université de BORDEAUX

THÈSE

PRÉSENTÉE POUR OBTENIR LE GRADE DE

DOCTEUR DE

L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE Sciences et Environnements

SPÉCIALITÉ : Épistémologie et histoire des sciences

par Marie PENAVAYRE

Enjeux épistémologiques de l'usage judiciaire des neurosciences : une analyse du pouvoir explicatif de la neurocriminologie

sous la direction de Thomas Boraud et Pascal Duris

Soutenue le 27 novembre 2020

Membres du jury :

Thomas BORAUD, directeur de recherche au CNRS, Université de Bordeaux	Co-directeur
Cédric BRUN, maître de conférences, Université Bordeaux Montaigne	Examinateur
Céline CHERICI, maîtresse de conférences HDR, Université de Picardie – Jules-Verne	Rapporteur
Steeves DEMAZEUX, maître de conférences, Université Bordeaux Montaigne	Examinateur
Sonia DESMOULIN-CANSELIER, chargée de recherche au CNRS, Université de Nantes	Examinatrice
Pascal DURIS, professeur des universités, Université de Bordeaux	Co-directeur
Luc FAUCHER, professeur titulaire, Université du Québec à Montréal	Examinateur
Marie GAILLE, directrice de recherche au CNRS, Université Paris Diderot	Présidente

Résumé

Enjeux épistémologiques de l'usage judiciaire des neurosciences :
Une analyse du pouvoir explicatif de la neurocriminologie

Ce travail de thèse propose une analyse épistémologique des études publiées dans le domaine de la neurocriminologie, vaste programme de recherche mené dans l'objectif de redéfinir le problème de la criminalité en termes neurobiologiques, et qui vise à isoler les spécificités cérébrales des individus représentant une menace pour la société. L'objectif principal est d'évaluer la légitimité scientifique des ambitions annoncées par les neurocriminologues, c'est-à-dire de déterminer si leurs résultats garantissent une explication pertinente pour apprécier la dangerosité d'un individu. Il s'agit à la fois de s'interroger sur la possibilité empirique d'une explication neurobiologique du comportement criminel et de déterminer si l'on peut donner un sens à l'ambition de fonder l'évaluation du risque de récidive sur des critères neurobiologiques.

Le premier niveau d'analyse propose une reconstruction historique de la neurocriminologie, depuis les premières études électroencéphalographiques (EEG) démontrant un lien entre la présence d'anomalies cérébrales et la manifestation d'un comportement violent, impulsif ou agressif. À partir d'une analyse conceptuelle et méthodologique de la littérature publiée depuis le début des années 1940, je montre que ce programme de recherche est entièrement structuré autour d'une logique neuro-interventionniste qui vise à promouvoir une stratégie de pathologisation et de médicalisation du criminel. Le principal obstacle épistémologique réside dans le fait que cette logique contraint la construction du modèle explicatif au prix d'une série de pétitions de principe : les finalités poursuivies par les neurocriminologues introduisent un biais systématique dans la construction des modèles expérimentaux et dans l'interprétation des données recueillies chez les criminels.

Le second niveau d'analyse propose de resituer le projet neurocriminologique dans l'histoire de la phrénologie et de l'anthropologie criminelle du XIX^{e} siècle. Si la neurocriminologie s'efforce de défendre le caractère inédit de ses recherches et de dissimuler les stigmates de son long héritage, l'analyse conceptuelle et méthodologique de la littérature montre qu'elle est condamnée à réactualiser des présupposés hérités de

l'organologie de Franz-Joseph Gall (1758-1828) et de la criminologie positive de Cesare Lombroso (1835-1909). Cette analyse permet en particulier de déconstruire le double discours de la neurocriminologie, laquelle se défend de négliger l'influence des facteurs socio-environnementaux dans le développement de la criminalité, tout en cherchant à confirmer un modèle préformationniste admis au préalable : l'interprétation des données recueillies chez les criminels vise à montrer que leur dangerosité s'exprime au travers d'une potentialité préexistante, d'un état cérébral porteur d'une anomalie particulière, et qui détermine toutes les conditions d'émergence du comportement criminel. Il en résulte une démarche parfaitement circulaire, condamnée à produire des biais de confirmation en vue de promouvoir les finalités poursuivies par les chercheurs, sans apporter aucune explication des phénomènes psychologiques étudiés.

Mots-clés

Neurocriminologie ; Neurodroit ; Neurodisciplines ; Neurosciences ; Droit pénal ; Cerveau criminel ; Cesare Lombroso ; Franz-Joseph Gall ; Adrian Raine

Summary

Epistemological issues of the use of neuroscience in the legal system :
An analysis of the explanatory power of neurocriminology

This thesis work proposes an epistemological analysis of publications related to the field of neurocriminology, a vast research program that aims to redefine the problem of crime in neurological terms, and which proposes to isolate the neurobiological characteristics of individuals representing a threat to society. The main objective is to evaluate the scientific legitimacy of the ambitions announced by neurocriminologists, that is, to determine whether these results ensure a relevant explanation to assess the dangerousness of an individual. The aim is both to question the empirical possibility of a neurobiological explanation of criminal behaviour, and to determine whether we can assign a meaning to the ambition to base the risk assessment on neurological criteria.

The first level of analysis proposes a historical reconstruction of neurocriminology, starting from the first electroencephalographic studies (EEG) that demonstrated a link between brain abnormalities and violent, impulsive or aggressive behaviour. Based on a review of the literature published since the early 1940s, I show that this research program is entirely built around a neuro-interventionist strategy, which aims to promote the pathologization and medicalization of criminals. The main epistemological obstacle lies in the fact that this logic constrains the construction of the explanatory model at the cost of a series of petitions of principle: the implicit purposes of the researchers introduce a systematic bias in the design of experimental models and in the interpretation of data collected on criminals.

The second level of analysis proposes to resituate this research program in the history of 19th century criminal anthropology. While neurocriminology seeks to defend the novelty of its research and to conceal the marks of its long legacy, the conceptual and methodological analysis of the literature shows that it is condemned to update presuppositions inherited from the organology of Franz-Joseph Gall (1758-1828) and the positivist criminology of Cesare Lombroso (1835-1909). In particular, this analysis allows us to deconstruct the double discourse of neurocriminology, which defends itself from neglecting the influence of socio-environmental factors in the development of criminality, while seeking to confirm a previously accepted preformationist model : the

interpretation of the data collected from criminals aims to show that their dangerousness is expressed through a pre-existing potentiality, which satisfies all the conditions for the emergence of criminal behaviour. The result is a perfectly circular approach, condemned to produce confirmation biases in order to promote the purposes of the researchers, without providing any explanation of the psychological phenomena they study.

Keywords

Neurocriminology ; Neurolaw ; Neurodisciplines ; Neuroscience ; Criminal law ; Criminal brain ; Cesare Lombroso ; Franz-Joseph Gall ; Adrian Raine

LE CHEF DE LA PÈGRE

Tu dois être mis hors d'état de nuire.

HANS BECKERT

Mais je ne peux pas m'en empêcher !
Je n'y suis pour rien ! Ce n'est pas de ma faute...

UN HOMME

Eh, on la connaît celle-là, on est tous innocents dans un tribunal !

HANS BECKERT

Qu'est-ce que tu en sais, qu'est-ce que tu racontes, qui es-tu d'abord ? Qui êtes-vous, tous ? Des criminels ! [...] Mais moi... est-ce que je peux faire autrement ?
Je porte en moi cette chose maudite...
Ce feu, cette voix, ce supplice...

LE CHEF DE LA PÈGRE

Tu prétends... que *tu dois* assassiner ?

HANS BECKERT

Quelque chose me pousse à errer dans les rues... Je sens que quelqu'un me suit sans arrêt – c'est moi-même – qui me poursuit, sans bruit, mais je l'entends quand même ! Oui, j'ai parfois l'impression de me poursuivre moi-même ! Je voudrais m'enfuir, mais je ne peux pas m'échapper ! Cette force qui me pousse... Je veux résister, je veux résister !

[...]

L'AVOCAT DE LA DÉFENSE

Mon honorable collègue [...] prétend que mon client qui cède à une force irrésistible ne mérite que la mort. Il se trompe. C'est précisément cette contrainte qui l'acquitte.

***M le maudit*, Fritz Lang, 1931**

Table des matières

Remerciements 1
Introduction 3

CHAPITRE 1 : L'ANTHROPOLOGIE CRIMINELLE AU XIX^e SIÈCLE 22

1.1. Franz-Joseph Gall et la localisation cérébrale des instincts individuels 26
1.1.1. Le paradigme de la phrénologie de Gall 26
1.1.2. Une cérébralisation des facultés 29
1.2. Les thèses de l'innéité et de la pluralité 31
1.2.1. Le cerveau comme « organe exclusif » des facultés 31
1.2.2. La pluralité des facultés correspond à une pluralité cérébrale 34
1.3. L'application de l'organologie à l'étude du phénomène criminel 38
1.3.1. La naturalisation de la morale et du crime 38
1.3.2. Une volonté de repenser le système pénal 42
1.4. Lombroso et l'étude anthropologique du criminel 43
1.4.1. La spécificité de la démarche lombrosienne 46
1.4.2. De la naturalisation du crime à l'étude anthropométrique du criminel 51
1.4.3. Une hiérarchisation des « types criminels » 56
1.5. La nature pathologique du criminel 59
1.5.1. Le criminel est un « fou moral » 59
1.5.2. L'hypothèse de « l'épilepsie larvée » 63
1.6. L'ancrage du crime dans le corps du criminel 66
1.6.1. La dynamique de « l'instinct criminel » 66
1.6.2. L'incorrigibilité du criminel-né 74

CHAPITRE 2 : VERS UNE DÉFINITION NEUROPHYSIOLOGIQUE DE LA DANGEROSITÉ 80

2.1. Des troubles du comportement pendant l'enfance à la psychopathie 84
2.1.1. La délinquance juvénile et le soupçon d'une origine pathologique 84
2.1.2. L'hypothèse d'une « immaturité corticale » chez le psychopathe 86

2.1.3. Des « anomalies constitutionnelles » chez le criminel psychopathe 90
2.2. La structuration d'un programme autour de trois principes 94
2.2.1. Une connexion directe du cerveau au comportement 94
2.2.2. Un modèle psychopathologique 95
2.2.3. Un principe d'identité psychophysiologique 96
2.2.4. Une finalité prédictive 98
2.3. La construction d'un « cerveau dangereux » 100
2.3.1. « Le prisonnier normal a un EEG normal » 101
2.3.2. Une approche prototypique de la criminalité 104
2.4. Les figures de la dangerosité criminelle 106
2.4.1. Le criminel qui commet des crimes violents sans raison 106
2.4.2. Le criminel violent et récidiviste 109
2.4.3. Le criminel présentant des « tendances agressives » 111
2.4.4. Le criminel souffrant de troubles mentaux ou de la personnalité 116
2.5. Le (criminel) normal et le (criminel) pathologique 119
2.5.1. La définition neurophysiologique de la dangerosité 119
2.5.2. Un projet normatif 121
2.5.3. Une preuve de l'irresponsabilité pour cause psychiatrique ? 123
2.6. L'hypothèse de l'épilepsie 128
2.6.1. Un « EEG paroxystique » comme symptôme d'une « épilepsie latente » .. 130
2.6.2. Une forte prévalence de l'épilepsie dans les populations carcérales 135
2.6.3. L'epilepsy defense ou « *not guilty by reason of epilepsy* » 139

CHAPITRE 3 : IMAGERIE CÉRÉBRALE ET « VISION NEUROCRIMINOLOGIQUE » 144

3.1. Un domaine de recherche vaste et hétérogène 148
3.1.1. Délimitation du champ 148
3.1.2. Sélection de la littérature 151
3.2. Revue de la littérature 154
3.3. La « vision neurocriminologique » 176
3.3.1. L'effacement du social 176
3.3.2. Trois stratégies éliminativistes 179
3.3.3. Quel modèle pour quelle finalité ? 183
3.3.4. Un « cerveau-type » pour chaque type de criminalité 186

3.4. La problématique du crime violent 189
3.4.1. Une classification dichotomique 189
3.4.2. Un criminel « biologiquement anormal » 193
3.5. Les nouvelles figures de la dangerosité 198
3.5.1. Le criminel psychopathe 198
3.5.2. Le pédocriminel 202
3.6. Le problème de l'opérationnalisation 206

CHAPITRE 4 : CERVEAU « MONSTRUEUX » OU CERVEAU « MALADE » ? 212

4.1. Une terminologie médicale 216
4.1.1. Un cerveau défaillant, déficient, irrégulier 216
4.1.2. Une double qualification 218
4.1.3. La question de la validité empirique de l'interprétation pathologique 220
4.2. Un modèle normatif et non descriptif 222
4.2.1. Le « *language of disorder* » 222
4.2.2. Un soupçon de raisonnement *ad hoc* 225
4.2.3. La construction d'un « cerveau à corriger » 228
4.3. Le « *neuroenhancement* » comme idéal de réhabilitation ? 234
4.3.1. La « neuroaugmentation » morale 234
4.3.2. La « neurocorrection » du comportement criminel 240
4.3.3. Quels enjeux éthiques ? 244
4.4. La question du pouvoir explicatif de la stratégie neuro-interventionniste 250
4.4.1. Le modèle mécaniste de l'explication 250
4.4.2. Le modèle interventionniste de James Woodward 255
4.4.3. Un interventionnisme explicatif ou médical ? 259

CHAPITRE 5 : RECONSTRUCTION HISTORIQUE DE LA CONCEPTION (NEURO)BIOLOGISANTE DU CRIMINEL 268

5.1. Le paradigme de la recherche en neurocriminologie 273
5.1.1. Cinq hypothèses explicatives 273
5.1.2. Le criminel ressemble à son cerveau 277
5.1.3. Les prémisses de la filiation 280

5.2. L'adoption d'un postulat innéiste 285
5.2.1. Trois conceptions de l'innéisme criminel 285
5.2.2. Le projet localisateur de la neurocriminologie contemporaine 289
5.2.3. Une « nouvelle phrénologie » ? 292
5.2.4. Un retour à la dynamique de l' « instinct criminel » 297
5.2.5. Une force irrésistible mais corrigible 301
5.3. Les échecs du modèle intégratif 306
5.3.1. Un double discours à l'égard de l'influence environnementale 306
5.3.2. L'éternel retour des thèses de Gall et Lombroso 310
5.4. Pourquoi le criminel *n'est pas* son cerveau 314
5.4.1. Etudier le cerveau *seul* ne suffit pas 316
5.4.2. Les cerveaux ne commettent pas de crimes 323
5.4.3. Les raisons de dépasser le rejet de principe 329

Conclusion 336
Annexes 339
Bibliographie 345

Remerciements

J'adresse d'abord toute ma reconnaissance à mes directeurs de thèse, Pascal Duris et Thomas Boraud, pour m'avoir donné l'opportunité inouïe de travailler sur un sujet aussi passionnant. Merci d'avoir contribué à affiner ce manuscrit par vos remarques avisées et vos relectures attentives et minutieuses. Je remercie également le laboratoire SPH (Sciences, Philosophie, Humanités) pour son soutien financier, sans qui ce travail n'aurait définitivement pas pu voir le jour.

Mesdames et Messieurs les membres du jury, Cédric Brun, Céline Cherici, Steeves Demazeux, Sonia Desmoulin-Canselier, Luc Faucher, Marie Gaille, je vous remercie de m'avoir fait l'honneur de participer à l'examen de ce travail de thèse.

Ma gratitude s'adresse tout particulièrement à Cédric Brun, qui a joué un rôle décisif dans mon éveil pour la philosophie des neurosciences en 2014, à l'issue de ma formation en biologie. Je te dois une reconnaissance immense pour ton investissement, ton soutien constant et rassurant, et je ne peux que saluer ta grande exigence intellectuelle qui m'a fait l'effet d'un *neurostimulant* tout au long de ces quatre années de recherche.

Je remercie Luc Faucher de m'avoir accueillie dans son département de philosophie à l'Université du Québec à Montréal (UQAM) en 2017. Votre patience et votre grande générosité ne font que raviver le regret de ne rendre compte que très partiellement de ces quatre mois de recherches menées dans le champ du neurodroit. Je comprends aujourd'hui qu'un travail de recherche est un chemin erratique qui trompe souvent nos attentes et nos hypothèses initiales...

J'ai également une pensée pour Pierre Poirier et Daniel Kostic, dont les séminaires respectifs sur les Sciences Cognitives et les « *epistemic norms in theories of explanation* » coïncidaient parfaitement avec les voies de recherche que j'empruntais pendant cette période.

Certains questionnements n'auraient pas pu être résolus sans l'aide d'Owen Jones, l'actuel directeur du « *Research Network on Law and Neuroscience* », qui m'a aidée à mieux comprendre les enjeux qui caractérisent le projet du neurodroit. Je remercie aussi Valerie Hardcastle pour sa bienveillance et le temps précieux qu'elle m'a accordé lors de sa venue à Bordeaux.

Je dois aussi beaucoup aux étudiants en troisième année de la Licence MIASHS de l'Université de Bordeaux, dont la curiosité et la vivacité intellectuelle m'ont définitivement convaincue que l'épistémologie avait un rôle à jouer sur le terrain empirique, dans la pratique même des neurosciences cognitives.

Sur un plan plus personnel, je tiens à remercier tou.te.s mes ami.e.s qui m'ont permis de déloger temporairement cette thèse de mon esprit, dans la joie et l'insouciance : Eloïse, avec qui j'ai eu les échanges intellectuels les plus stimulants ces dix dernières années ; mes ami.e.s d'enfance Raphaëlle, Maëva, Éric et Nassim ; mes très chers compagnons de route de Paris et d'ailleurs, Coline, Elsa, Chloé, Marine, Louis, Laura, Pierre, Charles et Shiran, qui ont su trouver le juste équilibre entre l'exigence du travail solitaire et *l'impérieuse* nécessité de m'extirper de ce dialogue en tête à tête avec mon cerveau. Je remercie Line pour sa présence lumineuse et sa finesse d'esprit grâce à qui j'ai tant appris ; Franck et Zina pour leurs rires aux éclats et leur immense talent pour ravir mes oreilles ; et enfin Judith, que j'appellerais volontiers « mon agent » : merci pour la juste dose de confiance, de persévérance et de sérénité que tu m'as insufflée pendant ces derniers mois de rédaction.

Je remercie ma voisine Suze pour ses questions toujours pertinentes, qui m'ont amenée à structurer mes idées et à creuser de nouvelles pistes de réflexion.

Je salue affectueusement mon frère Hélios, et je dédie cette thèse à ma mère, qui a accompagné chaque *mot* par la *pensée*...

Introduction

Les progrès réalisés par les neurosciences cognitives dans la compréhension des mécanismes de la décision ou du contrôle de soi ont jeté le trouble sur certaines notions fondamentales du droit pénal. Depuis le début des années 1990 dans le milieu anglo-saxon, on observe une multiplication des publications relatives aux applications judiciaires des neurosciences. En 2011, la très influente *MacArthur Foundation* a lancé un programme de recherche de grande ampleur, le « *Research Network on Law and Neuroscience* » (RNLN), qui mène une réflexion interdisciplinaire sur l'apport potentiel des neurosciences dans le système pénal américain. Ce programme de recherche constitue encore à ce jour le meilleur indicateur institutionnel de l'émergence du neurodroit (« *neurolaw* ») aux Etats-Unis.[1] Il s'est principalement développé autour de la possibilité d'utiliser les données d'imagerie cérébrale pour statuer sur la responsabilité d'un prévenu (Aharoni et al. 2008 ; Jones et al. 2013 ; Maoz and Yaffe 2015). Le système accusatoire anglo-saxon a en effet facilité le recours aux techniques de neuroimagerie comme l'imagerie par résonance magnétique fonctionnelle (IRMf) pour attester la preuve d'une lésion cérébrale et justifier ainsi une atténuation ou une exonération de la responsabilité pénale du prévenu. L'argument repose sur ce qui est communément appelé le « *my brain made me do it defense* » ou la défense « c'est pas moi, c'est mon cerveau » (Sternberg 2010). Il a favorisé un renouvellement des problématiques liées aux conditions requises pour bénéficier d'une irresponsabilité pour cause psychiatrique, et l'apparition d'une nouvelle jurisprudence sur l'introduction des données neuroscientifiques dans le système pénal américain (Denno 2015 ; Hardcastle 2015).[2]

[1] Ce réseau de recherche regroupe une quinzaine de membres actifs, juristes et de neuroscientifiques issus de prestigieuses universités comme *Yale* ou *Columbia*, mais dont les principales activités (colloques, séminaires, etc.) sont organisées par la *Vanderbilt University Law School,* à Nashville. Précisons que le RNLN constitue le prolongement du *Law and Neuroscience Project*, lancé en 2007 par la *McArthur Foundation*, et regroupant globalement le même réseau de chercheurs.

[2] Aux Etats-Unis, Deborah Denno (2015) a recensé 800 affaires pénales ayant eu recours aux neurosciences entre 1992 et 2012. Dans la majorité des cas, les données neuroscientifiques sont invoquées comme une preuve à décharge de l'accusé, en complément de l'expertise psychiatrique et lorsqu'une peine capitale est susceptible d'être prononcée.

Globalement, les recherches publiées dans le domaine du neurodroit sont menées dans l'objectif de pallier les insuffisances de l'expertise psychiatrique en matière pénale, jugée trop subjective et source de contradictions. La possibilité d'isoler les régions cérébrales impliquées dans la capacité à contrôler son comportement a par exemple ouvert de nouvelles perspectives sur l'apport potentiel des modèles neuroscientifiques du contrôle de soi, pour supplanter le *Control Test* traditionnellement utilisé dans le cadre de la procédure d'irresponsabilité pour cause psychiatrique (Redding 2006) : les techniques de neuroimagerie permettraient de déterminer si le prévenu répond à la condition dite « volitionnelle » de l'*Insanity Defense*, c'est-à-dire d'évaluer l'intégrité des régions cérébrales impliquées dans la capacité à contrôler son comportement.

En parallèle de ces recherches, on voit apparaître une série d'études qui se développent dans un espace disciplinaire relativement autonome par rapport au champ du neurodroit. Cette voie de recherche, qui définit le champ de la « neurocriminologie » (« *neurocriminology* »), vise à utiliser les techniques de neuroimagerie comme un instrument de connaissance ou de prédiction du comportement criminel. Elle bénéficie d'une attention croissante depuis le début des années 1990, car les études s'accordent à montrer l'existence d'un lien entre certaines anomalies cérébrales et la manifestation d'un comportement violent, agressif ou impulsif : le cerveau des criminels serait caractérisé par des déficits d'activité, des atrophies ou des irrégularités dans certaines structures cérébrales (e.g. Raine et al. 1994 ; Aigner et al. 2000 ; Yang et al. 2005 ; Leutgeb et al. 2015).

Cette biologisation du phénomène criminel n'est pas nouvelle. Elle s'inscrit dans une longue tradition qui s'enracine dans la phrénologie de Franz-Joseph Gall (1758-1828) et la criminologie positiviste de Cesare Lombroso (1835-1909). Dès le début du XIX^{e} siècle, à la faveur des thèses de Gall sur la localisation cérébrale des instincts et des penchants individuels, mais surtout à partir de la publication par Lombroso de la première édition de *L'Homme criminel* en 1876, le crime est perçu comme une espèce naturelle que l'on peut isoler, mesurer et éventuellement prévenir. L'historienne de la criminologie Nicole Rafter (2008) explique la résurgence du discours biologique sur le crime par deux facteurs : « l'émergence d'une culture imprégnée de croyances sur les racines biologiques du comportement d'un côté, et de l'autre, un intérêt florissant pour la prévention de préjudices de tous types, depuis le cancer jusqu'au terrorisme en passant par la criminalité. » (Rafter 2008, p. 199).

La biologisation des comportements et la culture de la prévention et du contrôle ont effectivement fourni un terrain propice à la réactualisation de l'approche biologique du phénomène criminel (Larregue 2020). Ces dernières décennies, elle a surtout bénéficié d'un contexte socio-politique favorable à la multiplication des lois pénales sécuritaires pour résoudre la problématique du crime violent et de la récidive. Depuis les années 1980 aux Etats-Unis, les mesures pénales en matière de lutte contre la récidive imposent une « tolérance zéro » pour les criminels souffrant de troubles sévères de la personnalité.[3] L'idéal préventif a ainsi favorisé le recours à des techniques innovantes pour reconnaître les signes d'une « dangerosité criminologique » et identifier les marqueurs biologiques d'un comportement à risque (Eastman and Campbell 2006).[4]

On situe généralement la neurocriminologie et le neurodroit dans le prolongement du « tournant neuroscientifique » qui a pris place au début des années 1990 (Littlefield and Johnson 2012). La résurgence du discours neurobiologique sur le crime s'inscrit dans un contexte où la fonction descriptive ou explicative des neurosciences cognitives s'est très rapidement doublée d'une fonction normative. La décennie du cerveau a en effet favorisé un déplacement du champ d'application des techniques de neuroimagerie fonctionnelle, d'une utilisation limitée au domaine clinique ou scientifique, à un usage prescriptif visant à construire des catégories normatives exploitables à des fins sociales, juridiques ou politiques. C'est dans ce contexte qu'ont émergé une vaste série de « neurodisciplines », disciplines hybrides qui attestent du pouvoir alloué au cerveau dans l'explication de nos comportements : la neuroéconomie (e.g. Glimcher 2003), la neuroéducation (e.g. Battro et al. 2010), le neuromarketing (e.g. Zurawicki 2010), la neuroéthique (e.g. Marcus 2002), ou encore la neurophilosophie (e.g. Churchland 1986). Le cerveau devient l'objet d'étude privilégié pour résoudre un large ensemble de

[3] En France, une procédure similaire a été adoptée en février 2008, dans le cadre de la loi relative à la déclaration d'irresponsabilité pénale pour cause de trouble mental. Largement controversé dans la communauté des juristes, le principe de la rétention de sûreté permet le placement en centre de soins des individus ayant exécuté leur peine mais présentant un risque élevé de récidive ; il concerne en particulier les auteurs de crimes sexuels souffrant d'un trouble sévère de la personnalité (voir à ce sujet Mbanzoulou et al. 2008 ; Danet 2008 ; Senon 2008).

[4] On oppose habituellement la « dangerosité psychiatrique » (celle du criminel porteur d'une maladie mentale) à la « dangerosité criminologique » (l'individu dangereux qui ne relève pas de la pathologie psychiatrique). La définition de celle-ci reste problématique du point de vue du droit pénal car elle recouvre une large gamme de troubles de la personnalité qui peuvent entretenir une certaine confusion avec la maladie mentale (Bourgeois and Bénézech 2001 ; Danet 2008 ; Senon et al. 2009).

problèmes relevant traditionnellement des sciences humaines et sociales, suscitant des craintes légitimes à l'intérieur même de la communauté scientifique. Cette rhétorique de la promesse contribue à fragiliser le discours neuroscientifique car il suscite des attentes déraisonnables et entretient une grande confusion sur ce que les données d'imagerie cérébrale peuvent décrire, expliquer ou prédire (Gonon, Konsman, et Boraud 2014).

Il est important d'insister sur le fait que ces neurodisciplines, au même titre que la neurocriminologie ou le neurodroit, ne peuvent se réduire à une simple extension du champ disciplinaire constitué par les neurosciences cognitives (Dupont 2015). Si elles sont perçues comme un phénomène symptomatique de la place accordée aux neurosciences dans nos sociétés, elles s'inscrivent dans des logiques qui leur sont propres et leur histoire institutionnelle, leur formation disciplinaire et leur statut épistémologique ne peuvent être appréhendés sans une analyse approfondie de leurs conditions d'émergence et du cadre théorique ou méthodologique dans lequel les recherches sont conduites. La question de leur pouvoir explicatif recoupe certains enjeux propres au champ de la philosophie des neurosciences, car elles invitent à déterminer les conditions pour lesquelles la neuroimagerie fournit des données exploitables à des fins de connaissance (Forest 2014). En outre, s'il existe de bonnes raisons d'être sceptique à l'égard du pouvoir explicatif des neurosciences, on peut également adopter une position sceptique vis-à-vis du « neuroscepticisme » et considérer que le problème ne concerne pas les neurosciences en tant que telles, mais l'exportation de leurs outils ou de leurs données hors de leur contexte d'élaboration.[5]

Sans perdre de vue leurs spécificités méthodologiques, on peut considérer que ces neurodisciplines se sont construites autour d'un même présupposé : le neuroessentialisme ou l'idée selon laquelle « je suis mon cerveau » (Roskies 2002 ; Racine, Bar-Ilan, and Illes 2005).[6] Ce présupposé définit l'idée qu'en étudiant le cerveau de l'individu, on étudie l'individu lui-même, dans toute sa complexité et ses dimensions (subjectives, intersubjectives, socio-écononomiques, etc.) : le consommateur, le sujet apprenant, comme le sujet criminel, relèveraient du domaine des neurosciences car leur cerveau

[5] Selon la définition proposée par le philosophe Denis Forest, le « neuroscepticisme » désigne toute attitude de l'esprit qui « interroge et met en doute la solidité, la portée ou l'innocuité de la connaissance que produisent les neurosciences » (Forest 2014, p. 13).

[6] Le neuroessentialisme fait écho aux analyses proposées par Alain Ehrenberg (2004) et Fernando Vidal (2005) sur l'avènement du « sujet cérébral » (voir aussi Castel 2010).

permettrait de rendre compte de l'ensemble de leurs manifestations comportementales, de leur subjectivité ou de leur identité personnelle. Le neuroessentialisme apparaît ainsi comme un présupposé théorique pour justifier l'intérêt exclusif porté au cerveau dans l'étude de nos comportements, et inscrire les techniques de neuroimagerie dans un nouveau régime de production de la preuve (Chamak et Moutaud 2014). Il s'accompagne en effet d'un autre présupposé, à savoir l'idée que les neuroimages (notamment les images IRMf) sont des « photographies du cerveau » et qu'elles permettent d'apporter la preuve objective et infaillible d'un phénomène intrinsèquement subjectif (Dumit 1999 ; Racine, Bar-Ilan, and Illes 2005 ; Roskies 2007).

C'est précisément ce double argument qui préside à la stratégie rhétorique du « *my brain made me do it defense* » : le neuroessentialisme conduit à adhérer à un nouveau dualisme « cerveau - reste du sujet » qui autorise un déplacement de la responsabilité de l'individu vers son cerveau. Il contribue à effacer les dimensions psycho-affectives et socio-économiques du crime, favorisant l'idée selon laquelle le cerveau *seul* suffit à expliquer le phénomène de la criminalité. Comment doit-on comprendre ce basculement théorique qui imprègne à la fois le discours neurobiologique sur le crime et les applications judiciaires des neurosciences ? Peut-on donner un sens à l'affirmation selon laquelle le sujet criminel (ou l'individu en général) est « d'essence cérébrale » ? Dans quelle mesure les données neuroscientifiques peuvent-elles informer les conditions de l'irresponsabilité pour cause psychiatrique, c'est-à-dire apporter la preuve d'une altération des capacités requises pour être tenu pénalement responsable de son acte ?

Jusqu'ici, et à notre connaissance, la question de l'usage judiciaire des neurosciences a été étudiée selon quatre approches. La première renvoie aux réflexions théoriques menées dans le champ du neurodroit lui-même, sur la question de l'apport des données neuroscientifiques pour informer la question de la responsabilité pénale. Il convient de préciser ce que recouvre cette question car certains chercheurs du *Research Network on Law and Neuroscience* (RNLN) sont eux-mêmes sceptiques quant à l'apport potentiel des neurosciences dans le système pénal. La seconde approche s'intéresse à l'influence exercée par les neuroimages sur une décision de justice, lorsqu'elles sont présentées comme une preuve d'atténuation ou d'exonération de la responsabilité pénale. La troisième est issue du champ de la sociologie des sciences et porte en particulier sur les modalités d'institutionnalisation des approches biosociales de la criminalité. On peut également mentionner une dernière approche, qui nous intéressera davantage car elle

puise ses arguments dans le domaine de la philosophie critique des neurosciences. Elle s'est développée dans la continuité des objections formulées à l'encontre de l'usage de la neuroimagerie pour étudier les phénomènes mentaux (e.g. Uttal 2001), et s'apparente en ce sens à un rejet de principe contre la prétention des neurosciences à investir le champ pénal.

La première approche s'est développée au tournant des années 2000, dans une dynamique de recherche interdisciplinaire regroupant des juristes et des psychologues cognitifs (Goodenough 2001 ; Waldbauer and Gazzaniga 2001 ; Bonnie 2002 ; Denno 2002 ; Morse 2003 ; Steinberg and Scott 2003). En posant les bases d'une réflexion théorique sur les interactions possibles entre les neurosciences et le droit pénal, ces recherches ont constitué l'amorce des premières investigations empiriques menées par le RNLN. Encore aujourd'hui, la plupart des enjeux tournent autour de la question de savoir sous quelles conditions la notion de responsabilité pénale pourrait être accessible à l'investigation neuroscientifique. Précisons que la question soulevée par les membres du RNLN est moins pragmatique qu'épistémologique, puisqu'ils emploient un certain effort pour résoudre les questions conceptuelles ou empiriques posées par l'introduction des neurosciences dans le système pénal : paradoxalement, la plupart des limites épistémologiques de ce projet proviennent du champ du neurodroit lui-même.

Force est de constater que certaines recherches publiées par le RNLN livrent une réflexion très fine sur la nécessité de proposer une clarification des concepts mobilisés dans le champ, afin de garantir la validité conceptuelle de leurs protocoles expérimentaux et d'éviter d'éventuelles erreurs inférentielles (e.g. Buckholtz, Reyna, and Slobogin 2016). Le juriste Stephen J. Morse, l'un des membres les plus actifs du réseau depuis sa création en 2007, a longuement discuté du problème de la naturalisation des états intentionnels (Morse 2007, 2010, 2015). Selon lui, le principal enjeu posé par l'intégration des neurosciences dans le système judiciaire est le problème de la traduction entre les concepts psychologiques mobilisés dans le droit pénal (par exemple, l'intention ou le contrôle de soi) et les concepts qui sont opérationnalisés puis traduits en termes neuroscientifiques. La question est de savoir s'il est possible de traduire en termes mécanistes un concept éminemment normatif qui s'intègre dans un environnement social complexe et qui dépend des catégories juridiques dans lesquelles il est invoqué.[7]

[7] Cette question ne concerne pas la notion de responsabilité pénale elle-même, mais les capacités requises pour être tenu pénalement responsable. Comme le précise Eyal Aharoni, un membre actif du RNLN, « il

Telle qu'elle est envisagée dans le droit pénal, la définition du comportement est elle-même une source de conflit conceptuel dans la mesure où les neurosciences fournissent des explications en termes de causes et non de raisons d'agir. Du point de vue du droit pénal américain, le sujet juridique est un être conscient, intentionnel, rationnel, c'est-à-dire doué d'une délibération consciente qui lui permet de contrôler son comportement en vertu de raisons morales (Morse 2006). Le problème de la naturalisation de la morale, de l'intentionnalité et du contrôle de soi fait donc partie intégrante des discussions portant sur l'introduction des neurosciences dans le système pénal, ce qui explique le fait que le RNLN s'est progressivement orienté vers une dynamique interdisciplinaire qui inclut des philosophes naturalistes comme Walter Sinnott-Armstrong ou Adina Roskies (e.g. Sinnott-Armstrong et al. 2008).

Précisons que le problème du libre arbitre pèse finalement peu dans les débats, non seulement parce que le concept n'est pas pertinent au regard de la notion juridique de responsabilité – le libre arbitre n'étant pas défini comme un critère d'attribution de la responsabilité pénale −, mais aussi parce que les chercheurs en neurodroit adoptent une position *compatibiliste*, considérant que le déterminisme causal est parfaitement compatible avec les notions de libre arbitre et de responsabilité (voir à ce sujet Morse 2007, 2015). De fait, les réflexions théoriques menées à l'intérieur du champ du neurodroit excluent la question de savoir si les neurosciences cognitives peuvent représenter une menace vis-à-vis des concepts de libre arbitre et de responsabilité morale ou pénale.[8] En revanche, la notion de contrôle de soi est centrale dans les débats, précisément parce que la capacité à contrôler son comportement est une condition requise pour être tenu pénalement responsable de son acte. La possibilité d'opérationnaliser

n'y a pas, et il n'y aura jamais, un corrélat neuronal de la responsabilité » (Aharoni et al. 2008, p. 145). Les chercheurs en neurodroit s'inspirent en effet de la conception dite « capacitarienne » (ou « volitionniste ») de la responsabilité, sur laquelle repose le système pénal américain (Hart 1968). Selon cette conception, l'attribution de la responsabilité repose sur la possession de certaines capacités, en l'occurrence, la capacité à contrôler son comportement et la capacité à apprécier la valeur morale de son acte.

[8] Cette question fait l'objet d'un vif débat depuis la publication d'une série d'études EEG ou IRMf qui invitent à reconsidérer le rôle de la volonté consciente dans la réalisation de nos actions (Libet et al. 1983, 1985 ; Wegner and Wheatley 1999 ; Soon et al. 2008, 2013). Ce débat est jugé peu fécond par les chercheurs en neurodroit car il soulève le problème de l'épiphénoménisme de la conscience, et non pas des états intentionnels en tant que tels : de leur point de vue, tant que les neurosciences n'ont pas montré que la délibération consciente et rationnelle ne joue aucun rôle causal dans la réalisation de nos actions, elles ne peuvent être considérées comme une menace vis-à-vis de la notion de responsabilité pénale (Morse 2015).

expérimentalement le concept (tel qu'il est défini dans le droit pénal) fait donc l'objet d'une discussion approfondie, la finalité étant de construire des modèles neuroscientifiques du contrôle de soi pour informer la condition volitionnelle de l'*Insanity Defense.*

S'il n'est pas dans notre propos de rendre compte de l'ensemble des enjeux discutés dans le champ du neurodroit, on peut enfin noter que des réflexions d'ordre technique ou méthodologique ont été développées par les membres du RNLN, notamment le problème de la variabilité interindividuelle entre les niveaux d'activité cérébrale − communément appelé le « *many-one problem* » (e.g. Jones et al. 2013) −, ou encore la difficulté d'appliquer sur des cas individuels des résultats obtenus sur des centaines d'individus et impliquant une multitude de paradigmes expérimentaux (e.g. Faigman, Monahan, and Slobogin 2013). Ce problème, nommé « *Group to Individual (G2i) Inference* », constitue l'une des difficultés majeures posées par l'intégration des neurosciences dans le système pénal, car il pourrait affecter la définition du seuil pathologique justifiant une irresponsabilité pour cause psychiatrique (voir à ce sujet Hardcastle 2020).

La deuxième approche concernant les problèmes soulevés par l'usage des neurosciences en contexte judiciaire s'intéresse aux éventuels biais interprétatifs induits par la présentation des neuroimages (notamment les images IRMf) dans le cadre d'une procédure d'irresponsabilité pour cause psychiatrique (Aspinwall, Brown, and Tabery 2012 ; Baker et al. 2017 ; Desmoulin-Canselier 2020). Elle fait écho à une longue série d'études en psychologie cognitive concernant l'influence exercée par les neuroimages sur une décision de justice, ou plus généralement, sur l'évaluation de la responsabilité d'un agent ayant commis un acte immoral (Gurley and Marcus 2008 ; Greene and Cahill 2012 ; Saks et al. 2014). Malgré quelques exceptions (Schweitzer and Saks 2011 ; Schweitzer et al. 2011), la majorité des études s'accordent à montrer que les neuroimages attestant la présence d'une lésion cérébrale introduisent un biais déresponsabilisant, même en l'absence de données relatives au profil psychiatrique ou psychologique du sujet. De manière générale, ces recherches s'inscrivent dans la continuité des réflexions initiées par l'anthropologue des sciences Joseph Dumit (1999, 2004) sur le postulat d'objectivité des neuroimages et l'influence persuasive qu'elles peuvent exercer par rapport à d'autres types de données (voir aussi McCabe and Castel 2008 ; Weisberg et al. 2008). Précisons que le RNLN s'est lui-même saisi du problème du statut juridique et des conditions

d'admissibilité de la preuve neuroscientifique, la question étant de savoir si les neuroimages pourraient à terme se substituer à l'expertise psychiatrique (e.g. Shen and Jones 2011 ; Steinberg 2013). En l'occurrence pour Joseph Dumit, le problème de la subjectivité de l'expertise subsiste lorsque les neuroimages sont présentées dans le cadre de la défense d'aliénation mentale, dans la mesure où l'expert se positionne en tant que « garant subjectif d'une preuve objective, en tant que témoin faillible d'un appareil infaillible. » (Dumit 1999, p. 186).

La troisième approche est issue de la sociologie des sciences et s'intéresse en particulier au champ de la criminologie biosociale (Fallin, Whooley, and Barker 2018 ; Larregue 2017, 2019, 2020). Celle-ci est définie comme la branche de la criminologie qui étudie « les effets de la génétique, de la physiologie, et des facteurs neurologiques, ainsi que les influences de la société et de la famille » (Eichelberger and Barnes 2015). Elle a fait l'objet d'une analyse sociologique très récente par Julien Larregue (2020), qui a examiné les facteurs institutionnels et structurels ayant favorisé son émergence aux Etats-Unis à partir des années 1940. Le cas de la neurocriminologie ne fait pas l'objet d'une analyse spécifique car il est étudié au travers des modalités de développement institutionnel de la criminologie biosociale. Précisons néanmoins que certains enjeux autorisent un recoupement entre les deux champs, notamment le phénomène de biologisation du social et le développement d'une prévention médicalisée du crime. Comme le montre Julien Larregue (2019), cette double stratégie s'inscrit dans la continuité logique des politiques d'économie pénale amorcées aux Etats-Unis dans les années 1990, dont les principales lignes directrices ont été formulées dans un article intitulé « *The New Penology : Notes on the Emerging Strategy of Corrections and Its Implications* » (Feeley and Simon 1992). Cette « nouvelle pénologie » promeut une gestion probabiliste du risque criminel : elle défend l'idée que le crime doit être traité selon une logique actuarielle, c'est-à-dire par une combinatoire de facteurs qui catégorise les individus dans des groupes « à risque » (voir à ce sujet Quirion 2006).

Enfin, la dernière approche s'inspire du champ de la philosophie critique des neurosciences. Elle ne relève pas à proprement parler de l'épistémologie des neurosciences, mais elle s'inscrit dans la continuité des critiques formulées à l'encontre de l'exportation des données d'imagerie cérébrale hors de leur contexte d'élaboration (voir à ce sujet Uttal 2001 ; Choudhury, Nagel, and Slaby 2009 ; Rose and Abi-Rached 2013 ; Vidal and Ortega 2017). C'est pourquoi la littérature critique sur l'usage judiciaire

des neurosciences s'articule autour du procès de ce qui est communément appelé la « nouvelle phrénologie », à savoir la réintroduction de l'idée selon laquelle les fonctions cognitives, affectives et comportementales sont localisées dans des régions individualisées du cerveau (Markowitsch 2008 ; Pustilnik 2009 ; Dartigues 2017).[9] Ici encore, cette critique n'est pas spécifique au champ de la neurocriminologie puisqu'elle prend pour cible toute tentative visant à intégrer les neurosciences dans le système pénal. On retrouve d'ailleurs le même emprunt à la philosophie critique des neurosciences dans un ouvrage consacré à l'analyse du champ du neurodroit, publié par les philosophes néo-wittgensteiniens Michael S. Pardo et Dennis Patterson (2010, 2013). Selon eux, ce projet s'épuise dans l'argument dit de « l'erreur méréologique » (ou de « l'erreur de catégorie »), tel qu'il a été proposé par Max Bennett et Peter Hacker (2003) dans le cadre de leur analyse critique des neurosciences cognitives : les neurosciences n'apporteraient aucune information pertinente vis-à-vis du droit pénal, car les critères d'attribution de la responsabilité se rapportent à des comportements et non à des états cérébraux.

Jusqu'ici, les recherches menées dans le champ de la neurocriminologie n'ont pas fait l'objet d'un traitement spécifique. Lorsqu'elles sont mentionnées dans la littérature secondaire, elles sont englobées dans le domaine général des applications judiciaires des neurosciences, et décrites comme le prolongement direct du projet du neurodroit. On observe d'ailleurs une confusion fréquente entre les deux champs, qui sont pourtant gouvernés par des logiques, des stratégies de recherche et des enjeux épistémologiques qui leur sont propres. Leur différence de statut tient en partie à l'appartenance disciplinaire de leurs principaux acteurs, qui révèle différentes manières de se saisir des problématiques liées aux applications judiciaires des neurosciences. En effet, le projet du neurodroit est essentiellement porté par des juristes, et non par des neuroscientifiques ; à l'inverse, la neurocriminologie s'est construite sous l'impulsion d'une petite communauté de psychiatres et de neurobiologistes qui ont progressivement investi le domaine de la criminologie pour repenser la compréhension du phénomène criminel.

[9] La référence à la « nouvelle phrénologie » est attribuée à William Uttal (2001), qui a formulé une vive critique à l'égard de l'usage de la neuroimagerie dans le champ des neurosciences cognitives. Selon lui, la démarche localisatrice qui sous-tend cet usage repose sur des hypothèses contestables, infondées ou erronées, dont la plupart sont héritées de l'organologie de Franz-Joseph Gall.

Cette orientation particulière a des conséquences sur le plan de la légitimité scientifique et du statut épistémologique de leur projet, compte tenu du fait que les études relatives au champ de la neurocriminologie ne sont pas publiées dans des revues de criminologie mais dans des revues de neurosciences cognitives, de neurosciences sociales ou de psychiatrie biologique (*NeuroImage* ; *Journal of Neuroscience* ; *Social Cognitive and Affective Neuroscience* ; *Biological Psychiatry,* etc.).[10] En l'occurrence, dans la mesure où la neurocriminologie mobilise les ressources théoriques et techniques des neurosciences cognitives, elle maintient une relative autonomie vis-à-vis du champ qu'elle a investi : elle peine à convaincre les criminologues, et ses partisans se situent plutôt du côté des neuroscientifiques réductionnistes et éliminativistes, lesquels proposent d'expliquer l'ensemble de nos comportements par des mécanismes neurobiologiques, dans l'objectif de redéfinir les problèmes sociétaux à la lumière de ces données.

Par ailleurs, on peut regretter que les critiques formulées à l'encontre de l'usage judiciaire des neurosciences ne soient pas suffisamment informées sur le plan empirique : il est rare que les analyses rendent compte du cadre méthodologique dans lequel les recherches sont menées. Cette limite peut être attribuée au fait que la plupart s'inspirent des réflexions issues de la philosophie critique des neurosciences cognitives, dont la méthodologie est elle-même centrée sur l'usage des techniques de neuroimagerie comme l'IRMf. Le procès de la « nouvelle phrénologie » est souvent invoqué comme un argument définitif pour déconstruire toute tentative visant à exporter ces techniques hors de leur champ d'application clinique ou scientifique. Il relève en ce sens davantage d'une critique « en bloc » des divers domaines d'application des neurosciences, que d'un examen empirique des conditions d'émergence et des spécificités méthodologiques du projet du neurodroit ou de la neurocriminologie : ce dernier est discrédité d'avance de par les origines historiques de la biologisation du phénomène criminel et la présomption d'une idéologie déguisée. Les fondements épistémologiques de cette filiation n'ont pourtant jamais été réellement démontrés, car les quelques recherches ayant étudié cette question échouent à fournir une analyse explicite des présupposés et des stratégies de recherche gouvernant les deux approches.

[10] Précisons également que certaines études ont été publiées dans des revues prestigieuses comme *Proceedings of the National Academy of Sciences* (Aharoni et al. 2013) ou *Nature Reviews Neurosciences* (Glenn and Raine 2014).

Il semble d'ailleurs que l'accusation de « nouvelle phrénologie » ait fini par masquer toute une série de recherches qui ne reposent pas sur l'usage de la neuroimagerie fonctionnelle. Les analyses critiques publiées à ce jour sous-estiment largement la profondeur historique de ce programme de recherche, lequel ne saurait se réduire à un phénomène symptomatique de l'« impérialisme neuroscientifique » qui se déploie depuis les années 1990 : l'émergence des techniques de neuroimagerie ne signe pas l'acte de naissance du projet du neurodroit ou de la neurocriminologie. Les premières tentatives visant à utiliser les outils d'exploration cérébrale comme un instrument de connaissance du sujet criminel remontent au début des années 1940, à la faveur d'un intérêt croissant des psychiatres pour l'étude électroencéphalographique (EEG) des troubles du comportement ou de la personnalité. De manière surprenante, aucune mention n'est faite de ces publications, aussi bien dans la littérature secondaire que dans la littérature en neurodroit ou en neurocriminologie elle-même. Ces études ont été manifestement oubliées dans l'histoire du champ disciplinaire, les origines du projet étant généralement situées au début des années 1990, c'est-à-dire à l'émergence des techniques de neuroimagerie fonctionnelle et aux premières publications signées par Adrian Raine, l'une des figures majeures de la neurocriminologie contemporaine (Raine, Venables, and Williams 1990 ; Raine et al. 1994).[11]

La question de la continuité historique entre ces projets est d'autant plus importante que le discours neurocriminologique entretient un rapport ambivalent à l'égard de ses influences et de son héritage théorique. Depuis ses premières études menées sur le cerveau des criminels au début des années 1990, Adrian Raine s'efforce de défendre le caractère inédit de ses recherches, tout en glissant quelques allusions à Cesare Lombroso dans ses interventions publiques. Dans son ouvrage de vulgarisation *The Anatomy of Violence : The Biological Roots of Crime* (2013), il imaginait une fiction dystopique dans laquelle les Etats-Unis se sont dotés d'un programme de détection de la criminalité appelé « LOMBROSO » (pour *Legal Offensive on Murder : Brain Research Operation for the Screening of Offenders*), combinant des facteurs de risque neurologiques, génétiques, psychologiques ou encore démographiques. Ce programme gouvernemental apparaît comme un idéal de prévention et de contrôle social puisqu'il

[11] Adrian Raine a largement contribué au développement de la neurocriminologie contemporaine, au travers d'une série d'études utilisant la TEP sur des criminels condamnés pour meurtre (Raine et al. 1994, 1997a, 1998a, 1998b).

permettrait d'assigner tous les hommes âgés d'au moins 18 ans dans la catégorie des « *Lombroso Negative* » ou des « *Lombroso Positive* », ces derniers étant placés en détention préventive jusqu'à ce qu'une normalisation de leurs données leur permette de réintégrer la société.

Il serait tentant d'y voir la revendication explicite d'une filiation si Adrian Raine ne s'appliquait pas à dissimuler les stigmates de cet héritage dans ses publications scientifiques : la littérature en neurocriminologie ne comporte aucune référence aux thèses de Cesare Lombroso. Les chercheurs se défendent de réactualiser la théorie du criminel-né et revendiquent une démarche scientifique inspirée de la méthodologie classique des neurosciences cognitives. Depuis la fin des années 1990, ils cherchent à asseoir la légitimité scientifique de leur projet en proposant de nourrir des réflexions propres au champ des neurosciences cognitives : l'usage de la neuroimagerie pour caractériser et localiser les anomalies cérébrales des criminels s'est très rapidement transformé en un projet d'identification des corrélats neuronaux du contrôle de soi, de la régulation émotionnelle ou de la formation des jugements moraux. Autrement dit, la neurocriminologie entend acquérir une reconnaissance institutionnelle sur le terrain des neurosciences cognitives, en montrant que sa portée épistémologique ne se limite pas à l'étude du phénomène criminel, mais ouvre des perspectives plus générales sur la compréhension des mécanismes neurobiologiques impliqués dans diverses fonctions cognitives.

