de la pensée et de l'action. La relation de causalité inférée par l'esprit serait alors erronée.

Dans le modèle, trois notions de bases vont être proposées pour que cette fausse relation de causalité puisse s'établir dans l'esprit des agents. Ces trois idées sont les suivantes :

La priorité : la pensée doit précéder l'action, et ce, dans un temps raisonnable

La cohérence : la pensée doit être compatible avec l'action

L'exclusivité : la pensée doit être considérée comme la seule cause possible de l'action.

Voyons à présent ce que recouvrent ces points.

La Priorité : Comme les auteurs, pour illustrer ce point, employons l'exemple des boules de billard. Soit une boule B_1 , en mouvement, frappant une boule B_2 . Pour que l'on attribue le mouvement de B_1 comme étant la cause du mouvement de B_2 après le choc, il faut que le mouvement de B_1 soit antérieur au mouvement de B_2 mais pas trop non plus. Si le mouvement de B_2 apparaît longtemps après le choc, la causalité avec B_1 aura été perdue. De la même manière, si le mouvement de B_2 précède le mouvement de B_1 , cela ne fonctionne pas. Pour la causalité, entre pensée et action, le problème est en tous points identique. Il faut que la pensée précède l'action mais suivant le bon timing. Une pensée consciente trop tôt (3 jours avant une action par exemple) ne sera plus considérée comme la cause première de l'action. De manière identique, si une pensée arrive après

une action, elle ne sera pas considérée comme une cause et le sentiment de causalité sera aussi perdu.

La Cohérence : Encore une fois, pour illustrer ce point, les boules de billard sont un très bon exemple. Pour que ce sentiment de causalité soit acceptable, il faut que B_2 ait un mouvement cohérent après le choc qu'elle a reçu de B_1 . Qu'est-ce qu'un mouvement cohérent ? Il semblerait, dans le cas de cet exemple, qu'il s'agisse d'un mouvement en accord avec les lois de la physique que nous expérimentons tous les jours. Il ne s'agit pas d'être un expert des mécanismes physiques pour induire une relation de causalité, mais juste que les mouvements soient cohérents avec ce que nous avons appris empiriquement en contemplant le monde. Pour la relation de causalité entre pensée et action, c'est la même chose. Il faut, quand il s'agit des individus extérieurs, qu'il y ait une cohérence entre leurs idées habituelles et leurs actions. La même chose est vraie également pour l'individu. Les auteurs suggèrent une association sémantique entre l'idée et l'action. C'est d'ailleurs ce qu'ils ont utilisé dans leur expérience en prononçant des mots pour évoquer l'idée. Un des exemples de ce point est le fait que les individus ont tendance à penser qu'ils maîtrisent la chance s'ils ont eu un certain nombre de succès dans leurs entreprises passées.

L'Exclusivité : Une dernière fois, l'exemple du billard est intéressant ici. En effet, pour supposer que B_1 a causé le mouvement de B_2 , il ne faut pas de causes concurrentes au moment du choc. Il y a fort à parier que le sentiment de causalité serait amoindri si une troisième boule, disons B_3 , entraînait en ligne de compte au moment du choc. Pour que la relation de causalité entre pensée et action soit conservée, de la même manière, il ne faut pas qu'une autre cause vienne concurrencer la première. Deux types de causes peuvent alors être envisagés.

Les causes internes : il s'agit des causes internes à l'individu : les compulsions par exemple.

Les causes externes : les forces extérieures à l'individu : la présence d'autres agents à qui on attribue un Libre Arbitre par exemple.

Des exemples de phénomènes où ce point est altéré sont cités par les auteurs. Par exemple, les célèbres expériences de MILGRAM. Dans les travaux de psychologie de MILGRAM², des individus font un véritable transfert de causalité sur la figure autoritaire. On peut aussi citer des cas d'hypnoses où l'agent attribue son mouvement à une autre personne dont il ne serait plus que l'intermédiaire.

Ainsi, le modèle présenté par WEGNER et WEATLEY est bien défini. De plus, leur expérience semble bien cohérente avec cette idée. Dans ce cas il a surtout été question de tester le premier point, la priorité, au travers de l'étude du timing dans un cas où l'exclusivité était ambiguë mais où la cohérence ne faisait pas de doute. Les auteurs ont ainsi pu révéler certains aspects de ce sentiment si commun qui consiste à croire que nos pensées précèdent nos actions alors qu'elles sont en fait artificiellement réintroduites dans la chaîne causale après avoir constaté l'action. Nous voyons donc que ce type de proposition alimente l'aspect inspirant de l'expérience en donnant des arguments à la thèse de l'épiphiénoménalisme.

2. Stanley Milgram, *Obedience to authority : an experimental view*, 1974

Critique d'un épiphénoménalisme fort

Que le phénomène, décrit par WEGNER et WHEATLEY, existe, cela ne fait aucun doute. Notamment, des dizaines d'expériences sont présentées par GAZZANIGA dans son ouvrage *Le Libre Arbitre et la Science du Cerveau* [Gaz13]. GAZZANIGA est un neuroscientifique qui est surtout connu pour avoir travaillé avec des agents dont le cerveau est dit scindé (split-brain). Il s'agit de patients dont les deux hémisphères du cerveau ne communiquent plus entre eux suite à une scission du corps calleux. Il a montré qu'un dispositif très important se logeait dans l'hémisphère gauche. Ce dispositif est ce qu'il appelle *le module interprète*. C'est ce dispositif cérébral qui, sans qu'on en ait conscience, donne du sens à nos actions. Certaines expériences remarquables où l'on mettait à contribution les deux hémisphères cérébraux de manière différente, chez des patients aux cerveaux scindés, montrent de leur part des réinterprétations très claires des actions qu'ils ont menées³. Ces réinterprétations sont alors de pures inventions cohérentes avec ce qui se trouve dans la pièce. Nous comprenons ainsi que, dans ce cadre, les résultats de WEGNER et WEATLEY ne sont pas très surprenants.

Malgré des résultats assez convaincants, semblant en accord avec le modèle proposé, l'interprétation de l'expérience comporte certains problèmes. En effet, quand les auteurs affirment l'épiphénoménalisme suite à ces résultats, il semble que cela soit un peu court. Listons rapidement quelques problèmes relatifs à cette expérience :

PB1 : le fait qu'il y ait une différence entre une expérience volontaire consciente (EVP) et la volonté consciente.

PB2 : comment doit-on comprendre l'idée de cohérence proposée par WEGNER ? Dans l'expérience, ce sont des mots qui sont induits dans l'esprit de l'individu pour provoquer une pensée. N'y a-t-il pas un problème à simplement suggérer qu'un contenu sémantique suffirait à produire une expérience de volonté consciente ?

PB3 : WEGNER montre que ce type de reconstruction de l'intention après l'action se produit, certes, mais cela est-il vrai de toutes les actions ?

En ce qui concerne le **PB1**, que les EVP ne soient pas la cause des actions est certainement vrai mais cela ne signifie pas que la volonté consciente ne le soit pas. Pour le **PB2**, un contenu sémantique semble pouvoir être activé dans le cadre d'une action et d'une idée sans pour autant induire le sentiment que l'action a été causée par notre conscience. Il ne suffit apparemment donc pas que les trois conditions que WEGNER propose soit réunies pour que ce sentiment que notre conscience a causé l'action apparaisse. Le **PB3** est celui qui nous semble le plus intéressant et c'est ce dernier qui va enclencher notre réflexion de la prochaine section.

Au travers de ce troisième problème, il nous semble que WEGNER prétend quelque chose de trop

3. Par exemple, dans ces expériences, il était demandé à l'agent de choisir une image avec le bras droit (hémisphère gauche) et une autre image avec le bras gauche (hémisphère droit). Il était ensuite demandé aux agents de justifier leurs choix. Dans le cas du choix de la main droite, le module interprète (situé dans l'hémisphère gauche) permettait de justifier le choix sans problème. Dans le cas du choix de la main gauche, la justification était une réinterprétation en cohérence avec le choix de la main droite.

fort : soit tous les états mentaux conscients sont épiphénoménaux, soit aucun ne l'est. A ce propos le philosophe Alfred R. MELE écrit que WEGNER, dans son ouvrage *The Illusion of Conscious Will*, évoque le fait qu'il soit possible que toute action soit produite par le même processus essentiel n'incluant pas d'intention consciente en son sein [Mel02]. Dans sa conférence, qui précède la remise du prix Jean Nicod⁴, Patrick HAGGARD précisera lui aussi que cette vision "confabulationniste totale" semble trop forte. L'interprétation de type "tout ou rien" de WEGNER est-elle tenable ? Nous pensons que non et ce pour au moins trois raisons.

La première chose qui nous fait douter de la proposition de WEGNER concerne une autre expérience dont nous avons parlé. Ce travail est celui concernant la stimulation électrique du cerveau réalisée par FRIED *et al.* Dans ces expériences, il semblait bien qu'une sensation de "urge to move" était parfois créée avant une action. Suite à la stimulation électrique, parfois, les agents ressentaient le besoin et l'envie de bouger l'un de leur membre. Ce sentiment ne pouvait pas être réintroduit après avoir vu une action puisqu'aucune action ne s'était produite. Qui plus est, en augmentant la stimulation, cela conduisait à un mouvement dans la région du corps concerné par l'envie de mouvement. Il apparaîtrait donc que le "urge to move" conscient d'effectuer une action, et le déclenchement du mouvement lui-même, ait parfois la même racine. Dans ce cas, aucune reconstruction *a posteriori* de l'intention d'effectuer une action ne semble recevable.

Le deuxième argument, contre la thèse de WEGNER, est un argument scientifique plus général. Si l'idée de la théorie de l'évolution initiée par Charles DARWIN dans son *Origine des espèces* est correcte, comment comprendre que l'évolution ait doté les êtres humains de cette conscience des intentions si cette dernière n'a aucun impact sur le déroulement des événements du monde ? Affirmer cela reviendrait à dire que toutes les actions des agents pourraient aussi bien se dérouler de la même manière en l'absence de toute conscience. Cela ne paraît pas très raisonnable. MELE à ce propos écrit :

« *Why then do we even have conscious intentions ? Why did we evolve in such a way as to have them ? Wegner's (341) reply is that we have conscious intentions because they give us a sense of which of the things we do we are responsible for.* » [Mel02]

Bien que WEGNER donne une raison à cette question, il semble que cela ne change rien au fond du problème. Dans le même ordre d'idée, dans son ouvrage *Liberté et Neurobiologie*, le philosophe John R. SEARLE écrit :

« *Dans l'état présent des connaissances, l'objection principale qui nous retient d'adhérer à l'épiphénoménalisme consiste à dire qu'il va à l'encontre de tout ce que nous savons de l'évolution.* » [Sea04]

Il précise ensuite que si la rationalité consciente, qui est un phénotype très important chez les êtres humains, ne devait jouer aucun rôle fonctionnel pour la vie et la survie de l'organisme alors toute la théorie de l'évolution elle-même serait contredite. Cet argument nous paraît très fort et semble se suffire, quasiment à lui seul, pour faire tomber la thèse d'un épiphénoménalisme trop radical.

Enfin, le troisième argument contre l'épiphénoménalisme que nous aimerions évoquer est lié au **PB3**. En fait, cet argument va à l'encontre de l'idée que nous avons qualifié de "tout ou rien" proposée par WEGNER. Dans son propos WEGNER semble se concentrer sur un seul type

4. Propos de Patrick HAGGARD, lors de la remise du Prix Jean-Nicod le mardi 17 mai 2016 : <http://www.institutnicod.org/seminaires-colloques/conferences-et-prix-jean-nicod/archives/article/prix-et-conferences-jean-nicod-1378>

d'intention qui est qualifiée d'intentions *proximales* (juste avant l'action). L'angle d'attaque de l'argumentaire que nous allons maintenant mener repose sur ce fait. Cet angle n'est d'ailleurs pas spécifique à l'expérience de WEGNER mais s'applique à quasiment toutes les expériences que nous avons abordées. Il nous semble qu'il y a un problème majeur à tirer des conclusions sur l'action volontaire en général si l'on ne prend pas soin de distinguer, comme SEARLE l'a fait, entre intentions dites *préalables* et intentions dites *en action*. Notons que MELE, ainsi qu'Eddy NAHMIAS, ont eux aussi remarqué ce problème [Mel02]. Cette distinction, entre ces deux types d'intentions, semble donc bien primordiale.