Cette logique de légitimation révèle toute l'ambivalence du discours neurocriminologique, lequel entretient un certain flou quant aux finalités poursuivies. Contrairement au champ du neurodroit qui a clairement défini sa ligne directrice au travers du RNLN, la neurocriminologie n'affiche pas explicitement ses ambitions et la nature réelle de ses finalités se confond dans un double discours. Les neurocriminologues se montrent par exemple très prudents en ce qui concerne les implications de leurs résultats sur la question de la responsabilité pénale ou de l'évaluation du risque de récidive ; pourtant, la littérature laisse parfois entrevoir des conclusions hâtives qui dépassent largement la portée explicative des données empiriques. Plus encore, l'intérêt exclusif porté au cerveau des criminels semble s'effacer au travers d'une dynamique de recherche intégrative, qui tient compte de l'influence des facteurs socio-environnementaux de la criminalité : bien que les neurocriminologues s'abstiennent de toute référence à l'histoire personnelle ou aux caractéristiques socio-économiques des

individus, ils multiplient les stratégies rhétoriques pour désamorcer toute accusation de « cérébro-centrisme » et défendre la rigueur méthodologique de leur approche.

Dans cette perspective, il est crucial d'interroger la scientificité du projet neurocriminologique, en rejetant toute hostilité de principe à l'égard de l'apport potentiel des neurosciences dans la compréhension du phénomène criminel. Les efforts déployés par les neurocriminologues pour affirmer leur autorité scientifique nous invitent à dépasser la critique *a priori* pour questionner le pouvoir explicatif réel de leurs résultats. La neurocriminologie marque-t-elle véritablement une rupture épistémologique dans l'histoire de la biologisation du phénomène criminel ? Peut-on encore donner un sens à l'ambition d'expliquer le phénomène de la criminalité en étudiant uniquement le cerveau des criminels, en dépouillant le crime de sa dimension socio-économique ou de l'histoire personnelle de l'individu ? Comment expliquer la résurgence de ce discours biologisant sur le crime dans des revues scientifiques majeures, compte tenu de fait que cette approche s'inscrit dans une longue tradition de recherches dont la scientificité a été invalidée depuis plus d'un siècle ?

Si nous proposons de retracer la généalogie de ce programme de recherche et de questionner sa continuité avec les études phrénologiques ou criminologiques du XIX^{e} siècle, le problème que nous posons est bien épistémologique : il concerne la portée explicative et normative de la neurocriminologie, telle qu'elle se développe depuis les premières études EEG menées sur le cerveau des criminels au début des années 1940. Il s'agit donc à la fois de s'interroger sur la possibilité empirique d'une explication neurobiologique du comportement criminel, et de déterminer si l'on peut donner un sens à l'ambition d'utiliser ces données pour informer les questions relatives à l'évaluation de la responsabilité pénale ou de la dangerosité d'un individu.

Compte tenu de l'importance de ces enjeux vis-à-vis des notions fondamentales du droit pénal, il convient de porter notre analyse sur le terrain empirique, en prêtant une oreille attentive aux arguments avancés par les chercheurs pour justifier la pertinence scientifique de leur démarche. Notre parti pris méthodologique est donc de prendre au sérieux ce programme de recherche, en renonçant à un anti-naturalisme de principe et en examinant la rationalité interne de ce projet. La question de sa légitimité scientifique est indissociable d'une enquête empirique sur la construction des études menées depuis les années 1940, car elle invite à analyser la structure argumentative qui définit l'arrière-plan du projet d'intégration des neurosciences dans le système pénal. Nous proposons ainsi

d'examiner les stratégies de recherche mobilisées par les neurocriminologues, les présupposés qui sous-tendent les choix méthodologiques opérés, ou encore les concepts employés dans l'interprétation des données. Il s'agit globalement de comprendre dans quel univers conceptuel et méthodologique s'inscrit la neurocriminologie, mais aussi d'expliciter le type de programme anthropologique ou idéologique qu'elle adopte, compte tenu du fait que le discrédit jeté sur ce projet provient d'un héritage inavouable mais paradoxalement peu étudié du point de vue épistémologique.

À ce titre, il convient de lever une confusion éventuelle sur l'objet de notre analyse. Elle concerne bien le projet neurocriminologique (et dans une moindre mesure, le projet du neurodroit), et non pas les neurosciences en tant que telles. Plutôt que de s'inscrire dans le prolongement de la philosophie critique des neurosciences, cette analyse propose une réflexion inédite sur un usage très spécifique des outils neuroscientifiques, qui est gouverné par ses propres présupposés théoriques et méthodologiques. C'est pourquoi nous nous intéresserons seulement de manière périphérique aux grands enjeux contemporains en philosophie des neurosciences : nous n'entrerons pas dans le cœur des débats concernant la nature de l'explication neuroscientifique, le problème de la naturalisation des phénomènes mentaux, du neuro-réductionnisme et de l'éliminativisme, ou encore la question de l'apport épistémique des techniques de neuroimagerie dans l'étude des phénomènes mentaux. Ces enjeux seront abordés dans une perspective interne au projet neurocriminologique, c'est-à-dire en explicitant au préalable le contenu scientifique des publications, la nature des finalités poursuivies et la logique des stratégies de recherche privilégiées par les neurocriminologues.

Pour cette même raison, on ne trouvera pas dans ce travail une discussion sur les implications de leurs résultats sur les notions de libre arbitre ou de responsabilité morale. Bien que prometteuse, nous verrons que cette question ne peut se poser à l'intérieur du champ de la neurocriminologie et qu'elle prendrait le risque de tomber dans le procès d'intention : elle nécessiterait de sortir du cadre empirique et de mobiliser des ressources théoriques issues de la philosophie de l'action ou de la philosophie morale. De même, nous laisserons de côté les enjeux propres à la philosophie de la psychiatrie, notamment la question des recoupements que l'on pourrait observer avec le projet neuro-réductionniste qui sous-tend les nouvelles méthodes de classification des maladies mentales comme le *Research Domain Criteria* (RDoC) (Demazeux et Pidoux 2015 ;

Faucher et Goyer 2016).[12] L'ambition de fonder des catégories normatives sur des critères neurobiologiques, mais aussi la question de la médicalisation de la violence telle qu'elle est envisagée dans le champ psychiatrique pourraient très certainement nourrir nos réflexions, mais elles ne nous diraient rien sur la dynamique interne au projet neurocriminologique.

Enfin, et si cette remarque peut sembler évidente au regard de notre objet de recherche et de l'approche empirique que nous privilégions, précisons que nous n'évoquerons pas les enjeux proprement juridiques sur le statut de preuve des données neuroscientifiques, ni la jurisprudence relative à l'introduction des neurosciences dans le système pénal (voir à ce sujet Desmoulin-Canselier 2018, 2019, 2020). Nous n'aborderons pas non plus le débat qui émerge au sein de la philosophie du droit depuis le développement grandissant des neurosciences cognitives, sur l'abandon de la conception rétributive de la peine (centrée sur l'acte criminel et la culpabilité du sujet) au profit de la conception utilitariste ou réhabilitative, centrée sur la prévention du risque de récidive.[13] Ce débat s'adresse, ici encore, au champ du neurodroit, et nous engagerait dans une discussion métaphysique sur la question de savoir si les neurosciences peuvent représenter une menace vis-à-vis des notions de libre arbitre ou de responsabilité morale (voir à ce sujet Cova 2011). Pour autant, nous verrons que les recherches récentes en neurocriminologie ont contribué à renouveler la question de la réhabilitation criminelle, soulignant l'impérieuse nécessité d'engager un débat sur la place des neurosciences cognitives dans les futures procédures préventives.

Ce travail d'analyse s'emploie à reconstruire la généalogie de ce programme de recherche en examinant trois moments-clés dans l'histoire de son développement : l'émergence du projet au début des années 1940 à partir des premières études EEG menées dans des centres de détention pour mineurs et des institutions psychiatriques ou

[12] Précisons brièvement que le projet RDoC, lancé en 2009 par le *National Institute of Mental Health* (NIMH), s'oppose au Manuel diagnostique et statistique des troubles mentaux (DSM) et accorde une place très importante aux neurosciences dans la définition et la classification des maladies mentales, puisqu'il conceptualise les maladies mentales comme des désordres neurologiques.

[13] Selon certains neuroscientifiques comme Joshua Greene et Jonathan Cohen (2004), les neurosciences pourraient engager une modernisation des principes fondamentaux du système pénal, en transformant nos intuitions morales concernant le libre arbitre et la responsabilité et en favorisant une conception conséquentialiste (et non plus rétributive) de la peine (voir à ce sujet Penavayre, Brun et Boraud 2019).

carcérales en Grande-Bretagne et aux Etats-Unis ; la neurocriminologie contemporaine telle qu'elle se développe depuis le début des années 1990 autour des techniques de neuroimagerie comme la tomographie par émission de positons (TEP), l'imagerie par résonance magnétique anatomique (IRMa) ou fonctionnelle (IRMf), ou la tomodensitométrie (CT-scan) ; et enfin, les nouvelles perspectives de recherche amorcées au milieu des années 2000, qui reposent sur l'usage des techniques de neurostimulation comme la stimulation cérébrale profonde (SCP).

Le premier chapitre vise à retracer les conditions d'émergence de la biologisation du phénomène criminel, en procédant à un examen attentif des thèses de Franz-Joseph Gall et Cesare Lombroso. L'objectif n'est pas de reconstruire l'histoire de la phrénologie et de la criminologie positiviste italienne, les enjeux entourant l'anthropologie criminelle du XIX^{e} siècle ayant été largement explicités et débattus dans le domaine de l'histoire et de la sociologie des sciences.[14] Il s'agit plutôt d'isoler les grands principes théoriques et méthodologiques qui régissent l'étude du cerveau des criminels au XIX^{e} siècle dans l'objectif de déterminer, à terme, si ces principes sont réinvestis dans le champ de la neurocriminologie. Certains principes exigent en effet une analyse interne au système organologique de Gall et à l'anthropologie criminelle de Lombroso, indépendamment de ce que nous réservera l'examen proprement comparatif. Plutôt que de nous attarder sur les limites épistémologiques des études phrénologiques et criminologiques du XIX^{e} siècle, nous chercherons surtout à reconstruire la structure interne des thèses de Gall et Lombroso, en nous approchant au plus près du raisonnement qui guide leurs recherches respectives sur les causes biologiques du crime.

Le deuxième chapitre propose d'examiner les prémices du développement de la neurocriminologie, au travers des premières études EEG menées sur les criminels au début des années 1940. Ces études signent les premières tentatives visant à démontrer un lien entre la perturbation de l'activité cérébrale et la manifestation d'un comportement violent, agressif ou impulsif. Elles méritent une analyse approfondie des protocoles expérimentaux car les stratégies mobilisées tout au long des années 1940-1950 vont structurer durablement le champ de la neurocriminologie. Les questions que nous

[14] Les contributions sont nombreuses du côté de l'histoire des sciences (Lantéri-Laura 1970 ; Villa 1985 ; Renneville 1994, 1995, 1997, 2000, 2003 ; Rafter 2008). On peut également citer les analyses fournies par Laurent Mucchielli (1994, 2000), Martine Kaluszynski (2008), et Laurent Clauzade (2008a, 2008b) en ce qui concerne spécifiquement la phrénologie de Franz-Joseph Gall.

poserons dans ce chapitre nous donneront ainsi l'occasion d'amorcer une réflexion critique sur la logique qui sous-tend l'ambition de mesurer l'activité cérébrale des criminels. Comment s'est structuré le projet, à la fois sur le plan théorique et méthodologique, mais aussi du point de vue des finalités pratiques concernant les applications judiciaires potentielles de ces résultats ? Comment les psychiatres ont-ils interprété les anomalies observées chez des criminels ayant été déclarés pénalement responsables et qui ne présentaient donc aucun trouble psychiatrique ? Quelles hypothèses ont été avancées pour expliquer l'origine de ces anomalies ?

Dans le troisième chapitre, nous déplacerons notre analyse sur la période 1990-2010, pour examiner les études reposant sur l'usage des techniques de neuroimagerie comme l'IRMa, l'IRMf ou la TEP. Sur la base d'un corpus regroupant 62 études menées dans diverses institutions carcérales aux Etats-Unis mais aussi en Europe, nous poserons la question de la continuité théorique et empirique de ces recherches avec le projet neurocriminologique qui s'est développé sur la base de l'EEG. Cette question est cruciale car la littérature contemporaine a manifestement délaissé cinquante années de publications qui rapportent pourtant la même association entre les anomalies cérébrales et le comportement criminel. Sans y voir nécessairement un déni ou une dénégation en raison des limites techniques de l'outil lui-même − l'EEG présentant une résolution spatiale très faible par rapport à l'IRMf ou la TEP − on peut aussi admettre que les neurocriminologues manifestent une grande méconnaissance de la profondeur historique de leur programme de recherche. La question de l'apport épistémique des techniques de neuroimagerie nous amènera surtout à analyser la pertinence du projet *explicatif* de la neurocriminologie contemporaine, à savoir l'ambition d'expliquer le comportement criminel par la présence d'anomalies cérébrales. Ce projet ne peut être questionné sans être mis en perspective avec le double discours qui est entretenu par les neurocriminologues pour affirmer leur autorité scientifique et garantir la validité de leur modèle neurobiologique de la criminalité. Il nous donnera ainsi l'occasion de mesurer la cohérence interne du projet au regard des finalités poursuivies, des stratégies de recherche mobilisées dans la construction des protocoles expérimentaux, et des inférences interprétatives qui en découlent.

Nous poursuivrons cette analyse dans le quatrième chapitre. Celui-ci porte en particulier sur la stratégie de pathologisation et de médicalisation du criminel qui sous-tend l'interprétation des données recueillies depuis le début des années 1940. Cette stratégie se déploie discrètement dans la littérature, et participe en ce sens à l'ambivalence

du discours neurocriminologique. C'est pourquoi le problème qui nous occupera dans ce chapitre se situe sur le plan empirique et épistémologique : notre parti pris méthodologique étant de prendre au sérieux le projet neurocriminologique en nous intéressant à son contenu scientifique, nous procèderons à un examen attentif des arguments empiriques sur lequel repose l'ambition de traiter la criminalité comme un problème médical. La conception psychopathologique du criminel est-elle justifiée du point de vue des données neurobiologiques recueillies depuis les années 1990 ? Sous quelles conditions épistémologiques ces données pourraient-elles affecter les questions relatives à la prise en charge de la criminalité ? À bien des égards, les enjeux liés à l'approche médicale du crime recoupent certains débats propres au champ du droit pénal. Ils ont été largement discutés dans le domaine de la sociologie et de la philosophie de la psychiatrie, dans la lignée des travaux menés par Michel Foucault (1972, 1975) et Robert Castel (1976) sur la psychiatrisation du droit pénal. La logique qui sous-tend la médicalisation du criminel est ici très différente puisqu'elle promeut une complète déjudiciarisation du phénomène de la criminalité : elle vise à favoriser une extension du champ d'intervention de la neurocriminologie, d'un usage des outils neuroscientifiques comme un instrument de *connaissance* du sujet criminel à un instrument de *modification* de ses états cérébraux.

Enfin, le dernier chapitre vise à prendre la mesure de l'héritage de Franz-Joseph Gall et Cesare Lombroso dans la recherche en neurocriminologie. Nous chercherons d'abord à reconstituer le cadre conceptuel et méthodologique commun à l'ensemble des études publiées depuis les années 1940, avant d'expliciter chacun des termes de la filiation. Nous voudrions à cet égard montrer que cette continuité historique s'étend au-delà d'une simple entreprise de naturalisation du crime, et qu'elle s'affirme surtout dans les présupposés, les inférences interprétatives et la structure argumentative qui sous-tend le projet neurocriminologique. La question du pouvoir explicatif de ce programme de recherche constitue donc l'arrière-plan de cette reconstruction historique. Celle-ci nous donnera notamment l'occasion de désamorcer le double discours mis en œuvre par les neurocriminologues pour asseoir la légitimité scientifique de leurs recherches et justifier l'intérêt exclusif porté au cerveau dans l'étude du phénomène criminel.

CHAPITRE 1

L'ANTHROPOLOGIE CRIMINELLE AU XIX^{e} SIÈCLE

On peut difficilement retracer les origines historiques des recherches menées sur le cerveau des criminels sans faire référence à la phrénologie et à la criminologie italienne du XIX^{e} siècle. Le discrédit jeté sur la neurocriminologie provient d'un soupçon de déterminisme biologique qui serait hérité de Franz-Joseph Gall : les quelques recherches ayant exploré la question de cette filiation montrent que la neurocriminologie s'apparente à une « nouvelle phrénologie », et qu'elle relève en ce sens bien plus d'un « neuromythe » que d'une véritable avancée scientifique (voir en particulier Pustilnik 2009 ; Dartigues 2017, 2018). Dans l'objectif qui est le nôtre de questionner la continuité du projet neurocriminologique avec la phrénologie et de la criminologie du XIX^{e}, il convient de retracer les conditions d'émergence de la biologisation du phénomène criminel, en procédant à un examen attentif des arguments théoriques et empiriques sur lesquels elle repose.

La phrénologie de Franz-Joseph Gall et la criminologie de Cesare Lombroso signent l'élaboration d'un savoir scientifique sur le crime, le criminel et son cerveau.[15] Ce déplacement du regard, du crime vers le cerveau du criminel, va introduire une nouvelle perspective dans l'étude du phénomène criminel, dans un contexte où le crime était essentiellement étudié par le prisme de sa relation avec la folie (Renneville 2003 ; Llorca 2005 ; Kaluszynski 2008). Franz-Joseph Gall et Cesare Lombroso partagent l'idée que le crime est inné et que sa cause réside dans les dispositions organiques du criminel : dans le déséquilibre des fonctions cérébrales chez Gall ; dans les disproportions, les asymétries et les arrêts du développement chez Lombroso.

Le choix méthodologique de remonter jusqu'à la phrénologie de Gall peut se justifier sur deux plans. Le premier tient à la démarche localisatrice qui sous-tend sa recherche des causes biologiques du crime. Bien que la phrénologie ne constitue pas le point de départ de l'histoire des localisations cérébrales, Gall a été le premier à attribuer

[15] Précisons que le terme « phrénologie » n'est pas attribué à Franz-Joseph Gall mais à son disciple Johann Gaspar Spurzheim (1776-1832). Gall nommait sa doctrine « organologie », « physiologie du cerveau », ou « doctrine du cerveau ».

au cerveau un « rôle exclusif » dans l'exercice des facultés morales ou intellectuelles.[16] Il est devenu l'acteur central de la cérébralisation des facultés, en se positionnant en rupture par rapport au modèle sensationniste qui dominait à la fin du XVIII[e] siècle (Cherici et Dupont 2008) : selon lui, les facultés morales ou intellectuelles ne sont ni le produit de l'action des sens extérieurs, ni réductibles à la stricte influence du milieu ; elles naissent d'après des dispositions innées qui présentent leur siège dans des « organes cérébraux » distincts et indépendants.

Par ailleurs, la phrénologie de Gall tient une place centrale dans le discours biologique sur la criminalité car elle signe les premières tentatives visant à faire du cerveau des criminels l'objet d'une connaissance scientifique. Dès le premier volume de son traité fondateur *Sur les fonctions du cerveau et sur celles de chacune de ses parties* (1825), Gall va chercher à appliquer les principes fondamentaux de son organologie à l'étude du phénomène criminel. Ce projet de naturalisation du crime découle directement de son système organologique : le cerveau étant le siège de toutes les facultés morales, c'est dans sa désorganisation qu'il faut rechercher la cause matérielle du « penchant au meurtre ». Autrement dit, c'est en posant l'innéité des facultés morales que Gall va chercher à isoler l' « organe cérébral » responsable de leur altération.

L'anthropologie criminelle de Lombroso partage avec la phrénologie de Gall l'idée que pour comprendre le crime, il faut étudier le criminel. Le phénomène de la criminalité ne peut être compris sans une étude anthropologique du criminel lui-même, comme Lombroso le précise dans la préface de la cinquième édition de son œuvre maîtresse, *L'Homme criminel* : « [j]e crois [...] que pour réconcilier tant de divergences, pour résoudre le problème s'il existe ou non une vraie nécessité dans le crime, et si l'homme criminel appartient à un monde tout à lui, il fallait laisser de côté toutes les théories philosophiques et étudier, en somme plus que le crime, les criminels. » (Lombroso, 1895, p. 6).

Ce renversement de perspective acquiert une dimension supplémentaire dans la pensée lombrosienne. Si la phrénologie de Gall a favorisé la représentation d'un individu *naturellement* porté au crime, l'anthropologie criminelle de Lombroso a définitivement ancré le crime dans le corps du criminel : elle a imposé l'idée que le criminel est porteur

[16] Précisons que si ses thèses ont été développées dès 1790, ses contributions à l'anatomie et à la physiologie du cerveau seront connues par ses contemporains à partir de 1808, date de la publication de ses *Recherches sur le système nerveux en général et sur celui du cerveau en particulier*.

d'un danger qu'il convient de mesurer. Ainsi, au-delà d'une simple théorie sur l'innéité du crime, *L'Homme criminel* marque l'émergence d'une science nouvelle − l'anthropométrie du crime − qui fait du corps du criminel non plus seulement un objet de connaissance, mais un véritable instrument de mesure. La dangerosité devient un « fait scientifique », dans le sens où elle est conçue comme une donnée objective et quantifiable (Jay Gould 1983) : l'essentiel de l'anthropologie criminelle italienne consistera à quantifier la dangerosité des individus, en l'indexant sur le degré de leurs anomalies anatomiques ou morphologiques.

Afin de préciser au mieux les spécificités du système organologique d'une part, et de l'anthropologie criminelle d'autre part, nous avons fait le choix de circonscrire notre analyse en deux volets en examinant successivement les thèses de Gall (dans les trois premières sections) et de Lombroso (dans les trois suivantes). Nous présenterons en premier lieu le système général de la doctrine de Gall et le contexte d'émergence de son mouvement de cérébralisation des facultés (section 1.1), avant d'expliciter dans le détail ses deux thèses fondamentales, à savoir, la thèse de l'innéité des facultés et la thèse de la pluralité des facultés et des organes (section 1.2). L'une des particularités de sa démarche réside dans le fait qu'elle ne repose pas tant sur des arguments anatomiques, que sur une analyse psychophysiologique des différences individuelles : la psychologie différentielle – c'est-à-dire l'hétérogénéité et la variabilité des facultés morales ou intellectuelles – constitue le point de départ de son raisonnement (Forest 2008 ; Clauzade 2008a). Nous montrerons ainsi comment le système organologique de Gall a progressivement investi le champ moral (section 1.3), à la fois pour expliquer la grande variabilité des penchants et des instincts individuels, mais aussi pour reconnaître les signes caractéristiques d'une altération du sens moral.

Nous verrons que c'est dans ce contexte qu'a émergé la notion de « déséquilibre cérébral », c'est-à-dire la représentation d'un cerveau plus ou moins vulnérable aux facteurs socio-environnementaux. Cette notion est primordiale dans la compréhension de l'organologie de Gall, car elle montre l'influence mutuelle des facteurs organiques et socio-environnementaux dans l'expression du « penchant au meurtre ». Sans entrer dans le cœur du débat sur les applications sociales et juridiques de la doctrine de Gall, nous évoquerons ses perspectives théoriques en ce qui concerne la sanction pénale. Gall défend l'intérêt de ses recherches sur l'anatomie et la physiologie du cerveau pour éclairer les pratiques punitives et refonder le système pénal selon le principe de l'individualisation

des peines : selon lui, il ne faut plus juger le crime, mais le criminel lui-même, et la détermination de la peine doit être adaptée aux dispositions organiques de l'individu.

Les thèses de Lombroso exigent un examen plus attentif car elles ont fait l'objet de nombreux remaniements au cours des rééditions successives de *L'Homme criminel*, entre 1876 et 1895. Tout au long de notre analyse, nous allons passer en revue les différentes thèses qui définissent les contours de son anthropométrie criminelle : la « théorie atavique du crime », la théorie dite de « l'arrêt infantile » et la théorie de « l'identité du criminel-né, du fou moral et de l'épileptique ». Partant du fondement théorique de sa démarche anthropométrique – à savoir, l'ambition de retrouver les lois de l'atavisme dans le comportement des animaux – nous montrerons comment l'analyse statistique des données recensées chez les criminels a été mise au service d'une théorie de la classification et de la hiérarchisation des « types criminels » (section 1.4).

Nous analyserons ensuite les arguments mobilisés en faveur de la conception pathologique du criminel, avant d'expliciter l'hypothèse dite de l'« épilepsie larvée » (section 1.5). Cette hypothèse revêt une importance capitale dans l'anthropologie criminelle italienne, car elle a contribué à faire du crime violent l'expression d'une « force irrésistible ». Nous verrons ainsi comment Lombroso parvient à justifier sa théorie de l'innéité criminelle, en favorisant la représentation d'un criminel physiquement et moralement monstrueux, porteur d'un « instinct » intrinsèque, irrésistible, et par conséquent incorrigible.

Enfin, nous montrerons que cette dynamique morbide révèle un présupposé central dans la pensée lombrosienne, qui consiste à établir un lien entre le criminel et son crime (section 1.6). Ce présupposé sera mis en perspective avec le problème des crimes monstrueux, problématique emblématique du XIX^{e} siècle qui se déploie autour de la question de la rationalité du crime et de la qualification juridique du criminel. Notre analyse prendra comme point de départ la critique foucaldienne de l'expertise psychiatrique en matière pénale. Elle nous permettra de comprendre en quoi la dynamique de l' « instinct criminel » a favorisé l'ancrage du crime dans le corps du criminel, et l'ambition de rechercher les signes pathologiques d'une dangerosité en devenir. Mais surtout, cette analyse nous amènera à saisir toute la circularité des arguments mobilisés par Lombroso pour garantir la validité de son anthropométrie du crime : l'ambition de relier le crime à son auteur définit le cadre théorique dans lequel se déploient tous les arguments invoqués à l'appui de ses thèses.

1.1. Franz-Joseph Gall et la localisation cérébrale des instincts individuels

Si l'on veut peut prendre la mesure de l'héritage présumé de Gall dans la neurocriminologie contemporaine, il convient d'examiner dans le détail les fondements de sa doctrine. La raison est double : d'abord, il paraît difficile de saisir toute la complexité de son approche organique du crime, sans poser en premier lieu les principes théoriques sur lesquels elle repose. Ensuite, le souci de questionner sa continuité historique avec le projet neurocriminologique nous engage à restituer les grands axes de l'organologie, en s'approchant au plus près du raisonnement adopté par Gall. Nous entendons en effet dépasser la posture critique *a priori*, qui se dispense d'analyser les ressorts méthodologiques et conceptuels de cette filiation.

1.1.1. Le paradigme de la phrénologie de Gall

Les thèses de Gall ont été énoncées pour la première fois dans une lettre adressée au baron M.J.F de Retzer, et publiée en 1798,[17] avant d'être développées dans ses deux ouvrages écrits en collaboration avec Johann Gaspar Spurzheim (1776-1832), *Recherches sur le système nerveux en général et sur celui du cerveau en particulier* (1808) et *Anatomie et physiologie du système nerveux* (1810). Elles ont ensuite été reprises et précisées dans les six volumes de son traité *Sur les fonctions du cerveau et sur celles de chacune de ses parties*, publiés entre 1822 et 1825.[18]

[17] Gall, Franz-Joseph : Lettre ouverte au baron M.J.F de Retzer publiée dans : *Neuer deutscher Merkur, Weimar*, éd. C. Wieland, volume 3, livre 12, déc. 1798. Lettre traduite dans Fossati : *Questions philosophiques, sociales et politiques, traitées d'après les principes de la physiologie du cerveau*, Paris, Amyot, 1869, pp. 277-302.

[18] Dans la suite de notre travail, nous nous référerons aux six volumes de son traité *Sur les fonctions du cerveau* (1825), car il y reprend les principales thèses développées dans son *Anatomie et physiologie du système nerveux* (1810), tout en y développant les perspectives sociales et juridiques de sa doctrine.

Les deux premiers volumes de ce traité présentent les fondements théoriques de son organologie, notamment sa thèse de l'innéité des facultés morales et intellectuelles. Gall y défend l'importance des recherches sur l'anatomie et la physiologie du cerveau pour étudier l'hétérogénéité et la variabilité des penchants et des instincts individuels. Il insiste en particulier sur l'intérêt de ces recherches pour mesurer le perfectionnement ou le dérangement des facultés morales : le cerveau n'est pas seulement le siège des facultés ; il est conçu comme un véritable instrument de mesure pour évaluer leur degré de développement. Le troisième volume décrit dans le détail sa thèse de la pluralité, et les principes permettant de reconnaître les dispositions morales par la conformation du cerveau et du crâne. Enfin, les implications philosophiques, morales et juridiques de sa doctrine sont présentées dans les trois derniers volumes, notamment en ce qui concerne la possibilité de mesurer l'expression du « penchant au meurtre ».

La thèse centrale de Gall consiste à poser d'une part, l'innéité des facultés morales et intellectuelles, et d'autre part, la pluralité des fonctions et des « organes cérébraux » : toutes les facultés et les qualités sont des dispositions innées qui présentent leur siège dans des « organes cérébraux » distincts et indépendants. Ces deux thèses fondamentales constituent les fondements de la doctrine phrénologique. Celle-ci se résume en six principes de base, énoncés pour la première fois en 1798 :[19]

> 1/ Toutes les facultés et les qualités sont innées et ne peuvent se manifester indépendamment de conditions matérielles.
> 2/ Les facultés et qualités présentent leur siège dans le cerveau.
> 3/ Le cerveau est composé d'autant d'organes qu'il existe de facultés spécifiques.
> 4/ Les facultés présentent leur siège dans des parties du cerveau distinctes et indépendantes.
> 5/ Plus un organe est développé, plus la faculté qui lui correspond l'est aussi.
> 6/ La conformation extérieure du cerveau détermine celle de la surface interne du crâne ; il est donc possible d'évaluer le développement des facultés d'un individu en palpant sa boite crânienne.

Avant d'entrer dans le détail de ces principes, il paraît important de reconstituer le système général de la doctrine phrénologique. Celle-ci est fondée sur un double

[19] Gall, 1869, *op. cit*.

principe de correspondance entre les facultés et les « organes cérébraux » d'une part, et le cerveau et le crâne d'autre part. Le premier principe constitue le fondement de l'organologie, laquelle est divisée en une thèse dite « générale », qui définit la spécialisation fonctionnelle du cerveau, et une thèse « spéciale », correspondant à la distinction et à la localisation cérébrale de chacune des facultés. Gall opère en parallèle une décomposition de l'esprit en facultés distinctes et indépendantes, et une division du cerveau en un ensemble d'organes spécifiques : il isole vingt-sept facultés fondamentales, correspondant à l'exercice de vingt-sept organes parfaitement individualisés dans le cerveau.[20] Ce principe de correspondance entre les facultés et leurs organes spécifiques permet de faire le lien entre sa psychologie phrénologique (la division de l'esprit en facultés distinctes et indépendantes) et son organologie proprement dite (la division du cerveau en un ensemble d'organes spécifiques) : la pluralité des facultés correspond à une pluralité des organes.

Ce principe organologique permet ainsi d'évaluer le degré de développement d'une faculté en fonction du volume de l'organe cérébral correspondant. Ces facultés s'expriment en effet dans les limites de leur organe propre, proportionnellement au degré d'activité de ces derniers : plus un organe est développé, plus la faculté qui lui correspond l'est aussi. C'est à partir de ce principe organologique que Gall va proposer un dernier postulat : la conformation externe du cerveau exerce une influence directe sur la boîte crânienne. Il devient alors possible d'évaluer le développement relatif de chacune des facultés, en mesurant les protubérances localisées à la surface du crâne.

D'après Gall, le principe organologique trouve une validation empirique dans l'examen cranioscopique, lequel permet en retour de confirmer la pluralité des facultés et

[20] La liste des vingt-sept facultés est la suivante : I. Instinct de la génération. II. Amour de la progéniture. III. Attachement, amitié. IV. Instinct de la défense de soi-même et de sa propriété. V. Instinct carnassier, penchant au meurtre. VI. Ruse, finesse, savoir-faire. VII. Sentiment de la propriété, penchant au vol. VIII. Orgueil, hauteur, fierté. IX. Vanité, ambition, amour de la gloire. X. Circonspection, prévoyance. XI. Mémoire des choses, éducabilité. XII. Sens des localités. XIII. Mémoire des personnes. XIV. Mémoire verbale, sens des mots. XV. Sens du langage, talent de la philologie. XVI. Sens des rapports des couleurs, talent de la peinture. XVII. Sens des rapports des tons, talent de la musique. XVIII. Sens des rapports des nombres, talent des mathématiques. XIX. Sens de la mécanique, talent de l'architecture. XX. Sagacité comparative. XXI. Esprit métaphysique, profondeur d'esprit. XXII. Esprit caustique, esprit de saillie, esprit d'induction. XXIII. Talent poétique. XXIV. Bonté, bienveillance, douceur, compassion, sensibilité, sens moral. XXV. Faculté d'imiter. XXVI. Dieu et la religion. XXVII. Fermeté, conscience, persévérance, opiniâtreté.

des organes : « [i]l ne peut être question d'interpréter les différentes formes de la tête ou du crâne, qu'autant qu'elles dénotent les formes du cerveau. » (Gall, 1825, t.2, p. 311). En somme, l'organologie et la cranioscopie phrénologique représentent une alliance complexe entre plusieurs niveaux d'organisation, comprenant une psychologie des facultés, une anatomie cérébrale et une méthode de mesure. On peut résumer ce couplage en quatre principes :

1/ Une psychologie phrénologique : l'esprit est divisé en un certain nombre de facultés distinctes et indépendantes (la pluralité des facultés) ;
2/ Une organologie phrénologique : le cerveau est divisé en « organes cérébraux » distincts et indépendants (la pluralité des organes) ;
3/ Une psychophysiologie : chaque faculté correspond à un organe cérébral spécifique (la pluralité des facultés correspond à une pluralité des organes) ;
4/ Une cranioscopie phrénologique : chaque faculté s'exprime par une protubérance localisée à la surface du crâne.

1.1.2. Une cérébralisation des facultés

Les thèses fondamentales de Gall s'inscrivent dans un mouvement de réaction au sensationnisme qui dominait à la fin du XVIII^{e} siècle (Cherici et Dupont 2008 ; Dupont et Cherici 2015). Dans le premier volume de son traité *Sur les fonctions du cerveau* (1825), Gall formule une vive critique à l'égard des physiologistes qui se hasardent « au-delà du monde matériel » et négligent l'importance des études sur les structures et les fonctions du cerveau (Gall, 1825, t.1, p. 190). D'après lui, « ce n'est qu'à l'aide de ces recherches, que les physiologistes pourront lever des difficultés et se tirer des contradictions dans lesquelles ils s'engagent sans cesse » (p. 198). Cette critique vise du même coup les aliénistes, notamment Philippe Pinel (1745-1826) et Jean-Étienne Esquirol (1772-1840) dont les recherches sont jugées « peu certaines », voire « contradictoires », parce qu'elles ignorent « le rôle exclusif » joué par le cerveau dans l'exercice des facultés morales et intellectuelles. Pour Gall, on ne peut étudier les maladies du cerveau sans avoir une connaissance exacte de sa structure et de ses fonctions : « [i]l est impossible de faire des observations pathologiques exactes sur une

partie dont on ne connaît pas la structure, et sur laquelle on n'a que des idées fausses » (Gall, 1825, t.2, p. 188).[21]

Cette critique vise en particulier le mouvement sensationniste car il échoue selon lui à expliquer un fait d'observation incontestable, à savoir la grande variabilité interindividuelle des facultés intellectuelles ou morales. La première partie de son traité s'ouvre en effet sur le problème de la psychologie différentielle : Gall fait le constat que des individus ayant reçu la même éducation témoignent d'une grande variété de facultés ou d'instincts. Selon lui, le modèle sensationniste ne permet pas de rendre compte de la diversité des facultés, car il nie l'influence exercée par les conditions organiques de l'individu. D'après ce modèle, le cerveau reçoit passivement les impressions extérieures et n'a pas d'activité propre ; les facultés sont produites par les sens extérieurs et donc réductibles à la stricte influence du milieu.[22]

En réaction à cette physiologie parfaitement aveugle à l'influence organique et au caractère dynamique de l'organe cérébral, Gall va proposer une approche psychophysiologique des facultés et imposer l'idée d'un cerveau actif et autonome : les facultés ne sont plus le « produit accidentel » du milieu ; elles résultent d'une activité organique qui s'exerce indépendamment de l'action des sens extérieurs. La rupture avec le modèle sensationniste s'opère ainsi au travers d'un renversement de perspective, d'une transmission passive qui se focalise sur l'influence extérieure, à un mouvement dynamique qui prend forme à l'intérieur du cerveau avant de se propager vers la périphérie : le cerveau n'est plus le récepteur passif des impressions extérieures, mais le siège d'une activité qui lui est propre (Clauzade 2008a, 2008b).

Ce renversement de perspective constitue le point de départ des deux thèses fondamentales de Gall. La thèse de l'innéité consiste à la fois à ancrer les facultés dans

[21] L'erreur des aliénistes consiste selon lui à confondre les troubles qui affectent les fonctions vitales (troubles fonctionnels sans lésion), et les troubles organiques associés à des lésions « visibles ». Pour Gall, les maladies mentales, et la manie en particulier, résultent d'un dérangement dans les fonctions du cerveau. Or, une altération des fonctions mentales ne se traduit pas nécessairement par un dérangement organique visible : « [l]es maladies mentales, en tant qu'elles constituent la manie ou la démence, ne sont qu'un dérangement dans les fonctions du cerveau, de même que les autres maladies ne sont qu'un dérangement dans les fonctions d'autres parties du corps. » (Gall, 1825, t.2, p. 317).

[22] Voir à ce sujet Clauzade, Laurent. 2008b. "Le cerveau chez Cabanis et Gall : la philosophie biologique du XVIIIe siècle en débat." In *Les querelles du cerveau : comment furent inventées les neurosciences*, 235–55. Vuibert.

les dispositions organiques de l'individu, et à affirmer la spontanéité fonctionnelle du cerveau. La thèse de la pluralité réfute l'unité de l'organe cérébral et postule l'existence d' « organes cérébraux » distincts, correspondant à une faculté particulière. C'est donc en conjuguant ces deux thèses que Gall parvient à fournir une explication physiologique des différences comportementales. Selon lui, la grande diversité des facultés correspond à une différenciation fonctionnelle qui s'opère à l'intérieur même du cerveau : c'est le caractère dynamique et pluriel de l'organe cérébral qui engendre les différences individuelles. L'activité spontanée du cerveau garantit à chaque organe la possibilité d'exercer ou non sa fonction. Dans la mesure où ces organes sont plus ou moins développés selon les individus, leur mise en action produit une « variété infinie du caractère moral et intellectuel des hommes » :

> « Le cerveau, l'instrument des qualités morales et des facultés intellectuelles est essentiellement le même chez tous les hommes bien constitués ; mais les diverses parties intégrantes du cerveau, ou les divers organes ne sont pas chez tous également développés. Les rapports de ces divers développemens varient à l'infini. De-là la variété infinie du caractère moral et intellectuel des hommes » (1825, t.1, p. 262).

Cette psychophysiologie différentielle qui plaide en faveur d'une primauté du cerveau dans l'exercice des facultés nous amène à la première thèse fondamentale de Gall, qui est résumée dans ces deux premiers principes :

> 1/ Toutes les facultés et les qualités sont innées et ne peuvent se manifester indépendamment de conditions matérielles.
> 2/ Les facultés et qualités présentent leur siège dans le cerveau.

1.2. Les thèses de l'innéité et de la pluralité

1.2.1. Le cerveau comme « organe exclusif » des facultés

Selon Gall, le principe d'innéité est « la base de la physiologie du cerveau » car il permet d'affirmer le « rôle exclusif » de ce dernier dans l'exercice des facultés

intellectuelles ou morales.[23] La recherche de leurs conditions matérielles est selon lui indissociable de l'étude du cerveau car c'est le seul organe qui rend possible leur exercice : « [l]e cerveau est la source de tous les sentiments, de toutes les affections, de toutes les passions ; leur manifestation quelconque doit donc dépendre du cerveau et être modifiée par lui. » (1825, t.4, p. 291). Autrement dit, le cerveau est le seul et unique siège des facultés dans le sens où il ne constitue pas seulement leur cause matérielle, mais leur condition même de possibilité :

> « Quand je dis que l'exercice de nos facultés morales et intellectuelles dépend de conditions matérielles, je n'entends pas que nos facultés soient un produit de l'organisation ; ce serait confondre les conditions avec les causes efficientes. Je m'en tiens à ce que l'on peut soumettre à l'observation. Ainsi je ne considère nos facultés morales et intellectuelles qu'autant qu'elles deviennent pour nous des phénomènes par le moyen des organes cérébraux. » (1825, t.1, p. 189-190).

Chaque faculté est conditionnée par la mise en action de son organe propre. Il en résulte « qu'aucun exercice de faculté n'est possible sans organe, mais que l'organe peut exister sans que la faculté à laquelle il est affecté soit mise en exercice » (1825, t.1, p. 292). C'est en ce sens que Gall définit les facultés morales ou intellectuelles comme des « forces fondamentales ». Les facultés individuelles naissent d'après des « dispositions intrinsèques », dans le sens où elles possèdent une « tendance déterminée » qui peut être mise en action par l'activité spontanée des organes cérébraux :

> « J'entends par dispositions innées, des aptitudes industrielles, des instincts, des penchants déterminés, des facultés, des talens déterminés. J'entends […] que chaque organe cérébral est empreint d'une tendance déterminée ; que chaque organe jouit d'un aperçu intérieur, d'une force, d'une faculté, d'une impulsion, d'un penchant, d'un sentiment particuliers. Ici rien n'est le résultat vague et incertain, ni d'une influence extérieure, ni d'une abstraction intérieure. Aussitôt que les organes relatifs ont acquis leur parfait développement et leur entière

[23] Comme nous le préciserons plus loin, la « physiologie » au sens où elle est employée par Gall renvoie à sa « psychologie des facultés » (ou « psychologie phrénologique »), c'est-à-dire à la division de l'esprit en facultés distinctes et indépendantes.

activité, les fonctions qui en résultent sont aussi déterminées que les dispositions elles-mêmes dont ces organes sont les dépositaires. » (1825, t.1, p. 63)

C'est donc en opposant au modèle sensationniste le principe d'innéité des facultés que Gall est amené à attribuer au cerveau un rôle premier dans leur manifestation. Nul autre organe n'entretient de tels rapports, précisément parce que ces facultés résident de manière intrinsèque dans l'organisation cérébrale de l'individu :

> « L'innéité des forces morales et intellectuelles fondamentales est la base de la physiologie du cerveau ; car si au lieu de pouvoir démontrer qu'elles sont innées, on pouvait prouver qu'elles ne sont que le produit accidentel des choses extérieures et des sens extérieurs, il serait inutile d'en chercher l'origine et le siège dans le cerveau » (1825, t.1, pp. 66-67)

Il paraît important de préciser la nature des arguments mobilisés en faveur de la thèse de l'innéité, dans la mesure où ce principe constitue le premier fondement théorique de son organologie. La thèse selon laquelle le cerveau est « l'organe exclusif » des facultés repose sur un certain nombre de preuves anatomiques, physiologiques et pathologiques, énoncées dès 1808 dans son *Introduction au cours de physiologie du cerveau*, et synthétisées dans le premier volume de l'ouvrage *Sur les fonctions du cerveau* publié en 1825.

Le principal argument consiste à montrer l'existence d'un « rapport direct » entre le développement du cerveau et le perfectionnement ou le dérangement des facultés morales ou intellectuelles (1825, t.2, p. 156). Il s'agit globalement de montrer que les différences anatomiques ou structurelles sont directement liées à des différences *fonctionnelles*, lesquelles s'expriment au travers du développement des facultés fondamentales. En s'appuyant sur des observations d'anatomie et de physiologie comparées, Gall constate par exemple que le perfectionnement graduel des animaux suit de près le développement de leur cerveau : les structures se complexifient à mesure que l'on remonte l'échelle des êtres, et la manifestation des facultés est d'autant plus complète que les organes cérébraux sont plus développés.[24] Il en déduit que la variabilité des caractères moraux ou intellectuels entre espèces ou individus tient à des différences dans

[24] L'anatomie comparée a été le principal moyen d'investigation de ces recherches. Elle s'inspire des travaux sur l'anatomie cérébrale des animaux de Vincenzo Malacarne (1744-1816) et de Félix Vicq-d'Azyr (1748-1794) : le cerveau humain se mesure à celui de l'animal (Cherici 2003).

la conformation de leur cerveau : « Toutes les fois que l'on observe dans les animaux des qualités et des facultés différentes, l'anatomie comparée ne découvre des différences essentielles que dans leurs cerveaux » (1825, t.2, p. 165). Inversement, deux individus présentant une conformation cérébrale semblable manifestent globalement les mêmes dispositions morales et intellectuelles : « [l]es facultés de l'animal sont essentiellement les mêmes, toutes les fois que la structure du cerveau est la même pour l'essentiel » (1825, t.2, p. 164).

Certaines preuves sont également basées sur l'observation de cas cliniques présentant un dérangement des facultés intellectuelles ou morales. Gall observe qu'une altération des facultés morales n'atteint aucun autre organe que le cerveau, et inversement, qu'une désorganisation du cerveau rend impossible l'exercice des facultés morales ou intellectuelles. Il remarque par exemple qu'un développement incomplet ou défectueux du cerveau entraîne « une imbécilité proportionnée à l'imperfection du cerveau » (1825, t.2, p. 168). La manie ou la démence sont pour lui des exemples probants car leur « siège immédiat » est situé dans le cerveau : « la cause prochaine de la manie n'existe et ne peut exister que dans l'encéphale » (1825, t.2, p. 312). Gall confirme cette idée par des examens *post-mortem*, car il retrouve dans le cerveau et le crâne « les traces les moins équivoques des changemens qui ont eu lieu » (1825, t.2, p. 180). Il en déduit que les causes organiques du dérangement des facultés résident exclusivement dans le cerveau : « C'est dans le cerveau qu'ont lieu toutes les fonctions morales et intellectuelles ; car les fonctions, dans leur état d'intégrité, ne peuvent avoir lieu que là où se manifeste leur dérangement » (1825, t.2, p. 177).

1.2.2. La pluralité des facultés correspond à une pluralité cérébrale

D'après la thèse de la pluralité, les facultés morales ou intellectuelles ne sont pas déterminées par le cerveau dans sa globalité, mais par chacun des « organes cérébraux » qui le constituent. Cette thèse, explicitée en particulier dans le troisième volume de *Sur les fonctions du cerveau* (1825), correspond aux troisième et quatrième principes :

> 3/ Le cerveau se divise en autant d'organes spécifiques qu'il existe de facultés spécifiques.

4/ Les facultés présentent leur siège dans des parties du cerveau distinctes et indépendantes.[25]

La thèse de la pluralité est indissociable du postulat innéiste, car c'est l'activité spontanée du cerveau qui garantit à chaque organe la possibilité d'exercer ou non sa fonction. Autrement dit, c'est en posant l'innéité des facultés que Gall parvient à les localiser dans des parties individualisées du cerveau. La spontanéité et l'hétérogénéité fonctionnelle du cerveau permettent ainsi de mesurer le développement d'une faculté selon le volume de l'organe cérébral correspondant :

> « Ce n'est pas la masse absolue du cerveau qui détermine seule, ni tel caractère moral, ni tel génie ; mais que chaque partie cérébrale, en particulier, en tant qu'elle a acquis un développement plus ou moins favorable, peut modifier le degré de manifestation de telle qualité morale, et de telle faculté intellectuelle particulière » (t.2, pp. 348-349).

Les arguments mobilisés en faveur de la thèse de la pluralité reposent sur des observations anatomiques effectuées sur des animaux. La plupart relèvent du même raisonnement que pour la thèse de l'innéité, à savoir l'existence d'un lien direct entre les différences anatomiques et fonctionnelles : « [l]es différences les plus marquées de la structure de l'encéphale [...] correspondent à des différences marquées dans ses fonctions » (t.2, pp. 394-395). Gall constate en particulier que les facultés se perfectionnent à mesure que les parties constituantes du cerveau sont plus nombreuses et plus développées :

> « [L]es cerveaux des animaux sont d'autant plus simples ou plus composés, selon que leurs instincts, leurs penchans et leurs facultés sont plus simples ou plus composés ; que les diverses régions du cerveau sont affectées à des catégories différentes de fonctions ; qu'enfin le cerveau de chaque espèce d'animal, et par conséquent aussi celui de l'homme, constitue une réunion d'autant d'organes particuliers, qu'il y a dans l'homme ou dans l'animal de qualités morales et de facultés intellectuelles essentiellement différentes » (t.6, p. 499).

[25] Précisons que si la formulation de cette thèse apparaît dans le premier volume de *Sur les fonctions du cerveau,* publié en 1825, elle fait l'objet d'une description complète dans le troisième volume.

Cette thèse revêt une grande importance dans la méthode cranioscopique, puisqu'elle sert d'assise théorique pour justifier l'influence du cerveau sur la configuration du crâne. Comme Gall le rappelle lui-même, de nombreux physiologistes admettaient avant lui la pluralité des organes. La particularité de sa démarche tient au fait qu'il a mis en relation son analyse anatomique avec sa théorie psychophysiologique d'une part, et sa mesure crânioscopique d'autre part :

> « Tous mes devanciers, même ceux qui admettent la pluralité des organes de l'âme, ont échoué dans leurs tentatives pour déterminer le siège de chacun d'eux, et c'est en examinant la forme des crânes et des têtes, et en la comparant aux qualités morales et aux facultés intellectuelles dont les sujets étaient doués, que j'y ai réussi. » (t.3, p. 2).

Une fois couplée au postulat innéiste, la thèse de la pluralité permet ainsi de mesurer le développement des facultés à partir des protubérances localisées à la surface du crâne. Cette méthode de mesure est résumée dans les deux derniers principes du système phrénologique :

> 5/ Plus un organe est développé, plus la faculté qui lui correspond l'est aussi.
> 6/ La conformation extérieure du cerveau détermine celle de la surface interne du crâne ; il est donc possible d'évaluer le développement des facultés d'un individu en palpant sa boite crânienne.

Si l'on s'en tient au système organologique en lui-même (c'est-à-dire aux trois premiers principes), la thèse de la pluralité est double : il s'agit de diviser le cerveau en autant d'organes distincts et indépendants qu'il existe de facultés distinctes et indépendantes. Autrement dit, Gall opère en parallèle une décomposition de l'esprit en un certain nombre de facultés, et une division du cerveau en un ensemble d'organes qui leur sont spécifiques. Cette double décomposition trouve une confirmation réciproque dans l'adoption d'un principe d'identité psychophysiologique. Ce principe constitue le point de convergence entre la psychologie des facultés et la division anatomique, puisque chacune des facultés isolées trouve son corrélat physiologique dans une partie individualisée du cerveau. L'importance des contributions de Gall à l'anatomie et la physiologie du cerveau réside dans ce principe qui fait le lien entre sa psychologie

phrénologique et son organologie phrénologique : la pluralité des facultés correspond, du point de vue anatomique et fonctionnel, à une pluralité cérébrale.

La thèse de la pluralité est donc basée sur un principe de correspondance nécessaire entre structure et fonction. Elle traduit l'existence d'un rapport étroit entre son anatomie et sa physiologie, puisque chaque faculté ne peut exercer sa fonction qu'au travers d'une structure organique spécifique. Il ne s'agit pas seulement de *corréler* une faculté à une structure cérébrale, car chaque faculté psychologique isolée lors de la décomposition de l'esprit est identifiée à la fonction exercée par son organe propre : chaque organe cérébral est conçu comme la « condition matérielle » nécessaire à l'exercice d'une faculté, et chaque faculté comme « fonction d'un organe » particulier. De ce point de vue, la recherche des causes matérielles d'une faculté revient à situer la faculté elle-même. Le cerveau constitue le support biologique des facultés de l'esprit, dans le sens où chacun des organes qui le constitue conditionne l'exercice d'une fonction spécifique.

À cet égard, il est important de noter que la physiologie au sens où elle est employée par Gall dans son organologie ne désigne rien d'autre qu'une « psychologie des facultés » : Gall définit les facultés psychologiques comme des fonctions cérébrales, précisément parce que leur exercice est conditionné par l'activité de l'organe cérébral correspondant. Le principe d'identité psychophysiologique constitue en ce sens la condition de sa biologisation des facultés psychologiques, et le fondement théorique de sa stratégie localisatrice (Clauzade 2008a). Il permet d'identifier les « fonctions de l'âme » à des « fonctions du cerveau », en instaurant un principe de correspondance nécessaire entre l'analyse psychologique et l'analyse anatomique.[26] Autrement dit, il vise à ancrer la diversité des facultés dans une pluralité organique, de telle sorte que l'analyse structurelle et fonctionnelle se trouvent « dans une concordance parfaite » (Gall, 1825, p. 25). Pour Gall, cette double analyse doit effectivement répondre à une exigence de confirmation réciproque, puisque « toute doctrine sur les fonctions du cerveau serait fausse, si elle se trouvait en contradiction avec sa structure » (Gall, 1825, p. 23).

Il est d'ailleurs intéressant de constater que malgré l'importance accordée à la pluralité cérébrale dans la description de son organologie, Gall défend l'antériorité de sa

[26] L'usage du terme « fonction du cerveau » est relativement fréquent chez Gall. Comme l'a montré Laurent Clauzade (2008a, p. 68), « c'est à partir de la phrénologie que l'on commence à parler de fonction du cerveau ».

psychologie des facultés sur son analyse anatomique : « [l]a connaissance des fonctions a toujours précédé celle des parties » (Gall, t.2, p. 21). Comme le souligne Laurent Clauzade (2008a), ses arguments anatomiques dérivent d'une physiologie déjà constituée, c'est-à-dire qu'ils n'ont de sens qu'en étant mis en rapport avec son analyse psychophysiologique des facultés : la base fondatrice de son organologie n'est pas la pluralité cérébrale, mais la pluralité des facultés elles-mêmes. De ce point de vue, la division anatomique du cerveau ne constitue rien de plus que le versant organique de sa théorie des facultés. Elle est pensée comme un moyen de confirmer la décomposition fonctionnelle initiale, c'est-à-dire qu'elle est entièrement tributaire de son analyse physiologique (entendue dans le sens d'une « psychologie des facultés ») :

> « La physiologie du cerveau ne doit pas être en contradiction avec l'anatomie du cerveau : son anatomie doit même venir à l'appui de sa physiologie. Mais la physiologie a été trouvée indépendamment de l'anatomie. Rien, dans la physiologie du cerveau, n'a été bâti sur une donnée anatomique quelconque ; rien n'a été interprété par la structure ou par l'arrangement des parties cérébrales ; l'anatomie du cerveau n'a servi que de confirmation aux découvertes physiologiques. » (1825, t.6, p. 118)

1.3. L'application de l'organologie à l'étude du phénomène criminel

1.3.1. La naturalisation de la morale et du crime

La doctrine de Gall marque l'existence d'une relation étroite entre l'organe et la fonction d'une part, et le physique et le moral d'autre part (Clauzade 2008b). Cette détermination réciproque affirme la possibilité de développer une approche organique des facultés morales, car elle postule un lien direct entre les différences fonctionnelles (entendues dans le sens de la théorie des facultés) et les différences structurales ou anatomiques. Pour Gall en effet, l'étude des structures et des fonctions du cerveau nous incline à reconnaître que les phénomènes moraux sont intimement liés à des phénomènes organiques qui prennent place à l'intérieur du cerveau : le perfectionnement ou le

dérangement des facultés morales est directement lié à l'organisation cérébrale des individus.

C'est dans ce contexte que les principes de Gall vont progressivement s'étendre à la question de la nature du criminel. En soutenant l'innéité et la pluralité des facultés, Gall va justifier l'intérêt des recherches sur l'anatomie et la physiologie du cerveau pour reconnaître les signes d'une altération du sens moral, et mesurer l'expression du « penchant au meurtre ». Les enjeux qui entourent l'organologie et la cranioscopie de Gall vont alors s'appliquer à des questions proprement sociales, morales et juridiques : le criminel est-il responsable de ses actes ? Est-il curable ou éducable ? C'est en ce sens que Gall définit sa doctrine comme une « véritable science de l'homme ». Elle permet à la fois d'expliquer l'hétérogénéité et la grande variabilité des facultés morales, et de faire du cerveau un instrument de mesure pour éclairer les questions relatives à la responsabilité ou à la prise en charge des criminels.

Les réflexions concernant l'application des principes organologiques à l'étude du phénomène criminel sont énoncées dès le premier volume de son traité *Sur les fonctions du cerveau* (Gall, 1825). Comme nous l'avons précisé plus haut, l'approche organique du crime, telle qu'elle est préconisée par Gall, découle directement de ses deux thèses fondamentales : le cerveau étant « l'organe exclusif » de tous les instincts et penchants individuels, il devient possible de localiser la structure cérébrale responsable du « penchant au meurtre ». Ce dernier est alors mesuré par le « développement excessif » de son organe, localisé « au-dessus des oreilles, dans la ligne perpendiculaire de la colonne vertébrale. » (Gall, 1825, t.5, p. 454).

La découverte de l' « organe du meurtre » repose sur des observations d'anatomie comparée. Elle est née du constat que le crâne des frugivores diffère de celui des carnassiers : ces derniers présentent une protubérance localisée au-dessus de l'oreille, dont les frugivores sont privés. À partir de ces observations, Gall va suggérer l'idée que le penchant au crime est lié au comportement carnassier, et qu'il est partagé par tous les animaux :

> « Par l'expression de penchant au meurtre, je n'entends nullement un organe qui porte immédiatement à l'homicide, mais simplement le penchant naturel de tuer d'autres animaux, penchant qui appartient à tout animal carnivore, et par conséquent à l'homme » (Gall, 1825, t.1, p. 298).

Gall précise que l' « organe du meurtre », comme l'ensemble des organes cérébraux, préexiste chez tous les Hommes avant l'exercice de toute faculté. Le penchant criminel ne fait donc pas l'objet d'un traitement spécifique dans son organologie, puisque l'organe peut subsister si la faculté ne s'exerce pas, au même titre que tous les autres organes cérébraux : « l'organe peut exister sans que la faculté à laquelle il est affecté soit mise en exercice » (Gall, 1825, t.1, p. 292). En clair, si cette protubérance constitue la « condition matérielle » nécessaire à l'exercice du penchant au meurtre, l'organe lui-même ne traduit pas la manifestation effective de l'instinct criminel.

Cette précision est importante car elle nous invite à relativiser l'importance accordée au cerveau dans l'expression du penchant criminel. Le développement excessif de l' « organe du meurtre » n'est pas une condition suffisante à l'exercice du penchant car celui-ci peut être réprimé par d'autres organes ou d'autres facultés (notamment intellectuelles), ou sous l'influence des circonstances extérieures (l'éducation, la morale, la religion, etc.). Autrement dit, l' « organe du meurtre » ne prédispose pas nécessairement la manifestation du penchant, dans la mesure où son exercice dépend du maintien de l'équilibre entre trois niveaux : le degré d'activité des autres organes cérébraux, l'exercice des facultés prédéterminées par cette organisation cérébrale, et l'influence des circonstances extérieures. Les dispositions organiques de l'individu peuvent par exemple être corrigées ou perfectionnées grâce à l'éducation ou la morale :

> « L'homme, doué de facultés intellectuelles supérieures, saura donner à son penchant une direction plus favorable que celui qui a l'esprit faible. L'éducation, les habitudes, l'exemple, la religion, la morale, les lois, etc., agissent dans l'homme doué de liberté morale, comme autant de motifs pour conformer ses actions à l'ordre social, même en dépit de ses penchans. » (Gall, 1825, t.4, p. 168)

Il faut bien comprendre que cette notion d'équilibre s'inscrit dans le cadre de la thèse de la pluralité des facultés et des organes : le cerveau possédant autant de divisions fonctionnelles qu'il existe de facultés, la manifestation du « penchant au meurtre » peut être contrebalancée par l'exercice de qualités plus nobles. Si les circonstances extérieures peuvent jouer un rôle dans l'expression du « penchant au meurtre », c'est en vertu du fait qu'elles peuvent contrebalancer l'activité de l'organe correspondant. Dans la mesure où l'individu ne peut agir sur le développement de ses organes cérébraux, l'influence extérieure garantit la possibilité de rectifier ses penchants de manière favorable. Le

développement excessif de cet organe n'est donc pas une condition suffisante à l'exercice du penchant criminel, « tout ce que l'on peut soutenir avec assurance, c'est que, toutes circonstances extérieures étant égales, un individu chez qui cet organe est très prononcé, commettra un homicide plus facilement qu'un autre que la nature n'y a pas disposé au moyen de son organisation » (Gall, 1825, t.4, p. 169)

Il est important de préciser que la prise en compte des facteurs socio-environnementaux n'est pas incompatible avec la conception innéiste des facultés. Pour Gall, c'est précisément parce que le penchant criminel est inhérent à la nature humaine que l'influence environnementale peut s'exercer. Les circonstances défavorables peuvent conduire au crime dans la mesure où elles interagissent directement avec les dispositions organiques de l'individu. Inversement, l'éducation ou la morale peuvent affecter ses penchants de manière favorable, en agissant sur le degré d'activité de ses organes cérébraux. Le criminel n'est donc pas entièrement soumis à ses instincts violents, car il peut les combattre en vertu des circonstances extérieures, dans les limites de son organisation cérébrale :

> « L'homme en proie au cruel penchant dont nous parlons, conserve encore la faculté de le vaincre ou de lui donner une direction qui n'est pas nuisible. Mais le pouvoir de dompter un penchant vicieux, s'affaiblit à proportion qu'il a reçu moins d'éducation, ou que les organes des qualités d'un ordre supérieur sont moins développés. » (Gall, 1825, t.4, p. 89)

Ces réflexions vont donner à Gall l'occasion de préciser les fondements théoriques de sa psychophysiologie différentielle. Selon lui, l'influence mutuelle des facteurs organiques et socio-environnementaux permet d'assurer des effets très variables entre les individus, en fonction de leur organisation cérébrale et des circonstances extérieures. Elle permet ainsi d'expliquer la grande variabilité des facultés morales, et la raison pour laquelle l'organe cérébral ne détermine pas à lui seul l'expression des instincts et des penchants individuels : « le même degré d'activité d'un organe doit produire des actions toutes différentes chez différens individus » (1825, t.4, p. 167). Cette psychophysiologie différentielle va amener Gall à s'interroger sur les apports de sa doctrine pour guider et orienter les pratiques punitives. Deux propositions en découlent : une volonté de fonder le système pénal sur le principe d'individualisation des peines, et la nécessité de protéger

la société contre les « incorrigibles », c'est-à-dire les individus dont « l'organisation malheureuse » ne peut être réprimée par les circonstances extérieures.

1.3.2. Une volonté de repenser le système pénal

Gall propose de repenser le système pénal à la lumière de ses recherches sur l'anatomie et la physiologie du cerveau, dès le premier volume de son traité *Sur les fonctions du cerveau*. Selon lui, les crimes et les délits ont été trop longtemps jugés en eux-mêmes, sans considérer l'individu agissant. Or, la méconnaissance des prédispositions organiques est source d'injustice car « [t]ous les individus ne jouissent pas de la liberté morale au même degré » (Gall, 1825, t.1, p. 288). Il faut donc procéder à une individualisation des peines en fixant le degré de responsabilité sur les spécificités anatomiques et fonctionnelles de l'individu : non seulement il faut juger le criminel et non le crime, mais surtout, la peine doit être proportionnée aux dispositions organiques de l'individu :

> « Le code pénal détermine la nature des délits et des crimes, et puis fixe la peine à infliger. C'est la matérialité de l'acte qui donne la mesure de la punition, sans égard à la personne agissante dans l'acte, ni à la personne qui doit l'expier. Sans doute on rencontra trop de difficultés pour procéder autrement, et on jugea que c'était le seul moyen d'obtenir une égalité et une impartialité parfaites dans l'exercice de la justice. Mais il est évident que c'est précisément par-là qu'on se rend coupable de l'injustice la plus affligeante, et que l'on manque presque toujours la juste estimation du crime, et par conséquent l'application proportionnée de la peine [...] Les délits et les crimes ne se commettent pas d'eux-mêmes ; ils ne peuvent donc pas être considérés comme des êtres abstraits. Les délits et les crimes sont des produits d'individus agissans ; ils reçoivent donc leur caractère de la nature et de la situation de ces individus ; et ils ne sauraient être estimés, déterminés que d'après la nature et la situation de ces mêmes individus. » (t.1, pp. 357-358).

Remarquons que ce principe reconnaît la détermination réciproque de la « situation » et de la « nature » des individus dans l'expression de leur penchant criminel : l'application de la peine doit être proportionnée aux dispositions organiques de l'individu,

mais elle doit aussi tenir compte des circonstances extérieures, dans la mesure où elles exercent une influence directe sur le développement de l' « organe du meurtre ». La notion de « déséquilibre cérébral » s'étend donc jusque dans l'application de la sanction pénale, puisqu'elle pose les termes de la juste mesure entre la nature du crime, la nature (corrigible ou non) de son auteur, et la nature de la peine : l'efficacité de la justice pénale doit reposer sur un principe de proportionnalité entre la gravité de l'acte criminel, la nature et la situation de l'individu, et la sévérité de la sanction.

Il est important de préciser que ce principe de proportionnalité répond à un impératif de défense sociale. Gall estime que l'efficacité des mesures punitives doit être mesurée en termes d'utilité sociale. Les recherches sur l'anatomie et la physiologie du cerveau doivent donc être mises au service de la protection de la société, en identifiant les individus dont l'organisation cérébrale ne peut être corrigée par les circonstances extérieures. C'est pourquoi l'individualisation des peines préconisée par Gall exclut la question du libre arbitre ou de « la détermination de la volonté ». Il faut selon lui rejeter la notion de « culpabilité intérieure » car il est avant tout question « des besoins de la société ; de prévenir les délits, de corriger les malfaiteurs et de mettre la société en sûreté contre ceux qui sont le plus incorrigibles » (t.1, p. 356).

1.4. Lombroso et l'étude anthropologique du criminel

L'histoire de l'anthropologie criminelle remonte au dernier quart du XIX^{e} siècle, avec l'émergence de l'École positiviste italienne menée par Cesare Lombroso (1835-1909), Enrico Ferri (1856-1929) et Raffaele Garofalo (1852-1934).[27] Cesare Lombroso

[27] Précisons que si nous situons les origines de l'anthropologie criminelle dans l'Italie de la fin du XIX^{e} siècle, ses prémices historiques prennent place en France dès le début du XIX^{e} siècle, notamment au travers de la doctrine physiologique de Pierre-Jean-Georges Cabanis (1757-1808), de la phrénologie de Franz-Joseph Gall, ou encore des recherches menées par Jean-Étienne Esquirol (1772-1840) sur la « monomanie homicide » (Renneville 1994). Pour autant, ces recherches ne renvoient pas à une étude anthropologique en tant que telle, car elles ne reconnaissent pas l'existence d'un « type criminel », et le crime lui-même est seulement étudié par le prisme de sa relation avec la folie (Renneville 2003).

est le représentant le plus actif et le plus prolifique de cette École anthropologique.[28] Son œuvre maîtresse, L'*Homme criminel* a été réédité quatre fois entre 1876 et 1895. Elle est largement consacrée à la description du « portrait-type » du criminel-né, véritable « monstre physique et moral », réputé incorrigible car voué au crime par son organisation biologique. C'est à partir de la seconde édition de *L'Homme criminel*, parue en 1878, que la théorie du criminel-né commence à se répandre en Europe occidentale (Renneville 1994, 1995). Sous l'impulsion des recherches menées par Enrico Ferri et Raffaele Garofalo – qui ont respectivement popularisé les concepts de « criminel-né » et de « criminologie » − l'anthropologie criminelle se revendique comme science autonome, et organise son premier Congrès à Rome en 1885.[29]

C'est en qualité de médecin que Lombroso mène ses premières recherches consacrées à l'étude anthropologique du criminel, alors qu'il travaille dans les prisons et asiles de la région de Pavie, au début des années 1870. En cherchant à isoler les critères anatomiques ou morphologiques permettant de distinguer les fous des criminels, Lombroso remarque chez ces derniers de nombreuses anomalies que l'on rencontre habituellement chez les « êtres inférieurs ». Il fait alors l'hypothèse que le crime trouve sa cause dans l'atavisme, c'est-à-dire le retour aux formes ancestrales : ces résurgences ataviques sont des stigmates innés qui font du criminel un être biologiquement inférieur et intrinsèquement dangereux.

La théorie atavique du crime sera confirmée par une vaste enquête anthropométrique qui portera sur un échantillon de plusieurs milliers de prisonniers. Pendant près de trente ans, Lombroso va dresser l'inventaire des caractéristiques physiques et morales des criminels, auxquelles s'ajouteront plusieurs séries

[28] L'historien Renzo Villa a recensé une cinquantaine de livres et plus de 1 600 articles publiés dans diverses revues ou recueils, en Europe et en Amérique (Villa 1985).

[29] Raffaele Garofalo, élève de Lombroso, a publié *La Criminologie : étude sur la nature du crime et la théorie de la pénalité* (1890), un traité scientifique et politique qui s'inscrit directement dans la continuité de la théorie du criminel-né. Enrico Ferri va développer dans sa *Sociologie criminelle* (1893) une conception juridico-politique de la criminalité, en analysant les données issues de l'anthropologie criminelle d'une part, et de la statistique criminelle d'autre part. Ferri accordera une place plus importante à l'influence du milieu social dans le développement de la criminalité, mais insistera également sur la nécessité d'adapter la sanction pénale à « la personnalité du criminel ». Selon lui, l'objectif de l'anthropologie criminelle est de « faire de la personnalité du criminel l'objet primaire et principal des règles de justice pénale, au lieu de la gravité objective du crime » (Ferri 1911, p. 251).

photographiques et une immense collection de documents relatifs à leur psychologie.[30] Cette quantité colossale de données sera interprétée par Lombroso comme le socle méthodologique de sa principale thèse : le criminel représente un « type anthropologique » que l'on peut reconnaître par ses caractéristiques anatomiques, morphologiques, et psychologiques.

Nous pouvons d'ores et déjà préciser qu'il est très difficile de dégager une représentation cohérente des thèses de Lombroso, tant ce dernier multiplie les remaniements méthodologiques et théoriques. Les rééditions de *L'Homme criminel* sont pour lui autant d'occasions d'affiner sa grille criminologique, en y ajoutant de nouveaux critères ou de nouvelles réflexions, qui ne parviendront jamais à convaincre ses contemporains. La valeur scientifique de ses thèses a été très rapidement invalidée, les critiques portant à la fois sur la réalité scientifique du « type criminel », les preuves de l'origine atavique du crime, ou plus généralement sur le cadre méthodologique de ses recherches.[31] Après avoir institué la figure anthropologique du « monstre criminel », Lombroso est devenu lui-même un « monstre épistémologique », l'objet d'une

[30] On peut citer, parmi les dizaines d'essais qui ont complété son œuvre maîtresse, *Les palimpsestes des prisons* (1894), qui recueille une large collection de documents écrits et illustrés sur les productions de prisonniers (poèmes, confessions, tatouages ou graffitis, anecdotes et extraits de conversations...).

[31] Les premières critiques ont été formulées dès les années 1880. Il paraît difficile de les mentionner tout en conservant la nuance qu'elles revêtent, car la plupart des objections portent sur des aspects très spécifiques de la théorie lombrosienne. Les Français Gabriel Tarde (1843-1904) et Alexandre Lacassagne (1843-1924) ont été les principaux adversaires de Lombroso. Ils rejettent la pertinence de l'approche anatomique et morphologique du criminel, et défendent l'influence du milieu social dans le développement de la criminalité. Pour autant, Tarde et Lacassagne ne nient pas l'influence exercée par Lombroso dans le développement de leurs thèses. Pour Gabriel Tarde (1890), la portée de la théorie lombrosienne réside dans le fait d'avoir attiré l'attention sur la relation entre le moral et le physique. Alexandre Lacassagne va quant à lui proposer une typologie criminelle largement inspirée de la méthode anthropométrique de Lombroso, basée sur le développement relatif des différentes parties du cerveau. Comme le souligne Marc Renneville (1994), il a été l'un des premiers défenseurs des thèses lombrosiennes en France, avant de fonder son « école du milieu social ». Celle-ci se résume grossièrement dans un certain nombre d'aphorismes que Lacassagne aimait rappeler : « le criminel est un microbe qui ne se développe que si le bouillon de culture est favorable » ou « la justice flétrit, la prison corrompt et la société a les criminels qu'elle mérite ».

dénégation à l'intérieur même du champ disciplinaire qu'il a contribué à fonder (Renneville 1997).[32]

Nous allons ici tenter d'éviter l'appréciation rétrospective, pour adopter une approche épistémologique centrée sur l'analyse conceptuelle et méthodologique de son œuvre maîtresse, *L'Homme criminel* (1887).[33] Loin de relativiser les critiques qui ont été formulées à son encontre, ni de faire l'impasse sur la circularité de ses arguments, ce travail d'analyse vise à retracer la logique de la pensée lombrosienne. Pour rappel en effet, notre objectif final est de déterminer si le projet neurocriminologique réactualise la théorie du criminel-né à la lumière des données neurobiologiques. En suivant la même démarche que celle employée dans l'analyse du système organologique de Gall, nous allons donc chercher à nous approcher au plus près du raisonnement adopté par Lombroso, en identifiant les présupposés et les choix méthodologiques qui sous-tendent son anthropométrie du crime.

1.4.1. La spécificité de la démarche lombrosienne

Avant d'entrer dans le détail des thèses de Lombroso, il convient de présenter quelques remarques générales sur la spécificité de sa démarche et le contexte (socio-politique et scientifique) dans lequel elle s'inscrit. Ces remarques nous donneront l'occasion d'évoquer les principales limites méthodologiques de son enquête anthropométrique, de manière à concentrer l'analyse qui suit sur le discours, les arguments et les stratégies de recherche mobilisés par Lombroso. Notre objectif étant de reconstruire la structure argumentative de ses thèses, nous n'insisterons pas davantage sur ses faiblesses méthodologiques.

En premier lieu, il est important de préciser que l'anthropométrie criminelle de Lombroso répond, comme chez Gall, à un impératif de protection sociale. Dans un contexte où, à la fin du XIX^{e} siècle, la récidive est définie comme une véritable

[32] Les thèses de Lombroso ont servi de caution scientifique aux politiques hygiénistes et eugénistes du début du XX^{e} siècle. Elles ont eu une forte influence sur le régime nazi, l'exclusion des « incorrigibles » ayant trouvé un fondement scientifique dans la théorie du criminel-né et l'exigence de protection sociale à laquelle elle répond (Rafter 2008).

[33] Précisons que notre analyse se rapporte à la traduction française de la quatrième édition, publiée en 1887.

« gangrène criminelle », le pouvoir politique italien fait appel à des médecins ou anthropologues pour repérer les criminels violents ou incorrigibles, et plus globalement, tous les individus qui constituent une menace pour l'ordre social (Châles-Courtine 2005, 2008).[34] La source de la criminalité n'est pas attribuée à des causes politiques ou économiques, mais à un problème d'ordre biologique et anthropologique (Llorca 2005) : il existerait des criminels « par nature », dont la dangerosité peut être isolée, évaluée et comparée. L'anthropologie criminelle de Lombroso répond précisément à cette exigence, puisqu'elle vise à caractériser la figure emblématique de la dangerosité, incarnée par le criminel-né. Elle n'est d'ailleurs pas seulement au service d'un impératif social ou politique, puisqu'elle s'inscrit elle-même dans une ambition réformatrice. Comme la phrénologie de Gall, l'école lombrosienne défend une hiérarchisation et une individualisation des peines, basée sur l'évaluation de la dangerosité sociale de l'individu (Renneville 1995 ; Llorca 2005). La détermination de la peine ne doit plus reposer sur la gravité de l'acte commis par le criminel, mais sur l'évaluation du niveau de danger qu'il représente pour la société.

Par ailleurs, la spécificité de la démarche adoptée par Lombroso tient au fait qu'il a étudié le sujet criminel comme une figure anthropologique en tant que telle, dans un contexte où la dangerosité était étudiée par le prisme des rapports entre crime et folie (Renneville 2003). Tout au long du XIX^{e} siècle, la dangerosité criminelle trouvait une explication dans le concept de « monomanie homicide » forgé par Jean-Étienne Esquirol en 1827, ou dans la notion de « dégénérescence » théorisée par Bénédict Augustin Morel en 1857.[35] Comme l'a bien montré Michel Foucault, le danger social était confondu dans l'entité juridico-scientifique du crime-folie, « d'un crime qui est tout entier folie, d'une folie qui n'est rien d'autre que crime » (Foucault 1981, pp. 407-408). Avec l'anthropologie criminelle de Lombroso, on s'affranchit de cette catégorie indistincte qui inclut une vaste série de dangers, pour poser la question de la nature fondamentalement anthropologique du criminel : Lombroso a été le premier médecin à dépasser le cadre de l'analyse clinique pour étudier le phénomène de la criminalité. Il n'entend pas relier le

[34] La référence à la « gangrène criminelle » est notamment attribuée à Joseph Reinach, qui a consacré un ouvrage au problème de la récidive (1882).

[35] Pour une analyse historique de la manière dont le concept de « monomanie homicide » a participé à l'étude anthropologique du criminel, voir Renneville (1994).

crime à un diagnostic de folie, puisqu'il cherche essentiellement à isoler les caractéristiques distinctives du criminel.

Pour autant, loin de renier les recherches menées par les aliénistes auxquels il fait référence à plusieurs reprises, Lombroso revendique lui-même l'influence du discours médical sur son étude anthropométrique. Dans un souci de légitimation de ses thèses, il se présente avant tout comme médecin, et partage avec les théoriciens de la dégénérescence le même emprunt au modèle médical et statistique (Renneville 1994 ; Hacking 2001) : il utilise la méthode statistique comme un outil de validation de sa théorie. Cette méthode garantit selon lui l'infaillibilité et la supériorité de ses thèses sur celles de ses contemporains, comme il le précise dans un ouvrage paru avant la publication de *L'Homme criminel*, consacré à l'étude clinique et anthropométrique de la microcéphalie :

> « [C]ombien d'erreurs et d'illusions s'éviteraient le criminologue et le psychiatre qui connaîtraient les résultats de l'anthropologie moderne, et qui sauraient, avec des instruments et des chiffres, se convaincre qu'à toutes les aberrations du sens moral et du psychisme correspondent des anomalies du corps et du crâne en particulier » (Lombroso, 1873).[36]

Lombroso expliquera par ailleurs que le véritable progrès de la civilisation réside dans le triomphe de la statistique : « pour moi, le vrai caractère qui distingue notre époque de l'époque ancienne est le triomphe du chiffre sur les opinions vagues, sur les préjugés, sur les vaines théories » (Lombroso, 1878, p. 386). Tout au long de *L'Homme criminel*, Lombroso ne cessera de rappeler que chacune de ses thèses a été « prouvée par la statistique ». Les données statistiques ne sont pas seulement mobilisées comme un argument d'autorité pour valider sa méthode anthropométrique ; elles doivent fournir la preuve définitive de ses thèses. Lombroso cultive d'ailleurs une certaine confusion entre les résultats de ses mesures et les faits eux-mêmes, puisqu'il se plait à rappeler qu'il est « esclave des faits » et « serviteur des chiffres » (Renneville 1995 ; Faugeras 2009). Le « fait scientifique » se révèle selon lui dans l'accumulation de données statistiques, dont on ne peut nier l'ampleur colossale : en multipliant les analyses chiffrées et les tableaux de synthèse, Lombroso cherche à asseoir la validité de ses thèses sur un modèle

[36] C. Lombroso, *Studi clinici ed antropometrici sulle microcefalie*, 1873, cité par R. VILLA, *Il deviante e i suoi segni. Lombroso e la nascita dell'antropologia criminale*, Franco Angeli, Milan, 1985, p. 134.

statistique, de sorte que toute critique ne peut finalement se placer que sur le terrain de la mesure et du chiffre.

Il serait fastidieux d'énumérer l'ensemble des biais, des résultats erronés, des incohérences, des généralisations inductives ou des confusions (entre corrélation et causalité notamment) qui parcourent son œuvre (e.g. Jay Gould 1983). Lombroso montre une certaine propension à réinterpréter les données à la lumière de sa théorie, ou à ne retenir que celles qui permettent de confirmer ses présupposés idéologiques. On retient également sa tendance à cumuler les hypothèses pour combler leurs lacunes, anticiper les critiques, ou s'inscrire dans les représentations dominantes de son époque. Lombroso cherche à rendre compte de l'ensemble des données statistiques recensées chez les criminels, et à expliquer, de la manière la plus exhaustive possible, toutes les formes de criminalités (Renneville 1997).

Cet effort de systématisation se manifeste par exemple dans la distinction opérée entre « les criminels nés » et les « criminels d'occasion ou d'habitude »,[37] ou encore dans la construction de la catégorie des « criminels de génie », pour rendre compte des données rapportant une plus forte capacité crânienne, et qui contredisent ainsi le postulat d'une « infériorité biologique ». Le souci de construire un modèle systémique s'exprime également dans la manière dont Lombroso a cherché à expliquer la fréquence plus faible de la criminalité féminine : cette donnée statistique l'amènera à considérer la prostitution comme l'équivalent féminin du crime, thèse qu'il s'efforcera de valider par l'observation de caractères régressifs chez les prostitués.[38]

De manière générale, Lombroso cumule les hypothèses de manière à effacer les incohérences ou la faiblesse de ses résultats. À partir du milieu des années 1880, il complète sa théorie de l'atavisme – qui ne convainc plus ses contemporains – par l'hypothèse d'une origine pathologique du crime, en défendant la fusion des concepts de criminel-né, de folie morale et l'épilepsie (Renneville 1995). L'expression de l'instinct criminel ne s'explique plus seulement par la résurgence de caractères ancestraux, mais aussi par la survenue de processus morbides comme l'épilepsie. Cet élargissement

[37] Comme nous le montrerons plus loin, Lombroso reconnaît l'influence des facteurs socio-environnementaux dans l'expression de l'instinct criminel dès la première édition de son ouvrage (1876).

[38] Cette thèse sera développée dans *La femme criminelle et la prostituée* (1896), écrit en collaboration avec Guglielmo Ferrero (1871-1942). En se basant sur la même méthode anthropométrique que dans *L'Homme criminel*, cet ouvrage recense les critères anatomiques, morphologiques ou psychologiques qui permettent de distinguer la « prostituée-née » de la « prostituée d'occasion ».

théorique lui permet ici encore de fédérer toutes les hypothèses avancées pour expliquer le phénomène de la criminalité, et de rendre compte de l'ensemble des données statistiques recensées dans ses tableaux de synthèse.[39] Il en résulte une théorie indéfiniment changeante, extensible, qui se renouvelle au gré de ses remaniements méthodologiques ou des nouvelles observations effectuées par ses contemporains.

Une dernière remarque s'impose avant d'entrer dans le détail de ses thèses. Bien qu'il revendique l'influence de la mesure et du chiffre dans le développement de sa théorie, Lombroso n'entre jamais dans les débats proprement scientifiques de la statistique médicale (Braunstein 2010). L'essentiel pour lui est de dégager les implications « philosophiques » de ses mesures, à savoir la réalité scientifique d'un nouveau « type anthropologique » : l'Homme criminel est un être « biologiquement inférieur » en raison des nombreuses anomalies qui le caractérise. Il n'est d'ailleurs jamais question d' « Homme normal » ou d' « Homme moyen », c'est-à-dire de l'homme qui n'est ni criminel, ni aliéné, et qui constituerait la référence statistique de son enquête anthropométrique.[40] Comme l'a justement remarqué l'historien Renzo Villa, Lombroso « ne travaille que sur le négatif » (Villa 1985, p. 266). Alors que les critères de définition de la normalité ne sont jamais explicitement définis, il s'évertue à énumérer les déviations et les aberrations qui définissent la marque constitutive d'un danger. Ses thèses révèlent donc un biais de sélection majeur, qui fait qu'il est très difficile de mesurer la fiabilité de ses données : il est impossible de vérifier si ces anomalies sont réellement caractéristiques du criminel, ou si elles se rencontrent chez les autres individus.[41]

---

[39] Lombroso n'aura de cesse de démultiplier les causes de la criminalité dans ses ouvrages postérieurs à *L'Homme criminel*. Dans son traité intitulé *Le crime politique et les révolutions*, quelques notes seront consacrées à l'influence des variations du milieu extérieur (température, climat, pression atmosphérique), Lombroso ne manquant pas de remarquer qu'en général, « les mois qui marquent le maximum ou le commencement de la chaleur, donnent pour les émeutes, comme pour les crimes, les chiffres les plus élevés. » (Lombroso, 1890, t.1, p. 67).

[40] Lombroso fait pourtant référence à Adolphe Quételet et à sa « physique sociale » (anciennement appelée « statistique morale ») créée en 1839, car elle complète selon lui les études sur la physiologie et la psychologie de l'homme.

[41] Précisons que cette critique a été formulée dès la fin des années 1880 par le médecin Charles Féré (1852-1907), dans son traité intitulé *Dégénérescence et criminalité*. Selon lui, l'anomalie psychologique et anatomique du criminel est « purement hypothétique », car elle n'est pas rapportée aux critères de définition de l'« Homme normal » : « [l]e type criminel n'est pas suffisamment défini, ni séparé des types qu'on peut considérer comme normaux ; il reste confondu par un grand nombre de caractères avec le type dégénéré

1.4.2. De la naturalisation du crime à l'étude anthropométrique du criminel

Le postulat fondateur de la thèse lombrosienne est l'idée d'une continuité de la nature et du crime : le crime est omniprésent dans la nature, et il est possible d'en observer l'expression dans le règne animal, et même végétal. Lombroso consacre en effet le premier chapitre de son ouvrage à la présentation d'une série d'observations démontrant « l'équivalent du crime chez les plantes et chez les animaux ».[42] Une lecture rétrospective de ce chapitre ne manquera pas d'y voir un remarquable sophisme ou, pour reprendre les termes de Stephen Jay Gould (1983), « la plus ridicule démonstration d'anthropomorphisme qui n'ait jamais été écrite ». On y apprend notamment que les plantes insectivores « commettent de véritables meurtres sur les insectes », que les chats sont spécialistes du « vol domestique », et les fourmis rouges de « l'enlèvement de mineurs ». Lombroso assimile le comportement des animaux à autant de catégories juridiques issus du droit pénal italien, relatant par exemple des cas de parricides chez le renard, ou d'infanticides chez le lièvre et le lapin.

Cette série d'observations vise à démontrer l'existence d'une continuité naturelle entre la férocité des animaux et les crimes commis par notre espèce. Il s'agit pour Lombroso de retrouver les traces d'une « criminalité naturelle » qui atteste de la permanence des instincts violents dans le règne animal et végétal :

> « Quoiqu'on s'efforce d'établir des distinctions, il faut bien l'avouer : il y a une continuité, un passage insensible de plusieurs de ces actes que nous appelons criminels à ceux qui ne constitueraient un crime que pour l'homme. Ainsi les vols des singes, les vengeances des chiens, les meurtres des fourmis peuvent, jusqu'à un certain point, rentrer dans la catégorie du meurtre accompli sur un champ de bataille, de celui qui a pour but la conquête des aliments, ou qui résulte de la lutte pour l'existence — tout comme le meurtre par cupidité, qui, par conséquent, doit être amené à la première catégorie. » (Lombroso, 1887, p. 49).

auquel ni Morel ni ceux qui l'ont suivi n'ont d'ailleurs essayé de fixer des limites précises ; il est donc impossible de baser sur lui une règle de conduite vis-à-vis des nuisibles, des antisociaux, qui restent sans caractère objectif spécifique. » (Féré 1888, p. 80).

[42] Précisons que ce chapitre a été introduit à partir de la troisième édition de *L'Homme criminel*.

Par ailleurs, Lombroso défend l'idée selon laquelle la civilisation marque une rupture dans cette continuité naturelle : l'homme civilisé hériterait d'un certain nombre de comportements lui permettant de s'extraire des lois de la nature, et ainsi réprimer l'expression des instincts violents. Cette idée, fortement influencée par l'évolutionnisme culturel qui se développe parallèlement aux thèses darwiniennes, va amener Lombroso à assimiler la criminalité aux comportements des hommes primitifs. Dans la continuité de son étude sur le comportement criminel des animaux, il va alors s'efforcer de rechercher dans les sociétés primitives les preuves d'une inclination naturelle au crime :

> « Quand on veut étudier le crime chez les sauvages et chez nos premiers ancêtres, on rencontre la même difficulté qui se présente dans l'étude des animaux. Comme chez ces derniers, le crime chez les sauvages n'est plus une exception, mais la règle presque générale. Aussi n'y est-il considéré par personne comme un crime et se confond-il dans ses origines avec les actions les moins criminelles. » (Lombroso, 1887, p. 55).

Lombroso voit notamment dans le sacrifice ou le cannibalisme l'expression des instincts les plus violents du règne animal, une bestialité qui démontre les tendances criminelles des peuplades primitives. Cette tendance portée au crime confirme selon lui leur infériorité biologique et morale : les hommes primitifs sont bloqués à un stade moins avancé du perfectionnement humain, ce qui explique la résurgence d'une criminalité aussi instinctive et cruelle que celle que l'on rencontre dans le règne animal.

Ces observations vont contribuer à définir le cadre théorique de l'anthropologie criminelle italienne, en spécifiant les conditions d'une naturalisation du crime et du criminel. D'abord, le crime est étranger à la civilisation, et l'on peut retracer son origine dans la nature. Il apparaît comme un « phénomène naturel », dans le sens où il répond aux lois de la nature et peut être mis en évidence chez tous les individus vivant à l'écart de la civilisation. L'homme criminel n'est donc pas une « aberration de la nature », puisqu'il concentre lui-même tous les mauvais instincts que l'on rencontre dans un milieu naturel :

> « Le crime, en somme, aussi bien d'après la statistique que d'après un examen anthropologique, apparaît comme un phénomène naturel ; et, si nous voulons emprunter la langue des philosophes, comme un phénomène nécessaire, aussi bien que la naissance, la mort, la conception, les maladies mentales, dont il est souvent une triste variété. — Et voilà que les actes instinctivement cruels des animaux et

même des plantes ne semblent plus séparés comme par un abîme de ceux de l'homme criminel ; chaque différence va diminuant sans cesse et tend à disparaître, si bien que, dans cette forme étrange qu'on dit de méchanceté brutale du criminel, se laisse entrevoir une analogie, je dirais presque une identité, à laquelle ne songeaient certainement pas ceux qui abusaient de ce mot. » (Lombroso, 1887, p. 464).

Par ailleurs, l'assimilation du crime à l'humanité primitive va appeler une autre remarque : les causes du crime sont biologiques et résident dans l'histoire évolutive des individus. C'est en supposant l'existence d'une identité évolutive entre l'homme criminel et l'homme primitif, que Lombroso va former l'intuition que la cause du crime réside dans l'atavisme : la persistance du crime dans les sociétés civilisées s'explique selon lui par la résurgence de caractères ancestraux. Cette théorie inspirée à la fois de l'évolutionnisme de Charles Darwin (1809-1882) et de la théorie de la dégénérescence de Bénédict Augustin Morel (1809-1873), fait du criminel un « sauvage » égaré dans la civilisation, un anachronisme marqué par la réapparition de caractères régressifs.[43]

Si la théorie de l'atavisme a été développée dans les cinq éditions de *L'Homme criminel*, elle a fait l'objet d'une première formalisation en 1870, à la suite d'une autopsie pratiquée sur le corps du célèbre brigand italien Giuseppe Villella. En cherchant à isoler les critères anatomiques permettant de distinguer les criminels des fous, Lombroso constate de nombreuses anomalies crâniennes caractéristiques des vertébrés inférieurs, notamment « une énorme fossette occipitale moyenne » et « une hypertrophie du vernis ». Ces anomalies vont l'amener à définir le criminel comme un « type anthropologique » qui doit être classé au rang des « êtres inférieurs » :

> « A la vue de ces étranges anomalies, comme apparaît une large plaine sous l'horizon enflammé, le problème de la nature et de l'origine du criminel m'apparut résolu : les caractères des hommes primitifs et des animaux inférieurs devaient se reproduire de nos temps » (Lombroso, 1908, p. 32).

Lombroso va alors mener une enquête anthropométrique d'envergure, qui portera sur un échantillon de 350 crânes de criminels morts, auxquels s'ajouteront les mesures de

[43] Précisons que la théorie de l'évolution de Darwin a été introduite en Italie par le zoologiste Filippo de Filippi (1814-1867), lors d'une conférence donnée en janvier 1864, intitulée *L'homme et les singes*.

3839 criminels vivants. Les indices anatomiques – notamment crâniens et céphaliques – sont pour lui un marqueur essentiel de criminalité. Parmi les innombrables caractères recensés par Lombroso, citons notamment la mesure de la capacité et de la circonférence crânienne, l'indice céphalique, l'indice céphalo-rachidien, l'angle facial, le diamètre frontal, le diamètre mandibulaire et zygomatique, l'indice frontal, vertical, nasal, orbitaire, etc. Lombroso mesure également la taille, le poids, la hauteur du tronc, la longueur du thorax, des mains et des pieds, etc. Il relève également une multitude de caractères physiologiques, parmi lesquels la sensibilité à la douleur, la réaction vasculaire et réflexe, la sensibilité tactile ou olfactive, ou encore l'acuité acoustique et visuelle.

Précisons que ces mesures sont le résultat d'une analyse statistique des données recueillies par Lombroso lui-même, mais aussi par des médecins qui lui sont contemporains, notamment par son disciple Enrico Ferri (1856-1929), les Français Arthur Bordier (1841-1910) et Armand Corre (1841-1908), ou encore l'Autrichien Moritz Benedikt (1835-1920). La synthèse de ces données confirme l'hypothèse atavique, car elles révèlent en tous points « l'infériorité biologique » des criminels : leurs anomalies se traduisent principalement par des « asymétries » crâniennes, faciales et sensorielles, et le « développement incomplet » de certains organes cérébraux. Elles confèrent au criminel un aspect tératologique qui le rapproche davantage de l'animal sauvage que de l'être humain.

Parmi les traits anatomiques qui caractérisent les criminels, Lombroso relève des anomalies de la face et du crâne « en proportions égales à celles des sauvages » : l'étroitesse du front, l'épaisseur plus grande de la boite crânienne, la persistance des poils sur le visage, le prognathisme, le développement disproportionné des mâchoires et des pommettes, le développement démesuré des canines, les orbites volumineuses et obliques, etc. Le cerveau lui-même révèle un certain nombre d'asymétries, notamment une « prépondérance du lobe droit du cerveau » et un « développement exagéré des sinus frontaux ». Il présente généralement un volume inférieur à la normale et ses circonvolutions « offrent de fréquentes anomalies ataviques comme la séparation de la scissure calcarone de l'occipitale, la formation d'un opercule du lobe occipital […] et des déviations absolument atypiques comme les sillons transversaux du lobe frontal. » (Lombroso, 1887, p. 199).[44]

[44] Nous renvoyons le lecteur à l'Annexe n°1 qui présente le tableau récapitulatif de l'ensemble des anomalies physiques et psychiques recensées chez le criminel-né.

Les anomalies au niveau des circonvolutions cérébrales sont d'ailleurs mises en perspective avec les observations effectuées par Franz-Joseph Gall, quelques décennies plus tôt. L'influence exercée par ce dernier n'est pourtant pas revendiquée par Lombroso, la phrénologie étant tombée en désuétude depuis le début des années 1850 :

> « Plusieurs auteurs croyaient avoir trouvé dans la disposition anormale des circonvolutions le secret des tendances criminelles. Sans remonter aux exagérations de Gall, qui n'étaient, cependant, pas toujours sans fondements, nous voyons encore, de nos jours, répéter la même affirmation. » (Lombroso, 1887, p. 153).

Précisons que la théorie atavique du crime n'a pas seulement été formulée pour rendre compte des anomalies crâniennes, faciales ou cérébrales observées chez les criminels, mais dans l'objectif d'expliquer leur émergence en tant que nouveau « type anthropologique ». Il s'agit pour Lombroso de montrer que le criminel est un « artefact » du darwinisme, aussi bien sur le plan anatomique que comportemental : c'est une « régression de l'évolution » qui reproduit les instincts violents des hommes primitifs et des animaux inférieurs.

Lombroso va alors chercher à confirmer sa théorie atavique du crime par l'examen comparatif de crânes préhistoriques, mais aussi par une analyse portant sur la *psychologie* des enfants : l'hypothèse d'un « arrêt infantile » va compléter la théorie atavique, en développant l'idée selon laquelle les caractéristiques physiques et psychiques du criminel indiquent une étape moins avancée de l'évolution humaine. Le criminel serait bloqué dans un « état d'enfance prolongée », en raison des nombreuses caractéristiques psychiques ou « sociologiques » qu'il partage avec les enfants et les peuplades sauvages : la paresse, la cruauté, le mensonge, l'oisiveté, l'absence de remords, l'imprévoyance, la coutume du tatouage, qui témoigne d'une « extrême vanité » et d'une « grande insensibilité morale et affective », ou encore l'argot, « qui reproduit, comme le tatouage, les tendances de l'homme primitif » et qui témoigne d'une grande faiblesse intellectuelle. Lombroso ne manquera en effet pas de relever que « le génie n'est, chez eux, qu'une exception » (1887, p. 456).

Si Lombroso accorde une grande importance aux caractéristiques psychologiques des criminels, son étude anthropologique repose essentiellement sur des critères anatomiques et morphologiques. Globalement, c'est dans une accumulation de disproportions, d'asymétries et d'arrêts du développement, que se révèlent les

caractéristiques physiques du criminel. Les asymétries crâniennes, faciales et cérébrales constituent pour Lombroso le caractère le plus remarquable, car elles confèrent au criminel un aspect véritablement monstrueux : « [l]e fait le plus important de tous, c'est-à-dire l'aspect complètement tératologique dû à la réunion de beaucoup d'anomalies dans un même crâne, se rencontre dans la proportion de 43 %, tandis que les anomalies simples, isolées, ne s'y présenteraient que dans la proportion du 21 %. » (Lombroso, 1887, p. 139). Une étude photographique de 424 criminels viendra confirmer cette analyse anthropométrique, en permettant « de fixer la fréquence du type physionomique du criminel dans la proportion de 25 %, avec un maximum de 80 % pour les assassins, et un minimum de 6 à 8 % pour les banqueroutiers, escrocs et bigames. » (1887, p. 200).

1.4.3. Une hiérarchisation des « types criminels »

Bien que l'ensemble des données anatomiques et morphologiques convergent vers la description d'un même « type anthropologique », elles révèlent néanmoins une certaine hétérogénéité à l'intérieur même de la population criminelle. Lombroso remarque en effet que la fréquence et la nature des anomalies diffèrent selon les catégories criminelles étudiées : chaque « type criminel » est anatomiquement et morphologiquement caractérisé. Ainsi, « [c]hez les violateurs [...] presque toujours l'œil est saillant, la physionomie délicate, les lèvres et les paupières volumineuses. La plupart sont grêles, blonds, rachitiques, et parfois bossus. » (Lombroso, 1887, p. 180). Les données anatomiques préciseront encore davantage ce portrait-type, et permettront de le distinguer des autres catégories criminelles : « [i]l semble que chez les voleurs soient plus fréquents les submicrocéphalies, les synostoses, les fronts fuyants, la trococéphalie, l'oxycéphalie, la sclérose du crâne, l'ampleur exagérée des cavités orbitaires ; chez les assassins on trouve nombre de mâchoires volumineuses, d'os wormiens, de platicéphalies et de sutures médio-frontales. » (Lombroso, 1887, p. 144).

Si les catégories criminelles étudiées par Lombroso sont généralement définies par la nature du délit ou du crime (voleurs, violateurs, escrocs, meurtriers...), sa typologie criminelle reconnaît également l'influence des facteurs socio-environnementaux. Il distingue en effet les « criminels-nés » − qui sont poussés au crime « à cause de leur nature innée » – des « criminels par passion » et des « criminels d'occasion ou

d'habitude », qui commettent des crimes « par la force des circonstances ». Chez ces derniers, l'environnement physique et social peut conduire au crime sous l'effet de l'imprévoyance ou de la faiblesse du sens moral de l'individu. Les criminels-nés désignent une classe anthropologique particulière, car ils sont voués au crime par leur organisation biologique. Leurs tendances criminelles sont innées et se manifestent « dès l'enfance, par l'onanisme, la cruauté, le penchant au vol, une excessive vanité, la ruse, le mensonge, leur aversion pour les habitudes de famille, leur rétivité à l'éducation, leur caractère impulsif. » (Lombroso, 1887, p. 16). Les criminels-nés représenteraient selon Lombroso plus de 40% de la population criminelle ; mais surtout, ils dénombrent la plus forte proportion d'asymétries, de difformités et d'arrêts du développement : « [b]eaucoup de ces caractères se réunissent, se soudent les uns aux autres, de façon à nous donner un maximum de monstruosité. » (Lombroso, 1887, p. 465).

La synthèse des données anatomiques et morphologiques recensées dans chacune des catégories criminelles va ainsi révéler l'un des faits les plus remarquables de l'anthropologie criminelle du XIX^{e} siècle : les criminels présentent un degré d'anomalies proportionnel à leur niveau de dangerosité. Le criminel-né devient la figure emblématique de la dangerosité, car il concentre à lui-seul « un maximum de monstruosité ». Ce constat introduit une dimension quantitative fondamentale dans la démarche typologique et classificatoire adoptée par Lombroso, qui perdurera tout au long du développement de l'anthropologie criminelle italienne. Comme le souligne son disciple Raffaele Garofalo dans la deuxième édition de son ouvrage *La Criminologie : étude sur la nature du crime et la théorie de la pénalité* (1890), la classification des « types criminels » ne doit pas reposer sur un partage dichotomique entre la catégorie des « anormaux » et des « normaux » ; elle doit être établie en fonction du degré des anomalies observées :

> « On ne peut donc pas séparer les criminels en deux classes distinctes, l'une d'êtres anormaux, l'autre d'êtres normaux ; on ne peut les classer que d'après le degré, plus ou moins grand, de leur anomalie. » (Garofalo, 1890, p. 98)

Cette gradation des anomalies va donner à Lombroso l'occasion de coupler sa théorie atavique du crime avec une conception hiérarchique des « types criminels » : la classification des criminels sera établie proportionnellement à leur niveau d'atavisme, c'est-à-dire selon le degré plus ou moins grand de leurs anomalies. La proportion d'asymétries, de disproportions ou d'arrêts du développement sera plus importante à mesure que l'on remonte l'échelle de l'immoralité. Au sommet de celle-ci, on retrouve

ainsi le criminel-né : le « type criminel complet », reconnaissable par son anatomie disproportionnée et son absence totale de sens moral. Ses résurgences ataviques font de lui un être physiquement et moralement monstrueux, c'est-à-dire porteur d'un danger proportionnel à son degré de difformité.

De là va découler un principe fondamental pour repérer la dangerosité « naturelle » des criminels : la monstruosité physique de l'individu constitue l'indice visible de sa monstruosité morale. Autrement dit, les asymétries, les disproportions et les arrêts du développement ne sont pas seulement le signe d'une « infériorité biologique », mais le symptôme d'une dangerosité que l'on peut *reconnaître* sur le plan anatomique et morphologique. Les mesures anthropométriques apparaîtront alors comme un outil statistique pour apprécier la dangerosité de l'individu, repérer le criminel-né, et le placer à l'écart de la société.