Dans la suite de ce chapitre, c'est cette piste que nous allons tenter de creuser. Dans un premier temps nous reviendrons sur cette distinction proposée par SEARLE (mais aussi par d'autres). Dans un second temps, nous affirmerons que ce modèle philosophique paraît solide pour expliquer l'action en collaboration avec les Neurosciences et leurs expériences. Nous montrerons notamment que certains comportements d'agents au cerveau lésé semblent rendre compte de cette proposition.

3.3 Une action, deux types d'intentions

3.3.1 Intentions préalables et intentions en actions

Le concept d'intention que nous avons évoqué jusqu'ici ne possédait pas une grande précision. C'est, la plupart du temps, ainsi qu'on le retrouve dans les articles neuroscientifiques. Pourtant, il semble qu'une précision concernant ce terme soit nécessaire pour éviter certaines interprétations fautives des résultats d'expériences. En effet, nous pouvons noter que les expériences de LIBET et de WEGNER sont souvent utilisées pour affirmer que le sens que nous donnons à l'action volontaire, et même à tout le Libre Arbitre, n'est qu'une illusion et que la science l'a montré.

Il convient d'être prudent face à la précédente affirmation car il semble que cette dernière ait pour origine une description naïve de l'idée d'action volontaire. Les neuroscientifiques qui affirment cette position semblent toujours se référer à l'idée du *fantôme dans la machine* dont nous avons parlé précédemment. L'idée serait alors que pour qu'une action soit libre, l'intention de cette action doit surgir d'une substance non matérielle. MELE, rapportant les propos du philosophe Daniel DENNETT, écrit que dans la question relative à l'action libre, tout dépend de ce que nous entendons par là. Il évoque le fait que le Libre Arbitre pourrait être moralement important sans être une capacité surnaturelle. En effet, DENNETT dit que si être libre nécessite une cause surnaturelle alors il est clair qu'aucune action n'est libre. Il s'empresse d'ajouter qu'il n'y a, à sa connaissance, aucune raison de comprendre l'action libre de cette façon [Mel02]. NAHMIAS partage cette position et critique le fait que les arguments des scientifiques impliquent toujours l'hypothèse que l'action libre suppose un esprit immatériel. L'auteur dit alors qu'ils font erreur lorsqu'ils présupposent cela⁵. Toute la subtilité est dans le propos que tient HAGGARD. Oui, le Libre Arbitre, et donc aussi l'action volontaire, au sens dualiste, ne semble pas crédible et nous avons essayé de le montrer dès le début de ce chapitre. Cependant, la philosophie a elle aussi fait du chemin et l'on ne peut plus considérer qu'elle propose encore sérieusement qu'une action volontaire se produise sur un mode dualiste cartésien.

Ce chemin, parcouru par la philosophie, nous pouvons le constater dans la description qui est

5. E. Nahmias, « Is Neuroscience the Death of Free Will? », *New York Times*, 13 Novembre 2011

faite de l'intention et de l'action intentionnelle en philosophie de l'action. Le sujet est vaste et assez pointu et il n'est donc aucunement question ici de le traiter en détail ; cependant, il nous semble qu'une distinction majeure mérite d'être citée ici dans la mesure où elle a de grandes répercussions vis-à-vis des conceptions des Neurosciences à propos de l'action volontaire. Cette distinction a été, entre autre, proposée par SEARLE [Sea85]. Elle sépare deux types d'intentions :

Des intentions distales : SEARLE appelle ces intentions : des *intentions préalables* (ce type de concept porte aussi d'autres noms en fonction des auteurs : intentions dirigées vers l'avenir (Michael E. BRATMAN), intentions prospectives (Myles BRAND)...). Il s'agit des intentions les plus éloignées dans le passé de l'action.

Des intentions proximales : SEARLE appelle ces intentions : des *intentions en actions* (d'autre auteurs les nomment : intentions dirigées vers le présent (Michael E. BRATMAN), intentions immédiates (Myles BRAND)...). Il s'agit des intentions qui se produisent pendant l'action, qui l'accompagne, et qui la rectifie si nécessaire. L'expression intention en action est assez parlante de ce point de vue.

Pour illustrer ces intentions, prenons l'exemple de l'action consistant à lever mon bras droit. Deux types d'intentions peuvent alors entrer en jeu :

Les intentions préalables. Elles sont le siège de la planification ou d'une préméditation de l'action. Leurs contenus se limitent à la présentation d'un schéma. Dans notre exemple, ce schéma est la représentation très générale de mon bras droit se levant. Ces intentions ne sont pas suffisantes pour produire une action intentionnelle. Pour cela, il faut qu'elles causent une intention en action. Un exemple typique de cet aspect est celui que l'on appelle : *le problème de la déviance causale*. Dans ce problème, une action préméditée peut arriver à son but mais pas du tout par le chemin causal que l'on avait prévu. Par exemple, imaginons que X projette de tuer Y. Dans ce cas, X prévoit les étapes de son action et il croit savoir où se trouve Y. Il prend donc sa voiture puis, nerveux à l'idée de ce qu'il va commettre, a un accident et tue un piéton qui s'avère être Y. La finalité est la même mais l'action ne semble pas pouvoir être qualifiée d'intentionnelle dans ce cas. Notons aussi que ces intentions ne sont pas nécessaires à toutes les actions. Cela prend en compte le fait que certaines actions soient intentionnelles et spontanées sans avoir nécessité de préplanification ⁶.

Les intentions en actions. Elles sont le siège du contrôle de l'action se déroulant. Ce type d'intention produit une action spécifique. Dans le cadre de notre exemple, si mon but est de lever mon bras droit, je dispose de mille manières d'atteindre ce but. Mon intention en action va produire un mouvement particulier : je vais lever mon bras d'une certaine manière, avec une certaine vitesse etc... Ces types d'intentions peuvent être causées par une intention préalable mais ce n'est pas toujours le cas. Comme nous venons de le voir, un

6. La définition d'action intentionnelle a un sens précis, mais ici, pour ne pas alourdir le propos, nous ne le préciserons pas plus.

cas où seul ce type d'intention est nécessaire est celui des actions spontanées qui n'ont pas besoin d'intentions préalables.