C'est donc sur cet arrière-plan typologique et classificatoire que se construit la figure du « monstre criminel ». Il est important de préciser ce que recouvre cette notion, car elle apparaît comme un référentiel méthodologique pour fixer les bornes d'une dangerosité ultime, permanente, et donc incorrigible. Sur le plan purement biologique, le criminel-né est d'abord un monstre de l'évolution : un être biologiquement « incomplet » ou « inachevé », dont les caractéristiques physiques le ramènent à une animalité primitive. Mais il est aussi un monstre moral : un individu dénué de tout sens moral, entièrement soumis à ses instincts violents. Précisons que la notion de « monstre moral » n'est pas réellement mobilisée par Lombroso, mais elle apparaît de manière substantielle dans le couplage établi entre la monstruosité du crime et celle du criminel. La typologie criminelle de Lombroso a favorisé la construction d'une « tératologie morale », qui définit l'extension dans le champ moral de la catégorie proprement biologique du « monstre physique ». Le « monstre moral » ne se réduit pas à une simple figure transgressive des lois morales ; il désigne l'expression directe et complète de la monstruosité (physique) du criminel.

Le couplage établi entre la monstruosité physique et morale du criminel est également soumis à des considérations d'ordre psychologique et juridique. Chez Lombroso, l'immoralité du criminel est toujours mise en rapport avec la nature du crime commis (et donc corrélée à l'appréciation de la dangerosité de l'individu) : le criminel-né est un monstre moral en vertu du fait que son crime est monstrueux. Les crimes violents,

impulsifs et sans raison constitueront les faits les plus emblématiques du criminel-né, le symptôme ultime d'une dangerosité absolue.

Comme nous le montrerons plus loin, l'association établie entre la monstruosité du criminel et la monstruosité de son crime est conçue comme un argument clé pour démontrer à la fois l'innéité du crime et l'incorrigibilité du criminel-né. C'est d'ailleurs en couplant la monstruosité physique et morale des criminels que Lombroso va s'assurer de la validité de son anthropométrie du crime. Il est en effet remarquable que ses observations sont relatées de manière à garantir la confirmation réciproque de ses thèses. La figure du monstre physique et moral, le postulat innéiste, et la présomption d'incorrigibilité, s'inscrivent dans un rapport de complémentarité mutuelle, un système circulaire qui réaffirme l'essentiel de sa thèse : l'idée que le criminel-né est porteur d'une pulsion criminelle *intrinsèque* et *irrésistible.*

Cette démonstration circulaire est particulièrement manifeste dans la manière dont Lombroso a progressivement introduit l'hypothèse d'une origine pathologique, pour combler les lacunes et les incohérences de son analyse statistique. Certaines anomalies recensées chez les criminels révèlent en effet les limites de l'explication atavique du crime : la maladie s'ajoute à sa monstruosité, car la théorie de l'atavisme ne suffit pas à expliquer toutes les anomalies observées, ni toutes les formes de criminalité. Précisons que cette thèse n'est pas seulement conçue comme une hypothèse alternative à la théorie atavique ; elle s'intègre dans un modèle explicatif qui conjugue à la fois l'atavisme et la pathologie. Comme le précise Lombroso dans les dernières lignes de *L'Homme criminel*, « [l]es études s'accordent […] pour nous faire voir dans le criminel l'homme sauvage et en même temps l'homme malade » (1887, p. 454).

1.5. La nature pathologique du criminel

1.5.1. Le criminel est un « fou moral »

Comme nous l'avons mentionné plus haut, l'anthropométrie criminelle de Lombroso s'est constituée sur la base d'une analyse comparative entre la figure du criminel et celle du fou. Lombroso s'est efforcé de recenser les critères distinctifs du

criminel, en s'appliquant à montrer que ses caractéristiques physiques et morales « rappellent l'homme sauvage plus que l'aliéné » (Lombroso, 1887, p. 384). Pourtant, à partir de la deuxième édition de *L'Homme criminel*, Lombroso va finir par reconnaître de nombreuses analogies avec ce qu'il appelle le « fou moral ».[45] Dans un souci d'anticiper les critiques qui pourraient souligner les incohérences de sa thèse, il affirme que le fou moral n'est pas un aliéné en tant que tel, parce que sa folie affecte uniquement son sens moral : « [l]e fou moral n'a rien de commun avec l'aliéné ; il n'est pas un malade, il est un crétin du sens moral. » (1887, p. 24). En outre, l'analogie établie entre le criminel et le fou moral reste compatible avec l'existence d'un « type criminel ». Le crime et la folie morale sont décrits comme deux conditions pathologiques de l'esprit, mais appartenant à la même catégorie naturelle : « la folie morale est un genre, dont le crime constitue une espèce. » (1887, p. 395).

L'analogie entre le crime et la folie morale est née du constat selon lequel les criminels-nés concentrent toutes les caractéristiques physiques et morales des fous moraux : « [l]eur déchéance morale correspond à leur déchéance organique » (1887, p. 16). Lombroso va alors s'efforcer de rechercher les preuves statistiques de l'identité entre les criminels-nés et les fous moraux, en dressant la synthèse des données anatomiques et morphologiques recensées par les aliénistes. S'ajouteront également des données relatives à leur répartition dans les institutions asilaires ou carcérales, qui confirmeront encore davantage sa thèse : non seulement les anomalies physiques et morales du criminel recoupent celles du fou moral, mais l'analyse statistique montre que ce dernier occupe davantage les prisons que les hospices d'aliénés. Ces données sont pour Lombroso la preuve que « les prétendues différences » rapportées par les aliénistes ayant étudié des cas de folie morale, proviennent en réalité d'une erreur d'appréciation des données :

> « À regarder de près, tous les caractères que, avec une analyse pénible, les aliénistes ont fini par découvrir pour tracer une différence entre le fou-moral et le

[45] Le concept de « folie morale » est emprunté aux Britanniques James Cowles Prichard (1786-1848) et Henry Maudsley (1835-1918). Ce cas particulier de folie a été étudié par Prichard dans son *Traité sur la folie et d'autres troubles affectant l'esprit* (1835) et par Maudsley dans *Le crime et la folie* (1874). Il désigne selon Prichard (1835) « la folie qui consiste en une perversion morbide des sentiments naturels, des affections, penchants, humeurs, habitudes, des dispositions morales et des impulsions naturelles, sans trouble ni déficit marqué de l'intellect ou des facultés de connaissance et de raisonnement, et en particulier sans aucune illusion délirante ou hallucination. »

> criminel-né, ne servent qu'à confirmer les analogies existant entr'eux. » (Lombroso, 1887, p. 399).

En assimilant le crime et la folie morale, Lombroso va être amené à postuler que le crime lui-même présente une origine pathologique. Cette hypothèse s'est progressivement imposée au vu des données anatomiques et physiologiques recensées chez le criminel, de par la nature des anomalies observées. Lombroso doit se résoudre à admettre que certaines de ces anomalies ne sont pas de simples artefacts de l'évolution, mais des stigmates pathologiques qui font du criminel « un vrai malade chronique » :

> « [L]e strabisme, l'asymétrie crânienne et les graves anomalies histologiques, les ostéomes, les lésions méningitiques, hépatiques et cardiaques, nous montrent [...] chez le criminel un homme anormal avant sa naissance, par l'arrêt du développement ou par maladie acquise de différents organes, surtout des centres nerveux, comme chez les aliénés ; et en font un vrai malade chronique. » (Lombroso, 1887, p. 201).

De toute évidence, l'hypothèse d'une origine pathologique répond à la volonté de constituer une théorie plus générale pour expliquer les asymétries, les disproportions et les arrêts du développement. Elle offre d'abord à Lombroso l'avantage de combler les limites de la théorie de l'atavisme, car elle permet de rendre compte des anomalies que le criminel ne partage pas avec l'homme primitif : certaines anomalies observées chez les criminels (notamment le strabisme, l'inégalité des oreilles et « les impulsions irrésistibles ») ne peuvent être imputées à l'atavisme, car elles n'existent pas chez l'homme primitif.

L'hypothèse pathologique permet également d'expliquer le caractère inné du crime, tout en justifiant la raison pour laquelle tous les enfants – ces « êtres inférieurs » qui abritent « les plus tristes tendances de l'homme criminel » – ne développent pas l'instinct criminel. Si les criminels-nés manifestent « des tendances très précoces à l'immoralité », c'est parce qu'ils souffrent d'un défaut *congénital* qui atteint les premières étapes de développement de ses organes : l'origine du criminel-né est exclusivement endogène car elle « remonte presque toujours au moment de la conception. » (1887, p. 384). Pour Lombroso en effet, seule la maladie peut expliquer l'origine constitutionnelle des tendances criminelles :

> « Ces tendances congénitales sont la meilleure explication de ce qui frappe nos yeux, et nous donnent la clef d'une telle diffusion ; car tout cela n'est, au fond, que le développement d'une maladie, d'un état physiologique. » (Lombroso, 1887, p. 405).

À l'inverse, l'absence de pathologie offre à l'individu les conditions favorables pour réprimer ses instincts violents ; du moins, elle l'épargne d'une mauvaise prédisposition organique qui pourrait faciliter l'expression de l'instinct criminel. C'est pourquoi tous les enfants ne manifestent pas leur penchant criminel à l'âge adulte. D'après Lombroso, l'absence de pathologie garantit à l'enfant les moyens de combattre ses instincts violents, d'autant plus lorsqu'elle est conjuguée à une bonne éducation. Enfin, il existe également des cas où l'instinct criminel se manifeste sous l'effet d'une pathologie acquise : un facteur exogène comme l'alcoolisme peut donner l'impulsion décisive à l'expression du penchant criminel. La pathologie permet ici de rendre compte de la criminalité exogène, et en particulier des anomalies observées chez le « criminel par passion » ou le « criminel d'occasion ». En somme, l'hypothèse d'une origine pathologique permet à la fois d'expliquer toutes les formes de criminalité (innée ou acquise), et la raison pour laquelle les individus sont plus ou moins prédisposés au crime. Elle permet également de rendre compte de la diversité des anomalies observées, en particulier lorsqu'elles ne peuvent être attribuées à l'atavisme :

> « Ainsi s'explique l'infinie variété des formes que revêtent la criminalité et la folie morale, selon qui est compromise l'une ou l'autre des régions de l'organisme, surtout des centres corticaux — tandis que les autres ne subissent aucune modification, du moins appréciable. » (Lombroso, 1887, p. 459).

Lombroso précise que l'atavisme « reste un des caractères les plus constants du criminel-né » (1887, p. 458). Il va alors chercher à concilier sa théorie atavique avec l'hypothèse pathologique, en défendant l'idée selon laquelle les anomalies atavistiques résultent d'un défaut organique inné. Certaines conditions morbides faciliteraient l'expression de l'atavisme, en causant un arrêt du développement des centres nerveux ou de certains organes. L'épilepsie joue un rôle explicatif majeur dans cette thèse, car elle constitue pour Lombroso le « trait d'union [...] qui réunit et fond les fous moraux et les criminels-nés dans une même famille naturelle. » (1887, p. 409). Les processus

dégénérescents qui la caractérisent favoriseraient l'expression de l'instinct criminel par un retour atavistique. L'épilepsie n'est donc pas une simple hypothèse alternative à la théorie atavique, mais l'explication privilégiée pour retracer l'origine des anomalies atavistiques :

> « [L]a théorie de l'atavisme du crime se complète et se corrige par l'adjonction de la nutrition défectueuse du cerveau, d'une mauvaise conductibilité nerveuse, d'un manque d'équilibre des hémisphères, par la condition épileptique. C'est, en un mot, la maladie qui vient s'ajouter à la monstruosité. » (Lombroso 1887, p. 458).

1.5.2. L'hypothèse de « l'épilepsie larvée »

L'hypothèse d'une relation entre l'épilepsie et l'expression de l'instinct criminel repose sur l'analyse statistique d'une série de données relatives à la physionomie et à la psychologie des épileptiques. En s'appuyant sur les recherches menées par des aliénistes italiens (notamment deux dénommés Cividalli et Tonnini), Lombroso va dresser la synthèse des données concernant la nature et la fréquence des anomalies observées. Les données anatomiques révèlent de nombreuses anomalies caractéristiques du « type criminel », notamment des asymétries cérébrales, crâniennes et faciales : exagération de l'index céphalique, prépondérance de la partie postérieure du crâne sur la partie antérieure, faible capacité crânienne, mandibule et pommettes volumineuses, etc. La plupart de ces anomalies se retrouvent exactement dans les mêmes proportions que chez les criminels, ce qui démontre pour Lombroso « la plus complète analogie avec les types criminels. » (Lombroso 1887, p. 492). Les traits de ressemblance se révèlent également dans les troubles de la sensibilité, et en particulier dans « l'asymétrie, le latéralisme, l'obtusité sensorielle » : une plus grande sensibilité tactile à gauche, une inégalité de l'ouïe ou de l'odorat, de la pupille, etc.

Bien que l'hypothèse d'une identité entre l'épileptique et le criminel-né bénéficie d'une longue série d'observations anatomiques ou morphologiques, il est important de noter que Lombroso fonde l'essentiel de sa thèse sur des critères psychologiques : selon lui, « c'est surtout dans l'étude psychologique des épileptiques qu'on peut démontrer leur identité avec les criminels. » (1887, p. 416). Parmi les traits psychologiques recensés par Lombroso, on peut citer l'indolence, la paresse, la vanité, l'irritabilité, l'impulsivité, la

contradiction, la grande émotivité, ou encore la tendance au suicide. Lombroso remarque en particulier que « tous les vices, ou mieux tous les penchants au crime prédominent [...] surtout cette impulsivité, cette irascibilité qui est la cause la plus fréquente des crimes contre les personnes. » (1887, p. 419). Il constate que les criminels souffrant d'épilepsie sont ceux qui commettent les crimes les plus graves et, en l'occurrence, que l'épilepsie est très fréquente chez les criminels-nés : « [u]n fait remarquable est que les épileptiques détenus sont précisément de tous les prisonniers ceux qui commettent le plus de méfaits dans la prison même. » (*ibid*).

De manière intéressante, Lombroso va chercher à recouper ces caractéristiques anatomiques et psychologiques, selon un raisonnement analogique qui permettra de les assimiler à un processus de même nature. Lombroso voit dans les asymétries cérébrales des épileptiques l'expression d'une relation directe entre leurs traits anatomiques et leurs caractéristiques psychologiques. La cause de leur impulsivité transparaît selon lui dans « l'extraordinaire asymétrie cérébrale révélée par la plagiocéphalie », et dans « le latéralisme des fonctions, qui rendent ces hommes, à la lettre, des gens sans équilibre » (p. 438). Autrement dit, l'impulsivité ou l'instabilité des individus trouveraient leur cause biologique dans le déséquilibre des fonctions cérébrales. Lombroso cherche à établir une équivalence en termes de propriétés entre les anomalies observées sur le plan anatomique et psychologique, dans l'objectif de démontrer un lien direct entre la dégénérescence physique et morale des épileptiques.

La notion de « déséquilibre » constitue le point d'ancrage de l'identité entre le criminel-né et l'épileptique : les criminels-nés sont pour Lombroso des individus « mal équilibrés », soumis à une violence impulsive caractéristique du sujet épileptique. Lombroso va d'ailleurs réaffirmer cette idée par le jeu d'une analogie entre l'accès épileptique et l'acte criminel. Les hallucinations des criminels-nés, leur amnésie et la soudaineté du crime témoignent selon lui d'une « remarquable ressemblance avec la décharge épileptique » (1887, p. 416). L'identité entre l'épileptique et le criminel-né va ainsi reposer sur un double principe de symétrie entre la « soudaineté » du crime qui évoque une « décharge épileptique », et l'« irritabilité maladive » des épileptiques qui rappelle celle des criminels-nés. En clair, le souci de confondre l'épileptique et le criminel-né l'amène à faire du crime violent un équivalent de l'épilepsie, et de l'épilepsie une véritable « caricature du crime » :

> « Les accès de fureur épileptique, ou, pour mieux dire, d'équivalent psychique, par leur caractère d'instantanéité soudaine, par leur férocité inutile et exagérée, forment aux yeux de l'observateur attentif, dans une période aigüe et courte, ce que l'épileptique est normalement — une caricature du crime, de même que l'obtusion sensorielle, la stupidité momentanée qui suivent et accompagnent les accès ne sont que l'excès de l'état habituel et souvent congénital de l'épileptique. » (Lombroso, 1887, p. 426)

C'est dans ce cadre que Lombroso réhabilite la notion d' « épilepsie larvée », forgée par Benedict Augustin Morel (1809-1873) en 1860.[46] Cette catégorie nosographique décrit les cas d'épilepsie qui ne s'accompagnent d'aucun symptôme convulsif, mais qui manifestent « les mêmes troubles de la sensibilité des facultés intellectuelles et affectives » que dans l'épilepsie proprement dite, à savoir : l'alternance périodique de phases d'excitation et de dépression, les accès de fureur, l'irritabilité, l'hallucination, l'amnésie, etc. (Morel, 1860).[47] La notion d' « épilepsie larvée » ne fait pas l'objet d'une description plus détaillée dans l'ouvrage de Lombroso. Elle est mobilisée à l'appui de sa thèse car elle permet selon lui de rendre compte des analogies observées entre l'épileptique et le criminel-né, à la fois sur le plan anatomique et psychologique. Les anomalies du criminel-né et de l'épileptique reflèteraient la présence d'une épilepsie maintenue dans un état latent, d'un « état épileptoïde » qui les confond dans une même origine constitutionnelle.

Plus encore, la notion d' « épilepsie larvée » sert d'argument théorique pour consolider l'analogie établie entre l'acte criminel et l'accès épileptique. Lombroso cherche à faire de l'épilepsie un signe de criminalité, et du crime violent un équivalent de la crise épileptique : il s'agit à la fois de démontrer les tendances criminelles des épileptiques, et de définir le criminel-né comme un individu à la constitution épileptoïde.

---

[46] B.A. Morel, « Epilepsie larvée. D'une forme de délire suite d'une surexcitation nerveuse se rattachant à une variété non encore décrite d'épilepsie », *Gazette hebdomadaire de médecine et de chirurgie*, 1860, t. VII, n° 48, pp. 773-775, n° 51, pp. 819-821, n° 52, pp. 836-841.

[47] On peut citer un passage de son traité consacré à l'épilepsie larvée : « [j]'ai donné le nom d'épilepsie larvée à cette situation particulière de souffrance du système nerveux, qui fait que pendant des temps plus ou moins longs, parfois même pendant des années, des individus sont signalés par toutes les perversions de l'intelligence et des sentiments propres aux aliénés par épilepsie, sans que leur mal se révélât au dehors par les crises, les convulsions et par les autres phénomènes symptomatiques du mal caduc. » (Morel, 1860, p. 15).

En somme, ce raisonnement analogique vise à montrer que l'instinct criminel est l'expression d'une pulsion aussi brutale et imprévisible qu'une crise épileptique. La notion d' « épilepsie larvée » permet alors d'assimiler le crime violent et la crise épileptique dans une seule et même condition pathologique : celle de la « force irrésistible », d'une pulsion incontrôlable qui fait du criminel-né ou de l'épileptique un être entièrement soumis à ses instincts violents.[48]

Il est intéressant de constater que la notion de « force irrésistible » parcourt tout l'ouvrage de Lombroso, indépendamment de l'hypothèse de l'épilepsie. Elle apparaît régulièrement dans la caractérisation du criminel-né et le récit de ses crimes, lorsque Lombroso s'applique par exemple à dresser le portrait psychologique d'un criminel qui « se sent entraîné irrésistiblement à tuer », ou d'un autre qui commet un acte brutal « sous le coup d'une impulsion irrésistible ». Comme nous le montrerons dans la suite de cette analyse, la notion de « force irrésistible » constitue le point de jonction entre la monstruosité physique et morale des criminels-nés. Elle permet de réaffirmer leur dangerosité, non seulement parce qu'elle témoigne du caractère incontrôlable de l'instinct criminel, mais aussi parce qu'elle résulte d'une dégénérescence intrinsèque, inscrite dans le corps du criminel : c'est cette « force irrésistible » qui fait du criminel-né un être aussi monstrueux sur le plan moral que sur le plan physique.

1.6. L'ancrage du crime dans le corps du criminel

1.6.1. La dynamique de « l'instinct criminel »

On peut difficilement interroger les notions de « force irrésistible » ou d' « instinct criminel », sans les mettre en perspective avec la problématique générale qui sous-tend le développement de la criminologie lombrosienne : celle des crimes sans raison qui ne s'accompagnent pourtant d'aucun symptôme visible de folie. Cette problématique

[48] Précisons que cette dynamique de la « force irrésistible » est intrinsèquement liée à la notion de « fureur épileptique » mobilisée par Henry Maudlsey pour expliquer les crimes sans raison : « [l]orsqu'un meurtre a été commis sans motif apparent et que sa raison semble inexplicable, il est possible que l'enquête révèle que son auteur était atteint d'épilepsie. » (Maudsley 1874, p. 244).

emblématique du XIX[e] siècle se déploie autour de la question des crimes monstrueux qui échappent aussi bien à l'assignation de raison qu'à l'assignation de folie, c'est-à-dire des crimes dépourvus de toute intelligibilité, mais qui ne sont pas imputables à la folie : le crime est véritablement fou, mais son auteur ne l'est pas.

Le crime monstrueux constitue, pour reprendre les termes de Michel Foucault, le « grand événement criminel, extrêmement violent et extrêmement rare » (Foucault, 1981, p. 406).[49] Il représente un problème nouveau pour l'institution pénale du XIX[e] siècle, car il pose la question de la rationalité du crime, et donc de la légitimité de la réponse punitive. Le problème est précisément celui de l'inintelligibilité de l'acte, qui ne permet pas de justifier l'exercice du pouvoir punitif. La raison ou le motif du crime étant la condition d'application de la loi, l'institution pénale doit se résoudre à déléguer une partie de ses pouvoirs – en l'occurrence, la question de la responsabilité de l'individu et de l'application de la peine − à la psychiatrie.

Le problème des crimes monstrueux a été analysé par Michel Foucault comme le point de départ d'une série de transformations qui ont affecté la pratique pénale au XIX[e] siècle.[50] En retraçant les origines historiques de l'expertise psychiatrique en matière

[49] Dans le cadre de son cours au Collège de France intitulé *Les anormaux* (1974-1975), Michel Foucault a analysé quelques crimes monstrueux ayant défrayé la chronique dans les années 1820-1830, notamment l'affaire Henriette Cornier. Cette affaire pénale est un exemple paradigmatique du crime sans raison, car l'acte criminel n'est motivé par aucun intérêt intelligible. En outre, Henriette Cornier ne montre aucun signe de folie, et témoigne au contraire d'une parfaite lucidité. Pour Michel Foucault, cette affaire révèle les limites de l'alternative binaire entre droit / psychiatrie et raison / folie, telle qu'elle est définie dans l'article 64 du Code pénal de 1810 : « [i]l n'y a ni crime ni délit, lorsque le prévenu était en état de démence au temps de l'action, ou lorsqu'il a été contraint par une force à laquelle il n'a pu résister. » L'affaire Henriette Cornier ne peut pas s'inscrire dans une logique pénale, car l'inintelligibilité du crime est incompatible avec l'exercice du pouvoir punitif : l'absence de raison exclut la punissabilité de l'acte. Elle ne peut pas non plus répondre à l'alternative psychiatrique, car l'absence de raison ne peut être démontrée. En clair, la folie du crime ne trouve aucune justification : ni dans la folie de son auteur, ni dans un quelconque motif ou intérêt intelligible.

[50] Pour Michel Foucault, il s'agit d'un problème nouveau, qui ne pouvait pas se poser avant le XIX[e] siècle. Jusqu'ici, on répondait à la violence d'un crime par la démesure de la punition : les crimes « atroces » ou « contre-nature » étaient l'occasion de réaffirmer la toute-puissance du souverain par la sévérité excessive du châtiment (Foucault, 1972, 1975). Au début du XIX[e] siècle, la question de la monstruosité criminelle s'est ensuite posée sous la forme d'une catégorie nosographique indistincte – la monomanie homicide − qui visait à rechercher le fond de criminalité chez les fous, et le fond de folie chez les criminels : ce domaine

pénale (entre 1810 et 1830), il a analysé la manière dont la psychiatrie naissante a contribué à déplacer le point d'application de la peine, du crime vers le sujet criminel lui-même. Il montre comment le pouvoir psychiatrique s'est saisi du problème de la monstruosité criminelle dès le début du XIX^{e} siècle, en cherchant à révéler les éléments pathologiques qui pourraient démontrer la dangerosité naturelle (intrinsèque) de l'individu. Il ne s'agit pas seulement pour les psychiatres de pathologiser le crime, mais d'établir un lien entre le criminel et son crime, de montrer « comment l'individu ressemblait *déjà* à son crime avant de l'avoir commis. » (Foucault, 1974-1975, p. 15). L'expertise psychiatrique en matière pénale vise à resituer l'acte criminel dans l'histoire du sujet, en recherchant les motifs ou les penchants qui garantissent l'imputabilité du crime à son auteur.

D'après Michel Foucault, le pouvoir psychiatrique s'est finalement détourné de la question de la responsabilité pour fonder son autorité sur l'appréciation de la dangerosité.[51] Il s'est donné pour fonction de rechercher les preuves d'une criminalité inhérente au criminel, c'est-à-dire d'identifier les déficiences physiologiques, psychologiques ou morales qui constituent les prémices d'une dangerosité en devenir. Ce rapport d'identité entre le criminel et son crime va ainsi offrir à l'expertise psychiatrique le moyen d'élucider le problème des crimes sans raison : il permettra de reconstruire une intelligibilité à l'acte, en recherchant dans l'histoire personnelle de l'individu les éléments précurseurs de son crime. Par ailleurs, en démontrant « la ressemblance du sujet à son acte », l'expertise psychiatrique justifie dans le même temps la nécessité d'adapter la détermination de la peine à la nature du criminel : il ne s'agit plus de punir le crime mais

---

indifférencié du crime-folie définit la maladie qui ne se manifeste que dans le crime, et qui n'a pour symptôme que le crime lui-même (Foucault, 1981).

[51] Comme le montre Foucault, l'expertise psychiatrique a fondé son autorité sur un ensemble de jugements appréciatifs et normatifs, dont l'objectif est de repérer l'individu représentant une menace pour la société. Elle a inventé la figure de « l'individu dangereux » pour affirmer sa propre légitimité dans le champ pénal. La fonction de l'expertise consiste alors à isoler le motif juridique, médical, psychologique et moral permettant de démontrer l'identité entre le crime et le criminel (Terrel 2003) : il s'agit à la fois de retrouver les germes d'une criminalité dans la psychologie du criminel, d'apprécier sa dangerosité, de déterminer la peine appropriée, et de fixer les termes d'une correction possible. Comme le précise Jean Terrel (2003), « [l]a psychiatrie ne se contente pas de dire quels sont les dangers pour une société. Elle nomme le danger, fait apparaître derrière des actes, des individus dangereux liés à leurs crimes, délits, par une psychologie de l'homme dangereux dont la construction est bien l'effet de la psychiatrie en tant que savoir. » (Terrel 2003, p. 41).

le criminel lui-même, et plus précisément « ce qui le rend criminel, ses motifs, ses mobiles, sa volonté profonde, ses tendances, ses instincts. » (Foucault, 1981, p. 401).

C'est donc autour du problème de la monstruosité criminelle, et en particulier des crimes sans raison, que s'est véritablement posée la question de la nature du criminel. L'impasse de la logique punitive a favorisé un basculement théorique dans la compréhension du phénomène criminel, de la question de la rationalité de l'acte à la caractérisation des instincts, des tendances ou des penchants individuels : on ne peut comprendre ou juger les crimes monstrueux sans tenir compte de la nature du criminel.[52] Comme le montre Michel Foucault, la psychiatrie du XIX^e^ siècle a alors cherché à résoudre ce problème par le recours à la notion d' « instinct criminel ». Cette notion répond à l'exclusion réciproque de la qualification juridique et psychiatrique des crimes monstrueux, car elle permet de les confondre dans une seule et même mécanique pathologique. Le crime monstrueux est désormais conçu comme un acte instinctif, c'est-à-dire comme résultant d'une impulsion irrésistible, déterminée par les dispositions organiques de l'individu. Le cas particulier des crimes sans raison trouvera dans cette notion l'intelligibilité qui faisait défaut : si le crime est sans raison, c'est précisément parce qu'il est instinctif. L'instinct criminel rend intelligible les crimes sans motif, car il permet de combler l'absence de raison par l'explication pathologique (la dégénérescence, le défaut de développement, etc.). Comme le précise Michel Foucault dans son cours au Collège de France sur *Les Anormaux* (1974-1975) :

> « L'instinct, en effet, permet de réduire en termes intelligibles cette espèce de scandale juridique que serait un crime sans intérêt, sans motif et, par conséquent, non punissable ; et puis, d'un autre côté, de retourner scientifiquement l'absence de raison d'un acte en un mécanisme pathologique positif. » (Michel Foucault, *Les Anormaux*, 1974-1975, cours du 12 février 1975, p. 95).[53]

[52] Précisons que ce basculement a été favorisé par le principe d'économie du pouvoir punitif, qui établit un lien de proportionnalité entre la nature du crime et le degré de la punition. À partir du XIX^e^ siècle, il ne s'agit plus de répondre aux crimes monstrueux par la démesure de la punition, mais de « trouver une certaine unité de mesure entre le crime et le châtiment » (Foucault, 1974-1975, p. 60) : l'essentiel, pour le pouvoir punitif, est d'adapter la punition à la nature du criminel, selon un juste équilibre qui est conçu de telle sorte à empêcher la récidive.

[53] L'élucidation des crimes monstrueux a permis à la psychiatrie d'acquérir sa pleine légitimité au sein de l'institution pénale, en fondant l'évaluation de la dangerosité sur un savoir médical et psychologique (Terrel

Cette pathologisation du crime par l'instinct a été systématisée par l'anthropologie criminelle italienne. Comme nous l'avons vu précédemment, elle s'est en particulier structurée autour de la dynamique de la « force irrésistible » : d'une violence irrépressible exprimant un danger permanent, et qui se cristallise dans la figure du « monstre physique et moral ». Ici encore, cette dynamique morbide contribue à expliquer le crime inexplicable, car elle marque l'ancrage du crime dans les dispositions organiques de son auteur : avec l'anthropologie criminelle italienne, le problème des crimes monstrueux trouve sa solution dans la monstruosité du criminel lui-même. Autrement dit, cette dynamique de la « force irrésistible », de même que la notion d'« instinct criminel », répondent à l'ambition de doubler le crime dans le corps du criminel. Elles visent à démontrer la dangerosité du criminel physiquement monstrueux, et relèvent en ce sens du même principe : l'idée que le criminel doit ressembler à son crime.

Il est intéressant de constater une certaine perméabilité entre les principes théoriques de l'anthropologie criminelle italienne, et l'effort déployé par la psychiatrie du XIX^{e} siècle pour démontrer un lien entre le criminel et son crime. Dans les deux cas, l'ambition de relier le crime à son auteur constitue la condition indispensable pour apprécier la dangerosité de l'individu. La notion d' « instinct criminel » répond précisément à cette finalité, dans la mesure où elle permet d'ancrer le crime dans la psychologie et/ou l'anatomie des criminels. Elle justifie la possibilité de rechercher l'ensemble des déficiences psychologiques ou anatomiques qui démontrent la dangerosité des individus : l'objectif est de faire apparaître l'irrégularité qui relie le sujet à son acte, de repérer l'instinct qui atteste d'une pulsion criminelle intrinsèque et irrésistible.

Pour autant, il est important de préciser que la notion d' « instinct criminel » n'a pas exactement la même fonction dans la criminologie de Lombroso et dans l'expertise psychiatrique du XIX^{e} siècle. Elle a été mobilisée par les psychiatres pour résoudre l'énigme des crimes monstrueux, en révélant la criminalité inhérente au criminel, et en faisant de lui « un sujet malade avant le crime » (Terrel 2003). Comme nous l'avons précisé plus haut, l'instinct criminel a été le moteur de l'intégration de la psychiatrie dans

2003). Elle s'est dotée d'un pouvoir normatif visant à apprécier la dangerosité des individus par le prisme de leurs irrégularités et de leurs anomalies. D'après Michel Foucault, cette pathologisation du criminel marque le passage de la figure du « monstre » à celle de « l'anormal », d'un nouveau registre du savoir-pouvoir qui fixe les termes du danger social : « ce n'est pas le crime qui est la maladie du corps social, mais le criminel qui, en tant que criminel, pourrait bien en effet être un malade » (Foucault 1974-1975, p. 62).

le champ pénal, car il permet à la fois d'apprécier la dangerosité de l'individu et de fixer les termes d'une correction possible : l'appréciation de la dangerosité est corrélative d'une possibilité de redressement ou de « normalisation » (Foucault 1974-1975).[54] Autrement dit, la notion d'instinct répond essentiellement à un effort de psychologisation et de pathologisation du criminel. Elle vise à garantir l'imputabilité du crime à son auteur, et recouvre nécessairement une qualification médicale.

Avec l'anthropologie criminelle de Lombroso, la notion d'instinct criminel s'est imposée comme un argument théorique pour rechercher le fond de monstruosité dans toutes les formes de criminalités. La question de la nature du criminel s'est en effet structurée autour d'un « soupçon de monstruosité » qu'il est possible de mesurer (Foucault, 1974-1975).[55] L'essentiel, pour Lombroso, n'est pas de pathologiser le criminel, mais de réaffirmer la dangerosité du criminel (physiquement) monstrueux. La notion d'instinct s'enracine dans l'objectivation anatomique et psychologique du crime et du criminel : c'est précisément autour de cette notion que s'articule le couplage entre la monstruosité physique et morale des criminels. En un mot, le criminel doit ressembler à son crime : si le crime est monstrueux (moralement), le criminel doit l'être tout autant physiquement.

L'idée selon laquelle le criminel ressemble à son crime constitue la ligne directrice de l'anthropologie criminelle de Lombroso, le présupposé implicite qui sous-tend le développement de son anthropométrie du crime. Si l'on considère l'ensemble des arguments invoqués à l'appui de ses thèses, on constate que ce présupposé a véritablement guidé sa recherche des causes organiques du crime : il définit le cadre théorique dans lequel se déploient les différentes hypothèses avancées pour expliquer l'origine des

[54] Comme le montre Michel Foucault dans son premier cours sur *Les anormaux* : « [l]'essentiel de son rôle, c'est de légitimer, dans la forme de la connaissance scientifique, l'extension du pouvoir de punir à autre chose que l'infraction. L'essentiel, c'est qu'elle permet de replacer l'action punitive du pouvoir judiciaire dans un corpus général de techniques réfléchies de transformation des individus. » (p. 14).

[55] Michel Foucault a montré comment l'économie du pouvoir punitif au XIX^e^ siècle a favorisé un basculement théorique vis-à-vis de la question de la monstruosité criminelle. Alors que, jusqu'au XVII^e^-XVIII^e^ siècle, l'individu monstrueux porte en lui-même un soupçon de criminalité, à partir du XIX^e^, le rapport s'inverse. Il y a un soupçon de monstruosité dans toute criminalité : « [t]out criminel pourrait bien, après tout, être un monstre, tout comme autrefois le monstre avait une chance d'être un criminel. » (Foucault, Cours du 29 janvier 1975, p. 55).

anomalies des criminels. La théorie atavique du crime, la théorie de l'arrêt infantile, et la théorie de l'épilepsie sont en effet dictées par le souci de confondre les anomalies physiques et psychiques recensées chez le criminel. Il s'agit pour Lombroso d'établir un principe d'équivalence en termes de propriétés entre les anomalies observées sur le plan anatomique et psychologique : le criminel doit ressembler à son crime, dans le sens où ses caractéristiques comportementales doivent refléter la nature de ses anomalies anatomiques ou morphologiques.

La théorie atavique du crime repose sur l'hypothèse d'un couplage entre la dégénérescence biologique et morale des criminels : leurs comportements primitifs et instinctifs constituent le miroir révélateur de leurs anomalies atavistiques, le pendant psychologique de leur dégénérescence biologique. Rappelons en effet que la théorie de l'atavisme s'est constituée dans l'objectif de démontrer l'infériorité biologique et morale des criminels et des hommes primitifs : il s'agissait pour Lombroso de prouver l'existence d'une relation directe entre leurs caractères régressifs et leur bestialité héritée des animaux inférieurs. La théorie de l'arrêt infantile perpétue ce même principe, car elle vise à montrer que cet « état d'enfance prolongé » s'exprime aussi bien sur le plan biologique que moral. D'après Lombroso, les traits infantiles du criminel (sa paresse, son oisiveté, sa cruauté...) sont le témoin visible d'un arrêt du développement de ses organes cérébraux : c'est en conjuguant l'immaturité comportementale et cérébrale des criminels que Lombroso justifie le bien-fondé de sa théorie de « l'arrêt infantile ». Enfin, la théorie de l'épilepsie introduit un double principe de symétrie entre l'instabilité comportementale et l'instabilité du système nerveux d'une part, et entre le crime violent et la crise épileptique d'autre part. Il s'agit non seulement d'établir un lien entre l'impulsivité des individus et le déséquilibre de leurs fonctions cérébrales, mais aussi de faire du crime violent un équivalent de la crise épileptique. Le criminel ressemble à son crime, dans le sens où son déséquilibre constitutionnel constitue le marqueur biologique d'une « pulsion irrésistible » aussi brutale et imprévisible qu'une crise épileptique.

Par ailleurs, il est intéressant de remarquer que l'appréciation de la dangerosité est elle-même fondée sur le présupposé d'une relation directe entre le criminel et son crime : ce lien est posé comme un principe heuristique pour repérer l'individu dangereux. Toute l'anthropologie criminelle de Lombroso s'est construite autour de l'ambition de rechercher la monstruosité (physique) du criminel derrière la monstruosité de son crime : il s'agit de montrer que le crime monstrueux trouve son origine dans la monstruosité de son auteur. La conception hiérarchique des types criminels a joué un rôle majeur dans la

consécration de ce principe, puisqu'elle a permis d'ériger la figure du « monstre physique et moral », mais aussi plus généralement, d'établir un lien de proportionnalité entre le niveau de dangerosité du criminel et le degré de ses anomalies anatomiques ou morphologiques. En clair, le présupposé selon lequel le criminel ressemble à son crime répond à la double ambition de reconnaître la dangerosité des individus en évaluant le degré de leurs anomalies, et de repérer le criminel-né afin de l'exclure de la société.

Ce principe offre à Lombroso l'avantage de garantir la validité de son anthropométrie du crime, car il permet de justifier la pertinence de son approche quantitative du crime : si les mesures anthropométriques confirment l'association établie entre le degré d'anomalies et le niveau de dangerosité des individus, alors il n'y a pas besoin d'aller rechercher la cause du crime en dehors du corps du criminel. L'objectif, pour Lombroso, n'est pas seulement de confirmer le caractère inné de ses anomalies (et de la criminalité elle-même), mais surtout de montrer que le crime est contenu de manière intrinsèque dans le corps du criminel : il s'agit de confondre la monstruosité du crime dans celle du criminel, afin d'apporter la preuve que ce dernier porte en lui les germes d'une dangerosité maintenue dans un état latent, et qui n'attend qu'à s'exprimer.

De ce point de vue, l'idée selon laquelle le criminel ressemble à son crime démontre à elle seule l'incorrigibilité du criminel-né : c'est précisément cet « instinct criminel » qui disqualifie le criminel-né de sa légitimité sociale et juridique. Lombroso démontre l'impuissance des mesures correctives (l'éducation, la morale, etc.) par le caractère inné et incontrôlable de l'instinct criminel. Ce danger permanent le prive de sa capacité à combattre ses instincts violents, et le soustrait à toute possibilité de redressement. En somme, comme le relève Michel Foucault dans une conférence prononcée en octobre 1977, intitulée *L'évolution de la notion d' « individu dangereux » dans la psychiatrie légale du XIX^e^ siècle* :

> « On voit bien la série des déplacements demandés par l'école anthropologique : du crime vers le criminel, de l'acte effectivement commis vers le danger virtuellement inclus dans l'individu, de la punition modulée du coupable à la protection absolue des autres. » (Foucault, 1981, p. 417).

L'exigence de protection sociale constitue l'arrière-plan idéologique et politique de l'anthropologie criminelle de Lombroso. Il justifie la pertinence de son approche anthropométrique pour repérer le criminel intrinsèquement dangereux, mais il constitue également un argument pour l'exclure de la sphère sociale. À ce titre, il convient

d'examiner les arguments mobilisés par Lombroso pour justifier l'incorrigibilité du criminel-né. La présomption d'incorrigibilité pose la question de la cohérence théorique entre la manière dont Lombroso caractérise le criminel-né et la réponse apportée à la question de sa prise en charge : elle souligne la nature équivoque du criminel-né − criminel-monstrueux et criminel-malade − dont le statut pathologique n'appelle pas une logique corrective, mais une politique répressive, d'exclusion sociale.

1.6.2. L'incorrigibilité du criminel-né

Lombroso a donc construit son anthropométrie du crime en réponse à un impératif de protection sociale. Tout au long de *L'Homme criminel*, il s'efforce de contester l'efficacité des politiques pénales pour résoudre le problème de la criminalité, en mobilisant les statistiques qui démontrent « la régularité et la fréquence toujours plus grande des récidives parmi les délinquants » (Lombroso, 1887, p. 278). Il note « l'insuffisance des mesures répressives » pour endiguer le phénomène de la récidive, et affirme que la prison en est bien souvent la cause principale : « ce n'est point, certes, le système pénitentiaire qui prévient les récidives ; les prisons en sont, au contraire, la cause principale. » (p. 281). Pourtant, si Lombroso insiste sur l'inefficacité des pratiques répressives, il ne voit pas d'autre alternative à la prise en charge du criminel-né. Il réclame pour lui « la détention perpétuelle » car la nature innée et incontrôlable de ses instincts le rend insensible aux mesures correctives comme l'éducation ou la morale :

> « L'éducation peut, en effet, empêcher une bonne nature de passer du crime infantile et transitoire au crime habituel − mais elle ne peut changer ceux qui sont nés avec des instincts pervers. » (p. 121).

La présomption d'incorrigibilité souligne toute l'ambivalence de la figure lombrosienne du criminel-né, qui caractérise à la fois un monstre à rejeter hors de la société, et un individu malade, dont les stigmates pathologiques font de lui un être « biologiquement inférieur ». Le criminel-né s'inscrit en effet dans le double registre de la pathologie et de la monstruosité ; pourtant, sa nature pathologique n'ouvre aucune perspective de correction de l' « instinct criminel ». Elle ne favorise pas non plus un renouvellement des questions relatives à la responsabilité (morale ou pénale), comme

c'est le cas dans l'organologie de Gall. En l'occurrence, Lombroso ne s'intéresse pas à la question de responsabilité des criminels, car sa seule exigence est de quantifier le niveau de danger qu'ils représentent pour la société. Alors que l'organologie de Gall était motivée par la volonté d'améliorer l'application des peines (en adaptant le degré de responsabilité aux dispositions organiques de l'individu), l'anthropométrie criminelle de Lombroso répond essentiellement à une exigence de protection sociale : l'enjeu est de repérer les individus jugés indésirables et irrécupérables, afin de les mettre hors d'état de nuire.

De fait, la priorité accordée à l'appréciation de la dangerosité l'amène à défendre une conception utilitariste de la peine. Selon lui, la peine doit être conçue en termes d'utilité sociale. Elle ne doit pas être pensée comme une réponse punitive, mais comme un mécanisme de défense sociale : la sévérité de la sanction doit donc être indexée sur le niveau de dangerosité de l'individu, afin de prévenir le risque de récidive.[56] C'est cette même exigence qui l'amène à rejeter toute possibilité de redressement du criminel-né. L'impératif de protection sociale rend légitime le glissement opéré entre une dangerosité « à risque » et une dangerosité définitive, immuable, et par conséquent incurable. Il est en effet important de préciser que dans la pensée lombrosienne (et plus globalement, dans la pensée du XIX^{e} siècle), la notion de « risque » n'est pas dissociée de la notion de « danger » (Le Bras 2018) : la présence d'un risque suffit à généraliser la représentation d'un danger social.[57]

C'est pourquoi Lombroso propose de substituer les catégories normatives du droit pénal par une typologie reposant sur la mesure quantitative du danger que représente l'individu pour la société. Il cherche à naturaliser le fonctionnement de l'appareil juridique, à partir des données statistiques relatives au degré d'atavisme attribué à chaque type de criminalité. Reprenant une formule employée par Alfred Maury, Lombroso

[56] Cette conception rejoint celle de Cesare Beccaria (1738-1794) qui avait, un siècle plus tôt, proposé une hiérarchisation des peines basée sur le niveau de dangerosité des individus : seule l'utilité sociale doit justifier l'exercice du pouvoir punitif. Elle se rapproche également du principe utilitariste de Jeremy Bentham (1748-1832), selon lequel la détermination de la peine doit permettre de maximiser le bien-être collectif.

[57] Comme le précise Anatole Le Bras (2008), le concept de « dangerosité » n'est lui-même pas employé au XIX^{e} siècle. Les criminologues et les aliénistes évoquent plutôt la notion d' « état de danger » ou « d'individu dangereux ». Pour une réflexion critique sur l'autonomisation de la notion de « risque de dangerosité », voir Castel (1983).

affirme que « nous sommes gouvernés par des lois muettes, mais qui ne tombent jamais en désuétude, et qui régissent la société avec plus d'autorité que les lois inscrites dans les codes » (Lombroso, 1887, p. 463). En clair, les lois de l'atavisme doivent selon lui dicter le principe d'individualisation et de hiérarchisation des peines. La sanction infligée aux criminels-nés, aux « criminels d'occasion », et aux « criminels par passion » répond donc à différentes logiques, selon le degré d'anomalies et le niveau de dangerosité qui lui est corrélatif. Dans la mesure où les criminels « d'occasion » et « par passion » peuvent rectifier leurs instincts grâce à l'éducation ou à la morale, il serait néfaste de les condamner à une longue peine de prison, cette « école du crime » qui ne ferait qu'accroître leur dangerosité. A l'inverse, la détention perpétuelle constitue la seule alternative pour les criminel-nés, car le caractère inné et irrésistible de leurs instincts les soustrait à toute possibilité de redressement.

Sur le plan théorique, les arguments mobilisés par Lombroso pour justifier l'exclusion sociale du criminel-né sont de deux ordres. Il s'agit d'une part, de faire de la monstruosité un qualificatif de dangerosité et d'incorrigibilité, et d'autre part, d'essentialiser le crime de manière à déresponsabiliser la société dans le développement et la prise en charge de la criminalité. Lombroso cherche à montrer que le criminel-né ne relève ni d'une qualification juridique, ni d'une qualification médicale, et qu'il résiste par conséquent à toute prise en charge. L'appréciation de sa dangerosité est telle qu'il est impossible de lui donner un sens médical (ni même un statut juridique) : cette appréciation ne s'opère en effet pas sur le terrain de la pathologie (de l'infirmité ou de la défectuosité), mais sur celui de la monstruosité.[58] Elle ne peut avoir pour conséquence qu'une expulsion hors de la société, car le statut social, juridique et médical du criminel-né est révoqué au profit d'une essentialisation de sa dangerosité.

Le premier argument consiste à reléguer toute forme de monstruosité à la catégorie de la « dangerosité incorrigible », c'est-à-dire à une construction abstraite qui combine

[58] Nous nous appuyons ici sur la différence établie par Michel Foucault entre l'infirmité et la monstruosité, dans son cours au Collège de France consacré à l'émergence de la figure du « monstre criminel » : « [l]'infirmité est bien, en effet, quelque chose qui bouscule l'ordre naturel aussi, mais l'infirmité n'est pas une monstruosité, parce que l'infirmité a sa place dans le droit civil ou dans le droit canonique. L'infirme a beau ne pas être conforme à la nature, il est en quelque sorte prévu par le droit. En revanche, la monstruosité est cette irrégularité naturelle, telle que, lorsqu'elle apparaît, le droit se trouve remis en question, le droit n'arrive pas à fonctionner. » (Foucault, cours du 22 janvier 1975, p. 43).

un excès d'inhumanité et d'inintelligibilité. Comme nous l'avons précisé plus haut, la figure du monstre lombrosien résulte d'une double transgression : une transgression des lois de la nature et une transgression des lois de la société. Il est un « monstre biologique » en raison des anomalies atavistiques qui le relèguent au rang des « êtres inférieurs », et un « monstre moral » car ses crimes expriment une immoralité absolue qui s'oppose aux règles de fonctionnement de la société. Cette transgression biologique et morale fait de lui un problème insoluble pour le droit pénal, car elle le rend étranger à l'humanité et aux lois qui régissent l'ordre social : en caractérisant le criminel-né sous les traits d'un monstre, Lombroso définit sa dangerosité comme un état permanent et incurable, et dissout définitivement son statut juridique. Non seulement il justifie son exclusion hors de la société, mais il rend également inintelligible la question de sa responsabilité pénale. Le criminel-né est disqualifié comme sujet de droit, car il témoigne d'une altérité radicale contre laquelle la société ne peut lutter.

Si Lombroso montre l'impuissance de la société dans la prise en charge du criminel-né, il nie également son rôle dans l'émergence de ce dernier. Rappelons que selon lui, la criminalité n'est pas un problème social mais un problème exclusivement biologique et anthropologique. En outre, bien que Lombroso reconnaisse l'influence des facteurs exogènes dans l'expression de l'instinct criminel (notamment le défaut d'éducation ou l'alcoolisme), ces facteurs n'agissent que sous l'effet d'une mauvaise prédisposition organique. La cause première du crime est endogène, et réside dans les asymétries anatomiques et morphologiques, les disproportions et les arrêts du développement. Autrement dit, l'origine du crime est imputable à la monstruosité (biologique) du criminel. Cette essentialisation du crime disculpe finalement la société de toute responsabilité dans le développement de la criminalité. Elle fait peser sur l'individu la fatalité d'un mal inscrit dans ses prédispositions organiques, qui le rend *définitivement* dangereux pour autrui.

Ainsi, si Lombroso cherche à justifier la validité de son anthropométrie du crime, le caractère tautologique de ses arguments paraît désormais évident. Comme nous l'avons mentionné précédemment, Lombroso mobilise une série de présupposés ou de concepts qui s'inscrivent dans un système de confirmation réciproque : le postulat innéiste, la figure du monstre physique et moral, la présomption d'incorrigibilité, ou encore les notions d' « instinct criminel » et de « force irrésistible ». Ce système théorique vise à démontrer la dangerosité et l'incorrigibilité du criminel présentant d'importantes anomalies anatomiques ou morphologiques : le criminel-né est incorrigible en vertu du

fait que sa dangerosité est innée, incontrôlable, et intrinsèquement liée à une monstruosité « de nature ». Autrement dit, Lombroso cherche à légitimer par un discours scientifique l'exclusion du criminel physiquement monstrueux. Il cherche à doter son anthropométrie d'une fonction prescriptive, pour garantir la possibilité de reconnaître le criminel porteur d'un danger permanent et incorrigible.

Avant de retracer la genèse du projet neurocriminologique, il convient de dresser le portrait général de l'étude du cerveau des criminels au XIX^{e} siècle. D'abord, si l'on s'en tient à la démarche qui sous-tend la recherche des causes biologiques du crime, on peut isoler deux paradigmes correspondant respectivement à l'organologie de Gall et à l'anthropologie criminelle de Lombroso : une démarche localisatrice qui consiste à identifier la condition matérielle nécessaire à l'exercice du « penchant au meurtre », et une démarche typologique et classificatoire, centrée sur la mesure quantitative du danger que représente l'individu pour la société. Ces deux grands principes méthodologiques s'inscrivent dans un cadre théorique commun qui défend l'essentialisation du criminel, c'est-à-dire la thèse selon laquelle le criminel, de par la nature de ses structures physiologiques, possède toutes les conditions matérielles nécessaires à l'exercice de ses instincts violents : de la phrénologie de Gall à l'anthropologie criminelle de Lombroso, la recherche des causes biologiques du crime est dictée par un postulat innéiste qui est conçu sur le modèle d'une disposition intrinsèque. Elle vise à montrer que la criminalité est une qualité immanente au sujet criminel, et qu'il est possible d'en reconnaître les signes dans la conformation de son cerveau ou dans sa structure morphologique.

De cette conception innéiste de la criminalité découle une ambition réformatrice qui défend l'individualisation et la hiérarchisation des peines. Précisons néanmoins que ce principe n'obéit pas tout à fait à la même logique chez Gall et chez Lombroso, car ce dernier propose de substituer la question de la responsabilité de l'individu au profit de l'appréciation de sa dangerosité : l'exigence de protection sociale est centrale chez Lombroso. Si le cerveau est conçu comme un marqueur objectif de la mesure de l'instinct criminel, c'est avant tout pour repérer les « incorrigibles » et les exclure de la société. Il ne s'agit pas seulement d'ancrer le crime dans les dispositions organiques du criminel, mais de faire de la monstruosité du criminel le stigmate caractéristique de sa dangerosité ; autrement dit, de fixer le critère de dangerosité de l'individu sur le degré de ses anomalies anatomiques ou morphologiques.

Nous avons également montré que l'organologie de Gall et la criminologie de Lombroso étaient gouvernées par un certain nombre de présupposés qui ont joué un rôle heuristique décisif dans la recherche des causes biologiques du crime : le principe d'identité psychophysiologique, qui a permis à Gall de faire le lien entre sa pluralité des facultés et sa pluralité des organes, le postulat d'une « infériorité biologique » et d'une conception psychopathologique du criminel, que l'on rencontre spécifiquement chez Lombroso, ou encore le présupposé selon lequel le criminel doit ressembler à son crime, qui imprègne toutes les hypothèses avancées pour expliquer l'origine des anomalies observées chez les criminels. La spécificité de la démarche lombrosienne est d'articuler son anthropométrie du crime autour du recoupement des données observées sur le plan anatomique, morphologique, et psychologiques : ses thèses sont fondées sur un principe de correspondance entre des traits comportementaux typiques (immaturité, instabilité, irritabilité, impulsivité, etc.), et un ensemble d'anomalies physiques dénotant le même type de propriété (défaut de développement des organes cérébraux, déséquilibre des fonctions cérébrales, asymétries et disproportions, etc.).

En outre pour la suite de notre analyse, il est important de garder à l'esprit que l'anthropologie criminelle de Lombroso s'est construite autour du problème des crimes monstrueux. En ce sens, l'hypothèse de l'épilepsie et la dynamique de la « force irrésistible » ont joué un rôle majeur dans la consolidation de la théorie de l'innéité criminelle, puisqu'elles ont contribué à faire du crime violent un acte *instinctif* résultant d'un déséquilibre constitutionnel : elles ont été mobilisées pour résoudre le problème des crimes violents qui ne pouvaient être imputés à la folie proprement dite, favorisant ainsi un renouvellement du débat sur la nature pathologique du criminel et la prise en charge de la criminalité.

CHAPITRE 2

VERS UNE DÉFINITION NEUROPHYSIOLOGIQUE DE LA DANGEROSITÉ

L'histoire de la neurocriminologie et du neurodroit s'enracine dans une même réflexion sur l'apport potentiel de l'électroencéphalographie (EEG) dans l'explication du comportement criminel.[59] Dès le début des années 1940 dans le milieu anglo-saxon, alors que les applications cliniques de l'EEG n'en sont qu'à leurs balbutiements, on voit apparaître une série de publications démontrant un lien entre la perturbation de l'activité électrique du cerveau et les troubles du comportement ou de la personnalité (e.g. Lindsley and Cutts 1940 ; Hill and Watterson 1942 ; Silverman 1943). En 1943, les psychiatres britanniques Denis Hill (1913-1982) et William Sargant (1907-1988) signent la première tentative visant à utiliser les données EEG comme une preuve de l'irresponsabilité pour cause psychiatrique.[60] Cet usage a été approuvé dans le cadre du procès d'un jeune britannique ayant commis un matricide, dont le crime a été décrit comme un acte impulsif, violent et sans motif apparent, et qui s'est suivi d'une amnésie pendant sept heures (Hill et Sargant 1943). L'expertise psychiatrique a montré que le prévenu ne présentait pourtant aucun antécédent de maladie mentale ou d'épilepsie, indiquant seulement qu'il était en

[59] Les termes de « neurodroit » (« neurolaw ») et de « neurocriminologie » (« neurocriminology ») n'apparaissent pas avant les années 1990, respectivement en 1991 (Taylor, Harp, and Elliott 1991) et en 1996 (Schmalleger 1996). L'usage de ces termes peut sembler anachronique dans un contexte où les neurosciences ne constituaient pas un champ disciplinaire en tant que tel. Nous les emploierons néanmoins par commodité tout au long de ce chapitre, pour rendre compte des études utilisant l'EEG pour informer la question de la responsabilité pénale (dans le cas du neurodroit) et la question de la dangerosité de l'individu (dans le cas de la neurocriminologie).

[60] Nous aurons l'occasion d'examiner dans le détail les recherches menées par Denis Hill car elles ont largement participé à la construction du projet neurocriminologique au début des années 1940. Bien que l'objet de notre analyse ne se réduise pas à ses incarnations historiques (notamment aux affiliations institutionnelles ou idéologiques des neurocriminologues), il est quand même important de préciser que le psychiatre William Sargant est devenu une figure très controversée dans la communauté scientifique. Il a notamment mené une série de recherches sur la psychochirurgie et les techniques dites du coma insulinique ou de la cure du sommeil, en Grande-Bretagne et aux Etats-Unis (Snaith 1994).

état d'hypoglycémie pendant les douze heures qui ont précédé le crime. À la demande de la défense, et dans l'objectif de reproduire les conditions dans lesquelles le crime avait été commis, une série d'enregistrements EEG a été réalisé après avoir induit chez le prévenu un état d'hypoglycémie. Les résultats de l'expertise menée par Denis Hill et William Sargant ont été publiés tels quels dans la revue *The Lancet* : « du point de vue juridique, l'accusé savait ce qu'il faisait au moment où il a tué sa mère et savait que ce qu'il faisait était immoral, mais nous avons estimé que son cerveau fonctionnait anormalement à ce moment-là, l'anomalie étant telle qu'elle altérait son jugement et le rendait incapable d'apprécier pleinement la nature de l'acte. » (Hill and Sargant 1943, p. 527). Les données EEG révélaient la présence d'une « décharge électrique anormale dans le cortex », ce qui a joué un rôle décisif sur le verdict rendu par le jury, puisque le prévenu a été reconnu « *guilty but insane* ».[61]

On peut s'interroger à la fois sur l'influence exercée par ces données sur l'évaluation de la responsabilité du prévenu et sur l'interprétation pathologique qui leur est donnée, dans un contexte où l'expertise psychiatrique n'a pas permis d'objectiver l'existence d'un trouble mental. Cette affaire pénale montre bien comment le pouvoir persuasif des données neuroscientifiques peut favoriser un diagnostic pathologique et une déresponsabilisation du prévenu, malgré l'absence de trouble mental caractérisé. Il est d'ailleurs probable que l'usage judiciaire de l'EEG ait été motivé par une volonté de pallier les insuffisances de l'expertise psychiatrique, de par la nature du crime commis par le jeune britannique : un acte monstrueux et sans raison, mais qui ne pouvait être imputé à la présence d'un trouble mental. De toute évidence, la stratégie de la défense a donc consisté à combler l'absence de raison par *l'explication neurophysiologique*. Elle s'est articulée autour de l'idée que les données EEG apportaient la preuve tangible d'un état de conscience altéré, et permettaient ainsi de justifier une atténuation de la responsabilité du prévenu.

Cette affaire pénale marque le tout premier cas de neurodroit, c'est-à-dire la première affaire ayant eu recours à des données neuroscientifiques pour informer la question de la responsabilité pénale. Pourtant, il est important de noter qu'elle n'a pas donné une impulsion immédiate à une généralisation de l'usage de l'EEG dans le système

[61] Ce verdict s'inscrit dans le cadre de ce qui est communément appelé les *M'Naghten Rules*. Relativement rare, il est considéré comme une alternative au verdict d'irresponsabilité pour cause psychiatrique (*Not Guilty by Reason of Insanity*) (voir à ce sujet Bonnie 1997).

pénal. Il faudra attendre le début des années 1960 pour voir apparaître les prémices du projet de neurodroit tel qu'il se développe actuellement aux Etats-Unis. Pour comprendre les raisons de cet immobilisme, mais aussi pour situer le contexte d'émergence de la première tentative visant à intégrer l'EEG dans le système pénal, il faut remonter à la genèse des premières études menées sur le cerveau des criminels. Il semble en effet que l'ambition d'utiliser les données EEG pour justifier une atténuation de la responsabilité pénale ait été favorisée par une série de publications démontrant un lien entre la présence d'anomalies cérébrales et la manifestation d'un comportement violent, impulsif ou agressif. Ces études vont progressivement introduire la possibilité de reconnaître les indicateurs neurophysiologiques d'un comportement potentiellement dangereux, et annoncer les prémices du développement de la neurocriminologie, aux Etats-Unis et en Grande-Bretagne.

Dans ce chapitre, nous proposons de reconstruire l'histoire de ce projet, en remontant jusqu'aux premières études EEG qui mettent en évidence un lien entre la perturbation de l'activité électrique du cerveau et les troubles du comportement ou de la personnalité. L'objectif principal est de retracer les conditions d'émergence du discours neurobiologique sur la criminalité, et d'isoler les principales stratégies de recherche qui seront mobilisées tout au long du développement de la neurocriminologie basée sur l'EEG, globlament jusqu'à la fin des années 1980. Nous analyserons dans un premier temps les premières publications relatives à la psychopathie et aux troubles du comportement pendant l'enfance, qui ont émergé au début des années 1940 dans le milieu anglo-saxon. Comme nous le montrerons dans les deux premières sections, ces études ont posé les bases d'une réflexion sur la possibilité d'utiliser l'EEG comme un outil de connaissance et de prédiction du comportement criminel. Elles ont contribué à définir l'agenda de la recherche en neurocriminologie, en structurant le projet autour de deux grands principes théoriques et méthodologiques.

Cette analyse nous amènera à identifier les principales ambitions du projet neurocriminologique, et les différentes formes que recouvre son objet de recherche, tout au long de la seconde moitié du XX^{e} siècle (section 2.3). Nous prendrons le temps d'examiner quelques-unes des études les plus emblématiques du champ, afin de préciser au mieux les conditions expérimentales dans lesquelles les recherches sont conduites (section 2.4). Enfin, dans les deux dernières sections, nous montrerons dans quelle mesure ces études ont favorisé l'apparition d'un discours médical sur l'impulsivité et

l'agressivité : tout au long du développement de la neurocriminologie, les données EEG sont mises au service d'une stratégie de pathologisation des crimes violents, encourageant par la même occasion une reconfiguration des problématiques liées à l'évaluation de la responsabilité ou de la dangerosité des individus.

Précisons que notre axe d'analyse n'est pas tant celui d'une critique méthodologique des études EEG publiées dans le champ de la neurocriminologie, que la reconstruction historique d'un projet qui perdurera tout au long de la seconde moitié du XX^{e} siècle. Ces études n'ont fait l'objet que d'une poignée de revues de la littérature, dont la plupart se gardent d'entrer dans le détail des paradigmes expérimentaux. La neurocriminologie telle qu'elle se développe aujourd'hui ne fait d'ailleurs aucune mention des publications des années 1940-1950. Le projet du neurodroit révèle lui-même une totale méconnaissance des premières tentatives visant à intégrer les outils ou les données neuroscientifiques dans le système pénal.[62] Nous allons donc chercher à nous engager sur le terrain empirique en analysant le cadre théorique et méthodologique de ces études, et en questionnant les arguments qui justifient l'ambition d'utiliser l'EEG pour évaluer la responsabilité pénale ou la dangerosité d'un individu.

[62] Dans une correspondance privée avec le juriste américain Owen Jones (l'actuel directeur du *Research Network on Law and Neuroscience* et acteur principal du domaine du neurodroit), nous avons pu constater qu'aucune des études EEG menées dans les années 1940-1980 n'était connue des membres du réseau de recherche en neurodroit. Ce constat est d'autant plus frappant que ces derniers ont publié des articles visant à retracer l'histoire du projet (Jones and Shen 2012 ; Shen 2016). Si le juriste Francis X. Shen admet que des réflexions sur l'usage judiciaire de l'EEG ont émergé au milieu du siècle dernier, elles sont considérées comme essentiellement spéculatives, et aucune mention n'est accordée aux publications que nous allons décrire dans ce chapitre.

2.1. Des troubles du comportement pendant l'enfance à la psychopathie

2.1.1. La délinquance juvénile et le soupçon d'une origine pathologique

Les investigations menées sur le cerveau des criminels s'inscrivent dans la continuité d'une longue série d'études EEG démontrant un lien entre les perturbations de l'activité cérébrale et les troubles du comportement et de la personnalité. Ces études sont menées dans des institutions psychiatriques et des centres de détention pour mineurs, et sont publiées dans des revues médicales de référence en psychiatrie et en neurologie, principalement dans le *American Journal of Psychiatry* ou le *Journal of Neurology and Psychiatry*. Dès 1938, une étude américaine montre que des troubles du comportement pendant l'enfance sont associés à un « fonctionnement cérébral anormal » qui évoque dans certains cas une « pathologie cérébrale fondamentale » (Jasper et al. 1938, p. 657). En comparant les enregistrements EEG au repos de 71 enfants et adolescents présentant des troubles comportementaux (dont certains ayant commis des actes de délinquance) avec un groupe-contrôle, les chercheurs ont observé une forte proportion d'anomalies se traduisant par une activité excessive d'ondes lentes (2-6 Hz), en particulier dans les régions frontales. En outre, certains enfants présentaient « une activité électrique épileptiforme » − c'est-à-dire un schéma d'activité cérébrale que l'on rencontre habituellement chez les patients épileptiques entre deux crises − malgré l'absence de signe clinique associé à l'épilepsie.

Cette étude est la première à corréler l'existence d'un trouble comportemental à une perturbation de l'activité électrique du cerveau. Ces résultats ont été rapidement confirmés chez des enfants ou adolescents présentant un large éventail de troubles, allant du comportement impulsif ou hyperactif à la délinquance (Lindsley and Bradley 1939 ; Lindsley and Cutts 1940 ; Secunda and Finley 1942 ; Brill et al. 1942 ; Miller and Lennox 1948 ; Kennard 1949 ; Forssman and Frey 1953).[63] L'ensemble de ces études décrit le

[63] Dans la mesure où notre objectif est ici de retracer le contexte d'émergence de la neurocriminologie, nous avons choisi de limiter ce corpus au début des années 1950. Précisons néanmoins que les études EEG sur la délinquance juvénile ou les troubles du comportement pendant l'enfance ont été publiées tout au long de la deuxième moitié du XX^{e} siècle (voir à ce sujet Hughes et al. 1965 ; Beshai 1971 ; Deckel, Hesselbrock,

même type d'anomalies, à savoir une « dysrythmie corticale » se traduisant par un excès d'ondes lentes, le plus souvent dans les régions frontales ou temporales. Il est par ailleurs important de préciser qu'à l'exception de l'étude de Brill et ses collaborateurs (1942), les résultats sont tous issus d'une comparaison avec un groupe contrôle (enfants ou adolescents ne présentant aucun trouble comportemental). La fréquence des anomalies observées est néanmoins très variable, puisque les études rapportent entre 35% (Miller and Lennox 1948) et 80% d'anomalies EEG (Lindsley and Cutts 1940), soit une proportion d'anomalies trois à huit fois supérieure à celle du groupe contrôle.

Il est intéressant de constater que la plupart de ces études évoquent l'existence d'une cause pathologique pour expliquer l'origine de ces anomalies. D'après Miller et Lennox (1948), les anomalies EEG des enfants présentant des troubles du comportement pourraient résulter d'une « pathologie cérébrale organique » en raison d'un lien rapporté dans la littérature entre ces anomalies et la présence de lésions cérébrales (Miller et Lennox 1948, p. 759). Globalement, les études assument la volonté d'inscrire ces troubles dans un modèle médical, puisqu'elles évoquent volontiers la possibilité d'utiliser l'EEG pour identifier les cibles d'éventuelles stratégies thérapeutiques. Comme le précisent Brill et Walker (1945), « les études EEG ont souligné l'importance d'une approche organique des troubles comportementaux, suggérant des approches thérapeutiques qui ne concernent pas entièrement les facteurs environnementaux » (Brill et Walker 1945, p. 548 ; voir aussi Walker et Kirkpatrick 1947). L'ambition d'utiliser l'EEG comme un outil de diagnostic et de recherche de traitement était d'ailleurs annoncée en conclusion de la première étude EEG menée sur les troubles du comportement pendant l'enfance :

> « *[A]bnormal brain function as revealed by the electroencephalogram is an important component in the aetiological picture of the majority of a group of problem children whose disorder had been considered as primarily psychogenic previous to using this method of diagnosis. The nature of the fundamental pathology of the brain indicated is not as yet known, but the findings have been found important in evaluating prognosis and treatment.* » (Jasper et al. 1938, p. 657).

and Bauer 1996 ; Niv et al. 2015). Pour une revue de littérature des études portant sur la délinquance juvénile, voir notamment Moffitt (1990).

Précisons que les études EEG menées sur des jeunes délinquants ont également recours à cette terminologie médicale. Un exemple en est donné par Forssman et Frey (1953), qui ont effectué des mesures EEG chez une centaine de garçons présentant des troubles du comportement et placés dans un centre de détention pour mineurs. Dans la mesure où l'examen EEG rapportait un degré d'anomalies nettement plus élevé chez les adolescents présentant « une personnalité agressive », les auteurs ont postulé l'existence d'une « origine physiopathologique » de l'agressivité :

> « *A more interesting observation is the distinct preponderance of electroencephalographic abnormality among the aggressive boys in the behavior problem group. This would indicate that aggressiveness is of pathophysiological origin and not a product of environment in the ordinary sense of the term.* » (Forssman and Frey 1953, p. 71).

2.1.2. L'hypothèse d'une « immaturité corticale » chez le psychopathe

Ces travaux ont été aussitôt repris dans une série de publications portant sur l'examen EEG de sujets souffrant d'un trouble de la personnalité psychopathique. Malgré une définition clinique encore incertaine, la psychopathie a fait l'objet de nombreuses recherches empiriques utilisant l'EEG, l'analyse des publications démontrant un intérêt croissant vis-à-vis de son étiologie ou des possibilités de traitement, à partir des années 1940 (Schirmann 2014 ; voir aussi McCord and McCord 1956).[64] L'une des finalités recherchées est de recourir aux données neurophysiologiques afin de préciser les critères de diagnostic et de classification définis notamment par David Henderson (1939) en Grande-Bretagne, ou Hervey Cleckley (1941) aux Etats-Unis. Mais si la psychopathie a retenu l'attention des physiologistes dans les années 1940-1950, c'est en raison de son association étroite avec le comportement criminel. La plupart des critères de diagnostic évoquent l'absence de sens moral ou de remords, l'impulsivité, l'agressivité ou encore l'égocentrisme (Curran and Mallinson 1944 ; Preu 1944 ; East 1945). Certains

[64] Pour une revue de littérature sur les études EEG menées sur les psychopathes, voir notamment Syndulko et al. (1978).

psychiatres iront jusqu'à considérer la délinquance comme un critère de diagnostic de la psychopathie (e.g. McCord and McCord 1956).

Cette notion équivoque va ainsi entretenir une grande confusion dans le domaine naissant de la neurocriminologie, et orienter une partie des recherches vers l'examen EEG de psychopathes admis dans des institutions psychiatriques ou carcérales, aux Etats-Unis et en Grande-Bretagne. En s'appuyant sur divers critères de diagnostic et de classification – généralement ceux définis par Henderson (1939), Cleckley (1941), ou Maughs (1941) – de nombreux psychiatres vont effectuer des enregistrements EEG sur des sujets présentant une personnalité psychopathique, dans l'objectif de mettre en évidence les marqueurs neurophysiologiques du comportement impulsif, agressif ou antisocial (Hill and Watterson 1942 ; Silverman 1943, 1944 ; Knott and Gottlieb 1943 ; Stafford-Clark et al. 1951 ; Ehrlich & Keogh 1956).[65]

L'intérêt porté à la psychopathie s'est rapidement concentré autour d'une des catégories proposées par Henderson (1939), à savoir les sujets « à prédominance agressive ».[66] La première étude ayant effectué des enregistrements EEG sur des psychopathes a été menée sur 151 patients (dont 66 « à prédominance agressive ») répartis dans plusieurs institutions psychiatriques londoniennes, et dont certains avaient été condamnés pour des actes de délinquance (Hill and Watterson 1942). En comparaison avec le groupe-contrôle (individus sains), les enregistrements EEG des psychopathes montrent une forte proportion d'anomalies se traduisant par une importante « dysrythmie corticale », en particulier chez les psychopathes « à prédominance agressive » (voir Annexe n°2).

Cette étude revêt une grande importance car non seulement elle a posé les bases de la recherche en neurocriminologie, mais elle a surtout orienté une grande partie des recherches vers l'étude neurophysiologique des sujets présentant une « personnalité

[65] Précisons que contrairement aux études EEG menées sur les troubles du comportement pendant l'enfance, celles-ci ne s'appuient généralement pas sur une comparaison avec un groupe contrôle. Les données concernant la fréquence des anomalies sont donc très difficiles à interpréter, d'autant plus qu'elles sont très variables selon les études : de 48% pour Hill et Watterson (1942) à 80% pour Silverman (1943).

[66] Pour le psychiatre britannique David Henderson (1939), les psychopathes « à prédominance agressive » (« *predominantly aggressive* ») sont caractérisés par l'absence de sentiment de culpabilité. Il définit également d'autres sous-catégories cliniques comme le psychopathe « à prédominance inadéquate » (« *predominantly inadequate* ») ou le psychopathe « à prédominance créative » (« *predominantly creative* »).

agressive ». Les conclusions de cette étude vont en effet annoncer l'un des résultats les plus significatifs du domaine de la neurocriminologie. Elles rapportent d'une part, une proportion d'anomalies EEG trois fois plus importante chez les psychopathes que dans le groupe-contrôle (49% contre 15%) et d'autre part, une corrélation directe entre le degré d'anomalies EEG et le niveau d'agressivité des sujets. Denis Hill et Donald Watterson résument ces résultats ainsi : « [p]lus le patient est agressif, plus l'EEG risque d'être anormal » (p. 64). Par ailleurs, il est important de noter que Hill et Watterson n'observent pas de différence significative dans le degré d'anomalies EEG chez les psychopathes ayant été poursuivis pour des actes de délinquance : « la délinquance seule ne semble pas être associée à un EEG anormal » (p. 61). Autrement dit, les résultats montrent que la forte proportion d'anomalies EEG observée chez les psychopathes n'est pas associée à la délinquance, mais uniquement au niveau d'agressivité des sujets.

Ces résultats vont être analysés à l'aune des études EEG relatives aux troubles du comportement pendant l'enfance. Les auteurs de l'étude relèvent en effet des similitudes dans la nature et la fréquence des anomalies observées, entre les psychopathes et les enfants présentant des troubles du comportement. Ce constat va amener les psychiatres à formuler l'hypothèse d'une « immaturité corticale » ou d'un « défaut de développement du système nerveux central » pour expliquer l'origine des anomalies EEG observées chez les psychopathes (Hill and Watterson 1942, p. 64 ; Hill 1944, p. 327).[67] Il est important de préciser qu'il s'agit d'une hypothèse purement théorique qui repose exclusivement sur des données neurophysiologiques, et non pas sur une approche neuro-développementale. C'est en observant des analogies entre les données EEG de jeunes enfants et de psychopathes « à prédominance agressive » que Denis Hill et Donald Watterson (1942) vont suggérer l'idée qu'un défaut de maturation du cortex pourrait être à l'origine de ces anomalies. Le même constat sera dressé dans des travaux ultérieurs, notamment par Stafford-Clark et al. (1951), à partir de l'analyse de données EEG recueillies chez 165 psychopathes ayant fait l'objet d'une condamnation criminelle. Les anomalies EEG seront interprétées comme l'expression d'une « immaturité physiologique » en raison du

[67] Hill et Watterson ont en effet très vite rejeté l'hypothèse selon laquelle ces anomalies EEG pourraient résulter de lésions cérébrales. Selon eux, « dans l'ensemble du groupe, et en particulier dans les cas agressifs, il y a très peu de preuves positives que le traumatisme cérébral soit responsable des modifications de l'E.E.G » (Hill and Watterson 1942, pp. 63-64).

fait qu'elles « se rencontrent habituellement chez les jeunes enfants » (Stafford-Clark et al. 1951, p. 125) :

> « *The correlations adduced between these variables suggest that a failure to mature is the fundamental abnormality in the subjects selected as psychopaths in our investigation.* » (Stafford-Clark et al. 1951, p. 126).

Une précision importante doit être apportée en ce qui concerne l'hypothèse de l' « immaturité corticale ». Outre les similitudes observées dans la nature et la fréquence des anomalies EEG, elle a été favorisée par le recoupement de ces données avec des critères d'appréciation relatifs aux traits psychologiques des enfants présentant des troubles du comportement. Une attention particulière a en effet été portée à l'une des observations rapportées par Jasper et ses collaborateurs (1938), à savoir « l'association entre l'humeur colérique et l'irritabilité des enfants, et la présence d'anomalies EEG » (Hill et Watterson 1942, p. 64). Dans la mesure où les psychopathes − en particulier ceux « à prédominance agressive » − présentaient le même type d'anomalie EEG, une association a été établie entre ce pattern d'activité cérébrale d'une part, et l'impulsivité ou l'irritabilité des individus d'autre part. En clair, telle qu'elle a été initialement formulée par Hill et Watterson (1942), l'hypothèse de l' « immaturité corticale » s'est articulée autour du constat que les psychopathes partagent des similitudes avec les enfants présentant des troubles du comportement, à la fois sur le plan neurophysiologique et psychologique. Comme le résument Hill et Watterson (1942) :

> « *In view of the similarity between the aggressive behaviour of psychopaths and the normal bad-temper response to frustration in young children on the one hand, and the similarity between the E.E.Gs. of aggressive psychopaths and those of young children on the other, the suggestion that the abnormality in the E.E.Gs. in these cases is produced by a failure of development in the central nervous system is very tempting.* » (Hill and Watterson 1942, p. 64).

Tout au long des années 1940-1950, l'hypothèse de l'« immaturité corticale » va s'imposer comme le modèle explicatif privilégié pour rendre compte des anomalies EEG observées chez les psychopathes. Cette hypothèse bénéficiera par ailleurs d'un autre argument, à savoir l'existence d'une corrélation négative entre l'âge et la fréquence de ces anomalies (Brill and Seidemann 1941 ; Secunda and Finley 1942 ; Silverman 1944 ; Gibbs et al. 1945 ; Hill and Pond 1952 ; Levy and Kennard 1953). Ces anomalies

disparaîtraient à mesure que l'âge augmente, ce qui amènera certains psychiatres à interpréter les données EEG des psychopathes comme « une apparente persistance de patterns infantiles (« *childhood patterns* ») à l'âge adulte » (Ehrlich and Keogh 1956, p. 291 ; voir aussi Hill 1955). En somme, le modèle interprétatif défendu par les psychiatres va favoriser l'idée selon laquelle les psychopathes sont « immatures », tant sur le plan comportemental que cérébral.

S'il n'est pas de notre propos d'expliciter davantage cette hypothèse, nous retiendrons néanmoins que tout au long des années 1940-1950, l'interprétation des données EEG collectées chez les psychopathes sera discutée en référence aux études menées sur la délinquance juvénile et les troubles du comportement pendant l'enfance. L'influence réciproque de ces deux voies de recherche va véritablement orienter les ambitions et stratégies de recherche privilégiées par la neurocriminologie. L'étude de Hill et Watterson (1942) sera d'ailleurs largement citée dans la littérature relative aux anomalies EEG des criminels. Bien qu'elle ne porte pas sur une population criminelle, elle a eu une grande influence sur la construction de ce programme de recherche, notamment en raison de l'association établie entre le degré d'anomalies EEG et le niveau d'agressivité des sujets.

2.1.3. Des « anomalies constitutionnelles » chez le criminel psychopathe

À la même période aux Etats-Unis, le psychiatre Daniel Silverman (1943) va mener une étude EEG sur 75 criminels psychopathes détenus dans un centre médical fédéral (le *Medical Center for Federal Prisoners* de Springfield). En s'appuyant sur les critères de diagnostic définis par Cleckley (1941), Maughs (1941) et Caldwell (1941), il va distinguer et classer différentes personnalités psychopathiques selon leur pattern d'activité cérébrale, mais aussi en fonction des données relatives à leur dossier judiciaire : les criminels seront classés en trois catégories (« hostilité », « hédonisme », et « autodestruction ») selon le type de crime commis (crime à caractère sexuel ou non, avec ou sans violence, etc.).

Les données ne montrent aucune relation entre le degré d'anomalies EEG et le type de personnalité psychopathique ou la sévérité des symptômes. En revanche, comme dans l'étude de Hill et Watterson (1942), Daniel Silverman note que la nature et la

fréquence des anomalies observées est très similaire aux données recueillies par Jasper et al. (1938) chez les enfants présentant des troubles du comportement. Il constate en effet le même type d'anomalie EEG (une activité continue d'ondes lentes dans les régions frontales), présente chez 80 % des sujets.[68] D'après lui, si les données EEG collectées chez les psychopathes et les enfants présentant des troubles comportementaux coïncident, c'est parce qu'elles traduisent la présence d'une « anomalie constitutionnelle » commune au développement d'un comportement impulsif, agressif ou antisocial. Il en déduit que la psychopathie est « une maladie mentale résultant d'un dysfonctionnement cérébral inné ou acquis très tôt pendant l'enfance » et suggère le développement d'un traitement agissant sur « le composant organique de la maladie » (Silverman 1943, pp. 30-31).[69]

L'étude de Silverman (1943) tient une place éminente dans l'histoire de la neurocriminologie car elle constitue la première analyse EEG menée sur des psychopathes ayant fait l'objet d'une condamnation criminelle.[70] Elle sera suivie par une étude d'envergure menée sur 411 criminels détenus dans le même centre médical fédéral, et regroupant un ensemble de données EEG, cliniques, sociologiques, des données relatives au dossier judiciaire des sujets, ou encore à leur distribution géographique (Silverman 1944).[71] Contrairement aux travaux précédents qui entretiennent un certain flou quant aux finalités recherchées, l'ambition de cette étude est annoncée dès les premières lignes de l'article :

> « *The tendency to view certain types of criminals as neuropsychiatric problems brings into focus the importance of a consideration of abnormalities of the central nervous system, as well as psychogenic factors. Electroencephalography seems to*

[68] Ce constat viendra renforcer l'hypothèse de « l'immaturité corticale » formulée par Hill et Watterson (1942). Nous reviendrons en détail sur ce point lorsque nous expliciterons les hypothèses avancées pour expliquer l'origine des anomalies observées, dans la première section du chapitre 5.

[69] Plus précisément, Daniel Silverman soutient que « presque tous les psychopathes de cette étude sont nés avec ou ont acquis dès les premiers stades de développement une défectuosité cérébrale, une anomalie physico-chimique (pour autant qu'elle puisse être déterminée) qui affecte principalement les zones frontales, ou les zones "silencieuses" du cerveau » (Silverman 1943, p. 30).

[70] Précisons que cette étude, contrairement à celle qui lui succèdera (Silverman 1944), ne s'appuie pas sur une comparaison avec un groupe contrôle.

[71] Bien que cette étude vise à identifier les anomalies EEG d'un large échantillon de criminels, on peut noter qu'elle souffre d'un biais de sélection lié à une forte proportion de criminels présentant un trouble psychotique (19%) ou psychopathique (51%).

be an effective method of detecting many such abnormalities. » (Silverman 1944, p. 38).

Ici encore, les criminels sont classés en plusieurs catégories selon les données cliniques et le type de crime commis, et comparés à un groupe contrôle constitué par 60 volontaires de l'armée américaine (voir Annexe n°3).[72] Les données cliniques permettent de classer les détenus en quatre catégories : 208 criminels psychopathes, 80 criminels psychotiques, 63 criminels présentant des maladies organiques du système nerveux central, et 60 criminels ne présentant aucun trouble mental ou organique.

Les résultats indiquent un degré d'anomalies pour chaque catégorie de criminels, mais aussi pour chaque sous-catégorie de criminels psychopathes (« hostiles », « hédonistes », « inadéquats » et « homosexuels »). Daniel Silverman constate que les criminels présentant la proportion la plus importante d'anomalies sont les psychopathes « hostiles » et « hédonistes », qui présentent des « EEG pathologiques » dans 80% des cas (Silverman 1944, p. 40). En revanche, il n'observe pas de différence significative entre la fréquence d'anomalies EEG chez les criminels dits « normaux » (c'est-à-dire ne présentant aucun trouble mental ou organique) et le groupe contrôle (sujets non-criminels) : la proportion d'anomalies cérébrales est finalement similaire à celle de la population générale.

Les conclusions de cette étude modèrent alors quelque peu l'ambition annoncée dès les premières lignes de l'article : « [l]'électroencéphalographie est un outil potentiellement important pour l'étude et la prise en charge du psychopathe, mais pas du criminel normal » (Silverman 1944, p. 42). Silverman admet en effet que « l'électroencéphalogramme ne peut pas être considéré comme une aide significative dans le diagnostic de la source de la criminalité » mais qu'il reste un outil précieux pour diagnostiquer la psychopathie, en particulier pour « le type hostile et hédoniste » (Silverman 1944, p. 41).

Les études de Denis Hill et Donald Watterson (1942) en Grande-Bretagne et Daniel Siverman (1943, 1944) aux Etats-Unis ont contribué à fonder le champ de la neurocriminologie, en enregistrant l'activité corticale de plusieurs centaines de psychopathes et/ou criminels admis dans diverses institutions psychiatriques ou

[72] A la différence de l'étude précédente (Silverman 1943), les mesures EEG ont été comparées à un groupe contrôle, constitué par soixante volontaires de l'armée américaine.

carcérales. Les travaux pionniers de Hill et Watterson (1942) vont en particulier favoriser le développement de la neurocriminologie sur le sol britannique, en donnant lieu à une série de publications décrivant les anomalies EEG de criminels psychopathes tout au long des années 1950 (Stafford-Clark et al. 1951 ; Hill 1952 ; Gibbens et al. 1955, 1959). Globalement, les mesures EEG révèlent des anomalies jusqu'à « quatre fois plus fréquentes » chez les criminels psychopathes par rapport aux autres criminels (Gibbens et al. 1955, p. 131), et un degré d'anomalies significativement plus élevé chez les sujets les plus agressifs.

Parmi les études que nous venons de citer, on retiendra en particulier celle de Stafford-Clark et al. (1951), car elle répond à la volonté d'utiliser l'EEG pour caractériser le criminel psychopathe sur le plan neurophysiologique, et le distinguer du reste de la population carcérale. Cette étude propose de comparer un groupe d'une soixantaine de criminels non-psychopathes avec 165 psychopathes ayant fait l'objet d'une condamnation pénale, en regroupant des données EEG, des données psychiatriques, psychologiques, mais aussi des données relatives à l'environnement familial et social des individus. Bien qu'elle constitue l'une des rares études à tenir compte du contexte environnemental des sujets, elle défend néanmoins une relation directe entre le degré d'anomalies EEG et le niveau d'agressivité ou d'instabilité émotionnelle des sujets : les criminels présentant des signes d'agressivité extrême (« *severe aggressiveness* ») expriment le plus haut degré d'anomalies EEG (83%, contre 50% chez les criminels ne présentant aucun signe d'agressivité). D'après Stafford-Clark et al. (1951), la combinaison de l'approche psychologique, psychiatrique et électroencephalographique permettrait de caractériser et reconnaître le « psychopathe impulsif, agressif, explosif », afin de lui assurer une prise en charge adaptée.

À ce stade, plusieurs remarques s'imposent concernant les recherches que nous venons de présenter. Dans la mesure où ces résultats vont définir les conditions d'émergence d'un discours neurobiologique sur la criminalité, il paraît important de préciser ce qui pourrait caractériser le paradigme de ce programme de recherche. Les résultats de ces études vont en effet structurer la neurocriminologie naissante autour d'un ensemble de principes qui, loin de se limiter à l'étude de la psychopathie ou des troubles du comportement pendant l'enfance, vont définir une véritable grille de lecture des données EEG recueillies sur les populations criminelles.

2.2. La structuration d'un programme autour de trois principes

2.2.1. Une connexion directe du cerveau au comportement

Le premier principe consiste à défendre l'existence d'une association directe entre les anomalies observées sur le plan neurophysiologique et comportemental : les études EEG menées sur la psychopathie et les troubles du comportement pendant l'enfance vont imposer l'idée que les comportements impulsifs, agressifs ou antisociaux sont *directement* liés à une perturbation de l'activité électrique du cerveau. La plupart des études EEG que nous avons mentionnées plus haut concluent à l'existence d'une « relation directe » (voire « causale ») entre les anomalies observées et la survenue des troubles du comportement ou de la personnalité. Elles attribuent généralement aux anomalies EEG leur entière « responsabilité » dans la manifestation des troubles du comportement, et il n'est pas rare qu'elles s'appuient sur une stratégie conceptuelle consistant à confondre l'anomalie cérébrale et comportementale sous la même dénomination. À titre d'exemple, d'après Brill et ses collaborateurs (1942), la forte proportion d'anomalies EEG associées à la délinquance juvénile indique « des désordres cérébraux sous-jacents aux désordres sévères du comportement » (p. 498). Pour Knott et Gottlieb (1943), les études menées sur les psychopathes et les enfants présentant des troubles du comportement suggèrent « qu'au moins 50% des deux groupes manifestent des désordres du comportement en conséquence de désordres du fonctionnement cortical » (p. 141).

Les tentatives visant à dégager des relations psychophysiologiques sur la seule base de données EEG vont contribuer à définir un cadre interprétatif dont la neurocriminologie bénéficiera tout au long de la deuxième moitié du XX^{e} siècle : il s'agira de décrire les perturbations de l'activité électrique cérébrale comme le marqueur neurophysiologique d'un comportement violent, impulsif ou agressif, et plus encore, comme l'expression causale d'un comportement aux propriétés équivalentes. Le déploiement de cette stratégie repose principalement sur l'association établie entre le degré d'anomalies EEG et le niveau d'agressivité des sujets, qui parcourt l'ensemble des études menées sur la psychopathie et les troubles du comportement pendant l'enfance.

Comme nous le préciserons plus loin, l'interprétation des anomalies sera ensuite gouvernée par une série de relations psychophysiologiques visant à conforter le postulat d'une relation directe entre l'anomalie cérébrale et comportementale. L'une d'entre elles consistera par exemple à associer la présence d'un « tracé EEG instable » avec la manifestation d'une « personnalité instable ».

2.2.2. Un modèle psychopathologique

Une autre remarque s'impose en ce qui concerne le modèle interprétatif adopté dans les études menées sur la délinquance juvénile et les troubles du comportement pendant l'enfance. La grande majorité de ces études évoquent l'existence d'une cause pathologique pour expliquer l'origine des anomalies EEG, alors même que le seuil de normalité n'est pas clairement défini. Le recours à une terminologie médicale est très fréquent dans la littérature, et la description d'un « EEG pathologique » s'accompagne généralement d'une réflexion sur les stratégies thérapeutiques à adopter (voir en particulier Brill et Walker 1945 ; Walker and Kirkpatrick 1947). Les données EEG renvoient donc à une lecture étiologique dont l'analyse repose exclusivement sur l'observation d'une forte proportion d'anomalies (comparable à celle que l'on observe chez des patients atteints de lésions cérébrales).

Comme nous l'avons vu, ce modèle interprétatif – qui associe l'excès d'anomalies EEG à un état pathologique – est également à l'œuvre dans les études portant sur les troubles de la personnalité psychopathique, en particulier en ce qui concerne le psychopathe « à prédominance agressive ». Pour Brill et Walker (1945), ces anomalies EEG confortent l'idée selon laquelle « *the psychopath may be a sick individual rather than a bad one* » (Brill and Walker 1945, p. 549). Le discours médical qui entoure l'interprétation des anomalies EEG va ainsi accompagner un mouvement général de pathologisation des troubles du comportement et de la personnalité : de l'observation d'une forte proportion d'anomalies EEG chez les individus les plus impulsifs ou agressifs, à l'hypothèse (empiriquement injustifiée) que l'impulsivité ou l'agressivité relève d'un état pathologique. Autrement dit, en l'absence de seuil pathologique, l'observation d'une forte proportion d'anomalies EEG autorise une double ambition : non seulement elle rend possible l'apparition d'un discours psychopathologique sur l'impulsivité et l'agressivité ;

mais surtout, elle contribue à définir l'EEG comme un moyen de reconnaître les signes pathologiques associés à un comportement potentiellement dangereux.

2.2.3. Un principe d'identité psychophysiologique

Une troisième remarque peut être émise au sujet de l'association établie entre la psychopathie et les troubles du comportement pendant l'enfance d'une part, et la nature des anomalies EEG observées d'autre part. Elle concerne l'intérêt porté aux traits de personnalité mis en évidence par l'examen psychologique et clinique des individus. En combinant leurs données cliniques, psychologiques et neurophysiologiques, les études que nous avons mentionnées proposent de mettre en relation les anomalies EEG avec ce qui est communément décrit comme des « déviations de la personnalité » (Knott and Gottlieb 1943, p. 139) ou des « personnalités anormales » (Hill and Watterson 1942, p. 47). L'interprétation des données est d'ailleurs marquée par une mise en relation systématique entre le degré de sévérité du trouble, la structure de la personnalité des individus, et la nature des anomalies EEG observées. Autrement dit, la lecture du tracé EEG est corrélée à la fois aux caractéristiques comportementales et aux traits de personnalité des individus. Citons par exemple un passage de l'étude pionnière de Jasper et al. (1938), qui défend l'idée d'une association étroite entre ces trois niveaux d'analyse :

> « *The behavior problems of the type presented vary in severity and form not only according to the severity and locus of electroencephalographic defect but also according to the general characteristics of the individual personality.* » (Jasper et al. 1938, p. 655).

Cette approche fait écho aux tentatives engagées dans les années 1930-1940 pour utiliser l'EEG comme un outil de diagnostic de la personnalité des individus. Effectivement, il est important de noter que les études EEG menées sur la psychopathie ou les troubles du comportement pendant l'enfance s'inscrivent dans le prolongement d'une série de travaux portant sur la relation entre les données EEG et les traits de personnalité des individus (Davis et Davis 1936 ; Loomis et al. 1937 ; Gottlober 1938 ; Williams 1941 ; Gallagher et al. 1942 ; Saul et al. 1949 ; Levy and Kennard 1953 ; Sisson and Ellingson 1955). Ces travaux proposent d'établir diverses typologies de la

personnalité selon la distribution des fréquences EEG, en particulier celle du rythme alpha. Les physiologistes Hallowell et Pauline Davis (1936) ont par exemple suggéré qu'il était possible de « classer les individus en fonction de leur EEG », après avoir observé une forme d' « individualité » dans le pattern d'activité cérébrale enregistrée chez chacun des sujets (Davis et Davis 1936, p. 1220).[73]

Pour autant, comme le souligne Margaret Kennard (1953) dans une revue de la littérature consacrée aux études psychopathologiques utilisant l'EEG, la plupart des études ne montrent pas d'association entre un « type de personnalité » et un « type d'ondes cérébrales », mais seulement entre « une bonne personnalité et un bon EEG » (Kennard 1953, p. 104). Autrement dit, la « régularité du pattern EEG » apparaît comme un indicateur d'une personnalité stable, tandis qu'une « instabilité de la personnalité » est associée à une « irrégularité » ou à une « instabilité du pattern EEG » (Kennard 1953 ; p. 111). Margaret Kennard résume ainsi les résultats portant sur la relation entre les mesures EEG et les traits de personnalité des individus :

> « *It is thus the opinion of many of the authors […] that there is some relationship between general personality fitness and normal pattern. There should be, then, a relationship between unfitness and the abnormal EEC Furthermore, regularity of pattern has been correlated with a well-integrated personality; and absence of alpha, or irregularity of pattern, has been correlated with psychological or emotional variability.* » (Kennard 1953, p. 98).

C'est ici que réside le point fondamental de notre troisième remarque. Les études EEG menées sur la psychopathie et les troubles du comportement pendant l'enfance vont toutes confirmer l'association établie entre « l'instabilité de la personnalité » et « l'instabilité de l'EEG ». Il est d'ailleurs intéressant de noter que les études EEG portant spécifiquement sur les typologies de la personnalité s'inscrivent elles-mêmes dans leur lignée. Ces deux voies de recherche font preuve d'une importante perméabilité, tant sur le plan des objectifs – établir une taxonomie des relations psychophysiologiques en recoupant les données EEG et cliniques des individus – que des résultats annoncés. Cette association psychophysiologique constitue en effet le dénominateur commun de

[73] Plus précisément, cette étude suggère que « le rythme alpha peut révéler des caractéristiques innées de l'organisation corticale » car ce pattern d'activité EEG apparaît comme une caractéristique individuelle qui reste inchangée au cours du temps.

l'ensemble des études EEG menées sur les troubles du comportement et de la personnalité. Quelle que soit la nature du trouble, le degré d'instabilité comportementale ou émotionnelle des individus est systématiquement associé au degré d'instabilité du tracé EEG.

L'association établie entre « l'instabilité de la personnalité » et « l'instabilité de l'EEG » permet alors de préciser le cadre interprétatif commun aux études menées sur la psychopathie et les troubles du comportement pendant l'enfance : elle apparaît comme le principe psychophysiologique sous-jacent à l'interprétation de la « dysrythmie cérébrale » observée chez ces individus. C'est d'ailleurs au travers de ce principe de correspondance que les psychiatres sont amenés à postuler l'existence d'une « relation directe » entre les perturbations de l'activité cérébrale et les comportements violents, impulsifs ou antisociaux. Si les données EEG autorisent une connexion directe du cerveau au comportement, c'est parce que l'instabilité du tracé EEG est perçue comme l'expression causale d'un comportement aux propriétés équivalentes. Cette association psychophysiologique soutient finalement le bien-fondé de la stratégie consistant à confondre l'anomalie cérébrale et comportementale sous la même dénomination. Elle répond à l'ambition de rechercher dans l'activité électrique du cerveau le corollaire neurophysiologique des divers troubles du comportement et de la personnalité : l'instabilité du tracé EEG apparaît comme l'indicateur neurophysiologique d'un comportement tout aussi instable.

2.2.4. Une finalité prédictive

Dans cette perspective, il semble que l'influence réciproque des études menées sur les psychopathes d'une part, et sur les enfants présentant des troubles du comportement d'autre part, révèle une partie des finalités poursuivies par la neurocriminologie naissante. Il est en effet important de rappeler que dès le début des années 1940, l'intérêt des psychiatres s'est concentré autour des individus présentant une « personnalité agressive », aussi bien dans les études menées sur les psychopathes que sur les enfants présentant des troubles du comportement. Par ailleurs, l'articulation étroite entre ces deux voies de recherche repose sur le constat selon lequel la nature et la fréquence des anomalies EEG est très similaire dans les deux populations. De ce constat

va naître l'hypothèse d'une « anomalie constitutionnelle » commune au développement d'un comportement antisocial, hypothèse qui sera bientôt confondue dans la notion d' « immaturité corticale » (Hill et Watterson 1942). Cette hypothèse théorique, ainsi que l'association établie entre « l'instabilité de la personnalité » et « l'instabilité de l'EEG », vont alors justifier l'idée qu'il existe un lien entre les troubles du comportement pendant l'enfance et le développement de la psychopathie à l'âge adulte (Williams 1941 ; Hill and Watterson 1942 ; Silverman 1943).[74]

De là apparaît le raisonnement sous-jacent aux hypothèses explicatives concernant les anomalies EEG des psychopathes et des enfants présentant des troubles du comportement. Si les psychiatres prennent le soin de souligner les similitudes dans la nature et la fréquence des anomalies EEG, c'est parce que leurs recherches sont motivées par l'ambition d'identifier les marqueurs neurophysiologiques d'une criminalité « à risque ». La manière dont les données EEG sont interprétées annonce la construction d'un programme orienté vers une approche préventive, qui vise à repérer le futur criminel derrière le jeune délinquant. De ce point de vue, on comprend également mieux les ressorts du discours psychopathologique sur l'impulsivité et l'agressivité. La pathologisation des troubles du comportement ou de la personnalité s'institue comme le principe directeur de l'interprétation des données EEG, afin de justifier l'usage de l'EEG pour rechercher les stigmates pathologiques démontrant la dangerosité des individus.

Si l'on résume les différentes remarques que nous venons d'adresser aux premières publications relatives à la psychopathie et aux troubles du comportement pendant l'enfance, on voit apparaître la structuration d'un programme autour de deux grands principes : un principe d'identité psychophysiologique, qui permet de circonscrire l' « instabilité de la personnalité » sous une définition neurophysiologique (l' « instabilité de l'EEG »), et un principe psychopathologique, qui consiste à faire de l'impulsivité ou

[74] Cette association a été défendue notamment par Denis Williams (1941) dans une étude EEG menée sur neuf cents soldats des forces armées britanniques, et qui montre que les individus présentant une activité corticale « paroxystique » sont les plus instables sur le plan psychologique. En commentant les études EEG menées sur les troubles du comportement pendant l'enfance – et en cherchant à faire un lien avec ses propres résultats − il soutiendra que « [c]es enfants fourniront une partie des psychonévrotiques constitutionnels et des psychopathes de leur génération, et il est évident que l'incidence élevée des anomalies EEG observées chez les individus instables [...] est liée aux anomalies similaires observées chez les enfants instables et constitutionnellement anormaux. » (Williams 1941, p. 264).

de l'agressivité un état pathologique résultant d'un excès d'anomalies EEG. C'est en recherchant les marqueurs neurophysiologiques des comportements impulsifs ou agressifs que les études EEG vont contribuer à les placer sous le signe de la maladie. En somme, l'identification des marqueurs neurophysiologiques associés à la psychopathie ou aux troubles du comportement pendant l'enfance est conçue de manière à justifier l'usage de l'EEG pour reconnaître les indicateurs prédictifs d'un comportement potentiellement dangereux.

L'intérêt porté aux sujets présentant une personnalité psychopathique (qui va subsister tout au long du développement de la neurocriminologie) va d'ailleurs entretenir un rapport très ambigu entre le champ criminologique et médical. De toute évidence, la figure du psychopathe telle qu'elle est décrite dans les premières publications cristallise la représentation d'une menace sociale qu'il faudrait repérer et contrôler : elle permettra à la neurocriminologie naissante de fixer les bornes d'une dangerosité justifiant le recours à une stratégie préventive. Il n'est donc pas surprenant que le développement de la neurocriminologie soit historiquement associé aux recherches menées sur le cerveau de psychopathes admis dans des institutions psychiatriques ou carcérales. Mais surtout, l'intérêt porté à la psychopathie révèle un métissage médico-judiciaire qui n'aura de cesse de bousculer les limites entre les questions criminologiques et médicales. Il va en particulier favoriser une grande confusion en ce qui concerne le terrain d'investigation de la neurocriminologie : tout au long de son développement, son objet de recherche va se situer dans le domaine indifférencié du crime et des troubles de la personnalité. Car l'ambition de ce programme ne s'articule pas autour du criminel en tant que tel, mais bien autour du criminel *dangereux*. L'étude de Daniel Silverman (1944) – qui rappelons le, a initié la recherche en neurocriminologie aux Etats-Unis – laissait déjà présager cette inflexion. Le désintérêt exprimé à l'égard de l'étude neurophysiologique du « criminel normal » a dans un sens annoncé l'ambition d'une recherche qui s'étendra sur plusieurs décennies...