Sous l'influence d'Elisabeth PACHERIE, ce modèle est aujourd'hui lié à un modèle à trois niveaux portant le nom de modèle DPM [Pac08, PH10]. Il se base sur les niveaux d'intention suivant :

Niveau D : ce sont les intentions distales comme nous les avons définies. Le temps caractéristique de constitution de ces intentions est celui de la délibération et de la planification bien avant l'action. Ces intentions portent une représentation conceptuelle de l'action. Elles mettent fin au raisonnement portant sur les fins d'un but pour enclencher le raisonnement sur les moyens.

Niveau P : ce sont les intentions proximales comme nous les avons définies. Le temps de constitution de ces intentions est celui de la contextualisation de l'action juste avant celle-ci. Elles portent une représentation d'une action spécifique adaptée à la situation présente.

Niveau M : ce sont les intentions moteurs. Elles viennent s'ajouter aux deux intentions que nous avons décrites précédemment. Le temps de ces intentions est le très court terme pendant l'action. Elles portent une représentation sensori-motrice de mouvement du corps dans l'espace. Elles sont le contrôle moteur de l'action.

Il s'agit donc ici d'une version complexifiée et affinée de la simple distinction que nous souhaitons introduire. Pour la suite, nous nous en tiendrons à la simple division searlienne et dès à présent regardons en quoi cette séparation de l'intention en deux entités peut être mise en évidence de façon empirique. Pour ce faire, nous nous baserons sur des cas pathologiques où l'un des types d'intentions semble altéré.

3.3.2 Indices du caractère double de l'intention

Certaines pathologies du cerveau, ou certaines lésions, semblent indiquer que la distinction entre deux types d'intention est loin d'être superficielle. En effet, cette différenciation semble bien pouvoir être observée [LS04, SDC⁺03, Hag08].

Les cas de défaillance des intentions préalables

Prenons le cas d'individu dont le cortex préfrontal est lésé. Dans ce cas, l'agent risque de présenter comportements d'utilisation d'objet particuliers. Le neurologue François LHERMITE a étudié ce type de comportement. Dans ces cas les personnes voyant un objet devant eux, ne peuvent s'empêcher d'avoir un comportement stéréotypé avec cet objet (voir refs précédentes). Par exemple, lorsqu'une paire de lunettes est posée sur la table devant eux, ils vont la chauffer et lorsqu'une deuxième paire leur est présentée ils vont la chauffer aux dessus de la première. Il semble bien que dans ce cas les agents ne forment plus d'intention préalables. Ils agissent par automatisme cohérent avec l'utilisation d'un objet. La vision de l'objet suffit à déclencher une intention en

action. Ne détectant aucune incohérence avec une intention préalable, ils ne s'en étonnent pas. Si l'expérimentateur leur demande de justifier leur acte, ils rationalisent a posteriori en affirmant qu'ils supposaient que c'est ce qui était attendu d'eux. Fait encore plus marquant, si on leur précise qu'il ne faut pas utiliser l'objet, ils le font quand même tout en se souvenant de la règle établie. Il semble donc bien que, dans ce cas, les agents soient incapables de produire des intentions préalables. Ils ne constatent donc pas d'incohérence avec leurs intentions en action et donc justifient, après coup, des actes automatiques. Ce type d'intention serait donc lié au cortex préfrontal.

Les cas de défaillance des intentions en actions

Le cortex pariétal, conjointement avec le cortex moteur, permettrait de contrôler une action vers un but. Le potentiel de préparation moteur, dont nous avons longuement parlé, apparaîtrait suite aux échanges électriques entre ces deux zones. Arrivée à un certain seuil, la conscience de l'action apparaîtrait à l'individu. La notion de contrôle semble confirmer que nous avons affaire ici à un processus proche de ce qui est désigné par SEARLE comme une intention en action. Ce que nous constatons c'est que la conscience du "urge to move" des agents semble faire référence à ce processus d'intention en action. Lorsque le cortex pariétal est lésé, c'est la précision des gestes qui est atteinte. L'individu peut toujours construire des intentions préalables mais, lors du passage à l'action, son déroulement est maladroit. Lorsque c'est le cortex moteur qui est touché, cette fois, il semble que ce soit le lien entre intention préalable et intention en action qui soit atteint. Les agents peuvent alors être atteints du syndrome de la main anarchique (voir refs précédentes). Ils ont alors le sentiment que leur main est mue par une volonté propre. Ils ont conscience que leur intention en action ne coïncide pas avec leur intention préalables, que le mouvement est inadéquat, mais ne ressentent pas cela comme venant d'eux.

Constatant l'existence de ces cas pathologiques, nous pouvons partir du principe que l'intention classique peut en fait s'étendre sur un processus possédant deux types d'intentions. Dans la prochaine section, voyons les répercussions que cela peut avoir sur les interprétations des expériences d'actions volontaires que nous avons abordées au chapitre 2.

3.3.3 Répercussions du double aspect de l'intention

Un point important est de revenir au **PB3** que nous avons évoqué concernant l'expérience de LIBET. Dû à ce problème, en prenant en compte le modèle du double aspect de l'intention que nous avons détaillé, les conclusions de l'expérience elles-mêmes pourraient être remises en question. En effet, LIBET ne demandait qu'une chose aux agents qu'il testait : "bougez le doigt quand vous le voulez". Par cette simple instruction nous pouvons considérer qu'il ruine une partie de ce qu'il prétend tester. Car, à quel moment peut-on considérer que l'agent effectue une action volontaire ? Quand ce dernier accepte l'instruction ou quand il effectue l'action réellement ? Ne peut-on pas supposer que l'agent souhaite simplement respecter l'instruction, et produit ainsi une intention préalable qui causera l'intention en action ? Ici, il se pourrait que l'intention distale soit la porteuse du caractère libre de l'action. Le potentiel observé par LIBET ne correspondrait qu'au processus cérébral menant de l'intention préalable vers la spécialisation en une intention en action qui produirait elle même l'action. Oui ce mécanisme serait, pour partie, inconscient mais il ne serait pas absurde de l'observer apparaître avant la conscience. Toute personne n'étant pas

dualiste reconnaît forcément que ce type d'activité doit être vu. Le problème de l'interprétation de LIBET c'est qu'elle reste sur une pensée naïve et toute puissante de l'esprit sur l'action volontaire : "je décide consciemment d'une action et je la fais dans l'instant". En conséquence, elle ne teste que l'intention en action en supposant que pour être libre l'action doit avoir une origine quasi "magique".