2.3. La construction d'un « cerveau dangereux »

2.3.1. « Le prisonnier normal a un EEG normal »

Les études EEG portant sur la psychopathie et les troubles du comportement pendant l'enfance ont progressivement introduit la question de l'usage de l'EEG comme un instrument de connaissance (voire de prédiction) du comportement criminel. Cette approche neurophysiologique de la criminalité a hérité du paradigme selon lequel les désordres du comportement sont *directement* liés à des désordres du fonctionnement cérébral. Plus précisément, cette approche a été rendue possible grâce à l'association établie entre la présence d'une forte proportion d'anomalies EEG et la manifestation d'un comportement impulsif, agressif ou antisocial. Les premières recherches relatives au champ de la neurocriminologie s'inscrivent d'ailleurs elles-mêmes dans la continuité des études EEG que nous avons citées précédemment, comme en témoignent les multiples références adressées à ces dernières :[75]

> « *The report by various workers of a correlation between electroencephalographic abnormality and behavior disorders raised the possibility that abnormal cortical function might be related to criminality.* » (Gibbs, Bagchi, and Bloomberg 1945, p. 294)

> « *The frequent finding of abnormal cortical rhythms in psychopaths and behavior disorders generally suggested that direct evidence of abnormal brain function in criminals might be obtained by electroencephalography.* » (Hill 1950, p. 350)

> « *The possibility of such relationship had been suggested by a number of previous studies, indicating that abnormal behavior may be related to abnormal EEG.* » (Levy and Kennard 1953, p. 832)

L'étude de Frederic Gibbs et ses collaborateurs (1945) fait suite aux recherches entreprises par Daniel Silverman (1943, 1944) aux Etats-Unis : elle constitue l'une des premières études publiées dans le domaine de la neurocriminologie.[76] Pour autant, il est

[75] Les études de Jasper et ses coll. (1938) ou Hill et Watterson (1942) font partie des références les plus souvent citées dans la littérature des années 1950.

[76] Précisons que Frederic Gibbs a largement contribué à l'étude et au diagnostic clinique de l'épilepsie, tout au long des années 1930-1940 (Cherici 2010).

important de préciser qu'elle tient une place à part dans l'histoire de la discipline, puisqu'elle a démontré l'absence de relation entre le comportement criminel et la présence d'anomalies EEG. Après avoir comparé les enregistrements EEG de plus de quatre cents criminels avec ceux d'un groupe-contrôle (sujets issus de la population générale) d'âge équivalent, les résultats n'ont rapporté aucune différence significative entre les deux groupes, ni même aucune relation entre les données EEG et le type de crime commis. D'après les auteurs de l'étude, l'interprétation des résultats nécessiterait un réajustement des données en fonction de l'âge des individus, car il existe une corrélation négative entre l'âge et la fréquence des anomalies EEG (e.g. Brill and Seidemann 1941).

Aussi selon eux, les études rapportant des résultats positifs reposent sur une mauvaise interprétation des données, car les anomalies observées pourraient tout aussi bien être attribuées à la présence d'un foyer épileptique ou d'une autre maladie neurologique, sans être directement liées au comportement criminel. Autrement dit, si les quelques études précédentes ont observé une forte proportion d'anomalies EEG dans les populations carcérales, c'est peut-être en raison d'un biais de sélection en faveur des individus présentant une atteinte cérébrale d'origine organique.

Des conclusions similaires avaient déjà été exposées à propos de la délinquance juvénile, notamment par Jenkins et Pacella (1943). D'après les auteurs de cette étude menée sur 37 adolescents placés dans un centre de détention pour mineurs, « l'électroencéphalographie ne contribue pas à l'explication de la plupart des cas de délinquance » (Jenkins et Pacella 1943, p. 119). Certains actes de délinquance (en particulier le vol) n'étaient pas associés à la présence d'anomalies EEG et, bien que les auteurs admettent l'existence d'un « facteur organique » à l'origine des comportements violents, ils excluent l'idée d'une relation directe entre la délinquance et la présence d'anomalies EEG. Ces anomalies reflètent selon eux la présence de lésions cérébrales, lesquelles favoriseraient indirectement la manifestation d'un comportement délinquant en provoquant une instabilité émotionnelle, une irritabilité ou une incapacité à contrôler son comportement.

Précisons d'emblée que les résultats négatifs font exception dans la littérature portant sur l'examen EEG des populations criminelles. Les quelques études démontrant l'absence de différence entre le groupe contrôle et les données EEG collectées chez des criminels ou des délinquants (à notre connaissance, Jenkins et Pacella 1943 ; Gibbs, Bagchi, and Bloomberg 1945 ; Olson et al. 1970 ; Driver, West, and Faulk 1974 ; Milstein

1982)[77] sont très rarement discutées à l'intérieur du champ de la neurocriminologie, ou bien largement discréditées par ses principaux acteurs. Ces derniers évoquent en particulier l'existence de biais de sélection ou des erreurs dans l'interprétation des données (e.g. Mednick et Volavka 1980).

Sans entrer dans le détail de cette discussion, il paraît important de revenir sur l'argument avancé dans les études rapportant des résultats négatifs : la nature et/ou la fréquence des anomalies EEG des criminels seraient finalement similaires à celle de la population générale. Les premières études EEG que nous avons citées précédemment (voir section 2.1) rapportaient déjà une grande hétérogénéité des données au sein même de la population criminelle. Pour rappel, l'étude pionnière de Silverman (1944) n'a démontré aucune différence significative entre les criminels dits « normaux » (ne présentant aucun trouble mental ou organique) et le groupe contrôle (sujets non-criminels). Hill et Watterson (1942) indiquaient également que le degré d'anomalies EEG était davantage associé à la « personnalité agressive » des individus qu'à un comportement délinquant. En clair comme le résume Denis Hill (1952), l'un des principaux acteurs de la recherche en neurocriminologie, « le prisonnier normal a un EEG normal » (Hill 1952, p. 438).

Il est intéressant de noter que si la figure du « criminel normal » apparaît fréquemment dans la littérature, les critères de définition de la normalité ne sont jamais explicitement définis. Le « prisonnier normal » apparaît en réalité comme une construction déterminée par son équivalent négatif : il représente la figure inversée du « prisonnier dangereux », c'est-à-dire du criminel souffrant de troubles de la personnalité, et visiblement incarné par le psychopathe « à prédominance agressive ». Autrement dit, si le « prisonnier normal » présente un « EEG normal », c'est parce qu'il existe une autre catégorie dont il est permis de spécifier sa valeur neurophysiologique en raison du danger qu'elle représente. À travers le constat selon lequel « le prisonnier normal a un EEG normal », la neurocriminologie naissante s'emploie ainsi à dessiner les contours de son objet de recherche. Elle révèle l'ambition d'un programme qui n'appelle pas tant à recenser les anomalies EEG observées dans une population homogène de criminels, mais plutôt à isoler les caractéristiques distinctives de l'individu dont le comportement constitue une menace pour la société.

[77] Il est possible qu'en raison d'un biais de publication, les études rapportant des résultats négatifs soient plus nombreuses.

2.3.2. Une approche prototypique de la criminalité

La caractérisation neurophysiologique du criminel dangereux va alors bénéficier d'une stratégie nouvelle par rapport aux études EEG menées sur les psychopathes admis dans des institutions psychiatriques. Alors que celles-ci proposaient essentiellement de classer différentes personnalités psychopathiques selon le degré d'anomalies EEG, la neurocriminologie va chercher à recouper les données EEG des individus avec les données relatives à leur dossier judiciaire, afin de faire correspondre le type d'infraction avec la fréquence des anomalies observées. Autrement dit, à partir du constat selon lequel les anomalies EEG ne sont pas associées à tous les types de comportements criminels, les neurocriminologues vont chercher à fonder des catégories juridiques sur des catégories neurophysiologiques, afin d'isoler les spécificités cérébrales des individus les plus dangereux.

Cette ambition va les conduire à assigner diverses proportions d'anomalies EEG à un ensemble de catégories mutuellement exclusives : « criminels récidivistes » vs. « criminels non-récidivistes » ; « crimes violents » vs. « crimes non-violents » ; « crimes motivés » ; « crimes sans motif », etc. Comme nous le verrons dans la section suivante, l'intérêt des neurocriminologues s'est très vite cristallisé autour de la figure du criminel récidiviste ou du criminel qui commet des crimes violents et sans motif apparent. Effectivement, comme le souligne Volavka (2002) dans une revue de la littérature consacrée aux corrélats neurobiologiques du comportement violent :

> « *Electroencephalographic abnormalities were particularly pronounced in motiveless, impulsive, violent crime. Furthermore, they were associated with recidivistic (rather than onetime) violence.* » (Volavka 2002, p. 119).

En somme, le constat selon lequel « le prisonnier normal a un EEG normal » va permettre à la neurocriminologie naissante de fixer ses principales stratégies de recherche, tout en précisant les contours de son objet d'étude. C'est ainsi que le statut de l'individu dangereux, jusqu'ici essentiellement incarné par le criminel psychopathe, va se doubler d'un ensemble de figures de la dangerosité définies d'une part, par la nature ou le nombre de crimes commis, et d'autre part, par le degré de leurs anomalies cérébrales. Le criminel dit « normal » occupera quant à lui une place secondaire dans la démarche comparative

adoptée par la neurocriminologie, puisqu'il sera systématiquement relégué au rang du groupe contrôle : il ne s'agira plus de dégager les différences neurophysiologiques entre les criminels et les sujets issus de la population générale, mais d'isoler les caractéristiques distinctives des individus représentant une grande menace pour la société.

Afin de mieux saisir la logique des stratégies privilégiées dans la caractérisation neurophysiologique du criminel dangereux, il convient de présenter quelques exemples paradigmatiques de la recherche en neurocriminologie basée sur l'EEG. Ces études que nous allons décrire figurent parmi les plus citées dans la littérature. L'intérêt que nous allons leur porter tient surtout au fait qu'elles ont concentré leur analyse autour de quatre figures emblématiques de la dangerosité criminelle, qui seront convoquées tout au long du développement de la neurocriminologie : le criminel qui commet des crimes violents et sans raison (Stafford-Clark and Taylor 1949), le criminel violent et récidiviste (Levy and Kennard 1953), le criminel présentant des « tendances agressives » (Williams 1969), et le criminel présentant des troubles mentaux et de la personnalité (Howard 1984). Ces études ont largement participé à la définition des principales ambitions et stratégies de recherche de la neurocriminologie. Elles figurent d'ailleurs parmi les rares à afficher clairement les finalités recherchées. Certaines ont véritablement posé la question de la nature pathologique de certaines formes de criminalité, contribuant par la même occasion à mener une réflexion inédite sur la possibilité d'utiliser l'EEG pour évaluer la responsabilité pénale d'un prévenu.

Dans la prochaine section, nous allons prendre le temps de décrire chacune de ces études, de manière à préciser au mieux les ambitions affichées par la neurocriminologie et les différentes formes que recouvre son objet de recherche. Nous réserverons nos commentaires pour la section suivante, afin d'isoler les grands principes théoriques et méthodologiques qui structurent la neurocriminologie basée sur l'EEG. Néanmoins, avant d'entrer dans le détail de ces travaux, quelques remarques préalables s'imposent concernant la stratégie globale adoptée par les neurocriminologues. Nous pouvons d'ores et déjà préciser que les résultats rapportent le même type d'anomalies EEG, à savoir un « excès d'ondes lentes » dans les régions frontales et temporales. Cette observation sera décisive dans la définition et la mise en place des stratégies de recherche, puisqu'elle amènera les neurocriminologues à classer les criminels selon des variations quantitatives

d'un même état neurophysiologique. En clair, les critères de distinction des criminels ne reposeront pas sur la nature mais sur le *degré* de leurs anomalies EEG.[78]

Par ailleurs, ces études combinent à la fois des données cliniques et neurophysiologiques, et il n'est pas rare que des tests de personnalité soient effectués en complément des mesures EEG. L'ensemble de ces données va ainsi favoriser la construction d'une démarche typologique et classificatoire consistant à qualifier les individus selon leur niveau de dangerosité et le degré de leurs anomalies EEG. Comme nous le montrerons plus loin, cette démarche vise essentiellement à construire une définition neurophysiologique de la dangerosité criminelle. Il s'agit d'une part, de dégager des relations psychophysiologiques sur la base du recoupement des données collectées au niveau neurophysiologique et clinique, et d'autre part, d'associer ces catégories psychophysiologiques à des catégories juridiques préalablement définies. De là va apparaître le véritable enjeu de la recherche en neurocriminologie, qui subsistera tout au long de son développement : fonder le critère de dangerosité sur le degré d'anomalies EEG, afin d'isoler les caractéristiques neurophysiologiques d'un cerveau « naturellement dangereux ».

2.4. Les figures de la dangerosité criminelle

2.4.1. Le criminel qui commet des crimes violents sans raison

L'étude de Stafford-Clark et Taylor (1949) est intéressante pour deux raisons. D'une part, elle représente la première tentative visant à explorer la relation entre le degré d'anomalies EEG et le caractère motivé ou non d'un crime. Ses conclusions ont sensiblement infléchi la recherche en neurocriminologie, puisque c'est dans ce contexte qu'est apparue pour la première fois la question de la nature éventuellement pathologique du comportement criminel. D'autre part, elle constitue la première étude EEG à discuter des implications judiciaires de ses résultats, notamment en ce qui concerne l'apport

[78] Les données EEG sont généralement classées en trois ou quatre catégories selon le degré des anomalies : « normal », « borderline », « anormal », « très anormal » (ou « paroxystique »).

potentiel d'un outil comme l'EEG pour combler les insuffisances de l'expertise psychiatrique. Elle a ainsi contribué à poser les prémices du développement du neurodroit, en s'engageant en faveur d'une continuité entre les réflexions d'ordre criminologique et judiciaire : de la question de la nature ou des causes du comportement criminel, à la question de la responsabilité pénale des individus présentant d'importantes anomalies cérébrales.

Cette étude repose sur l'examen EEG et clinique de 64 détenus condamnés pour homicide, et classés en cinq catégories selon le type de crime commis : 1/ homicide involontaire ou légitime-défense (11 individus) ; 2/ meurtre avec motif apparent (16) ; 3/ meurtre sans motif apparent (15) ; 4/ meurtre à caractère sexuel (8) ; 5/ meurtre commis par un sujet reconnu d'aliénation mentale (14).[79] Après avoir classé les données EEG selon le degré de leurs anomalies, Stafford-Clark et Taylor ont constaté que les crimes qui n'étaient pas motivés par un intérêt intelligible (e.g. aucune relation avec la victime) étaient associés à des anomalies nettement plus fréquentes que les crimes motivés : 73% des détenus ayant commis un crime sans motif présentaient des anomalies EEG, contre 25% lorsque le crime était motivé. L'importance accordée à ce résultat réside dans le fait que les auteurs d'un crime sans motif ne présentaient pourtant aucun signe clinique de maladie mentale : ils ont tous été reconnus pénalement responsables et la plupart ont été condamnés à la peine capitale.

Par ailleurs, les résultats ont rapporté une « prépondérance d'EEG normaux » dans les deux premiers groupes : les détenus condamnés pour homicide involontaire ou légitime-défense présentaient moins de 10% d'anomalies, c'est-à-dire dans une proportion similaire à celle de la population générale. D'après les auteurs, le caractère significatif de ces résultats ne tient pas tant à la nature des anomalies observées, qu'à la fréquence d'occurrence associée à chaque catégorie criminelle :

> « *[W]hatever the nature of the cerebral dysfunction which such abnormal records indicate, it is significantly correlated with a capacity to commit violent acts of an apparently motiveless character.* » (Stafford-Clark et Taylor 1949, p. 329).

À ce titre, une observation majeure a émergé de la comparaison des données correspondant à chaque catégorie isolée. Les psychiatres ont remarqué que les crimes

[79] Ces données ont été complétées par trente nouveaux cas étudiés par les psychiatres F.H. Taylor et J.C. Matheson, l'ensemble des données seront interprétées et résumées par Hill et Pond (1952).

violents et sans raison étaient associés à une proportion d'anomalies comparable à celle des criminels reconnus d'aliénation mentale, alors que les individus ne présentaient pourtant aucun signe clinique de maladie mentale. En clair, les 15 criminels reconnus pénalement responsables présentaient presque autant d'anomalies EEG (73%) que les 14 individus ayant bénéficié d'une irresponsabilité pour cause psychiatrique (86%) (voir Annexe n°4). Dans la mesure où ces criminels étaient, d'après les auteurs, « parfaitement normaux sur le plan clinique », ce constat a posé de manière implicite la question de l'origine pathologique de certaines formes de criminalité. En l'occurrence, des crimes violents et sans raison – bien que pénalement répréhensibles − seraient associés à d'importantes anomalies EEG, et dans une proportion semblable à celle que l'on rencontre chez des sujets atteints de troubles mentaux.

Stafford-Clark et Taylor vont alors engager une discussion sur les implications judiciaires de ces résultats, en insistant sur le fait que la plupart des criminels présentant un haut degré d'anomalies EEG ont été condamnés à la peine capitale. Les auteurs de l'étude accordent à leurs résultats une « valeur pratique » car ils y voient les bases d'une refondation du système pénal. Ils évoquent en particulier « la nature insatisfaisante du critère psychiatrique généralement admis dans les cours de justice pour déterminer la culpabilité d'un prévenu » (Stafford-Clark and Taylor 1949, p. 329) et défendent l'apport potentiel de l'EEG pour informer les conditions d'une irresponsabilité pour cause psychiatrique. Le passage ci-dessous illustre bien les enjeux d'une réflexion engagée en faveur d'une continuité entre les réflexions d'ordre criminologique et judiciaire. Il apparaît comme un tournant annonciateur du programme du neurodroit, tant il révèle l'ambition d'un programme qui prendra véritablement forme quelques décennies plus tard :

> « *Yet the possibility that in the E.E.G. we have at least an objective indication of innate propensities affecting responsibility − an indication moreover which is likely to be of maximal value in precisely those cases where clinical observation and the operation of reason are least able to elucidate the basis for the individual's actions − is one worthy of serious consideration. These cases may display no clinical evidence of insanity − either legal or medical − whatever, yet unless, as seems highly unlikely, the correlation between apparently motiveless violence and an abnormal E.E.G. is purely coincidental, they are the very subjects who most of all require further investigation and study over a prolonged period,*

as well as detention for the protection of the community. » (Stafford-Clark et Taylor 1949, p. 330)

On voit ainsi se profiler sous une forme inédite les prémices d'une redéfinition du projet neurocriminologique, qui conçoit l'EEG non plus seulement comme un outil de connaissance du comportement criminel, mais aussi et surtout comme un outil prescriptif : les données EEG recueillies chez certains criminels inviteraient à définir des catégories normatives utiles à l'expertise médico-légale. Ce passage évocateur cultive d'ailleurs l'ambivalence d'un projet qui se situe dans un entremêlement des questions relatives à l'évaluation de la responsabilité et de la dangerosité de l'individu. L'usage de l'EEG semble être envisagé non pas seulement pour appuyer (voire substituer) l'expertise psychiatrique, mais également pour justifier le recours à la détention préventive des sujets présentant d'importantes anomalies EEG.

Dans la mesure où cette étude ne précise pas davantage les conditions pour lesquelles l'EEG pourrait être utile à l'évaluation de la responsabilité ou de la dangerosité des sujets, il serait encore prématuré de questionner la légitimité scientifique des ambitions annoncées. Retenons seulement que si les crimes violents et sans raison ont retenu l'attention des psychiatres, c'est vraisemblablement en raison de leur caractère inintelligible et de leur rapport ambivalent avec la folie. La nature pathologique des anomalies qui leur sont associées (si elle était confirmée) favoriserait sans aucun doute un élargissement du champ d'intervention de la neurocriminologie. Non seulement elle tendrait à concevoir l'EEG comme un outil plus fiable que l'examen médico-légal pour rechercher la preuve d'une irresponsabilité pour cause psychiatrique, mais elle offrirait également à la neurocriminologie le moyen de déployer son action dans l'élaboration d'une réponse médicale adaptée.

2.4.2. Le criminel violent et récidiviste

L'étude que nous allons maintenant présenter s'intéresse à une autre figure de la dangerosité criminelle, qui concentrera une bonne partie des recherches en neurocriminologie : celle du criminel violent et récidiviste. Les neurologues Margaret Kennard et Sol Levy (1953) ont analysé les données EEG de cent détenus américains répartis en un « groupe violent » et un « groupe non-violent », en utilisant l'inventaire

multiphasique de la personnalité du Minnesota (*Minnesota Multiphasic Personality Inventory* ou M.M.P.I) comme test de personnalité.[80] De l'aveu des auteurs de cette étude, l'objectif est d'étudier la relation entre d'une part, « le type d'enregistrement EEG et la structure de la personnalité » des détenus, et d'autre part, entre « le type d'EEG et le nombre de crimes commis par un individu » (Levy and Kennard 1953, p. 832). Ils ont également étudié la relation entre les données EEG des détenus et l'âge de leur première condamnation, dans l'objectif de confronter leurs résultats aux conclusions des études relatives à la délinquance juvénile. En somme, les données EEG ont été mises en relation avec trois paramètres : la structure de la personnalité des individus, le nombre de condamnations pénales, et l'âge de la première condamnation.

Globalement, 30% des enregistrements EEG collectés chez les 100 détenus sont décrits comme « anormaux », c'est-à-dire dans une proportion trois à six fois supérieure à celle de la population générale (5-10%).[81] Par ailleurs, les neurologues constatent que la proportion d'anomalies est « également distribuée entre les crimes violents et non-violents », révélant ainsi l'absence de relation entre la fréquence d'anomalies EEG et le niveau de violence des individus (Levy and Kennard 1953, p. 838). Pourtant, en observant sur la base des études précédentes que « des actions violentes et non-préméditées peuvent être corrélées à une activité EEG paroxystique », les neurologues admettent qu'une analyse plus fine « aurait pu mettre en évidence une relation entre un comportement paroxystique et un EEG paroxystique » (p. 834). Le recoupement des données relatives à la fréquence des anomalies EEG avec l'âge de la première condamnation confirme quant à lui les conclusions des études EEG relatives à la délinquance juvénile. Les données montrent que les anomalies sont plus fréquentes chez les sujets ayant commis leur premier crime avant l'âge de seize ans (voir Annexe n°5).

Le résultat le plus significatif concerne la relation entre le degré d'anomalies EEG, la structure de la personnalité des individus, et le nombre de condamnations criminelles. Les données confirment globalement l'idée selon laquelle « une structure de personnalité relativement bonne est liée à un EEG normal » (Levy and Kennard, 1953, p. 838). Elles montrent que les détenus « les plus stables » et ayant commis un crime isolé présentent le

[80] Le M.M.P.I. est un test psychométrique développé par le psychologue Starke R. Hathaway et le psychiatre J.C. McKinley (McKinley and Hathaway 1944).

[81] Levy et Kennard s'appuient ici sur les données recensées dans l'ouvrage de référence de Frederic et Erna Gibbs, *Atlas of electroencephalography* (1950).

taux le plus bas d'anomalies (17%) ; inversement, la proportion d'anomalies EEG est deux fois plus importante chez les criminels récidivistes et présentant une personnalité instable (voir Annexe n°6). Les conclusions rapportent ainsi une corrélation positive entre d'une part, le degré d'anomalies EEG et d'autre part, le nombre de condamnations criminelles et l'instabilité psychologique des individus.

Contrairement à l'étude de Stafford-Clark et Taylor (1949), celle-ci ne pose pas directement la question de la nature pathologique de ces anomalies ou de leurs conséquences sur la notion de responsabilité pénale. En revanche, les implications judiciaires de ces résultats sont discutées sur le plan de la prise en charge et du traitement médical des individus potentiellement dangereux. D'après les auteurs de cette étude, ces résultats doivent engager une réflexion sur les modalités de prise en charge de certains criminels, car l'EEG garantit la possibilité d'identifier les individus nécessitant un traitement médico-psychologique. Ils proposent alors de distinguer deux catégories de criminels sur la base de leurs données EEG et de leur profil M.M.P.I., afin de leur permettre un traitement adapté à la fois à leur profil de personnalité et à leurs caractéristiques neurophysiologiques :

> « *If these data are confirmed by further investigation it may eventually be possible, with the assistance of such tests as above, to separate two types of transgressors of the law who have reached imprisonment for totally different, almost opposite, reasons and should from every point of view have different treatment.* » (Levy and Kennard 1953, p. 838).

2.4.3. Le criminel présentant des « tendances agressives »

La troisième étude que nous proposons de décrire a été publiée dans la revue *Brain* (Williams 1969). Elle vise à préciser l'association établie depuis les premières publications relatives aux troubles de la personnalité psychopathique, entre le degré d'anomalies EEG et la manifestation de comportements agressifs. Comme le souligne son auteur, en citant notamment l'étude pionnière de Denis Hill et Donald Watterson (1942), « [l]es groupes d'individus chez qui l'EEG est si souvent anormal sont les individus agressifs, les psychopathes agressifs ou les délinquants agressifs. » (Williams, 1969, p. 504).

Denis Williams a effectué des enregistrements EEG chez 333 détenus condamnés pour des crimes violents et admis dans plusieurs institutions carcérales britanniques. Dans l'objectif d'isoler les caractéristiques distinctives du « criminel agressif », il a comparé le degré et la fréquence des anomalies EEG entre 206 détenus présentant des antécédents d'agression (« *habitual aggressives* » ou « *persistently aggressives* ») et 127 détenus condamnés pour un acte violent isolé (« *normal aggressives* »). Après avoir exclu de sa sélection les sujets épileptiques ou présentant des lésions cérébrales, la proportion d'anomalies EEG atteint 57% dans le premier cas, et 12% dans le second.[82] Williams précise que ce dernier chiffre correspond à la proportion d'anomalies rencontrée dans la population générale. Le groupe des criminels ayant commis un acte isolé est généralement associé à un EEG « normal et stable », tandis que la grande majorité des criminels récidivistes présentent un EEG « anormal » ou « très anormal ». En outre, certains enregistrements EEG évoquent la survenue d'une « décharge épileptique », malgré l'absence de signe clinique associé à l'épilepsie.

Dans la mesure où Denis Williams a exclu de sa sélection les sujets présentant des troubles organiques, ces anomalies ont été interprétées comme étant « directement liées » à la manifestation de leur comportement agressif. Selon lui en effet, « [i]l est raisonnable de conclure que dans ces cas, l'anomalie de l'EEG est directement liée à la perturbation du comportement, et qu'elle n'est pas le résultat d'une pathologie cérébrale structurelle » (Williams 1969, p. 513). Il est d'ailleurs intéressant de noter que la terminologie employée pour décrire les anomalies EEG de certains de ces criminels tend à démontrer une équivalence en termes de propriétés entre la nature du crime et la nature de l'anomalie observée. En évoquant par exemple le cas d'un crime violent sans motif apparent, Williams fait apparaître un lien entre « la brutalité effroyable du crime » et la « qualité explosive » du rythme alpha de l'EEG (p. 508).[83]

[82] L'étude montre que cette différence est statistiquement significative. La plupart des anomalies sont localisées dans les régions frontales et temporales, avec une plus forte proportion d'anomalies frontales dans le groupe des criminels présentant des antécédents d'agression.

[83] Il est difficile de ne pas y voir une réactualisation de la stratégie conceptuelle mobilisée dans les premières publications des années 1940, qui consistait à confondre l'anomalie cérébrale et comportementale sous la même dénomination : l'interprétation des données reste marquée par le souci de dégager des relations psychophysiologiques sur la seule base des données EEG.

Dans un souci de préciser les caractéristiques distinctives des deux groupes isolés, Denis Williams va porter une attention particulière aux circonstances dans lesquelles chacun des crimes a été commis. Il remarque que les sujets présentant un EEG « normal et stable » ont généralement commis leur crime dans le cercle familial, le plus souvent intentionnellement mais sous l'influence d'un « stress extrême ». À l'inverse, les criminels présentant un EEG « anormal » ou « très anormal » n'ont souvent aucune relation avec leur victime, leur crime est généralement impulsif et sans motif apparent, et caractérisé dans certains cas par un « déchaînement de violence » et une « absence de remords ». Compte tenu de l'écart significatif entre la fréquence d'anomalies EEG des deux catégories isolées, Williams en déduit que la différence entre les deux groupes s'explique par l'influence causale de deux types de facteurs mutuellement exclusifs : des facteurs « endogènes » ou « inhérents au sujet » d'un côté, et des facteurs « exogènes » ou « accidentels » de l'autre.

> « *A broad distinction between the two groups is that in the first the persistent aggressive response to life situations is largely due to the kind of person, and in the second, largely to the kind of life situation. This view is supported by the fact that in just half the instances of murder as a solitary crime, the victim was the wife or consort. The case histories of many of these men showed how potent indeed was the environmental stress which provoked the act. We are therefore in the main contrasting aggression which is the result of endogenous or exogenous factors […] Another simple distinction is that in the first group the aggression is usually unexplained, in the second usually explicable. Whilst a small majority of the first group had abnormal EEGs, the incidence of similar abnormality in the second was no more than in the population at large.* » (Williams 1969, pp. 513-514).

En cherchant à expliquer les différences observées entre les données EEG des deux groupes de criminels, Denis Williams est donc amené à proposer deux modèles explicatifs du développement d'un comportement criminel. Les criminels agressifs et récidivistes et les criminels ayant commis un acte isolé sont selon lui « biologiquement différents » car les anomalies EEG observées chez les premiers sont « de nature

constitutionnelle ».[84] Elles ne sont pas le résultat d'une lésion cérébrale acquise puisqu'elles proviendraient d'une « déviation congénitale d'une norme stable dans sa nature, et déterminée très tôt dans la maturation du cerveau » (Williams 1969, p. 515). Ces anomalies favoriseraient ainsi la manifestation d'un comportement criminel, en rendant les sujets plus sensibles (vulnérables) aux circonstances défavorables :

> « *These findings indicate that a major factor in the etiology of pathological persistent aggression is disturbance of cerebral physiology. This must be a potent factor in rendering these subjects prone to this abnormal behaviour—a response to psychological or environmental factors.* » (Williams 1969, p. 519).

S'il n'est pas de notre propos ici de démontrer la faiblesse des arguments évoqués, il est néanmoins important de commenter brièvement ces conclusions. L'étude de Denis Williams occupe une place particulière dans l'histoire de la neurocriminologie, car elle répond explicitement à un modèle innéiste du comportement criminel. Les seules données EEG l'amènent à postuler l'existence d'une criminalité « endogène » dont la cause réside dans la physiologie cérébrale de l'individu, et s'opposant à une criminalité « exogène » qui trouverait son explication dans les dynamiques socio-environnementales. Autrement dit, Williams défend l'idée selon laquelle certains individus deviendraient criminels « par la force des circonstances », tandis que d'autres seraient condamnés à une criminalité « inhérente » en raison d'une forte proportion d'anomalies cérébrales.

Dans la mesure où cette étude constitue l'une des seules à défendre explicitement une conception innéiste de la criminalité, il convient de s'interroger sur sa réception à l'intérieur du champ de la neurocriminologie. Largement citée à la fois dans les études ultérieures et dans les quelques revues de littérature consacrée à la neurobiologie du comportement criminel, elle est souvent présentée comme la première étude ayant mis en évidence les anomalies EEG associées à la récidive violente (e.g. Volavka 2002). Les conclusions de Williams ont été admises sans réserve, suscitant même un regain d'intérêt à l'égard du criminel violent et récidiviste. Elles ont en effet inspiré une série d'études EEG portant spécifiquement sur des criminels violents admis dans des institutions psychiatriques (parfois dans des unités de haute sécurité) aux Etats-Unis, en Grande-

[84] Denis Williams (1969) évoque indifféremment les termes « *constitutional in nature* » (p. 512) et « *constitutional in origin* » (p. 515) pour décrire ces anomalies.

Bretagne, mais aussi en Allemagne (Howard 1984 ; Convit et al. 1991 ; Wong et al. 1994 ; Pillman et al. 1999 ; Lindberg et al. 2005 ; Keune et al. 2012, 2018 ; voir aussi Sayed et al. 1969).[85]

Ces études portent sur des échantillons très différents en termes de taille et de population ; il paraît donc difficile de les appréhender de manière unifiée. Précisons néanmoins qu'elles confirment toutes l'association établie entre le degré d'anomalies EEG d'une part, et le nombre d'infractions violentes ou le niveau de violence des individus d'autre part : les plus fortes proportions d'anomalies sont observées chez les individus les plus violents.[86] Ces études ont en particulier permis de préciser la nature des anomalies associées, puisque la plupart ont rapporté la présence d'« asymétries hémisphériques » dans les régions frontales ou temporales. D'après Convit et ses collaborateurs (1991), « le degré de violence est associé de manière significative à l'asymétrie hémisphérique de l'activité EEG dans les régions frontotemporales » (p. 360). Une autre étude a conclu sur la base de l'examen EEG de 222 individus que « l'altération fonctionnelle de l'hémisphère gauche pourrait augmenter la propension à la violence chez certains criminels » (Pillman et al. 1999, p. 454).[87]

Précisons que la plupart de ces études souffrent d'un important biais de sélection en faveur des criminels présentant des troubles mentaux. Les échantillons étudiés ne sont pas représentatifs de la population carcérale car ces études ont été menées dans des institutions psychiatriques, sur des criminels ayant bénéficié d'une irresponsabilité pour cause psychiatrique. Malgré ce biais de sélection majeur, il paraît important de ne pas les passer sous silence car elles représentent une part non-négligeable de la littérature en neurocriminologie. Elles marquent en effet un élargissement de son terrain d'investigation, lequel prend pour objet non plus seulement le criminel incarcéré, mais également le criminel admis dans des institutions psychiatriques.

---

[85] Il convient également de noter que ces résultats EEG ont donné naissance aux premières réflexions sur les possibilités de traitement de la récidive violente (e.g. Yeudall et al 1987).

[86] En complément des examens EEG, la plupart de ces études ont utilisé différents tests de personnalité ou des échelles d'évaluation de la violence.

[87] Précisons que contrairement aux études que nous avons décrites jusqu'ici, celles-ci n'ont pas été effectuées « au repos » mais en utilisant la technique dite des potentiels évoqués (EEG-ERP). Dans ce cas, l'activité électrique du cerveau est enregistrée en réponse à des stimuli aversifs (généralement visuels ou auditifs).

C'est ainsi que l'on voit apparaître une autre figure de la dangerosité, nettement plus confuse que les précédentes : celle du criminel présentant des troubles mentaux et de la personnalité. Cette figure de la dangerosité criminelle mêle un ensemble de catégories nosographiques très diverses − allant du trouble de la personnalité borderline et antisociale jusqu'à la maladie psychiatrique aliénante − mais surtout caractérisées par des comportements typiques : impulsivité, irritabilité, instabilité affective, anxiété, agressivité, etc. Il faut en effet bien comprendre qu'en investissant les institutions psychiatriques, la neurocriminologie ne cherche pas à classer les criminels selon des catégories nosographiques bien caractérisées. Son ambition n'est pas de recourir à l'EEG pour préciser des critères de diagnostic, mais bien de rechercher dans ces populations psychiatriques les marqueurs neurophysiologiques d'un ensemble de comportements estimés dangereux.

À cet égard, il convient de noter que si cette figure de la dangerosité criminelle apparaît relativement tardivement dans l'histoire de la neurocriminologie (au milieu des années 1980), elle témoigne d'une résurgence des choix méthodologiques adoptés dans les publications des années 1940. Alors que celles-ci situaient leur objet de recherche dans le domaine indifférencié du crime et du trouble de la personnalité psychopathique, cette autre figure de la dangerosité appartient à un champ encore plus diffus, où se côtoient à la fois les maladies mentales et les troubles de la personnalité. Elle recouvre une combinaison abstraite de comportements potentiellement dangereux – le plus souvent évocateurs d'un passage à l'acte imprévisible et violent – et pouvant faire l'objet d'une évaluation clinique quantitative. Il semble en effet que la neurocriminologie trouve dans ces populations la cible privilégiée d'une stratégie préventive, puisqu'en combinant les données EEG avec les outils actuariels, elle va chercher à isoler les facteurs neurophysiologiques et cliniques du risque de récidive.[88]

2.4.4. Le criminel souffrant de troubles mentaux ou de la personnalité

[88] Les outils actuariels utilisés dans le domaine de la neurocriminologie comprennent par exemple les instruments d'auto-évaluation de l'anxiété ou des affects négatifs comme la *Welsh Anxiety scale* (Welsh 1956).

Parmi les études menées sur des criminels admis dans des institutions psychiatriques et que nous avons citées plus haut, celle de R.C. Howard (1984) mérite une attention particulière pour deux raisons. Premièrement, et au même titre que les trois études que nous avons décrites précédemment, elle figure parmi les rares à afficher clairement les finalités recherchées. L'ambition n'est pas ici de définir des catégories neurophysiologiques utiles à l'évaluation de la responsabilité ou à la prise en charge d'un individu, mais plutôt d'utiliser l'EEG pour évaluer le risque de récidive d'un individu. Cette ambition est restée jusqu'ici silencieuse ; elle apparaît d'ailleurs en filigrane dans l'étude de Howard (1984). Deuxièmement, cette étude cristallise une grande partie des enjeux de la neurocriminologie, car son objet de recherche apparaît comme la somme de toutes les figures de la dangerosité mobilisées dans la littérature : elle s'intéresse au cas particulier du criminel récidiviste ayant été condamné pour des crimes violents et sans raison, et qui présente des troubles sévères de la personnalité.

L'étude de R.C. Howard (1984) s'inscrit dans la continuité directe de celle publiée par Denis Williams (1969), qui s'articulait autour de la figure du criminel présentant des « tendances agressives ». D'après lui, les deux groupes de criminels étudiés par Denis Williams (les « *habitual aggressives* » et les « *normal aggressives* ») correspondent à deux types de personnalité définis par le psychiatre R. Blackburn dans sa « typologie des criminels » (1971) : le « *paranoid-aggressive type* », caractérisé par une forte impulsivité et un risque important de récidive, et le « *overcontrolled-repressor type* », qui a davantage tendance à commettre un crime isolé et intrafamilial. Dans l'objectif de préciser les caractéristiques neurophysiologiques de ces deux groupes de criminels, Howard va effectuer des examens EEG sur 265 criminels admis à l'hôpital Broadmoor (un hôpital psychiatrique de haute sécurité situé au Royaume-Uni), et répartis dans diverses catégories selon leurs données cliniques. Cet échantillon comprend des sujets admis sous la catégorie juridique de « maladie mentale », de « trouble de la personnalité psychopathique », et des sujets présentant différents scores d'impulsivité et de sociabilité sur l'échelle du M.M.P.I.

Le principal objectif est d'étudier la relation entre les données EEG et trois variables liées aux traits de personnalité des criminels ou à l'acte criminel lui-même : le score d'impulsivité et de sociabilité sur l'échelle du M.M.P.I., le type de crime commis (crime contre la personne ou contre la propriété) et la nature de la relation avec la victime (connue ou inconnue). Comme le précise Howard, l'objectif est d'abord « d'associer des construits de haut-niveau dérivés de la théorie de la personnalité (e.g. impulsivité et

sociabilité) avec des construits dérivés de l'électrophysiologie comportementale (e.g. stimulation corticale) » (p. 569). En clair, il s'agit d'associer les données EEG avec un ensemble de catégories relatives aux traits de personnalité de l'individu, avant de les mettre en relation avec la nature du crime commis.

En accord avec les études précédentes, les résultats montrent une « association significative entre le type de personnalité et le pattern EEG » (p. 577). Howard identifie en effet deux types d'enregistrement EEG selon le score d'impulsivité et de sociabilité : un « EEG normal » chez les individus les moins impulsifs et les plus sociables (les « *approachers* »), et un « EEG atypique »[89] chez les plus impulsifs et les moins sociables (les « *withdrawers* »). Il en conclut que « différents aspects de la personnalité antisociale peuvent être reflétés dans différents patterns EEG » (p. 579). Mais surtout, les résultats montrent que les auteurs de « crimes violents contre des inconnus » présentent le plus haut degré d'anomalies EEG, à savoir « une activité dysrythmique avec des caractéristiques paroxystiques » : 70% des sujets présentant cette anomalie n'avaient aucune relation avec leur victime. Selon Howard, l'observation d'un « EEG paroxystique » chez ces criminels permet de conforter l'hypothèse selon laquelle « *the degree of directionally non-specific violence correlates with the degree of EEG abnormality* » :

> « *Our attempts to relate paroxysmal EEG features to personality, legal diagnosis and offence variables resulted in only one significant relationship, between the presence of a bilateral paroxysmal EEG abnormality and a tendency to attack strangers. The statistical significance of this result, despite a small sample size, indicates that it deserves serious consideration. Insofar as one might view a bilaterally paroxysmal EEG as the most abnormal pattern, in a rather gross quantitative sense, of all our atypical EEG features, this finding perhaps suggests that the degree of directionally non-specific violence correlates with the degree of EEG abnormality.* » (Howard 1984, p. 580).

Enfin, Howard remarque que la proportion d'anomalies rencontrées chez les criminels les plus impulsifs et les moins sociables (les « *withdrawers* ») fait écho aux

[89] Il est intéressant de noter que Howard s'interdit lui-même d'employer les termes « anormal » (« *abnormal* ») ou « anomalie » (« *abnormalities* »), leur préférant des notions qu'il juge plus neutres : « *atypical* » ou « *anomalies* ».

données recueillies par Denis Williams (1969) concernant les criminels présentant des antécédents d'agression (« *persistently aggressives* »). Dans les deux cas, les anomalies sont observées dans une proportion d'environ 60 %, et les individus sont décrits comme présentant un fort risque de récidive. Ces observations vont amener Howard à défendre l'intérêt potentiel de l'EEG pour « l'évaluation de la dangerosité » des individus, en complément des tests de personnalité comme le M.M.P.I. L'utilisation combinée de ces outils permettrait selon lui d'identifier les individus les plus à risque de récidive :

> « *It may be possible in the future to use the clinical EEG and MMPI personality assessments together to identify those patients most at risk with regard to the quality and quantity of violent re-offending.* » (Howard 1984, p. 580).

2.5. Le (criminel) normal et le (criminel) pathologique

2.5.1. La définition neurophysiologique de la dangerosité

Si les études que nous avons citées méritent une analyse approfondie des présupposés sous-jacents aux stratégies mises en œuvre, nous allons ici tenter de retracer la manière dont elles ont posé la question de la nature pathologique du criminel dangereux. Tout au long de son développement, la neurocriminologie propose d'isoler les caractéristiques neurophysiologiques de l'individu dont le comportement constitue une menace pour la société, en prenant pour objet différentes figures de la dangerosité criminelle. Comme nous l'avons vu depuis les premières publications relatives à la psychopathie et aux troubles du comportement pendant l'enfance, sa principale stratégie consiste à construire une taxonomie des relations psychophysiologiques : il s'agit de recouper les données EEG, cliniques, psychologiques, et les données relatives au dossier judiciaire des individus, afin de fonder la notion de dangerosité sur des critères neurophysiologiques.

Pourtant, si ces figures de la criminalité répondent à diverses catégories juridiques ou psychologiques (crimes violents, crimes sans raison, etc.), il convient de relever l'échec des tentatives visant à isoler sur le plan neurophysiologique leurs caractéristiques

distinctives. Si l'on s'en tient aux résultats rapportés par la neurocriminologie, la dangerosité criminelle répond à une définition neurophysiologique relativement abstraite, puisque son appréciation se mesure essentiellement par l'excès d'anomalies EEG : le criminel violent, agressif, récidiviste et qui commet des crimes sans raison est systématiquement associé au plus haut degré d'anomalies EEG, et la nature de ces anomalies n'apparaît pas comme un critère discriminant. Autrement dit, on assiste à la mise en œuvre d'un programme centré sur l'analyse de « profils criminels », mais qui ne parvient pas à leur attribuer de caractéristiques neurophysiologiques. Si la neurocriminologie prétend construire une typologie de catégories juridiques et neurophysiologiques, elle tend seulement à faire de l'excès d'anomalies EEG la condition de possibilité de toutes les formes de dangerosité. Le criminel dangereux apparaît en définitive comme une figure abstraite et confuse, réduite à une seule et même catégorie neurophysiologique (l'excès d'anomalies EEG), dont les critères de définition ne sont eux-mêmes jamais clairement explicités.

Cette remarque nous donne l'occasion de préciser le sens de la démarche typologique adoptée par la neurocriminologie. La caractérisation neurophysiologique du criminel dangereux est le fruit d'une classification établie selon la structure de la personnalité, l'âge de la première condamnation, le motif, le nombre ou encore la sévérité des crimes commis. Une personnalité instable ou agressive, une condamnation précoce, l'absence de motif, la récidive et le crime violent constituent autant d'éléments démontrant la dangerosité de l'individu. Tout au long de ces études, on voit ainsi s'instituer un partage constant entre deux catégories de criminels : le criminel dit « normal », relativement « stable » sur le plan psychologique, le plus souvent condamné pour homicide involontaire ou légitime-défense, et ayant commis un acte isolé ; et le criminel récidiviste, agressif, ou qui commet des crimes violents sans raison. Ces deux catégories sont associées à des valeurs neurophysiologiques exactement opposées : un EEG relativement « normal » et « stable » pour la première, et un EEG « anormal » et « très instable » pour la seconde.

On retrouve ici le constat dressé par Denis Hill (1952) selon lequel « le prisonnier normal a un EEG normal » et son corollaire négatif qui, bien qu'il ne soit pas explicitement mentionné dans la littérature, transparaît dans l'association établie entre le degré d'anomalies EEG et le niveau de dangerosité de l'individu. Cette association fait d'ailleurs apparaître sous une nouvelle forme le principe d'identité psychophysiologique qui structurait l'interprétation des anomalies EEG des psychopathes ou des jeunes

délinquants. Alors qu'au début des années 1940, l'observation d'un « EEG instable » était perçue comme l'indice neurophysiologique d'une « personnalité instable », elle est désormais révélatrice d'une instabilité proportionnelle à la dangerosité de l'individu.

Le meilleur exemple de cet effort constant de reconnaître dans l'excès d'anomalies le signe d'une dangerosité quantitativement équivalente s'illustre dans le souci d'associer l'observation d'un « EEG paroxystique » avec la manifestation d'un « comportement paroxystique ». Comme nous l'avons vu, cette association a été établie en vue de décrire le cas particulier des crimes violents et/ou sans raison (e.g. Howard 1984). L'observation d'une activité cérébrale « paroxystique » apparaît comme la signature neurophysiologique d'un comportement aux propriétés équivalentes, à savoir d'un crime impulsif (non motivé) et commis dans une violence extrême. À l'inverse, la caractérisation neurophysiologique du « criminel normal » repose sur le principe selon lequel « une structure de personnalité relativement bonne est liée à un EEG normal » (Levy et Kennard 1953, p. 838). Elle montre que les auteurs d'un crime isolé présentent un EEG aussi stable que leur personnalité, et qu'ils ne diffèrent donc pas fondamentalement de la population générale.

Finalement, les diverses tentatives visant à accorder la nature du crime commis avec la nature de l'anomalie observée ont pour effet d'accentuer le partage entre le « criminel normal » et le « criminel dangereux », puisqu'elles contribuent à les opposer aussi bien sur le plan comportemental, psychologique et neurophysiologique. Ces deux catégories relèvent d'ailleurs bien plus d'une opposition de nature que d'une différence de degré, dans la mesure où l'excès d'anomalies EEG à un niveau individuel se double d'une surreprésentation (en termes de fréquence d'occurrence) dans la catégorie des criminels les plus dangereux. Non seulement ces derniers sont systématiquement associés au plus haut degré d'anomalies EEG, mais ces anomalies sont observées dans une proportion statistiquement significative par rapport aux autres criminels.

2.5.2. Un projet normatif

Il est important d'insister sur le sens de cette démarche typologique car c'est dans ce contexte qu'est apparue la question de la nature pathologique de certaines formes de criminalité. L'assimilation du crime au domaine de la pathologie s'est formée autour du

constat que le criminel dangereux présente un excès d'anomalies EEG, en l'absence de signes cliniques associés à un quelconque trouble mental ou organique : ces anomalies ont été observées chez des criminels reconnus pénalement responsables, et n'ayant développé aucune pathologie psychiatrique ou organique depuis leur détention. Le soupçon d'une origine pathologique de certains comportements criminels a en particulier bénéficié du constat dressé par Stafford-Clark et Taylor (1949), selon lequel la proportion d'anomalies EEG est similaire à celle que l'on observe chez des criminels reconnus d'aliénation mentale. Le rapprochement établi entre la catégorie des criminels dangereux et celle des criminels aliénés repose donc sur des similitudes observées uniquement sur le plan neurophysiologique, et non sur le plan clinique. En clair, l'excès d'anomalies EEG est interprété comme le seul indice visible d'un état pathologique cliniquement silencieux. Il est perçu comme le stigmate d'une dangerosité qui s'exprime au travers de comportements impulsifs et violents, mais dont le caractère pathologique échappe à une évaluation purement clinique.

Le soupçon d'une origine pathologique a ainsi considérablement orienté les ambitions de la neurocriminologie, puisqu'il a permis de concevoir l'EEG comme un outil privilégié pour repérer des signes pathologiques inaccessibles à l'examen clinique. Comme nous l'avons suggéré plus haut, ces observations ont favorisé l'usage de l'EEG comme un outil prescriptif, et non plus seulement comme un instrument de connaissance du comportement criminel. La fonction descriptive des données neurophysiologiques s'est en effet très vite doublée d'une fonction normative, articulée autour de la question de l'usage judiciaire de l'EEG. Les études que nous avons citées témoignent de la diversité des usages envisagés, puisque ces derniers concernent aussi bien l'évaluation de la responsabilité pénale ou du risque de récidive (Stafford-Clark et Taylor 1949 ; Howard 1984) que la problématique de la prise en charge et du traitement des individus (Levy et Kennard 1953).

Notons également que ce mouvement de pathologisation s'est parfois accompagné d'une conception innéiste de certains comportements criminels. Dans l'étude de Denis Williams (1969), les données EEG ont été mobilisées pour proposer deux modèles explicatifs de la criminalité (endogène et exogène), correspondant respectivement à la catégorie des « criminels dangereux » (récidivistes, agressifs et impulsifs) et des « criminels normaux » (ayant commis un crime isolé). Plus précisément, le modèle

explicatif proposé dans le premier cas consistait à faire de l'excès d'anomalies EEG un défaut « congénital » permettant d'expliquer les crimes violents et sans motif apparent.

Cette tendance consistant à faire de l'excès d'anomalies un défaut « inné » ou « constitutionnel » apparaissait déjà dans les premières publications relatives aux troubles de la personnalité psychopathique (Silverman 1943). Mais en ce qui concerne les crimes violents et sans raison, l'extension de la portée explicative de l'EEG est très nette puisque les données neurophysiologiques sont mobilisées pour expliquer le crime inexplicable, et par voie de conséquence, pour justifier une conception innéiste des comportements les plus dangereux. En outre, il est difficile de ne pas y voir l'expression d'une stratégie consistant à combler l'absence de raison par l'explication neurophysiologique. Il semble en effet que l'intérêt porté aux crimes sans raison révèle une ambition qui dépasse largement la portée explicative des données EEG : faire de l'excès d'anomalies un état pathologique démontrant l'irresponsabilité du sujet, afin de poser la question de la responsabilité pénale comme un problème médical et non plus juridique.

À ce stade, il est surtout important de comprendre le sens de la stratégie globale adoptée par la neurocriminologie basée sur l'EEG. Tout au long des études que nous avons présentées, la portée explicative des données EEG est pensée de telle sorte à favoriser une extension du champ d'action de la neurocriminologie. Bien qu'elle soit entièrement structurée autour de la problématique de la dangerosité, elle s'est très vite instituée comme un moyen de repenser le système pénal, en particulier pour informer la question de la responsabilité pénale. En clair dès son développement, la neurocriminologie a bénéficié d'un ensemble de stratégies visant à doubler sa portée normative. Il ne s'agit pas seulement de déterminer les critères de définition neurophysiologique de la dangerosité criminelle, mais aussi de construire des catégories normatives permettant d'informer des questions fondamentales du droit pénal, comme l'évaluation de la responsabilité d'un prévenu.

2.5.3. Une preuve de l'irresponsabilité pour cause psychiatrique ?

Dans la mesure où les premières publications relatives au champ de la neurocriminologie répondent à une volonté de repenser certaines notions fondamentales du droit pénal, il serait intéressant de s'interroger sur la réception de ces travaux dans la

communauté des juristes. Durant la période de développement de la neurocriminologie, la question de l'usage judiciaire de l'EEG n'a pas été discutée à l'intérieur du champ pénal, à l'exception d'un article publié par un juriste américain, intitulé « *The electroencephalograph (EEG) as evidence of criminal responsability* » (Conrad 1959). Il convient de s'arrêter un instant sur cet article car il répond pleinement aux ambitions fixées par la neurocriminologie des années 1940-1950. Pour autant, il est important de préciser qu'il ne fait aucune mention des études relatives au champ de la neurocriminologie. Cet article ne nous renseigne pas sur la réception de ces études dans le champ du droit pénal, mais il a le mérite de montrer, au terme d'une analyse de la jurisprudence, que le système pénal anglo-saxon n'est pas tout à fait opposé à l'usage de l'EEG pour statuer sur la responsabilité de certains individus.

Dans cet article, Edwin Conrad discute de l'admissibilité de l'EEG comme une preuve de l'irresponsabilité pour cause psychiatrique, en analysant la jurisprudence issue de l'arrêt *Frye*. Cet arrêt prononcé en 1923 définit les critères d'admissibilité des preuves scientifiques dans les cours de justice américaines. Il impose que la technique utilisée soit issue de « connaissances généralement admises » par la communauté scientifique. Edwin Conrad admet d'abord des réserves quant à l'usage judiciaire de l'EEG, en citant quelques-unes de ses limites techniques et méthodologiques. Il évoque notamment le problème de la standardisation des paramètres de l'enregistrement, ou encore l'absence de consensus sur les critères de définition d'un EEG « normal » ou « anormal ». Néanmoins, s'il invite à une grande prudence quant à l'interprétation des données, il admet que l'usage judiciaire de l'EEG s'inscrit dans la continuité logique des recherches empiriques entreprises depuis les années 1920-1930 :

> « *The use of the EEG in criminal cases to prove or disprove the sanity of an accused was a logical development arising out of the studies and demonstrations of Hans Berger.* » (Conrad 1959, p. 409).

Tout au long de cet article, Edwin Conrad cite une dizaine d'affaires pénales et civiles ayant eu recours à l'EEG pour attester la présence d'un trouble affectant la responsabilité du prévenu, et dont les données EEG ont parfois joué un rôle décisif dans le verdict. Les quelques affaires pénales ayant abouti à un verdict d'irresponsabilité pour cause psychiatrique ont concerné des crimes violents et sans motif apparent. Les données EEG ont été principalement mobilisées pour prouver l'existence de lésions cérébrales, ou dans le cadre de ce qui est communément appelé l'« *epilepsy defense* » (ou « *not guilty*

by reason of insanity because of epilepsy »), une disposition légale exceptionnelle qui s'applique si un lien causal peut être établi entre l'acte criminel et la survenue d'une crise épileptique (voir à ce sujet Weinberg 1977 ; Treiman 1986, 1999).[90] Les conclusions que tire Conrad quant à l'admissibilité de l'EEG comme une preuve de l'irresponsabilité pour cause psychiatrique sont les suivantes :

> « *[W]e may say with absolute confidence that the EEG, when used to prove or disprove organic brain damage or epilepsy, has passed the test established by the Frye case. […] However, the EEG, though persuasive, does not ipso facto prove or disprove organic brain damage or epilepsy and, by inference, sanity or insanity. The EEG is only a diagnostic tool used by the medical man to reach a preliminary conclusion whether there is or is not organic brain damage or epilepsy, and a final conclusion whether the individual is sane or insane.* » (Conrad 1959, p. 413).

Le juriste insiste en effet sur la nécessité de mettre en relation les données EEG avec les données cliniques, car si l'EEG peut informer la question de la responsabilité pénale, il ne peut en aucun cas remplacer l'expertise psychiatrique. Il doit selon lui être utilisé comme un indicateur pour confirmer la présence d'un trouble affectant la responsabilité pénale d'un individu. Par ailleurs, il est intéressant de constater que malgré ces quelques réserves, et au même titre que l'étude de Stafford-Clark et Taylor (1949), cet article tient un discours annonciateur du projet de neurodroit. Compte tenu des avancées réalisées dans le domaine de la neurophysiologie, le juriste estime que « l'EEG continuera d'assumer un rôle dominant dans le jugement des affaires pénales dans lesquelles la question de la folie est en jeu. » (p. 414). Plus encore, il s'interroge sur la possibilité d'utiliser l'EEG en complément des polygraphes classiques pour évaluer la fiabilité d'un témoignage. Sans pour autant y voir le signe précurseur d'un programme de recherche qui ne débutera qu'au début des années 1990, cet élément de réflexion n'est pas négligeable dans l'histoire du neurodroit puisqu'il préfigure les premières tentatives

[90] Nous discuterons en détail de cet usage dans la section 2.6. Précisons néanmoins que cette disposition particulière s'inscrit généralement dans le cadre de l'*Insanity Defense* et de la règle *M'Naghten* (Weinberg 1977). La forme d'épilepsie reconnue dans les juridictions de *common law* est l'épilepsie temporale ou psychomotrice, car elle peut produire des pulsions agressives incontrôlables qui altèrent la conscience morale du sujet.

visant à reconnaître dans l'activité EEG d'un individu la « trace mémorielle » d'un mensonge (e.g. Farwell and Donchin 1991)[91] :

> « *It also seems quite logical to suggest that brain wave patterns of persons subjected to the standard lie detection tests now in use should be studied on an experimental basis to determine whether there is any correlation between conscious deception and the form of the brain wave*. » (Conrad 1959 ; p. 414).

Ce bref détour par les enjeux proprement juridiques de l'intégration de l'EEG dans le système pénal nous invite à préciser le contexte scientifique dans lequel s'est inscrit le projet d'utiliser cet outil comme une preuve de l'irresponsabilité pénale. Comme nous venons de le voir, ce projet s'est concrétisé autour de ce qui est appelé l'*epilepsy defense*, c'est-à-dire la recherche d'une preuve que l'acte criminel résulte de la survenue d'une crise épileptique. Cette disposition légale signe les premières tentatives visant à intégrer l'EEG dans le système pénal anglo-saxon, dès le début des années 1940 (Conrad 1959). Bien que controversée dans la communauté des juristes, elle a bénéficié d'une attention particulière à partir des années 1960, dans le cadre d'affaires pénales impliquant des crimes violents et pour lesquels l'expertise psychiatrique laissait subsister un doute sur l'état mental du prévenu (Weinberg 1977 ; Treiman 1986, 1999).

L'*epilepsy defense* mérite un examen attentif des fondements empiriques sur lesquels elle repose, car elle illustre la manière dont les résultats de la neurocriminologie se sont progressivement diffusés dans le champ pénal : l'usage de l'EEG dans ce cadre juridique très spécifique marque un tournant dans le développement et l'extension du champ d'application de la neurocriminologie. Il signe l'aboutissement d'une longue série de recherches visant à démontrer un lien entre l'épilepsie et la manifestation d'un comportement criminel, et dont le principal argument repose sur l'observation d'un « EEG paroxystique » caractéristique du sujet épileptique, chez des criminels ne présentant pourtant aucun signe clinique associé à l'épilepsie. L'hypothèse de l'épilepsie a véritablement nourri l'ambition de reconnaître dans le tracé EEG les stigmates d'un état pathologique cliniquement silencieux. Elle a contribué à définir les anomalies

[91] Comme nous le préciserons dans le chapitre 5, cette ambition s'est formée à partir des recherches menées par Lawrence Farwell sur ce qu'il a appelé lui-même le *Brain Fingerprinting Test.*

paroxystiques comme la principale condition pathologique favorisant l'émergence d'un comportement violent, impulsif ou agressif.

L'hypothèse d'une relation entre l'épilepsie et le comportement criminel a ainsi parcouru le champ de la neurocriminologie et du neurodroit, en s'exprimant sous diverses formes depuis le début des années 1940 : d'une hypothèse explicative conçue pour rendre compte des anomalies EEG des criminels, à une disposition légale qui redéfinit la question de la responsabilité pénale en un problème médical. Il paraît important de reconstruire les arguments avancés en faveur de cette hypothèse, dans la mesure où elle cristallise les principaux enjeux de la neurocriminologie basée sur l'EEG, notamment la conception innéiste et psychopathologique de la criminalité. Elle montre en particulier comment une réflexion d'ordre criminologique (relative à la question de l'origine biologique du comportement criminel) s'est progressivement imposée dans le champ pénal, jusqu'à poser la question de la responsabilité des individus présentant d'importantes anomalies EEG.

Dans la section suivante, nous allons tenter d'expliciter les fondements empiriques de l'hypothèse de l'épilepsie telle qu'elle a été formulée par les neurocriminologues au début des années 1940, jusqu'aux diverses tentatives visant à utiliser l'EEG dans le cadre de l'*epilepsy defense*. Nous insisterons en particulier sur les conditions d'émergence de l'hypothèse *explicative*, à savoir l'hypothèse avancée pour expliquer l'origine des anomalies paroxystiques observées chez les criminels. Cette hypothèse est loin d'être anecdotique dans l'histoire de la neurocriminologie, puisqu'elle s'inscrit dans le cadre de la stratégie de pathologisation du crime amorcée au début des années 1940 : elle a contribué à définir l'épilepsie comme la principale condition pathologique facilitant l'expression d'un comportement criminel. À cet égard, il convient d'ores et déjà de préciser qu'elle repose sur les deux grands principes gouvernant l'interprétation des données EEG recueillies chez les criminels : le principe d'identité psychophysiologique et le modèle psychopathologique.

L'hypothèse de l'épilepsie est d'abord paradigmatique de la stratégie consistant à dégager des relations psychophysiologiques sur la seule base de données EEG, car elle s'est formée autour du postulat d'une relation directe entre l'observation d'un « EEG paroxystique » et la manifestation d'un « comportement paroxystique ». Comme nous l'avons vu plus haut, ce modèle interprétatif a été élaboré en vue de démontrer une équivalence en termes de propriétés entre la nature de l'anomalie observée et la nature du

crime commis. Du point de vue des neurocriminologues, l' « EEG paroxystique » apparaît comme le marqueur neurophysiologique d'un comportement aux propriétés équivalentes, à savoir d'un crime impulsif, commis dans une violence extrême.

C'est à partir de ce cadre théorique que les neurocriminologues vont chercher à promouvoir une conception psychopathologique des comportements violents, impulsifs ou agressifs. L'association établie entre l' « EEG paroxystique » et le « comportement paroxystique » va jouer un rôle central dans la stratégie de pathologisation du comportement criminel, car elle va amener les neurocriminologues à confondre le sujet épileptique et le criminel violent dans la notion de « pulsion irrésistible ». Il s'agira de montrer que les comportements violents, impulsifs ou agressifs résultent d'une seule et même origine pathologique : d'un « EEG épileptique » qui reflète la personnalité instable des individus, et qui favorise l'émergence d'un comportement dangereux.

2.6. L'hypothèse de l'épilepsie

L'hypothèse d'une relation entre l'épilepsie et la manifestation du comportement délinquant ou criminel a été discutée tout au long de la deuxième moitié du XX^{e} siècle (voir en particulier Bonkalo 1967 ; Seward 1987 ; Milstein 1988 ; Volavka 1990, 2002, 2011). Dans la mesure où le développement de l'EEG a largement participé à l'étude et au diagnostic clinique de l'épilepsie (e.g. Gibbs et al. 1935),[92] il n'est pas étonnant qu'elle se soit invitée dans le débat concernant l'interprétation à donner aux anomalies EEG des criminels. Rares sont les études qui ne font pas référence à l'épilepsie, notamment en vue d'exclure les criminels épileptiques de l'échantillon étudié pour ne pas perturber les données. Pourtant, si l'épilepsie a été évoquée dans l'interprétation des anomalies EEG des criminels, ce n'est pas pour s'affranchir d'un éventuel biais de sélection en faveur des sujets présentant des lésions cérébrales, mais bien pour expliquer l'origine de ces anomalies.

[92] Pour une analyse historique du rôle joué par l'EEG dans l'étude de l'épilepsie, nous renvoyons le lecteur à l'article publié par Céline Cherici (2010) intitulé *La définition d'une entité clinique entre développements techniques et spécialisation médicale : épilepsie et épileptologie au XX^{e} siècle.*

Telle qu'elle a été initialement formulée, l'hypothèse de l'épilepsie repose sur un double constat : l'observation chez les criminels d'une « dysrythmie cérébrale » ou d'une « activité paroxystique » caractéristique du tracé EEG enregistré chez les patients épileptiques,[93] et la description de cas cliniques décrivant des accès agressifs immédiatement après une crise épileptique (Hill 1944, 1952 ; Gibbs et Gibbs 1951, 1964). Nous verrons que ce double constat va très vite poser la question des « tendances criminelles » des épileptiques, en favorisant une série d'analyses statistiques visant à évaluer la prévalence ou l'incidence de l'épilepsie dans les populations carcérales, en Grande-Bretagne et aux Etats-Unis (Gunn 1969 ; Gunn and Fenton 1969 ; King and Young 1978 ; Lewis et al. 1982 ; Whitman et al. 1984 ; Wong et al. 1994 ; Fazel et al. 2002 ; Durand et al. 2016). Les données EEG joueront un rôle majeur dans la vérification de cette hypothèse, puisqu'elles permettront à la fois de confirmer le diagnostic d'épilepsie chez les criminels, et d'associer la nature des anomalies observées avec le type d'infraction ou le degré de violence associé. En parallèle de ces recherches, certains psychiatres ayant participé au développement de la neurocriminologie − parmi lesquels Frederic Gibbs, qui a largement contribué à l'étude et au diagnostic clinique de l'épilepsie − vont suggérer que certains crimes peuvent être commis pendant ou immédiatement après une crise épileptique, annonçant les premières tentatives visant à utiliser l'EEG comme une preuve de l'irresponsabilité pour cause psychiatrique (« *not guilty by reason of insanity because of epilepsy* »).

Avant d'aborder la question très spécifique de l'usage de l'EEG dans le cadre de l'*epilepsy defense* − qui relève véritablement du champ du neurodroit, et non de la neurocriminologie en tant que telle – nous préciserons dans un premier temps les fondements empiriques de l'hypothèse d'une relation entre l'épilepsie et le comportement criminel. Nous verrons que cette hypothèse a d'abord été évoquée pour expliquer l'origine des anomalies observées chez les psychopathes et les enfants présentant des troubles du comportement, avant d'être mobilisée dans l'interprétation des données EEG des criminels. Il est donc important de comprendre qu'à ce stade (globalement, jusqu'au début des années 1960), l'hypothèse d'un lien entre l'épilepsie et le comportement criminel répond uniquement à la volonté d'expliquer l'origine des anomalies EEG des

[93] Plus précisément, les études montrent que le comportement délinquant ou criminel est associé à un complexe de pics positifs de 14 et de 6 cycles par seconde, un schéma d'activité cérébrale que l'on rencontre habituellement dans l'épilepsie thalamique et hypothalamique (Gibbs et Gibbs 1951).

criminels. Elle ne concerne pas le problème médicolégal des crimes commis pendant la survenue d'une crise épileptique, mais la recherche d'une condition pathologique qui pourrait favoriser le développement d'un comportement criminel.

2.6.1. Un « EEG paroxystique » comme symptôme d'une « épilepsie latente »

L'hypothèse de l'épilepsie a été évoquée dès les premières publications consacrées aux troubles du comportement pendant l'enfance, à la fin des années 1930. Comme nous l'avons vu au début du chapitre, elle a été suggérée dans la première étude qui a démontré un lien entre les troubles comportementaux et les perturbations de l'activité électrique du cerveau (Jasper et al. 1938). L'analyse des données EEG de certains enfants indiquait la présence d'un « désordre épileptoïde » ou d'une « activité électrique épileptiforme », c'est-à-dire un schéma d'activité cérébrale que l'on rencontre habituellement chez les patients épileptiques entre deux crises (Jasper et al. 1938, p. 647 ; voir aussi Strauss et al. 1940 ; Knott et Gottlieb 1943). L'intérêt porté à cette observation tient au fait que les enfants examinés dans ces études ne présentaient pourtant aucun signe clinique associé à l'épilepsie : les enregistrements EEG évoquaient la survenue d'une « décharge épileptique », malgré l'absence de convulsions. Le même constat a été dressé chez certains psychopathes ou criminels admis dans des institutions psychiatriques, les résultats révélant un « EEG paroxystique » caractéristique du sujet épileptique (Gibbs, Gibbs and Lennox 1938 ; Hill and Watterson 1942 ; Silverman 1943 ; Hill 1944 ; Hill 1952 ; Williams 1969).

Dans la mesure où ces observations ont été effectuées sur des individus témoignant d'une grande agressivité, elles ont très vite posé la question d'une relation entre l'épilepsie et la manifestation d'un comportement agressif. Le psychiatre Denis Hill a été l'un des principaux défenseurs de cette hypothèse. Dans une étude publiée en 1944 − soit aux prémices du développement de la neurocriminologie − il a comparé les données EEG de 80 sujets épileptiques et 80 psychopathes présentant un « comportement agressif dysrythmique » (« *dysrhythmic aggressive behaviour* » ou D.A.B.), dans l'objectif

d'expliquer les anomalies paroxystiques observées chez ces derniers.[94] Selon lui, « l'examen EEG fournit de bonnes raisons de penser qu'il existe une parenté entre le patient D.A.B et l'épileptique » car la fréquence et la nature des anomalies observées est très similaire dans les deux cas (Hill 1944, p. 321).

Une autre observation est venue compléter l'argument d'une relation entre l'épilepsie et les troubles du comportement ou de la personnalité, cette fois-ci sur le plan clinique. Les psychiatres ont remarqué que les enfants présentant des troubles du comportement manifestaient des « troubles mentaux et émotionnels caractéristiques, semblables à ceux qui sont généralement associés à la personnalité épileptique » (Jasper et al. 1938, p. 656). Les notions de « personnalité épileptique » ou de « personnalité épileptoïde » seront alors régulièrement évoquées à propos des enfants présentant des troubles du comportement, offrant une dimension supplémentaire aux similitudes observées sur le plan neurophysiologique. Tout au long des études EEG menées sur les enfants présentant des troubles du comportement, l'interprétation des résultats sera en effet guidée par le souci de faire correspondre leur « EEG épileptique » avec leur apparente « personnalité épileptique ». Comme le résument Denis Hill et Donald Watterson (1942) dans leur étude EEG menée sur des sujets souffrant d'un trouble de la personnalité psychopathique :

> « *These [EEG] abnormalities are very similar to those found in the psycho¬motor equivalents of epilepsy, and the children, by their irritability, tempers, egocentricity, and hyperkinesis, are in some respects reminiscent of the epileptic personality*. » (Hill and Watterson 1942, p. 51).

Dans la mesure où la psychopathie était également associée à un « EEG paroxystique », la même approche comparative a été défendue en vue de confirmer sa parenté avec l'épilepsie, à la fois sur le plan neurophysiologique et clinique. En clair, l'observation d'une « activité électrique épileptiforme » ou d'un « EEG paroxystique » a conduit les psychiatres à postuler l'existence d'une relation directe avec l'instabilité psychologique des individus. Cette double association entre les données EEG et le tableau

[94] Cette typologie correspond à la catégorie psychiatrique de la psychopathie « à prédominance agressive », proposée par le psychiatre britannique David Henderson (1939).

clinique de l'épilepsie et de la psychopathie est résumée ainsi par Denis Hill et Donald Watterson (1942) :

> « *Both from the clinical and from the electro-encephalographic points of view, there is plenty of evidence suggesting an association between certain forms of psychopathy and epilepsy. The episodic disturbance of behaviour, the explosive dangerousness of the aggressive psychopath who takes no thought for the consequences of his behaviour and the alleged amnesia, have led many psychiatrists to speak of an epileptoid personality*. » (Hill and Watterson, 1942, p. 60)

L'attention portée à ces observations tient au fait qu'elles ont été effectuées sur des sujets non épileptiques, mais dont les manifestations comportementales et la nature des anomalies EEG sont généralement associées à l'épilepsie. Les troubles du comportement pendant l'enfance et les troubles de la personnalité psychopathique ont été décrits comme des équivalents d'une épilepsie sans convulsions, car associés à « une activité électrique corticale de type épileptique en l'absence de phénomène convulsif. » (Brill and Walker 1945, p. 545). Ce constat a alors amené les psychiatres à évoquer l'existence d'une « épilepsie latente » pour expliquer l'origine de ces anomalies paroxystiques (Williams 1941, 1944 ; Brill and Walker 1945).

Cette notion a été introduite par Denis Williams (1941) pour rendre compte de cette « dysrythmie épisodique de type épileptique » et éviter toute confusion avec sa manifestation clinique. La présence de ces anomalies EEG malgré l'absence de symptômes convulsifs indique selon lui « un défaut constitutionnel inné pouvant s'exprimer par l'épilepsie ou un trouble du comportement tel qu'une psychonévrose constitutionnelle, une psychopathie ou une psychose, chez le sujet lui-même ou chez ses enfants » (Williams 1941, p. 266). Il s'agit de montrer que les anomalies paroxystiques ne résultent pas d'une lésion cérébrale acquise, mais d'un défaut congénital s'exprimant sous la forme d'un profil EEG caractéristique.

La notion d' « épilepsie latente » a donc été évoquée en vue de démontrer l'origine constitutionnelle des anomalies paroxystiques observées chez les psychopathes et les enfants présentant des troubles du comportement. Elle permet de caractériser ces derniers sous la même condition pathologique, mais aussi d'apporter un fondement théorique au postulat d'une relation directe entre l' « EEG épileptique » et la « personnalité

épileptique » des individus : d'après les neurocriminologues, les analogies observées sur le plan neurophysiologique et comportemental reflèteraient la présence d'une épilepsie maintenue dans un état latent. Autrement dit, si les données recueillies chez les psychopathes et les enfants présentant des troubles du comportement révèlent des anomalies de type épileptique, c'est parce que ces troubles résultent d'une même cause pathologique.

L'hypothèse de l' « épilepsie latente » signe le début d'une longue série d'investigations empiriques visant à étudier la question de la relation entre l'épilepsie et la manifestation d'un comportement criminel. L'observation d'un « EEG épileptique » en l'absence de phénomène convulsif a en effet été rapportée dans les populations criminelles, en particulier chez les détenus condamnés pour meurtre (e.g. Hill and Pond 1952 ; Driver et al. 1974 ; Okasha et al. 1975). Plus généralement, l'hypothèse d'une relation entre l'épilepsie et le comportement criminel a bénéficié d'une observation constante depuis le début des années 1940, à savoir la présence d'une forte proportion d'épileptiques parmi les criminels examinés (e.g. Silverman 1944 ; Stafford-Clark et Taylor 1949 ; Mundy-Castle 1955). L'hypothèse de l'épilepsie s'est alors progressivement imposée dans l'interprétation des anomalies EEG des criminels, dans la continuité directe des résultats concernant la psychopathie et les troubles du comportement pendant l'enfance.

L'idée d'une association entre l'épilepsie et le meurtre a été explicitement défendue par Hill et Pond (1952), à l'issue d'une série d'examens EEG de 105 meurtriers condamnés à la peine capitale. Les données EEG ont révélé la présence d'une « activité épileptique » chez certains criminels (en particulier ceux ayant commis des crimes sans motif) ne présentant pourtant aucun signe clinique d'épilepsie. Après avoir observé une proportion d'épileptiques « 32 fois plus importante » que dans la population générale, et tout en précisant que certains d'entre eux n'avaient aucun antécédent de crise d'épilepsie, Hill et Pond concluent :

> « *The evidence, therefore, for some relationship existing between murder and epilepsy in some murderers is undoubted, but this is not to say that such murders are committed in epileptic seizures or iir post-epileptic automatisms*. » (Hill et Pond 1952, pp. 32-33).

Cet argument revêt une grande importance en ce qui concerne le sens à donner à l'hypothèse d'une relation entre l'épilepsie et le comportement criminel, car il révèle l'influence exercée par la notion d'« épilepsie latente » telle qu'elle a été formulée par Denis Williams (1941) à propos de la psychopathie et des troubles du comportement pendant l'enfance. Il ne s'agit pas d'expliquer la manifestation d'un comportement criminel par la survenue d'une crise épileptique mais, pour reprendre les termes utilisés par les psychiatres, d'admettre la possibilité d'une relation « indirecte » ou « plus générale » permettant de rendre compte des anomalies de type épileptique malgré l'absence de symptômes convulsifs (Hill et Pond 1952 ; Mundy-Castle 1955).

Une attention particulière a en effet été portée à l'absence de jurisprudence concernant les crimes commis en relation directe avec la survenue d'une crise épileptique. L'association établie entre le meurtre et l'épilepsie ne concerne donc pas le problème médicolégal des crimes commis pendant ou immédiatement après une crise, mais bien la recherche d'une « anomalie constitutionnelle » commune au développement de l'épilepsie et du comportement criminel. Autrement dit, l'argument avancé pour expliquer les anomalies de type épileptique des criminels consiste à confondre l'épilepsie et le comportement criminel dans une seule et même origine pathologique. Ces anomalies traduiraient un défaut constitutionnel facilitant l'expression d'un acte criminel, indépendamment de la survenue d'une crise épileptique.[95]

Les données EEG récoltées depuis le début des années 1940 sur les populations carcérales vont donner une impulsion décisive dans la recherche des signes pathologiques contribuant au développement du comportement criminel. La mise en évidence d'un « EEG épileptique » chez les criminels les plus instables va en particulier poser la question de savoir si l'épilepsie augmente le risque de commettre un crime violent. Les études EEG vont alors inspirer une longue série d'études visant à préciser la relation entre l'épilepsie et le comportement criminel, réactualisant l'hypothèse de l'épilepsie autour d'une double ambition : l'évaluation de la prévalence ou de l'incidence de l'épilepsie dans les populations carcérales (Gunn 1969 ; Gunn and Fenton 1969 ; King and Young

[95] Il convient de spécifier que les études ne proposent pas davantage de précisions concernant la nature de la relation entre les anomalies de type épileptique et la manifestation d'un comportement criminel. Certains psychiatres évoquent néanmoins l'hypothèse que ces anomalies impliquent une plus grande sensibilité au stress et une altération du contrôle de soi (Woods 1961 ; voir aussi Mundy-Castle 1955).

1978 ; Lewis et al. 1982 ; Whitman et al. 1984 ; Wong et al. 1994 ; Fazel et al. 2002 ; Durand et al. 2016), et l'étude de la relation entre la survenue d'une crise épileptique et la manifestation d'un comportement agressif ou violent (Serafetinides 1965 ; Ervin, Mark, and Sweet 1969 ; Delgado-Escueta et al. 1981 ; Devinsky and Bear 1984 ; Marsch and Krauss 2000 ; Shih et al. 2009).[96]

Bien que les résultats de ces études ne soient pas toujours positifs, elles illustrent l'importance accordée à la question de la relation entre l'épilepsie et le comportement criminel. L'évaluation de la prévalence de l'épilepsie dans les populations carcérales va en effet favoriser une série d'analyses statistiques qui poseront véritablement la question des tendances criminelles des épileptiques. L'enjeu central de ces recherches ne sera pas tant d'évaluer la proportion d'épileptiques dans les populations carcérales, mais surtout de déterminer si les détenus épileptiques commettent des crimes plus violents que les autres détenus. L'utilisation de l'EEG permettra alors d'une part, de confirmer l'association établie entre le degré d'anomalies EEG et le degré de dangerosité des individus, et d'autre part, de tester l'hypothèse de « l'épilepsie latente » sur un large échantillon de criminels.

2.6.2. Une forte prévalence de l'épilepsie dans les populations carcérales

Les études cherchant à évaluer la prévalence ou l'incidence de l'épilepsie dans les populations carcérales se sont développées dans le prolongement des premières études EEG menées sur les criminels. Elles proposent d'analyser et de préciser une observation constante depuis le début des années 1940, à savoir la forte proportion d'épileptiques dans les populations carcérales.[97] Les données EEG vont d'ailleurs jouer un rôle fondamental dans les stratégies de recherche mobilisées pour étudier la relation entre l'épilepsie et le risque de commettre un crime violent. Elles permettront notamment de corréler les

[96] Ces études, qui s'appuient généralement sur des enregistrements EEG, s'accordent à montrer que des comportements violents ou agressifs peuvent survenir pendant la phase post-critique, c'est-à-dire immédiatement après une crise. Pour une analyse critique de ces travaux, voir Stevens and Hermann (1981) ; Treiman and Delgado-Escueta (1983) ; Volvaka (1995) ; Fessler and Treiman (2009).

[97] Déjà en 1924 au Royaume-Uni, une étude avait rapporté une forte prévalence de l'épilepsie parmi les criminels internés dans un hôpital psychiatrique de haute sécurité (Sullivan et al. 1924).

différentes formes d'épilepsie diagnostiquées, avec le type d'infraction ou le degré de violence associé.

Le psychiatre britannique John Gunn a initié ce programme de recherche, en analysant les données épidémiologiques de plusieurs institutions carcérales en Angleterre et au Pays de Galles (e.g. Gunn 1969). Ses premières publications tendent à confirmer l'hypothèse d'une relation entre l'épilepsie et le comportement criminel. Elles rapportent en moyenne une prévalence de l'épilepsie deux à quatre fois plus importante dans les populations carcérales que dans la population générale (Gunn 1977). Dans une étude comparant la prévalence de l'épilepsie dans une vingtaine d'institutions carcérales britanniques avec les données épidémiologiques de la population générale, les conclusions de John Gunn sont sans équivoque : « les épileptiques ont une plus forte probabilité de se retrouver en prison que les autres membres de la communauté » (Gunn and Fenton 1969, p. 328).

Il est intéressant de noter que John Gunn inscrit ses recherches − ainsi que les études EEG menées sur les criminels et psychopathes − dans la lignée des thèses de Cesare Lombroso et Henry Maudsley. Selon lui, les études rapportant la présence d'un « EEG épileptique » chez les psychopathes ou criminels illustrent une réactualisation de l'hypothèse de l'épilepsie telle qu'elle a été formulée au XX^{e} siècle : « l'idée d'une relation entre certaines formes d'épilepsie et certaines formes d'agressivité a persisté, même si le phénomène de l'épilepsie est mieux compris aujourd'hui » (Gunn 1969 ; p. 60).

D'ailleurs, les stratégies de recherche privilégiées dans l'étude de la relation entre l'épilepsie et le risque de commettre un crime violent rappellent à certains égards la démarche typologique et classificatoire de Lombroso. L'une des stratégies employées par Gunn consistera à comparer les criminels épileptiques et non-épileptiques selon le type de condamnation criminelle (Annexe n°7, table I), le degré de violence et d'impulsivité associé au crime (Annexe n°7, tables IV et VI), ou encore selon leurs antécédents judiciaires (Gunn and Bonn 1971). Les données EEG vont être mobilisées pour préciser la relation entre l'épilepsie et le risque de commettre un crime violent. Elles permettront non seulement de confirmer le diagnostic d'épilepsie et préciser la localisation de la lésion épileptogène, mais surtout d'associer les différentes formes d'épilepsie observées (idiopathique, temporale, etc.) avec le type d'infraction (violente ou non) commise (Annexe n°7, table II).

Les résultats de cette étude ne montrent aucune différence statistique entre les détenus épileptiques et les détenus non-épileptiques en ce qui concerne leur niveau de violence ou d'impulsivité ou le type d'infraction commis (voir aussi Gunn and Fenton 1971). Après avoir constaté que la plupart des détenus épileptiques ont été condamnés pour des infractions mineures (généralement pour des vols sans violence), Gunn et Bonn (1971) concèderont « qu'il est manifestement inexact de penser que les détenus épileptiques sont particulièrement violents. » (p. 341). Pourtant, il est intéressant de noter que les auteurs défendent la compatibilité de leurs résultats avec les études EEG soutenant l'hypothèse de « l'épilepsie latente ». Selon eux, ces résultats « ne contredisent pas les travaux de Hill et Pond (1952) qui ont montré une association entre l'épilepsie et le meurtre » (p. 341) car les différences observées peuvent être attribuées à une différence de taille de l'échantillon.

En outre, Gunn et Bonn (1971) admettent que cette association vaut pour certaines formes d'épilepsie. Les mesures EEG révèlent que « les différentes sous-catégories diagnostiques du groupe des [criminels] épileptiques sont associées à différents patterns de criminalité » : les criminels présentant une épilepsie temporale sont ceux qui ont été les plus fréquemment condamnés, tandis que l'épilepsie idiopathique représente « une part disproportionnellement grande des condamnations pour crimes violents. » (Gunn et Bonn 1971 ; p. 342). Ainsi pour Gunn et Bonn, si la relation entre l'épilepsie et le comportement criminel doit être analysée en fonction des différentes formes cliniques de l'épilepsie, elle confirme globalement l'association observée au début des années 1950 entre le degré d'anomalies EEG et le degré de dangerosité des individus :

> « *The rating scale for impulsivity produced an interesting between-group difference which would fit well with the observation reported by Stafford-Clark and F. H. Taylor (1949), and by Hill and Pond (5952), that individuals with abnormal EEG's are more likely to commit motiveless violence (homicidal violence in their series).* » (Gunn et Bonn 1971, p. 341).

L'étude de Gunn et Bonn (1971) constitue un exemple paradigmatique de la manière dont les données statistiques ont été mobilisées pour tester l'hypothèse de « l'épilepsie latente » sur un large échantillon de criminels. Depuis le début des années 1960, les analyses statistiques concernant la prévalence ou l'incidence de l'épilepsie dans les populations carcérales constituent l'arrière-plan d'un programme de recherche visant à étudier la relation entre l'épilepsie et le crime violent. Elles ont véritablement renouvelé

la question des tendances criminelles des épileptiques, en cherchant à recouper les données cliniques, neurophysiologiques, et les données relatives au dossier judiciaire des sujets. Le souci d'associer la nature des anomalies EEG avec la nature du crime commis va d'ailleurs persister jusqu'à la fin du XX^{e} siècle, comme en témoigne cet extrait d'une étude consacrée à la relation entre l'épilepsie temporale et les comportements violents :

> « *Although epilepsy was not significantly associated with violent behaviour, the incidence and severity of E.E.G. abnormalities were directly related to incidence, severity, and bizarreness of the aggressive act.* » (Stevens et Hermann 1981, p. 127).

L'évaluation de la prévalence ou de l'incidence de l'épilepsie continuera de faire l'objet de nombreuses analyses statistiques tout au long de la deuxième moitié du XX^{e} siècle (King and Young 1978 ; Lewis et al. 1982 ; Whitman et al. 1984 ; Wong et al. 1994). La plupart de ces études observent une proportion non négligeable d'épileptiques parmi les criminels, mais concluent à l'absence de relation entre l'épilepsie et le degré de sévérité ou de violence du crime. Ainsi, si la question de la relation entre l'épilepsie et le comportement criminel bénéficie d'un traitement attentif dans la littérature médicale contemporaine, les travaux présentant des résultats positifs restent minoritaires.

Par ailleurs, les données concernant la prévalence de l'épilepsie dans les populations carcérales sont particulièrement difficiles à interpréter dans la mesure où les critères diagnostics de l'épilepsie ne sont pas toujours explicitement définis, et le groupe contrôle n'est pas toujours représentatif de l'échantillon étudié. Comme le relève le neurologue David M. Treiman (1986), la prévalence observée dans les populations carcérales est similaire à celle que l'on rencontre dans les milieux défavorisés, dont sont issus la plupart des prisonniers britanniques ou américains. Ces données pourraient donc simplement refléter une association entre l'épilepsie et le statut socio-économique des individus, plutôt qu'un risque accru de criminalité.

Devant l'abondance de résultats confus et souvent contradictoires, des méta-analyses récentes ont été réalisées en vue de trancher la question de l'association entre l'épilepsie et le risque accru de criminalité (e.g. Fazel et al. 2002). La question de savoir si l'épilepsie augmente le risque de commettre un crime violent a été explicitement posée dans une étude populationnelle portant sur 35 années de registres de la population suédoise (Fazel et al. 2011 ; voir aussi Fazel et al. 2009). Ici encore, ces études concluent

à l'absence de relation entre l'épilepsie et le risque de commettre un crime violent. Elles évoquent en particulier l'existence d'un biais de publication dans les études précédentes, qui tend à surestimer le taux de prévalence des épileptiques dans les populations criminelles.

Jusqu'ici, l'hypothèse de l'épilepsie s'est développée sans faire référence à la notion de responsabilité pénale, en excluant le problème proprement juridique des crimes commis pendant la survenue d'une crise épileptique. Ce problème sort du cadre de la neurocriminologie proprement dite, car il concerne les affaires pénales qui ont eu recours à l'EEG pour confirmer le diagnostic d'épilepsie chez des criminels ayant commis des crimes violents et/ou sans motif apparent. L'hypothèse de l'épilepsie s'inscrit ici dans le contexte très particulier de l'*epilepsy defense,* c'est-à-dire la recherche d'une preuve d'un lien causal entre l'acte criminel et la survenue d'une crise épileptique. Comme nous l'avons mentionné plus haut, cette disposition légale marque les premières tentatives visant à intégrer l'EEG dans le système pénal anglo-saxon. Il paraît important de préciser brièvement le contexte et la portée de cet usage, dans la mesure où il interroge les conditions pour lesquelles l'observation d'un « EEG épileptique » peut être considérée comme une preuve d'irresponsabilité (*The Insanity Defense*) ou d'atténuation de la responsabilité pénale (*The Diminished Capacity Defense*).

2.6.3. L'epilepsy defense ou « *not guilty by reason of epilepsy* »

Certaines affaires pénales impliquant un crime violent ont bénéficié d'une « irresponsabilité pour cause d'épilepsie », en particulier dans les années 1970-1980 (Weinberg 1977 ; Treiman 1986, 1999). Le problème que nous voudrions évoquer ici ne concerne pas tant la question de la responsabilité pénale des sujets épileptiques ayant commis un crime pendant ou immédiatement après une crise, mais plutôt la question du rôle joué par les données EEG dans la recherche d'une preuve de l'irresponsabilité ou de l'atténuation de la responsabilité. Bien que l'*epilepsy defense* soit invoquée à titre exceptionnel et reste encore aujourd'hui controversée dans la communauté des juristes, elle nous renseigne sur la manière dont les données EEG ont été mobilisées dans le champ pénal pour informer la question de l'état mental d'un prévenu. Par ailleurs, elle apporte

d'une certaine manière une réponse judiciaire à l'ensemble des études EEG démontrant un lien entre l'épilepsie et la manifestation d'un comportement criminel. Précisons en effet que les premières affaires pénales ayant eu recours à l'EEG dans le cadre de l'*epilepsy defense* coïncident avec la série de publications consacrées à l'évaluation de la prévalence ou de l'incidence de l'épilepsie dans les populations carcérales (e.g. Gunn 1969), et à la relation entre la survenue d'une crise épileptique et la manifestation d'un comportement agressif ou violent (e.g. Ervin, Mark, and Sweet 1969).

L'affaire Jack Ruby, auteur du meurtre de Lee Harvey Oswald (l'assassin présumé de J.F. Kennedy) a véritablement lancé le débat sur l'admissibilité de l'EEG comme une preuve de « l'irresponsabilité pour cause d'épilepsie » (Livington 1964). Elle constitue en ce sens l'une des premières affaires marquant l'usage de l'EEG dans le cadre de l'*epilepsy defense*.[98] La defense de Ruby s'est en effet organisée autour de l'argument de l'épilepsie psychomotrice dont il souffrait, preuve attestée par des « anomalies EEG » traduisant une « forme très rare d'épilepsie », selon les termes de Frédéric Gibbs qui a réalisé l'expertise.[99] Depuis, plusieurs études de cas ont été discutées dans la communauté psychiatrique (Fenton and Udwin 1965 ; Gunn 1978 ; Bartholomew et al. 1978 ; Cope and Donovan 1979 ; Hindler 1989), y compris récemment (Eisenschenk et al. 2014). On compte également quelques revues de littérature concernant l'usage de l'*epilepsy defense* aux Etats-Unis (Treiman 1986 ; Pandya et al. 2013) ou en Grande-Bretagne (Mackay 1990 ; Reuber and Mackay 2008).

La synthèse la plus exhaustive (bien que relativement ancienne) est celle de David Treiman (1986), qui a analysé la jurisprudence américaine relative à l'*epilepsy defense* entre 1889 et 1984. Durant cette période, elle aurait été utilisée dans au moins 75 affaires pénales impliquant un crime violent survenu pendant la phase critique ou post-critique (Treiman 1986). Comme le souligne Treiman, un seul cas s'est conclu par une

[98] Précisons à cet égard que si la jurisprudence relative à l'*epilepsy defense* remonte à la fin du XIX^{e} siècle (voir à ce sujet Treiman 1986), l'usage de l'EEG dans ce contexte est relativement récent. La première affaire pénale ayant eu recours à l'EEG dans le cadre de l'*epilepsy defense* remonte vraisemblablement à 1960, et concernait un homicide involontaire suivi d'une amnésie (*Government of the Virgin Islands v. Smith*, 278 E2d 169, 1960). Comme le précise le juriste David Treiman (1986), la défense s'était appuyée sur des examens EEG qui révélaient la présence d'une « activité épileptiforme ».

[99] Bien que Ruby ait été reconnu pénalement responsable, l'expertise menée par Gibbs a joué un rôle notable dans le procès puisque ses conclusions évoquaient une association entre ces anomalies EEG et une « perte de contrôle émotionnel » (Gutmann 2007 ; Brigo 2013).

irresponsabilité pour cause psychiatrique (« *not guilty by reason of insanity* »).[100] Globalement, la preuve d'une relation entre l'épilepsie et le comportement criminel se limite à l'observation d'une simple « anomalie EEG » chez le prévenu. En Grande-Bretagne, 13 affaires pénales auraient bénéficié du « *not guilty by reason of epilepsy* » entre 1975 et 2001, ce qui représenterait 7.3 % de tous les verdicts d'irresponsabilité pour cause psychiatrique (Reuber and Mackay 2008).

D'après les auteurs de ces deux études, ces affaires bénéficieraient d'un flou juridique en ce qui concerne les critères permettant de déterminer si l'acte criminel résulte d'une crise épileptique (voir à ce sujet Delgado-Escueta et al. 1981 ; Rodin 1982 ; Hindler 1989). Les controverses autour l'*epilepsy defense* concernent également son usage dans le cadre de l'*Insanity Defense* (ou défense d'aliénation mentale). De nombreux juristes ou psychiatres s'accordent sur la nécessité de l'inclure dans un cadre juridique plus approprié, afin d'éviter l'internement de sujets épileptiques ne présentant aucun trouble psychiatrique (Weinberg 1977 ; Paul and Lange 1992 ; Mackay and Reuber 2007).[101]

En résumé, si l'hypothèse de l'épilepsie a été largement discutée tout au long de la deuxième moitié du XX^{e} siècle, la relation entre un comportement criminel et la survenue d'une crise épileptique a été rarement démontrée, aussi bien dans la littérature scientifique que dans la jurisprudence criminelle. Il n'est d'ailleurs pas exclu que certains verdicts d'irresponsabilité pénale soient justifiés par le soupçon d'une autre pathologie que l'épilepsie. Comme le relèvent certains psychiatres ou neurologues, l'hypothèse d'une relation entre l'épilepsie et le comportement criminel pourrait s'expliquer par la seule présence de lésions cérébrales (e.g. Treiman 1986). Autrement dit, l'observation de comportements violents ou agressifs chez les sujets épileptiques serait imputable à des déficits neurologiques ou psychiatriques indirectement associés à l'épilepsie. L'hypothèse d'une relation entre le comportement criminel et la présence de lésions cérébrales fera d'ailleurs l'objet d'une analyse plus approfondie à partir du début des années 1990, à la faveur du développement des techniques de neuroimagerie fonctionnelle comme l'IRMf.

---

[100] Matter of Torsney, 47 NY2d 667, 394 NE2d 262, 420 NYS2d 192 (1979).

[101] Pour les neuropsychologues Klaus Lange et Geraldine Paul (1992) : « *[t]he law of insanity is inappropriate when applied to epileptic automatisms. A change in the law is therefore necessary to remove epileptic offenders from the legal ambit of insanity, and to save them the threat of detention in a mental hospital as a consequence of their crime.* » (Paul and Lange 1992, p. 160).

Dans ce chapitre, nous nous sommes efforcés de retracer la genèse du projet neurocriminologique, depuis les premières publications relatives aux troubles du comportement ou de la personnalité, jusqu'aux tentatives plus récentes visant à utiliser l'EEG comme une preuve de l'irresponsabilité pénale. La construction de ce projet laisse apparaître quelques fondements épistémologiques communs avec l'anthropologie criminelle de Cesare Lombroso et, dans une moindre mesure, avec l'organologie de Franz-Joseph Gall. On peut d'ores et déjà relever quelques affinités du point de vue des présupposés, des stratégies de recherche et des finalités poursuivies : l'adoption d'un principe d'identité psychophysiologique et d'un modèle psychopathologique de la criminalité, la mise en œuvre d'une démarche typologique et classificatoire, ou encore l'émergence d'une ambition réformatrice fondée sur un impératif de protection sociale et dans laquelle le critère de dangerosité est fixé par le degré d'anomalies cérébrales. Les affinités conceptuelles et méthodologiques se dessinent également dans les inférences interprétatives mobilisées pour rendre compte des anomalies observées chez les criminels : lorsque ces anomalies ne sont pas attribuées à un défaut de développement du cerveau, elles sont reléguées à une condition « épileptoïde », c'est-à-dire à une épilepsie maintenue dans un état latent, et qui permettrait d'expliquer les crimes violents et sans motif apparent.

Un autre point essentiel à retenir de cette analyse est l'indistinction du champ de la neurocriminologie et du neurodroit. Tout au long de la seconde moitié du XX^{e} siècle, l'usage de l'EEG sur les populations criminelles s'est développé dans un *continuum* de réflexions d'ordre criminologiques et pénales. Il a été motivé par l'ambition de définir des catégories neurophysiologiques utiles à l'évaluation de la dangerosité de l'individu, mais aussi pour informer la question de la responsabilité pénale ou de la prise en charge des criminels. Ce point est important car les études menées sur les populations criminelles se sont par la suite détournées des réflexions d'ordre judiciaire. Les études que nous avons présentées dans ce chapitre sont les seules à poser de manière explicite la question de la prise en charge ou de la responsabilité (morale ou pénale) des criminels présentant d'importantes anomalies cérébrales. Depuis, les neurocriminologues font preuve d'une grande prudence quant à la portée judiciaire de leurs résultats, en particulier en ce qui concerne leurs implications sur la notion de responsabilité pénale.

Cette réserve laisse présager une rupture entre le projet neurocriminologique tel que nous l'avons présenté jusqu'ici, et celui qui se développera à partir des années 1990

sur la base de la neuroimagerie anatomique ou fonctionnelle (IRMa, IRMf, TEP, etc). La neurocriminologie contemporaine va en effet centrer son analyse sur la question de la dangerosité. L'ambition d'en proposer une définition neurobiologique sera centrale et la question de la responsabilité des criminels présentant d'importantes anomalies cérébrales sera très rarement discutée à l'intérieur du champ.[102] Cette restructuration du projet neurocriminologique ne va pas pour autant effacer les traits caractéristiques du projet initial, puisque les principes théoriques et méthodologiques qui sont à l'œuvre dans l'étude neurophysiologique des criminels vont perdurer tout au long du développement de la neurocriminologie contemporaine. Les données neurobiologiques recueillies chez les criminels seront interprétées selon une approche typologique, c'est-à-dire sur la base d'un recoupement avec des catégories psychologiques emblématiques d'un comportement dangereux : tout l'enjeu de la neurocriminologie sera d'isoler des catégories neurobiologiques associées à des comportements typiques comme l'impulsivité, l'agressivité, l'insensibilité affective ou morale, dans l'objectif de définir des critères de mesure (quantitative) de la dangerosité criminologique.

[102] Il faudra attendre le développement du neurodroit au début des années 1990 pour réintroduire une réflexion sur l'usage des données et des outils neuroscientifiques pour informer le système pénal.

CHAPITRE 3

IMAGERIE CÉRÉBRALE ET « VISION NEUROCRIMINOLOGIQUE »

L'émergence des techniques de neuroimagerie anatomique et fonctionnelle a marqué un tournant dans le développement et la structuration du projet neurocriminologique. Dès le début des années 1990, on observe une réduction significative du nombre de publications relatives aux anomalies EEG des criminels, par rapport aux études utilisant les techniques de neuroimagerie comme l'IRMf.[103] Dans la mesure où l'IRMf offre une résolution spatiale nettement plus fine que l'EEG, elle apparaît comme un outil plus fiable pour *localiser* les anomalies cérébrales dans des régions individualisées du cerveau des criminels, et ainsi isoler les états cérébraux qui pourraient être associés aux comportements violents, impulsifs ou agressifs. Le projet neurocriminologique s'est ainsi affiné sur le plan méthodologique, formulant des exigences nettement plus précises pour caractériser la relation entre les anomalies cérébrales et le comportement criminel. Contrairement aux études EEG qui sont pour la plupart effectuées « au repos », l'usage de l'IRMf nécessite de définir une tâche cognitive permettant d'isoler les régions cérébrales associées au comportement d'intérêt.[104] La construction des protocoles expérimentaux pose donc un certain nombre de questions conceptuelles et empiriques, notamment la question de savoir si ces tâches cognitives sont pertinentes vis-à-vis du comportement que l'on cherche à étudier.

Par ailleurs, et si cette remarque peut sembler évidente au regard de l'histoire des neurosciences cognitives, il est quand même important de préciser que la

[103] Une recherche effectuée sur *PubMed* concernant la période 1990 - 2020 recense 94 études EEG, contre 190 études utilisant l'IRMa, l'IRMf ou la TEP.

[104] Il est important de savoir que l'IRMf ne permet pas de détecter l'activation absolue d'une région cérébrale, mais seulement des différences entre deux états d'activité. C'est pourquoi l'enregistrement de l'activité cérébrale impliquée dans la réalisation d'une tâche cognitive repose sur la méthode dite de soustraction. Cette méthode, empruntée à la psychologie cognitive, consiste à soustraire les niveaux d'activité mesurés pendant l'exécution de la tâche aux niveaux d'activité mesurés pendant une tâche contrôle (correspondant généralement à l'état de repos) (e.g. Sergent 1994).

neurocriminologie contemporaine ne s'inscrit pas dans le même contexte que les premières publications des années 1940-1950. Depuis le début des années 1990 et l'émergence de l'IRMf, on assiste à une démultiplication des champs d'application des neurosciences, dont la plupart répondent à une visée normative : les neurosciences sont de plus en plus sollicitées pour répondre à une demande sociale ou politique (Ehrenberg 2018).[105] La neurocriminologie contemporaine a donc bénéficié d'une généralisation du « neurodiscours » et d'une tendance éliminativiste dont la finalité est de redéfinir en termes neurobiologiques des problèmes que l'on ne peut plus résoudre à l'intérieur de son champ disciplinaire (Frazzetto and Anker 2009 ; Littlefield and Johnson 2012 ; Rose and Abi-Rached 2013 ; Vidal and Ortega 2017). Cette tendance consistant à voir dans les neurosciences la solution à tous les problèmes sociétaux a été décrite sous les traits d'une « pathologie » par le juriste américain Stephen J. Morse (2006) : le « Brain Overclaim Syndrome » (BOS), dont il s'est attaché à décrire la symptomatologie, les principaux critères de diagnostic, et les possibilités de traitement. Selon lui, le BOS a créé un terrain particulièrement propice à l'émergence du neurodroit. Le pouvoir alloué aux neurosciences cognitives a permis de légitimer l'ambition de tirer des conclusions normatives (notamment en ce qui concerne la question de la responsabilité pénale) sur la seule base des données neurobiologiques, avant même que le projet ne soit empiriquement fondé.