Pour étayer ce point il y a une expérience, que nous n'avons pas détaillée dans le Chapitre 2, qui a été mise au point par HAGGARD et EIMER en 1999 [HE99]. Comme dans l'expérience de LIBET, c'est le potentiel de préparation moteur que l'on observait. Mais, cette fois, il était demandé aux agents testés de choisir de bouger d'index gauche ou l'index droit. Le potentiel de préparation moteur peut être scindé dans ce cas en deux signaux : un signal dit ipsilatéral et un signal dit contralatéral. Le premier correspond à l'activité électrique dans l'hémisphère du même côté que la main réalisant l'action, le second du côté opposé à l'action. Par exemple, si l'action est exécutée par la main droite, c'est le signal de l'hémisphère gauche (signal contralatéral) qui est plus prononcé, et inversement si l'action est exécutée de la main gauche. Ce que cette expérience montre c'est, qu'au départ, les deux signaux composant le potentiel de préparation moteur se suivent. En revanche, environ une demi-seconde avant l'action, les signaux commencent à se séparer. On pourrait donc proposer que, avant cette demi-seconde, l'action future n'est pas spécialisée. En revanche ensuite elle se spécialise pour donner soit action à droite, soit action à gauche. On aurait alors la représentation typique d'une intention préalable, non spécifique, qui se transforme en intention en action qui spécifie le geste. La conscience, que dit avoir le sujet de ce qu'il va faire, serait alors le point culminant où le processus de séparation est arrivé à maturité. Cela ne dit rien vis-à-vis de l'intention préalable qui pourrait être le véritable support du caractère libre de l'action. Voilà qui vient encore confirmer cette interprétation de la double intention.

En quoi l'expérience de SOON *et al.* produit-elle aussi une amélioration ? Premièrement, comme dans l'expérience de HAGGARD et EIMER, elle introduit une notion de choix qui n'était pas présente dans le cas de l'expérience de LIBET. Deuxièmement, l'IRM fonctionnelle permet d'observer bien plus précisément la dynamique du cerveau et sur des temps beaucoup plus longs. Or que voit-on ? Que des processus se référant à l'action à venir sont visibles bien avant l'action. Une activité neuronale inconsciente semble être porteuse d'une information qui va participer à la détermination du choix : main gauche ou main droite. Ce que montre cette expérience c'est qu'une forte part inconsciente agit bien avant l'action, avant que les agents ne ressentent le "urge to move". Mais si cette sensation est produite par un processus proche de l'intention en action, comme nous l'avons supposé, peut-on pour autant qualifier l'action de non-libre parce que prédictible ? Encore une fois, la distinction introduite par SEARLE mitige cette affirmation. Tout aurait pu se jouer lors de l'intention distale. D'ailleurs, on parle ici d'action très simple. Tirer des conclusions de ces résultats à toutes les actions paraît de toute façon périlleux. Comme l'écrit aussi MELE, à propos de l'expérience de LIBET :

« Generalizing from results obtained in this domain to, for example, a view about distal decisions made about important issues in situations of a very different kind would be extremely bold, to say the least. » [Mel02]

Pour avoir une information plus précise, il faudrait pouvoir étudier de la même façon tout le processus menant à l'action. Et, plus particulièrement, un processus plus complet possédant une intention préalable. De notre point de vue, le moment de la réflexion, de la délibération, est ce qui

semble le plus important relativement à une action dite libre. Il est un moment autrement plus complexe.

A présent, que dire de l'expérience de WEGNER à propos de cette distinction ? Le problème de l'interprétation que fait WEGNER de ses résultats c'est, comme nous l'avons dit à plusieurs reprises, la position du "tout ou rien". Que son expérience montre des phénomènes inconscients que l'on croit maîtriser ne fait aucun doute. En tirer la conclusion que toute conscience n'a aucune répercussion dans l'action est audacieux pour reprendre les mots que MELE employait à propos de l'expérience de LIBET. A priori, à aucun moment, WEGNER ne parle de ce type de phénomène concernant les intentions préalables. A ce propos, MELE écrit aussi :

« If Wegner intends his "illusion" thesis to apply even to distal intentions, he has done little to support that application. » [Mel02]

Nous pouvons donc conclure de ces résultats que l'expérience de WEGNER et WHEATLEY montre l'importance énorme jouée par l'inconscient dans les processus d'action. En ce sens, ce type de travaux est louable pour contribuer à éliminer l'hypothèse dualiste des substances. Cependant, conclure de ces résultats que toute conscience est absolument inopérante vis-à-vis des actions humaines nous semble être un pas trop grand à franchir tant que les processus de construction des intentions conscientes préalables n'ont elles pas été étudiées dans le détail.

Justement les derniers travaux que nous avons évoqués, ceux de 2018, semblent faire une avancée sur ce terrain. En effet, l'utilisation de la "skip response", employée par KHALIGHINEJAD *et al.*, permet d'observer un processus de délibération dans toute sa complexité. Ici, le choix n'était plus simpliste, ou d'un type binaire, mais il s'agissait pour les individus de peser le pour et le contre d'une situation complexe : allait-ils passer à l'essai suivant au risque d'avoir un gain moindre ou allait-ils faire preuve de patience ? C'est donc ici toute une stratégie qui était mise en place par les individus. C'est ce qui pourrait se rapprocher le plus de ce que l'on qualifie d'action volontaire. Cette action concernerait les agents personnellement, et aurait des répercussions à la fin de la session d'expérience. Nous pensons que ce type d'expérience, pour étudier l'action volontaire, est remarquable. Nous proposons donc que c'est ce type d'approche qu'il faut adopter si l'on souhaite approfondir les connaissances concernant l'action volontaire libre. Que montrent les résultats ? Principalement, ils permettent de faire la distinction entre deux modèles concurrents :

1. **Modèle de la décision tardive (late decision) :** Nous en avons un peu parlé concernant l'hypothèse du veto de LIBET. Ce modèle propose que l'action se déroule de manière inconsciente et que l'agent décide tardivement d'autoriser l'action à se produire.
2. **Modèle de la décision précoce (early decision) :** Ce modèle propose que le processus de décision commence très tôt dans la préparation de l'action.

Cette expérience montre qu'elle serait plutôt en accord avec un modèle précoce. En effet le signal que HAGGARD appelle "signal de volition"⁷, qui est constitué d'une réduction du bruit neuronal, semble commencer à croître bien avant l'action. Tout se passe comme si le cerveau, en ébullition,

7. Propos de Patrick HAGGARD, lors de la remise du Prix Jean-Nicod le mardi 17 mai 2016 : <http://www.institutnicod.org/seminaires-colloques/conferences-et-prix-jean-nicod/archives/article/prix-et-conferences-jean-nicod-1378>

possédait un potentiel de préparation moteur bruité et que ce signal venait calmer ce potentiel en le figeant. Le point important c'est que ce résultat est encore plus présent dans les actions auto-initiées que dans les actions déclenchées. On peut donc bien parler ici d'un signal reflétant un processus délibératif. Dans cette expérience tous les types d'intentions nous semblent donc bien pris en compte. Ce type de schéma est donc d'un intérêt majeur pour les recherches futur sur la question.

3.4 Conclusion

Dans ce chapitre nous avons tenté d'apporter un éclairage philosophique des interprétations des expériences menées en Neurosciences concernant l'action volontaire. Ce que nous avons voulu démontrer, dans un premier temps, c'est que l'hypothèse dualiste des substances n'était aujourd'hui plus tenable. Pour ce faire, nous nous sommes basés sur ce que nous avons appelé "mur explicatif". Cette idée consistait à dire que les sciences, de part leur mode d'appréhension du monde, se retrouveraient face à un mur explicatif si un processus métaphysique jouait un rôle dans le fonctionnement d'un phénomène. Ne constatant pas que ce problème se pose dans le cas de l'étude de l'action volontaire, nous en avons conclu que l'hypothèse du dualisme des substances semble s'éloigner de plus en plus de la boîte à outil théorique permettant d'expliquer l'action. Un problème apparaissait cependant. Ce problème consiste en une décohérence entre les pensées philosophiques actuelles et les conceptions neuroscientifiques de l'action volontaire. En effet, le dualisme semble adopté par nombre de neuroscientifiques lorsqu'ils conçoivent le schéma de l'action libre. Souvent, ils étendent d'ailleurs cette idée à tout le concept de Libre Arbitre. Ne constatant rien qui pourrait faire penser cette forme de dualisme, ils affirment que l'action libre, comme le Libre Arbitre, n'est qu'une illusion. Un exemple typique est constitué des arguments basés sur l'expérience de LIBET *et al.*

Un autre type de dualisme était lui par contre confirmé par les sciences du cerveau. Ce dualisme est un dualisme des propriétés et prend la forme d'un épiphénoménalisme des états mentaux sur les états physiques. Dans ce cas de figure, la conscience d'un acte ne jouait plus aucun rôle dans le déroulement d'une action. Une expérience représentait particulièrement bien cette position. Cette expérience était celle menée en 1999 par WEGNER et WEATLEY. Dans le second temps de ce chapitre nous avons tenté de démontrer que cette dernière position ne semblait pas être tenable non plus. Pour ce faire nous avons proposé plusieurs types d'arguments :

1. **Les autres résultats scientifiques** : d'autres résultats neuroscientifiques semblaient aller à l'encontre de la thèse de WEGNER (ceux de FIED *et al.*).
2. **Critique du "tout ou rien" de WEGNER** : selon nous l'interprétation de WEGNER qui consistait en une idée de "tout ou rien" est abusive.
3. **La théorie de l'évolution contre l'épiphénoménalisme** : l'argument repose sur l'idée que si la théorie de l'évolution fonctionne, ce que nous avons aujourd'hui toutes les raisons de croire, alors il est incompréhensible que l'évolution ait doté les humains de cette conscience

de l'action si celle-ci n'avait aucune utilité évolutive.

4. **L'idée naïve de l'intention** : l'argument qui consiste à dire que c'est la notion d'intention naïve, proposée par WEGNER, qui est fautive dans son interprétation. En conjonction avec le second argument et en distinguant deux types d'intentions nous avons montré qu'il était impossible de prouver, en utilisant les résultats de WEGNER, que toutes les intentions conscientes étaient inopérantes sur le monde.

Enfin, nous avons terminé ce chapitre en évoquant la dernière expérience que nous avons présentée au chapitre 2. Celle-ci nous semble extrêmement prometteuse, dans son protocole, pour au moins deux raisons. Premièrement, via le principe de la "skip response" c'est un processus de délibération complexe qui est étudié. Ce processus contient donc à la fois la notion d'intention préalable et la notion d'intention en action. Cela le différencie grandement des études basées sur des instructions du type : "appuyez quand vous le désirez", "appuyez quand vous ressentez un besoin impérieux de mouvement", "appuyez à droite ou à gauche"... Ici une véritable délibération, basée sur des raisons pour ou contre, apparaît. La seconde illustration de l'aspect prometteur de cette expérience est qu'elle utilise des modélisations mathématiques afin de rendre compte de ses résultats. L'apparition de ces modèles, fonctionnant de façon plutôt impressionnante, est un plus qui permet d'avoir un nouveau regard sur la question.

En conclusion ce qui ressort de ce chapitre c'est une position dualiste, qu'il s'agisse d'un dualisme des substances ou d'un dualisme des propriétés, concernant l'action volontaire, ne nous semble pas tenable dans l'état actuel des connaissances. Pour le premier cas, il semble que cette position soit totalement obsolète au vue des résultats contemporains des Neurosciences. Pour le second cas, la conclusion semble partielle. Oui des processus inconscients sont fortement présents dans l'action volontaire. Il est même probable que cela soit le cas d'une partie des intentions en action. Non, il n'est pas possible d'en conclure que toutes les intentions conscientes ne jouent aucun rôle dans le processus menant à une action.

Conclusion

Dans cette étude, nous avons choisi de nous poser la question de la pertinence de l'approche des sciences contemporaines pour éclairer un débat philosophique d'ampleur, à savoir, la question du Libre Arbitre. Que pouvons-nous tirer de ce travail ?

Dans un premier temps, nous pouvons dire que les sciences physiques ne semblent pas adéquates à éclairer significativement le débat. Bien que ces dernières donnent des arguments susceptibles d'être utilisés par les différentes thèses philosophiques concernant la question, dans une perspective pragmatique, il semble que seule une étude systématique du cerveau soit vraiment fructueuse. Cette étude peut apporter des arguments aux tenants des différentes thèses : libertarianisme, compatibilisme et déterminisme "dur". Constatant ce fait, nous nous sommes concentrés sur les Neurosciences dès le second chapitre.

Dans un second temps, des protocoles expérimentaux ont été décrits. Ces protocoles avaient pour ambition de présenter les découvertes impliquant le problème de l'action volontaire. Cela permettait de réduire notre champ d'investigation tout en conservant une clef de voûte vis-à-vis de toute la problématique du Libre Arbitre. De plus, autre argument de poids, le processus de l'action volontaire est très étudié par les Neurosciences. Notre description des protocoles a commencé avec l'expérience princeps concernant l'action volontaire qui a eu lieu en 1983 et qui a été réalisée par Benjamin LIBET. Au total, 5 expériences ont été décrites se distribuant temporellement homogènement sur près de 40 années. Nous avons alors pu constater que les résultats de ces expériences défiaient souvent la pensée commune que l'on pouvait avoir concernant la question de l'action volontaire.

Dans un troisième temps, il s'agissait d'analyser les résultats que nous venions de décrire et d'essayer de leur trouver une place dans un modèle philosophique cohérent. D'abord, nous avons éliminé une position qui nous a semblé totalement obsolète suite aux résultats obtenus par les Neurosciences. Le dualisme des substances s'est ainsi vu évincer de la boîte à outils théorique dont dispose la philosophie de l'action. Ces résultats apportaient en revanche une autre vision qui suggérait que l'intention consciente, détectée par les sujets, n'était en fait qu'un épiphénomène de l'action. Cette idée était particulièrement étayée par l'expérience de WEGNER et WEATLEY de 1999. Cette proposition porte en elle une position de dualisme des propriétés et nous l'avons aussi mise de côté, du moins provisoirement, car elle nous semblait trop radicale à la vue des résultats obtenus. Qu'un doute puisse être émis quant à l'efficacité des intentions conscientes sur l'action dans certaines circonstances, cela semble évident après les expériences décrites au chapitre 2, que cela soit le cas de tout les types d'intentions pose un véritable problème.

Grâce à une précision, proposée par John SEARLE, qui consiste en la distinction d'au moins deux types d'intentions, nous avons vu que les expériences neuroscientifiques pouvaient être inter-

prêtées d'une autre manière qu'une approche purement épiphénoménaliste. En effet, nous avons proposé que ces expériences se concentraient sur des intentions conscientes de type proximales et qu'elles n'abordaient que très peu les intentions conscientes de type distales. Que des intentions proximales conscientes ne jouent pas de rôle sur l'action semble envisageable mais qu'aucune intention distale consciente n'en joue semble improbable. L'argument le plus marquant contre l'idée de l'épiphénoménalisme des intentions étant l'argument se centrant sur la loi de l'évolution des systèmes biologiques.

Pour conclure, il apparaît que si l'on adopte le point de vue de la séparation des intentions, les sciences contemporaines, et plus particulièrement les Neurosciences, ont un grand rôle à jouer pour préciser le fonctionnement du processus d'action volontaire. Nous avons d'ailleurs fait remarquer que la dernière expérience que nous avons décrite est très intéressante, de part son protocole, dans la mesure où le processus qu'elle étudie semble prendre en considération tous les types d'intention. Continuer à donner de nouvelles réponses scientifiques à ce type de questionnement est donc primordial pour péreniser les vocations inspirante et modératrice des sciences. Ce fait est particulièrement vrai concernant des problématiques qui sont d'un intérêt majeur vis-à-vis de la question philosophique que constitue le problème du Libre Arbitre.

Bibliographie

- [App14] Krystèle Appourchaux. *Un nouveau libre arbitre. CNRS Editions*, 2014.
- [BA02] Robert C. Bishop and Harald Atmanspacher. The causal closure of physics and free will. In R. Kane, editor, *The Oxford Handbook of Free Will*, chapter 5, pages 101–111. Oxford University Press, Oxford, 2002.
- [Bis02] Robert C. Bishop. Chaos, indeterminism, and free will. In R. Kane, editor, *The Oxford Handbook of Free Will*, chapter 4, pages 84–100. Oxford University Press, Oxford, 2002.
- [CTL07] Claude Cohen-Tannoudji and Diu Bernard Laloë, Franck. *Mécanique Quantique Tome I. Hermann*, 2007.
- [Don69] Franciscus Cornelius Donders. Over de snelheid van psychische processen. *Nederlandsch Archief voor Genees- en Natuurstudie*, (4), 1869.
- [Don69] F.C. Donders. On the speed of mental processes. *Acta Psychologica*, 30 :412 – 431, 1969.
- [Esf12] M. Esfeld. *La philosophie de l’esprit : Une introduction aux débats contemporains*. Philosophie. Armand Colin, 2012.
- [FKM⁺91] I Fried, Amiram Katz, G McCarthy, K.J. Sass, P Williamson, S.S. Spencer, and D D Spencer. Fried i, katz a, mccarthy g.functional organization of human supplementary motor cortex studied by electrical stimulation. *J Neurosci* 11 :3656-3666. 11 :3656–66, 12 1991.
- [Gaz13] Michael Gazzaniga. *Le Libre Arbitre et la science du Cerveau*. Science. Odile Jacob, 2013. trad. fr : Pierre Kaldy.
- [GS07] Joshua I. Gold and Michael N. Shadlen. The neural basis of decision making. *Annual Review of Neuroscience*, 30(1) :535–574, 2007.
- [Hag08] Patrick Haggard. Human volition : towards a neuroscience of will. *Nature Reviews Neuroscience*, 9 :934–946, Dec 2008.
- [HE99] P. Haggard and Martin Eimer. On the relation between brain potentials and the awareness of voluntary movements. *Experimental Brain Research*, 126(1) :128–133, Apr 1999.
- [Hel50] Hermann von. Helmholtz. Vorläufiger bericht über die fortpflanzungsgeschwindigkeit der nerventhätigkeit. *Archive of the Berlin Brandenburgische Akademie der Wissenschaften*, pages 541–543, 1850.

- [Hod02] David Hodgson. Quantum physics, consciousness, and free will. In R. Kane, editor, *The Oxford Handbook of Free Will*, chapter 3, pages 57–84. Oxford University Press, Oxford, 2002.
- [KD65] H.H. Kornhuber and L. Deecke. Hirnpotentialänderungen bei willkurbewegungen und passiven bewegungen des menschen : Bereitschaftspotential und reafferente potentiale. 284 :1–17, February 1965.
- [KSD⁺18] Nima Khalighinejad, Aaron Schurger, Andrea Desantis, Leor Zmigrod, and Patrick Haggard. Precursor processes of human self-initiated action. *NeuroImage*, 165 :35 – 47, 2018.
- [LGWP83] Benjamin Libet, Curtis A. Gleason, Elwood W. Wright, and Dennis K. Pearl. Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential) the un conscious initiation of a freely voluntary act. *Brain*, 106(3) :623–642, 1983.
- [Lor72] E. N. Lorenz. Predictability : Does the Flap of a Butterfly’s Wings in Brazil Set off a Tornado in Texas? *American Association for the advancement of science*, 130th meeting, 1972.
- [LS04] Gilles Lafargue and Angela Sirigu. La volonté d’agir est-elle libre ? *Cerveau et Psycho*, (6), 2004.
- [Mel02] Alfred R. Mele. Free will and science. In Robert Kane, editor, *The Oxford Handbook of Free Will*, chapter 26, pages 499–514. Oxford University Press, Oxford, 2002.
- [Mic11] Cyrille Michon. *Qu’est-ce que le libre arbitre ?* Vrin, 2011.
- [MWW99] Daniel M. Wegner and Thalia Wheatley. Apparent mental causation : Sources of the experience of will. 54 :480–92, 08 1999.
- [OF02] Hiroshi Okamoto and Tomoki Fukai. Neural bases of accumulator models. 2002(2) :400–402, February 2002.
- [OH04] Sukhvinder S. Obhi and Patrick Haggard. Free will and free won’t : Motor activity in the brain precedes our awareness of the intention to move, so how is it that we perceive control? *American Scientist*, 92(4) :358–365, 2004.
- [Pac08] Elisabeth Pacherie. The phenomenology of action : A conceptual framework. *Cognition*, 107(1) :179 – 217, 2008.
- [PH10] Elisabeth Pacherie and Patrick Haggard. What are intentions? In L. Nadel & W. Sinnott-Armstrong, editor, *Conscious Will and Responsibility. A tribute to Benjamin Libet*, 7, pages 70–84. Oxford University Press, 2010.
- [Poi90] H. Poincaré. Sur le problème des trois corps et les équations de la dynamique. *Acta mathematica*, 13 :1–270, 1890.
- [Pos05] Michael I Posner. Timing the brain : Mental chronometry as a tool in neuroscience. *PLOS Biology*, 3(2), 02 2005.
- [Pé12] François Pépin. Le déterminisme, le mot et les concepts. In Editions Matériologiques, editor, *le déterminisme entre sciences et philosophie*, pages 17–36. 2012.
- [San13] Alain Sans. Libre arbitre et neurosciences cognitives. 2013.

- [SBHH08] Chun Siong Soon, Marcel Brass, Hans-Jochen Heinze, and John-Dylan Haynes. Unconscious determinants of free decisions in the human brain. *Nature Neuroscience*, 11 :543–545, Apr 2008.
- [Sch02] Henning Schmidgen. Of frogs and men : the origins of psychophysiological time experiments, 1850–1865. *Endeavour*, 26(4) :142 – 148, 2002.
- [Sch18] Aaron Schurger. Specific relationship between the shape of the readiness potential, subjective decision time, and waiting time predicted by an accumulator model with temporally autocorrelated input noise. 5 :ENEURO.0302–17.2018, 01 2018.
- [SDC⁺03] Angela Sirigu, Elena Daprati, Sophie Ciancia, Pascal Giraux, Norbert Nighoghossian, Andres Posada, and Patrick Haggard. Altered awareness of voluntary action after damage to the parietal cortex. *Nature Neuroscience*, 7 :80, Nov 2003.
- [Sea85] John R. Searle. *L’Intentionnalité - Essai de philosophie des états mentaux*. Les Editions de Minuit, 1985. trad. fr : Claude Pichevin.
- [Sea04] John R. Searle. *Liberté et Neurobiologie*. Nouveau Collège de Philosophie. Grasset & Fasquelle, 2004. trad. fr : Savidan, Patrick.
- [SH06] Hiroshi Shibasaki and Mark Hallett. What is the Bereitschaftspotential? *Clinical Neurophysiology*, 117(11) :2341 – 2356, 2006.
- [SSD12] Aaron Schurger, Jacobo D. Sitt, and Stanislas Dehaene. An accumulator model for spontaneous neural activity prior to self-initiated movement. 109(42) :E2904–E2913, 2012.
- [Wal02] Henrik Walter. Contributions of neuroscience to the free will debate : From random movement to intelligible action. In R. Kane, editor, *The Oxford Handbook of Free Will*, chapter 27, pages 515–529. Oxford University Press, Oxford, 2002.