Ainsi, depuis l'émergence des techniques de neuroimagerie, on observe un accroissement exponentiel des recherches menées sur le cerveau des criminels (Larregue 2019). Or, la démultiplication de ces études a contribué à brouiller les pistes du projet initial : le principal problème posé par la neurocriminologie contemporaine est qu'elle entretient un certain flou quant aux finalités recherchées. L'analyse du discours neurocriminologique ne suffit pas à comprendre la logique qui gouverne l'étude neurobiologique des criminels car les neurocriminologues n'affichent pas explicitement leurs ambitions, préférant recourir à des stratégies rhétoriques pour asseoir leur légitimité scientifique et s'implanter dans le champ des neurosciences cognitives. Il serait d'ailleurs

[105] Rappelons que les neurodisciplines ont des statuts disciplinaires, des histoires institutionnelles et des finalités essentiellement différentes. Elles ont néanmoins comme point commun d'utiliser les techniques de neuroimagerie dans une visée essentiellement prescriptive, par exemple pour définir des pratiques pédagogiques plus adaptées au développement cognitif des enfants (neuroéducation), identifier les stratégies marketing les plus efficaces (neuromarketing), ou pour résoudre des problèmes philosophiques et trancher un débat théorique (neurophilosophie).

erroné de voir la neurocriminologie contemporaine comme un simple prolongement du projet qui s'est dessiné depuis les années 1940-1950. Les études publiées depuis le début des années 1990 dans le champ de la neurocriminologie ou du neurodroit ne font pas mention de ces publications, ce qui pose la question de savoir s'il existe une continuité entre ces projets.

Ce chapitre vise essentiellement à clarifier les ambitions de la neurocriminologie contemporaine, à préciser les différentes formes que recouvre son objet de recherche, et à expliciter la logique des stratégies mobilisées dans l'étude neurobiologique des criminels. Nous nous intéresserons en particulier à la question de la légitimité scientifique de ce que nous appelerons le « programme fort » de la neurocriminologie, à savoir l'ambition de montrer que le comportement criminel est *entièrement expliqué* par la présence d'anomalies cérébrales. C'est pourquoi la plupart des problèmes épistémologiques que nous soulèveront s'articulent autour de la question de la relation entre la nature des finalités poursuivies par les neurocriminologues, les stratégies de recherche qui sous-tendent la construction de leurs modèles expérimentaux, et la manière dont les données empiriques sont interprétées. Comme nous le montrerons tout au long de cette analyse, la relation entre ces trois niveaux de description a des implications majeures sur la nature de l'explication fournie par la neurocriminologie contemporaine. Il est donc crucial de questionner les choix méthodologiques opérés par les neurocriminologues, si l'on veut déterminer si leurs ambitions sont empiriquement justifiées.

Nous expliciterons d'abord les raisons pour lesquelles il est nécessaire de délimiter le champ de la neurocriminologie basée sur la neuroimagerie, et de préciser au mieux les critères de sélection de notre corpus (section 3.1). Contrairement aux études EEG qui reposent globalement sur le même paradigme expérimental, celles utilisant les techniques de neuroimagerie comme l'IRMf recouvrent un champ de recherche nettement plus vaste et hétérogène. Ces différentes voies de recherche concentrent plusieurs centaines études depuis le début des années 1990, aux Etats-Unis principalement, mais aussi en Allemagne (e.g. Poeppl et al. 2013) et en Autriche (e.g. Leutgeb et al. 2015). Elles ont incontestablement nourri le projet neurocriminologique, mais elles peuvent entretenir une grande confusion concernant son objet de recherche et ses ambitions réelles. À notre sens, cette confusion introduit d'importants biais de sélection dans les quelques revues de la littérature déjà existantes. Nous montrerons donc la nécessité de

procéder à une sélection rigoureuse des études sur lesquelles nous appuierons notre réflexion, notamment pour lever toute ambiguïté quant au statut psychiatrique des criminels examinés.

Les résultats de notre revue de la littérature seront présentés sous la forme d'un tableau de synthèse résumant 31 des 62 études que nous avons sélectionnées (section 3.2). Cette revue de la littérature constitue la base empirique de notre analyse conceptuelle et méthodologique, qui se prolongera jusque dans le chapitre suivant. La section 3.3. vise à poser les termes du problème qui nous occupera tout au long de notre analyse épistémologique, à savoir : la question de la relation entre la nature des finalités poursuivies par les neurocriminologues, les choix méthodologiques opérés et le type d'explication qu'ils proposent. Cette réflexion nous donnera l'occasion de préciser ce qui différencie le projet neurocriminologique tel qu'il s'est dessiné depuis les années 1940, des études récentes utilisant la neuroimagerie. Nous montrerons que l'émergence de ces techniques a contribué à réactualiser la démarche typologique et classificatoire qui présidait à l'étude neurophysiologique des criminels, pousuivant l'ambition de définir autant de « profils neurobiologiques » qu'il existe de comportements criminels.

Dans la mesure où notre objectif général est de questionner la continuité entre cette démarche localisatrice et typologique d'une part, et l'organologie et la criminologie du XIX^{e} siècle d'autre part, nous ne pouvons faire l'impasse sur l'analyse des méthodes et des résultats décrits dans la littérature contemporaine. Nous examinerons donc quelques études paradigmatiques du champ, notamment celles publiées par Adrian Raine, l'une des principales figures de la neurocriminologie contemporaine (section 3.4). Nous montrerons dans la section 3.5 que l'intérêt porté à certaines catégories criminelles (en l'occurrence, les pédocriminels et les criminels atteints de psychopathie) nous renseigne sur la logique qui sous-tend la construction de ces études : spécifier et caractériser sur le plan neurobiologique des comportements typiques, c'est-à-dire isoler les anomalies cérébrales associées à une altération du contrôle de soi ou du sens moral, dans l'objectif de reconnaître les indicateurs prédictifs d'un comportement potentiellement dangereux. Une brève description de ces études nous permettra de mener une réflexion plus générale sur les contraintes épistémologiques propres aux stratégies de modélisation des comportements impulsifs (section 3.6). La question portera en particulier sur la pertinence explicative des tâches cognitives utilisées pour opérationnaliser le contrôle de soi : dans quelle mesure ces tâches experimentales permettent-elles de rendre compte du comportement tel qu'il est exercé dans un environnement social complexe ?

3.1. Un domaine de recherche vaste et hétérogène

3.1.1. Délimitation du champ

Si les premières études EEG que nous avons décrites dans le chapitre précédent ont été manifestement oubliées dans l'histoire du champ disciplinaire, les recherches en neuroimagerie bénéficient d'une attention particulière en raison des enjeux posés par leurs éventuelles applications judiciaires. Elles sont menées en parallèle des recherches en neurodroit, ce qui a naturellement nourri les réflexions sur l'apport des neurosciences dans l'explication du comportement criminel. Malgré un intérêt croissant vis-à-vis de ces travaux, il existe pourtant peu de revues de littérature consacrées au champ de la neurocriminologie contemporaine. La plupart sont publiées par des acteurs emblématiques du domaine (e.g. Glenn and Raine 2014), ce qui laisse peu de place à une réflexion épistémologique concernant le pouvoir explicatif des résultats.[106] En outre, la grande majorité des revues de littérature s'en tiennent à recenser la nature des anomalies observées dans les populations criminelles, sans entrer dans le détail des paradigmes expérimentaux.

À ce jour, la plus récente analyse (et certainement la plus instructive) est celle publiée par les sociologues américains M. Fallin, O. Whooley et K.K. Barkeret (2018). Nous présenterons dans le détail les résultats de cette revue de la littérature qui repose sur la sélection de 89 articles (dont 28 revues de littérature). Pour autant, nous pouvons d'ores et déjà préciser que cette analyse ne rend pas compte des spécificités méthodologiques des études en neurocriminologie. Comme nous le montrerons plus loin, il est primordial d'expliciter la logique des stratégies de recherche employées par les neurocriminologues, car le choix des modèles expérimentaux diffère selon les catégories criminelles étudiées. Par ailleurs, bien que l'étude de Fallin et coll. (2018) soit particulièrement pertinente vis-à-vis des questions épistémologiques que nous posons, elle souffre d'un biais de sélection

[106] Voir à ce sujet Raine (1997b, 2014, 2019) ; Pallone and Hennessy (1998) ; Volavka (1999, 2002) ; Brower and Price (2001) ; Bufkin and Luttrell (2005) ; Nordstrom et al. (2011) ; Poldrack et al. (2018).

en faveur des études menées sur des populations psychiatriques, et qui n'appartiennent pas véritablement au champ de la neurocriminologie.[107]

Effectivement, dans la mesure où les limites du champ n'ont jamais été explicitement définies, il existe une grande confusion quant au statut et à l'appartenance disciplinaire des études. Le domaine de la neurocriminologie recouvre par exemple une importante littérature consacrée à la neurobiologie des comportements violents, impulsifs et agressifs, mais ne portant pas sur des populations criminelles.[108] Ces études sont généralement menées sur des individus souffrant de troubles de la personnalité (notamment antisociale) admis dans des institutions psychiatriques, mais n'ayant jamais fait l'objet d'une condamnation criminelle (e.g. Blair 2001). Les revues de littérature consacrées à la neurocriminologie intègrent généralement ces recherches, car la plupart mettent en évidence un lien entre le degré d'anomalies cérébrales et le niveau de violence, d'impulsivité ou d'agressivité des individus.

La question de savoir si elles appartiennent au champ de la neurocriminologie est discutable, dans la mesure où elles portent exclusivement sur des populations psychiatriques. Pour autant il est important de préciser qu'elles tiennent une place privilégiée à l'intérieur du champ, puisque les neurocriminologues s'inscrivent explicitement dans leur continuité. Elles ont par exemple fait l'objet d'une méta-analyse publiée par deux grandes figures de la neurocriminologie contemporaine, et menée dans l'objectif de confirmer le lien entre les comportements antisociaux et les altérations structurelles et fonctionnelles du cerveau (Yang et Raine 2009). Après avoir analysé 43 études utilisant la neuroimagerie sur des patients caractérisés par des comportements violents, impulsifs ou agressifs, les chercheurs ont montré que l'essentiel des anomalies se traduisaient par des déficits fonctionnels et structurels au niveau du cortex préfrontal.[109]

---

[107] Nous prendrons néanmoins le temps de préciser cette analyse dans la prochaine section, car elle a le mérite de fonder une critique épistémologique très instructive en ce qui concerne les ambitions du programme.

[108] Pour quelques revues de la littérature consacrées à ces études, voir notamment Brower et Prince (2001) et Bufkin et Luttrell (2005).

[109] Bien que la grande majorité des études sélectionnées portent sur des patients psychiatriques ou des enfants ou adolescents présentant des troubles du comportement, cette méta-analyse est largement citée dans la littérature relative au champ de la neurocriminologie. Elle constitue en effet la première tentative visant à construire une cartographie des lésions cérébrales associées aux comportements antisociaux.

Globalement, la neurocriminologie s'appuie en grande partie sur les études rapportant la présence d'anomalies cérébrales sur des sujets souffrant de troubles sévères du comportement ou de la personnalité. Les recherches démontrant un lien entre les comportements antisociaux et la présence de lésions cérébrales sont fréquemment citées dans la littérature, à l'instar de la célèbre étude effectuée par Antonio et Hanna Damasio (1994) sur le cas de Phineas Gage.[110] Citons également « *The Vietnam Head Injury Study* », une étude longitudinale menée sur près de 300 vétérans cérébro-lésés, qui a démontré un lien entre la localisation des lésions cérébrales et le niveau de violence ou d'agressivité des sujets (Grafman et al. 1996). Cette étude a d'ailleurs confirmé le rôle du cortex préfrontal dans la régulation des émotions et des comportements antisociaux : elle a montré que les lésions préfrontales étaient associées à des comportements nettement plus violents et agressifs, par rapport aux autres types de lésions. D'autres études ont montré que les antécédents de traumatisme crânien étaient très fréquents dans les populations criminelles ou délinquantes, en particulier chez les individus les plus violents (Lewis et al. 1986 ; Hux et al. 1998 ; Carswell et al. 2004 ; Perron and Howard 2008).

Enfin, le champ de la neurocriminologie a également bénéficié d'une longue série d'études menées sur des populations carcérales, mais utilisant des tests neuropsychologiques pour révéler la présence de dysfonctionnements cérébraux (Langevin et al. 1988, 1989a, 1989b ; Joyal et al. 2007 ; Young et al. 2010, 2012 ; Seruca and Silva 2015).[111] Bien que ces études ne reposent pas sur l'usage de la neuroimagerie, elles ont largement contribué au développement de la neurocriminologie en révélant les déficits neuropsychologiques associés à divers comportements criminels. On peut d'ailleurs considérer qu'elles appartiennent bien au domaine de la neurocriminologie, car elles partagent avec les études de neuroimagerie un certain nombre de stratégies de recherche caractéristiques du champ. Citons par exemple l'étude de Young et ses collaborateurs (2012), dont le titre évocateur – « *A comparison of rape and molest offenders in prison psychiatric treatment* » – illustre parfaitement la démarche

[110] Le cas de Phineas Gage (1823-1860), devenu socialement et moralement inapte en raison d'une lésion préfrontale, est sans doute le plus illustre exemple de « sociopathie acquise ». D'autres cas sont fréquemment mentionnés dans la littérature en neurocriminologie, notamment celui de Charles Whitman (1941-1966), auteur d'une fusillade à l'Université du Texas en 1966 (voir à ce sujet Batts 2009 ; Bufkin & Luttrell 2005).

[111] Ces études permettent d'identifier (indirectement) les régions cérébrales affectées en fonction des conséquences psychologiques qui leur sont associées.

comparative adoptée par la neurocriminologie pour isoler les spécificités propres à chaque type de criminalité. Cette étude avait pour objectif d'identifier « les différences en termes d'antécédents judiciaires, de fonctionnement cérébral et de fonctionnement psychologique » entre des individus condamnés pour viol et des individus condamnés pour agression sexuelle. Elle a montré que les premiers étaient associés à des antécédents de violence et des déficiences intellectuelles nettement plus fréquents, ainsi qu'à « un dysfonctionnement cérébral diffus associé principalement aux régions temporales et préfrontales » (Young et al. 2012 ; p. 1103).

3.1.2. Sélection de la littérature

Dans la mesure où l'objet de notre analyse concerne l'usage de la neuroimagerie sur des populations criminelles, l'ensemble des études que nous venons de mentionner n'ont pas été retenues dans notre revue de la littérature : notre sélection a exclu les recherches menées sur des populations psychiatriques, ainsi que les études de cas privilégiant une approche lésionnelle ou neuropsychologique. L'analyse du champ de la neurocriminologie contemporaine comprend donc les études menées sur des populations criminelles, et défendant un lien entre la présence d'anomalies cérébrales (structurelles et fonctionnelles) et la manifestation d'un comportement criminel.

Précisons également que notre sélection a exclu les études dites longitudinales, conduites sur plusieurs années et menées dans l'objectif d'explorer la relation entre la fréquence de récidive et les données neurobiologiques des individus (e.g. Raine, Venables, and Williams 1990 ; Aharoni et al. 2013). Bien que ces études méritent une analyse approfondie des résultats et des méthodes utilisées, il est nécessaire de les distinguer des études rétrospectives, qui concentrent la majeure partie de la littérature en neurocriminologie. Les études longitudinales font appel à des stratégies qui leur sont propres, lesquelles nécessitent une analyse individualisée en raison de leur grande hétérogénéité méthodologique. C'est pourquoi tous les articles que nous avons retenus dans notre sélection sont des études rétrospectives, c'est-à-dire qu'elles ont été effectuées sur des criminels détenus dans des institutions carcérales, plusieurs années après leur procès.

Cette sélection a été effectuée à partir des bases de données *PubMed*, *Web of Science*, *Scopus* et *Google Scholar*, sur une période allant de 1990 à 2019. Afin de constituer un échantillon représentatif du champ, notre recherche a combiné des mots clés relatifs à l'outil de neuroimagerie utilisé (exemple : « *neuroimaging* » ; « *brain imaging* » ; « *fMRI* » ; « *PET* », etc.) et un ensemble de concepts relatifs à la criminalité ou aux comportements qui lui sont associés (exemple : « *criminals* » ; « *offenders* » ; « *crime* » ; « *violence* » ; « *aggression* » ; « *antisocial* », etc.).[112] Des articles supplémentaires ont été ajoutés à cette sélection sur la base des revues de littérature consacrées à la neurobiologie des comportements antisociaux (Raine 1997b, 2014, 2019 ; Pallone and Hennessy 1998 ; Volavka 1999, 2002 ; Brower and Price 2001 ; Bassarath 2001 ; Bufkin and Luttrell 2005 ; Raine et Yang 2006 ; Yang et Raine 2009). La sélection des articles les plus pertinents a ensuite été établie en fonction du nombre de citations et des articles les plus mentionnés dans les revues de littérature. En somme, trois critères principaux ont été retenus : le nombre de citations, l'utilisation d'une technique de neuroimagerie (anatomique ou fonctionnelle), et la sélection d'une population criminelle.

Il est important de préciser que les recherches menées sur des criminels admis dans des institutions psychiatriques ont été majoritairement exclues de notre sélection, afin d'éviter une trop grande variabilité populationnelle. Bien que ces travaux représentent une part importante de la recherche en neurocriminologie, il paraît essentiel de limiter au mieux les biais de sélection en faveur des criminels présentant des troubles mentaux.[113] Précisons à cet égard que les neurocriminologues cultivent une certaine ambiguïté quant au statut psychiatrique des sujets examinés, et leur origine institutionnelle (psychiatrique ou carcérale) n'est pas toujours mentionnée dans les études. En clair, il existe un certain flou sur la question de savoir si les criminels examinés ont ou non bénéficié d'une irresponsabilité pour cause psychiatrique. Dans la mesure où cette question a des conséquences évidentes sur la relation établie entre le comportement

[112] Plus précisément, cette recherche a été effectuée selon les combinaisons suivantes : ((« neuroimaging » OR « brain imaging » OR « magnetic resonance imaging » OR « fMRI » OR « MRI » OR « positron emission tomography » OR « PET ») AND (« criminal* » OR « offender* » OR « crime » OR « murder* » OR « violen* » OR « aggress* » OR « antisocial » OR « impulsive* »)).

[113] Comme nous le préciserons plus loin, ce biais de sélection est inévitable. Les études menées sur des criminels admis dans des institutions psychiatriques sont finalement majoritaires, vraisemblablement pour des raisons juridiques. La législation allemande interdit par exemple toute investigation empirique sur des détenus.

criminel et la présence d'anomalies cérébrales, nous avons pris soin de privilégier les études menées sur des criminels condamnés et détenus dans des institutions carcérales, ou ayant par la suite été transférés dans des centres psychiatriques de haute-sécurité.

Au total, 62 études ont été retenues, dont la moitié sont détaillées dans le tableau de synthèse figurant dans la section suivante.[114] Toutes démontrent la présence d'une forte proportion d'anomalies cérébrales (structurelles et/ou fonctionnelles) parmi les criminels examinés. La plupart reposent sur l'utilisation de l'IRM fonctionnelle (n = 33) ou anatomique (n = 20), mais aussi sur la TEP (n = 4) ou sur l'utilisation combinée de l'IRM et CT-scan et/ou EEG et/ou SPECT (n = 5). La répartition géographique de ces études est aussi très large, puisqu'elles ont été menées par des équipes neuroscientifiques réparties aux Etats-Unis (n = 25), en Allemagne (n = 19), mais aussi au Canada (n = 3), en Autriche (n = 4), Espagne (n = 2), Angleterre (n = 1), Suède (n = 1), Finlande (n = 2), Danemark (n = 1), Pays-Bas (n = 3), ou encore Taiwan (n = 1).

Si l'on tient compte du nombre de sujets inclus dans chacune de ces études, cette sélection représente un échantillon de plus de 2000 criminels examinés lors de ces deux dernières décennies. En moyenne, les études sont effectuées sur une vingtaine de détenus. La taille des échantillons étudiés est cependant très variable puisqu'elle fluctue entre une petite dizaine (e.g. Müller et al. 2003) et plusieurs centaines, lorsque l'analyse repose sur les bases de données recueillies lors de précédents examens médicaux (e.g. Schiltz et al. 2013). Par ailleurs, la grande majorité des études ont été réalisées en présence d'un groupe contrôle constitué par des sujets issus de la population générale, n'ayant jamais fait l'objet d'une condamnation criminelle, et généralement d'âge équivalent aux criminels examinés. Comme nous le préciserons plus loin, les quelques études qui ne font pas appel à des sujets témoins ont recours à une stratégie que l'on retrouvait déjà dans les publications des années 1940-1980, consistant à comparer les données neurobiologiques associées à deux types de criminalités (e.g. crime de nature sexuelle vs. crime contre la propriété).

Dans un souci de cohérence, et pour garantir la possibilité de comparer les résultats concernant la nature des anomalies observées, nous avons choisi de distinguer

[114] Dans un souci de constituer un échantillon représentatif de l'ensemble des études menées sur le cerveau des criminels, la sélection présentée dans le tableau de synthèse respecte globalement leur répartition selon l'outil de neuroimagerie utilisé : IRMa (n = 13), IRMf (n = 11), TEP (n = 4), IRM et CT-scan (n = 2) et IRM et SPECT (n = 1).

ces études selon leur objet de recherche, c'est-à-dire en fonction des catégories de criminalité les plus souvent étudiées par les neurocriminologues.[115] En parcourant la littérature publiée depuis le début des années 1990, on ne peut en effet pas manquer de relever la faible diversité des populations criminelles étudiées. Plus précisément, si les participants inclus dans ces études correspondent à un large éventail d'infractions (du vol à l'homicide volontaire, en passant par le trafic de drogue), on voit clairement émerger les quelques figures de la dangerosité qui concentrent l'essentiel de la recherche en neurocriminologie contemporaine : les criminels violents (e.g. Raine et al. 1994), les auteurs d'infractions sexuelles (e.g. Walter et al. 2007), et les criminels souffrant d'un trouble de la personnalité psychopathique (e.g. Leutgeb et al. 2015).

Ces trois grandes catégories de criminels définissent l'objet de recherche de la neurocriminologie contemporaine : elles font l'objet d'une caractérisation neurobiologique, dans l'objectif d'isoler les anomalies associées aux diverses formes de criminalité. Si l'on se réfère à l'ensemble des études publiées ces deux dernières décennies, leur répartition est globalement respectée dans notre revue de la littérature, ainsi que dans le tableau de synthèse figurant dans la section suivante. Ainsi au total, la répartition des études de neuroimagerie retenues pour notre analyse se présente comme suit : 22 études menées sur des criminels violents (généralement condamnés pour meurtre), 15 études menées sur des criminels sexuels (poursuivis pour la plupart pour des actes de pédophilie) et 25 études menées sur des criminels psychopathes.

3.2. Revue de la littérature

Une partie de notre revue de la littérature est résumée dans le tableau ci-après, qui présente dans le détail une sélection de 31 études parmi les plus citées dans la littérature en neurocriminologie.

[115] Précisons que ce choix méthodologique contraste avec les quelques revues de littérature consacrées à la neurocriminologie, lesquelles proposent plutôt de distinguer les études selon l'outil de neuroimagerie utilisé. Une telle distinction ne nous semblait pas pertinente dans la mesure où notre premier objectif est de préciser les différentes formes que recouvre l'objet de recherche de la neurocriminologie contemporaine.

Référence de l'étude	Echantillon étudié	Objectif(s)	Technique, méthode	Principaux résultats	Conclusions
Raine et al. (1994) "Selective reductions in prefrontal glucose metabolism in murderers"	- **22 criminels violents** ayant plaidé "NGRI" ou "IST" (20 accusés de meurtre, 2 de tentative de meurtre)[116] - **22 sujets témoins** (âge et genre équivalents) Critères d'exclusion : traitement médicamenteux, maladie psychiatrique, antécédents d'épilepsie ou de traumatisme crânien, d'alcoolisme ou de toxicomanie	Déterminer si les criminels violents ayant plaidé "NGRI ou "IST" sont caractérisés par un dysfonctionnement préfrontal	TEP *Continuous performance task*	1/ **Réduction du métabolisme du glucose** dans le cortex préfrontal (médial et latéral) 2/ Pas de différence observée dans le cortex frontal postérieur, temporal et pariétal	« *[D]eficits localized to the prefrontal cortex may be related to violence in a selected group of offenders*. » (p. 365)
Raine et al. (1997a) "Brain abnormalities in murderers indicated by positron emission tomography"	- **41 criminels violents** accusés de meurtres ayant plaidé "NGRI ou "IST"[117] - **41 sujets témoins** (âge et genre équivalents)	Déterminer si les criminels violents ayant plaidé "NGRI ou "IST" présentent des dysfonctionnements corticaux et sous-corticaux	TEP *Continuous performance task*	1/ **Réduction du métabolisme du glucose** dans le cortex préfrontal, le gyrus pariétal supérieur, le gyrus angulaire gauche et le corps calleux 2/ **Asymétries fonctionnelles** (activité moins importante dans	« *These preliminary findings provide initial indications of a network of abnormal cortical and subcortical brain processes that may predispose to violence in murderers pleading NGRI*. » (p. 495)

Abréviations : TEP tomographie par émission de positons ; IRM imagerie par résonnance magnétique ; SPECT tomographie par émission monophotonique ; CT-scan tomodensitométrie ; VBM morphométrie à base de voxel ; ROI définition des régions d'intérêt ; NGRI *not guilty by reason of insanity* ; IST *incompetent to stand trial*.

[116] Précisons que tous ces criminels n'ont pas bénéficié d'une irresponsabilité pour cause psychiatrique car certains (leur proportion n'est pas mentionnée dans l'étude) ont été condamnés à des peines de prison.

[117] Une grande partie de l'échantillon est issue de Raine et al. (1994).

	Critères d'exclusion : traitement médicamenteux, maladie psychiatrique, antécédents d'épilepsie ou de traumatisme crânien, d'alcoolisme ou de toxicomanie			l'hémisphère gauche) au niveau de l'amygdale, du thalamus et du lobe temporal médial	
<u>Raine et al. (1998a)</u> "Prefrontal glucose deficits in murderers lacking psychosocial deprivation"	- **20 meurtriers ayant souffert de carences psychosociales** : abus physiques et/ou sexuels pendant l'enfance, négligence parentale, antécédents familiaux de criminalité, etc. - **26 meurtriers n'ayant pas souffert de carences psychosociales**[118] Critères d'exclusion : traitement médicamenteux, maladie psychiatrique, antécédents d'épilepsie ou de traumatisme crânien, d'alcoolisme ou de toxicomanie	Déterminer si les criminels ayant bénéficié d'un environnement familial stable présentent des déficits au niveau du cortex préfrontal	TEP *Continuous performance task*	**Hypométabolisme** préfrontal chez les meurtriers n'ayant pas souffert de carences psychosociales : réduction du métabolisme du glucose dans les régions préfrontales	« *These results suggest that murderers lacking psychosocial deficits are characterized by prefrontal deficits. It is argued that among violent offenders without deprived home backgrounds, the "social push" to violence is minimized, and consequently, brain abnormalities provide a relatively stronger predisposition to violence in this group*. » « *One implication of these findings, in conjunction with the previous literature, is that they challenge the assumption often made in forensic settings that violent offenders from normal home backgrounds are biologically "normal", whereas those offenders exposed to abnormal rearing conditions are more likely to be biologically dysfunctional.* » (p. 7)

[118] Une grande partie de l'échantillon est issue de Raine et al. (1997a).

<u>Raine et al. (1998b)</u> "Reduced prefrontal and increased subcortical brain functioning assessed using positron emission tomography in predatory and affective murderers"	- **15 meurtriers "prédateurs"**[119] (« *predatory murderers* ») : individus ayant commis leur crime de sang-froid - **9 meurtriers "affectifs"** (« *affective murderers* ») : individus ayant commis leur crime dans un état de détresse émotionnelle - **41 sujets témoins** (âge et genre équivalents) Critères d'exclusion : traitement médicamenteux, maladie psychiatrique, antécédents d'épilepsie ou de traumatisme crânien, d'alcoolisme ou de toxicomanie	Identifier les mécanismes cérébraux sous-tendant la violence impulsive et affective d'une part, et la violence préméditée et prédatrice d'autre part	TEP *Continuous performance task*	1/ Chez les meurtriers "prédateurs" : activité du cortex préfrontal équivalente à celle des sujets témoins, **hyperactivité sous-corticale** au niveau de l'hémisphère droit 2/ Chez les meurtriers "affectifs" : **déficit fonctionnel** au niveau du cortex préfrontal gauche, **hyperactivité sous-corticale** au niveau de l'hémisphère droit	« *Results support the hypothesis that emotional, unplanned impulsive murderers are less able to regulate and control aggressive impulses generated from subcortical structures due to deficient prefrontal regulation. It is hypothesized that excessive subcortical activity predisposes to aggressive behaviour, but that while predatory murderers have sufficiently good prefrontal functioning to regulate these aggressive impulses, the affective murderers lack such prefrontal control over emotion regulation.* » (p. 319)
<u>Aigner et al. (2000)</u> "Brain abnormalities and violent behavior"	**82 criminels (dont 50 délinquants sexuels)** admis dans une prison de haute sécurité, répartis en un « **groupe très violent** » » (e.g. homicide, viol) et un « **groupe peu ou pas**	Déterminer s'il existe une association entre le comportement violent et la présence d'anomalies cérébrales	IRM	1/ **48.8% d'anomalies** parmi les 82 criminels examinés 2/ 46% d'anomalies parmi les criminels sexuels	« *[T]here is an association between unspecified brain anomalies and high violent behaviour in the whole sample as well as in the sex offender group. The basis for diagnosis of such dramatic forms of behavioral*

[119] Cette classification correspond à des critères de psychiatrie légale définis par Meloy (1988 ; 1997) et Dodge (1991). La catégorie des « meurtriers affectifs » se réfère à des agressions impulsives et non préméditées, et la catégorie des « meurtriers prédateurs » à un acte délibéré et de sang-froid, survenant le plus souvent chez des sujets souffrant d'un trouble de la personnalité psychopathique.

	violent » (e.g. exhibitionnisme, tentative de viol)[120] Critères d'exclusion : maladie neurologique, trouble psychotique, troubles mentaux organiques sévères			3/ **65,5% d'anomalies** parmi les criminels « très violents » 4/ 16,6% d'anomalies parmi les criminels « peu ou pas violents »	*disorders should therefore include modern brain imaging techniques such as magnetic resonance imaging in order to allow patients appropriate treatment and legal protection.* » (p. 63)
<u>Soderstrom et al. (2000)</u> "Reduced regional cerebral blood flow in non-psychotic violent offenders"	- **21 criminels** (16-51 ans) **auteurs de crimes violents et impulsifs** (e.g. meurtre, viol) admis dans plusieurs centres de haute sécurité en Suède - **11 sujets témoins** (31-50 ans) Critères d'exclusion : schizophrénie, trouble bipolaire ou psychotique, antécédents de traumatisme crânien ou de maladie neurologique	Déterminer si les criminels violents et impulsifs sont caractérisés par une diminution du flux sanguin cérébral régional (rCBF) au niveau des régions frontales et/ou temporales	SPECT et IRM (ROI)	1/ **Hypoperfusion** dans les régions temporales et/ou frontales chez 16 des 21 criminels 2/ **Réduction significative du flux sanguin** dans le gyrus angulaire droit, gyrus temporal médian droit, hippocampe (bilatéralement) et matière blanche frontale gauche 3/ **Augmentation du flux sanguin** dans le parietal association cortex (bilatéralement)	« *The MRI scans revealed no consistent pattern of major structural pathology but showed a higher frequency of nonspecific minor changes than is generally expected in the age range of our index subjects. SPECT demonstrated functional pathology by means of both visual and quantitative assessments*. » « *SPECT imaging has, as no other diagnostic tool used in forensic psychiatry today, demonstrated a consistent pattern of aberration from the normal in non-psychotic perpetrators of crimes of extreme aggression and lack of impulse control. […] Should the brainimaging findings of fronto-temporal abnormalities in violent offenders be further replicated and shown to remain stable over time, as well as being closely related to violent criminal behavior, research on functional*

[120] L'étude ne precise pas le nombre d'individus répartis dans les deux groupes.

					consequences and possible treatment strategies would be of great interest to society. » (p. 40)
<u>Tiihonen et al. (2008)</u> "Brain anatomy of persistent violent offenders: more rather than less"	- **26 criminels violents et récidivistes** souffrant d'un trouble de la personnalité antisociale et de toxicomanie - **25 sujets témoins** (d'âge équivalent) Critères d'exclusion : antécédents de psychose ou de schizophrénie	Déterminer si les criminels violents, récidivistes et antisociaux sont associés à « une structure cérébrale anormale »	IRM (VBM)	1/ **Augmentation du volume de matière blanche** dans les lobes pariétaux et occipitaux et dans le cervelet gauche 2/ **Augmentation du volume de matière grise** dans le cervelet droit 3/ **Atrophie bilatérale** dans le gyrus postcentral, le cortex fronto-polaire et le cortex orbitofrontal	« *In conclusion, the results of the present study indicate that persistent violent offenders with ASPD present markedly abnormal white and grey matter regional brain volumes as compared with healthy men. Substantially larger regional volumes in the posterior brain areas were observed among violent offenders with a history of antisocial behavior going back to childhood. These marked differences in brain structure may reflect atypical neurodevelopmental processes that are involved in the initiation and maintenance of persistent aggressive behavior.* » (p. 210)
<u>Bertsch et al. (2013)</u> "Brain volumes differ between diagnostic groups of violent criminal offenders"	**25 criminels condamnés pour des crimes violents** (incluant meurtre, viol, braquage) admis dans plusieurs institutions pénales et psychiatriques de haute sécurité : - **13 criminels souffrant d'un trouble de la personnalité borderline** (ASPD-BPD) - **12 criminels souffrant d'un trouble de la personnalité psychopathique** (ASPD-PP)	Isoler les différences structurales entre les criminels violents souffrant de psychopathie, et ceux souffrant d'un trouble de la personnalité borderline	IRM (VBM et ROI)	**Réduction du volume de matière grise** au niveau du cortex frontal et occipital : 1/ Chez les criminels borderline : réduction significative au niveau du cortex préfrontal ventromédial, orbitofrontal et du cortex temporal (impliqués dans la régulation émotionnelle et l'interprétation des motivations d'autrui)	« *The findings of this exploratory study thus indicate that the comparison of subgroups of offenders may be important to gain more information about specific deficits and underlying brain mechanisms, which might then also be relevant for selecting adequate treatments in forensic facilities*. » (p. 604)

	14 sujets témoins (âge et QI équivalents) Critère d'exclusion : antécédents de troubles psychotique ou bipolaire, dépression sévère			2/ Chez les criminels psychopathes : réduction significative au niveau du cortex préfrontal dorsomédian, cortex cingulaire postérieur, précuneus et gyrus postcentral (impliqués dans la reconnaissance des émotions, la formation des jugements moraux et les processus d'autoréflexion)	
<u>Schiltz et al. (2013)</u> "High prevalence of brain pathology in violent prisoners: a qualitative CT and MRI scan study"	- **287 détenus dont 162 condamnés pour des crimes violents** (incluant pour la plupart des homicides volontaires) et **125 condamnés pour des actes de délinquance non-violents** (vol, fraude, trafic de drogue) - **52 sujets témoins** (d'âge équivalent) Aucun critère d'exclusion	Déterminer la fréquence et l'étendue des anomalies cérébrales dans un large échantillon de criminels violents et ne présentant aucun antécédent de trouble neuropsychiatrique	CT-scan et IRM (ROI)[121]	Chez les criminels violents : **forte proportion d'anomalies morphologiques** dans le cortex frontal/pariétal, temporal médial, troisième ventricule et ventricule gauche	« *The remarkable prevalence of brain pathology in convicted violent prisoners detectable by neuroradiological routine assessment not only highlights the importance of frontal and temporal structures in the control of social, and specifically of violent behaviour, but also raises questions on the legal culpability of violent offenders with brain abnormalities. The high proportion of undetected presence of structural brain damage emphasizes the need that in violent criminals, the comprehensive routine neuropsychiatric assessment usually performed in routine forensic psychiatric expertises should be complemented with brain imaging*. » (p. 607)

[121] Précisons que les scans cérébraux sont issus des bases de données de plusieurs institutions pénales allemandes, réalisés dans le cadre d'une expertise médico-légale ou d'un examen médical de routine, entre 1993 et 2007.

<u>Cope et al. (2014)</u> "Abnormal brain structure in youth who commit homicide"	**155 jeunes délinquants ou criminels** admis dans un centre de détention pour mineurs de haute sécurité (Nouveau-Mexique)[122] : - **20 adolescents condamnés pour homicide** (volontaire ou involontaire) - **135 adolescents condamnés pour d'autres infractions** (violentes : agression, viol, etc. ou non violentes : vol, fraude, possession de drogue) Groupe contrôle : - **20 adolescents condamnés pour d'autres infractions** et matchant sur le plan démographique et psychométrique - **21 sujets témoins** (d'âge équivalent) Critères d'exclusion : trouble bipolaire ou psychotique	Déterminer si les adolescents condamnés pour homicide sont caractérisés par des anomalies de matière grise dans les régions paralimbiques	IRM (VBM)	Chez les adolescents condamnés pour homicide : **réduction du volume de matière grise** dans les lobes temporaux (médial et latéral), hippocampe et insula postérieur	« *Our results indicate that brain structural differences may help identify those at the highest risk for committing serious violent offenses*. » (p. 800) « *This work has important implications for the development of biomarkers for youth on a high-risk trajectory of antisocial behavior and it also has important implications for developing novel interventions to help reduce the enormous fiscal and emotional toll homicidal behavior has on society*. » (p. 806)
<u>Bueso-Izquierdo et al. (2016)</u>	41 délinquants ou criminels recrutés dans un centre de	Comparer le fonctionnement cérébral des auteurs de violence	IRMf Stimuli visuels neutres et émotionnels (images de	1/ Pendant la présentation des images IVC (par rapport aux images neutres) : **augmentation de l'activité**	« *In sum, our results have shown that batterers have different brain functioning, as compared to other criminals, when they observe both*

[122] Précisons que cette étude porte sur un sous-échantillon étudié par Ermer et al. (2013), qui comprenait au total 191 jeunes délinquants.

"Are batterers different from other criminals? An fMRI study"	réinsertion sociale à Grenade (Espagne) : - **21 auteurs de violence conjugale** - **20 criminels ayant été condamnés pour d'autres infractions** Critères d'exclusion : antécédents de troubles psychologiques ou de la personnalité sévères, traumatisme crânien, maladies neurologiques, maladies infectieuses, antécédents de toxicomanie	conjugale par rapport à d'autres criminels	violence conjugale, IVC ; images violentes, IV)	du cortex cingulaire antérieur et postérieur 2/ Pendant la présentation des images IV (par rapport aux images neutres) : **augmentation de l'activité** de l'insula et des régions pariétales 3/ Pendant la présentation des images IVC et IV : **Augmentation de l'activité** du cortex préfrontal médian **Réduction de l'activité** dans le cortex préfrontal supérieur	*intimate partner violence and general violence images as compared to neutral images. Future studies should replicate our results in batterers who have committed more severe offenses.* » (p. 862)
<u>Witzel et al. (2016)</u> "Increased frequency of brain pathology in inmates of a high-security forensic institution: a qualitative CT and MRI scan study"	**148 criminels** détenus dans un hôpital psychiatrique de haute sécurité en Allemagne : - 51 individus condamnés pour des crimes de nature sexuelle (abus sexuels sur mineurs) - 80 individus condamnés pour des crimes violents (meurtre, homicide, voies de fait graves) - 9 individus condamnés pour des actes de pyromanie	1/ Mesurer la fréquence et l'étendue des pathologies cérébrales dans un large échantillon de criminels n'ayant jamais souffert de lésions cérébrales 2/ Déterminer si les pathologies cérébrales sont plus fréquentes chez les criminels que chez des individus non-criminels	IRM et CT-scan[124]	1/ 40 à 70% des criminels (selon les types de condamnation) présentent des **signes de pathologie cérébrale** (contre 8% chez les sujets témoins), y compris après réajustement des données en fonction de l'âge des individus 2/ Nature des anomalies observées : - **atrophie** du cortex frontal/pariétal et/ou temporal	« *The astonishingly high prevalence of brain pathology in institutionalized inmates of a high-security mental health institution who previously had not been considered to be suffering from an organic brain syndrome raises questions on whether such neuroradiological assessment might be considered as a routine procedure in newly admitted patients. Furthermore, it highlights that organic changes, detectable under clinical routine conditions, may play a role in the*

[124] Précisons que les scans cérébraux proviennent d'une base de données recueillies entre 1992 et 2008 lors d'examens médicaux réalisés pendant l'incarcération des sujets.

	- 8 criminels non-violents (vols, fraude, trafic de drogue) Diagnostics à l'admission : schizophrénie, trouble délirant, trouble schizo-affectif (n = 50), troubles de la personnalité (n = 41), pédophilie (n = 36), retard mental (n = 22), troubles du comportement pendant l'adolescence (n = 12), toxicomanie et alcoolisme (n = 55)[123] **52 sujets témoins** Pas de critère d'exclusion			- **asymétries, malformation ou élargissement** des ventricules latéraux et du troisième ventricule - élargissement des sillons corticaux et de la scissure interhémisphérique au niveau du lobe frontal - élargissement des sillons corticaux et de l'espace sous-arachnoïdien au niveau du lobe temporal - malformation de l'hippocampe 3/ Pas de relation statistiquement significative entre le type de condamnation criminelle et la nature ou la fréquence des anomalies observées	*development of legally relevant behavioral disturbances which might be underestimated*. » (p. 533) « *With respect to the future treatment of perpetrators, it appears of paramount importance to diagnose neurological impairments, for mental incapacity due to neurological defects may prove much more resistant to psychopharmacologic and psychotherapeutic treatment as compared to other psychiatric disease. Therefore, if identified early in the process of forensic treatment, patients suffering from such neurological impairments might be assigned to a treatment context which is more committed toward optimizing social and circumstantial factors in order to prevent further delinquent or even violent behavior of the patients*. » (p. 539)
<u>Da Cunha-Bang et al. (2017)</u> "Violent offenders respond to provocations with high amygdala and striatal reactivity"	- **18 criminels condamnés pour des crimes violents** (e.g. homicide, viol, voies de fait graves), incarcérés dans plusieurs prisons du Danemark, et présentant divers troubles de la personnalité (e.g. antisociale, borderline, schizoïde)	Examiner les réponses neurales des criminels à des stimuli de provocation	IRMf Paradigme d'agression par soustraction de point : tâche de récompense monétaire qui mesure la réaction agressive à la frustration	1/ **Réactivité cérébrale anormalement élevée** au niveau de l'amygdale et du striatum 2/ **Réduction de la connectivité fonctionnelle** entre ces deux régions	« *These data provide novel evidence of aberrant brain function in a unique cohort of individuals with a history of extremely violent behavior. The findings suggest that an exaggerated neurobiological sensitivity to provocations or*

[123] Les auteurs précisent que certains criminels ont fait l'objet de diagnostics multiples, ce qui explique leur surreprésentation par rapport à l'échantillon étudié.

	- **26 sujets témoins** (d'âge équivalent) Critères d'exclusion : antécédents de troubles psychiatriques, maladies neurologiques, traumatisme crânien, consommation de médicaments psychotropes			(amygdale et striatum) et le cortex préfrontal 3/ Corrélation positive entre la réactivité de l'amygdale et du striatum d'une part, et l'agressivité des sujets d'autre part	*frustrations and lack of prefrontal control are key features of pathological aggression. Prevention and treatment of aggressive behaviors would benefit from interventions targeting this type of vulnerability*. »

Référence de l'étude	Echantillon étudié	Objectif(s)	Technique, méthode	Principaux résultats	Conclusions
Walter et al. (2007) "Pedophilia is linked to reduced activation in hypothalamus and lateral prefrontal cortex during visual erotic stimulation"	- **13 pédophiles** coupables d'abus sexuels sur mineurs (< 10 ans) admis dans un hôpital psychiatrique en Allemagne - **14 sujets témoins** Critères d'exclusion : sujets sous traitement, maladie psychiatrique ou neurologique, antécédents d'alcoolisme ou de toxicomanie	Identifier les corrélats neuronaux des déficits d'excitation sexuelle et émotionnelle chez les pédophiles	IRMf Stimuli visuels émotionnels / neutres / sexuellement explicites (adultes uniquement) Tâche impliquant un contrôle attentionnel (non précisée)	**Réduction de l'activité** de l'hypothalamus et du cortex préfrontal dorsolatéral pendant la présentation des images sexuellement explicites	« *Hypothesized regions relevant for processing of erotic stimuli in healthy individuals showed reduced activations during visual erotic stimulation in pedophilic patients. This suggests an impaired recruitment of key structures that might contribute to an altered sexual interest of these patients toward adults.* »
Sartorius et al. (2008) "Abnormal amygdala activation profile in pedophilia"	- **10 pédophiles homosexuels** coupables d'abus sexuels sur mineurs, admis dans plusieurs instituts psychiatriques de haute sécurité en Allemagne - **10 sujets témoins hétérosexuels** Critères d'exclusion : antécédents de dépression, troubles obsessionnels compulsifs, troubles neurologiques ou psychiatriques, toxicomanie	Isoler les différences d'activité de l'amygdale entre des pédophiles et un groupe contrôle pendant la présentation d'images sexuellement implicites	IRMf Stimuli visuels sexuellement implicites (femmes et enfants en maillot de bain) et neutres (formes géométriques) Tâche impliquant un contrôle attentionnel (*oddball task*)	1/ Chez les sujets témoins : augmentation de l'activité de l'amygdale pendant la présentation des images de femmes 2/ Chez les pédophiles : augmentation de l'activité de l'amygdale pendant la présentation des images d'enfants (en particulier pour les images de jeunes garçons)	« *Our data show enhanced activation to children's pictures even in an incidental context and suggest the provocative hypothesis that a normally present mechanism for reduced emotional arousal for children relative to adults is reversed in pedophilia, suggesting a neural substrate associated with deviant sexual preference in this condition*. » (p. 271)

<u>Habermeyer et al. (2013)</u> "Response inhibition in pedophilia: an fMRI pilot study"	**11 pédophiles** admis dans un hôpital psychiatrique à Bâle (Suisse) : - 5 individus coupables d'abus sexuels sur mineurs - 6 individus coupables de détention d'images pédopornographiques **7 sujets témoins**	Identifier les corrélats neuronaux de l'inhibition comportementale chez des sujets pédophiles	IRMf *Go/NoGo task* (alternance de lettres)	1/ Temps de réaction plus lent et discrimination visuelle moins précise 2/ Augmentation de l'activité dans les régions préfrontales et pariétales pendant la condition No-Go 3/ Augmentation de l'activité du cortex cingulaire antérieur gauche, précuneus et gyrus angulaire pendant la condition Go 4/ Désactivation atténuée du précuneus et gyrus angulaire (réseau du mode par défaut) pendant la condition No-Go	« *Our pilot study therefore indicates that in pedophilia, cognitive control might also be impaired due to a failure to stay focused on the task rather than representing a result of frontal pathology. Further studies on the interplay between attentional and frontal control networks in relation to the default network might be a promising approach to further our understanding of the neurobiology of pedophilia.* » (p. 236)
<u>Poeppl et al. (2013)</u> "Association between brain structure and phenotypic characteristics in pedophilia"	20 criminels admis dans trois instituts psychiatriques de haute sécurité en Allemagne : - **9 pédophiles** (7 homosexuels, 2 hétérosexuels) coupables de plusieurs abus sexuels sur mineurs - **11 criminels non sexuels** (hétérosexuels)	Déterminer si les pédophiles sont caractérisés par des anomalies structurelles au niveau des régions préfrontales et limbiques, et si ces anomalies sont associées à leurs caractéristiques phénotypiques	IRM (VBM)	1/ **Réduction du volume de matière grise** au niveau de l'amygdale droite chez tous les pédophiles 2/ **Réduction du volume de matière grise** au niveau du cortex préfrontal dorsolatéral gauche chez les pédophiles récidivistes 3/ **Réduction du volume de matière grise** dans le cortex orbitofrontal et le gyrus angulaire	« *These neuroanatomical deficits might account for distinct features of pedophilia such as strong pedophilic fixation, sexual recidivism, and low victim age. The observed relationship of specific grey matter reductions with certain phenotypic characteristics might account for the heterogeneous findings in previous neuroimaging of pedophilia. Furthermore, it suggests that neuroanatomical abnormalities in pedophilia are of a dimensional rather than a categorical nature, supporting the*

	Critères d'exclusion : sujets sous traitement, maladie psychiatrique ou neurologique, antécédents d'alcoolisme ou de toxicomanie			chez les pédophiles ayant agressé de très jeunes victimes (< 7 ans)	*notion of a multifaceted disorder*. » (p. 7)
<u>Chen et al. (2016)</u> "Abnormal white matter integrity in rapists as indicated by diffusion tensor imaging"	- **15 délinquants sexuels condamnés pour viol** et détenus dans une prison de Taiwan - **15 sujets témoins** (âge, genre et niveau d'éducation équivalents) Critères d'exclusion : antécédents de maladie psychiatrique ou neurologique, sujets suivant un traitement psychiatrique, antécédents d'alcoolisme et de toxicomanie	Déterminer si les violeurs présentent des anomalies de l'intégrité (corticale et sous-corticale) de la matière blanche	IRM de diffusion	1/ Augmentation significative de l'anisotropie fractionnelle de la matière blanche au niveau du pole frontal, gyrus angulaire, cortex cingulaire postérieur, cortex occipital latéral, dans la capsule interne du thalamus, noyau caudé et globus pallidus 2/ Réduction significative de l'anisotropie fractionnelle au niveau du cortex cingulaire postérieur et du faisceau fronto-occipital inférieur	« *Findings indicate abnormalities in white matter connectivity in brain regions involved in reward/motivation and moral judgment, which may predispose rapists to be both over-responsive to sexual reward stimuli and also to make inappropriate moral decisions*. »

Référence de l'étude	Echantillon étudié	Objectifs	Technique	Principaux résultats	Conclusions
Kiehl et al. (2001) "Limbic abnormalities in affective processing by criminal psychopaths as revealed by functional magnetic resonance imaging"	- 16 criminels admis dans une prison de haute sécurité en Colombie-Britanique (Canada) : **8 criminels psychopathes** (PCL-R : 28-36) **8 criminels non-psychopathes** (PCL-R : 8-23) - **8 sujets témoins** (d'âge, statut socioéconomique des parents et QI équivalents) Pas de critères d'exclusion	Identifier les systèmes neuronaux sous-tendant le traitement émotionnel chez les criminels psychopathes	IRMf, (ROI) *Affective memory task*	Chez les criminels psychopathes (par rapport aux deux autres groupes) : 1/ **Réduction significative de l'activité** de l'amygdale, formation hippocampique, gyrus parahippocampique, striatum ventral, et gyrus cingulaire antérieur et postérieur 2/ **Hyperactivité** au niveau du cortex fronto-temporal bilateral	« *These data suggest that the affective abnormalities so often observed in psychopathic offenders may be linked to deficient or weakened input from limbic structures.* » « *In summary, we have shown that processing of affective stimuli is associated with less limbic activation in criminal psychopaths than in criminal nonpsychopaths and noncriminal control participants. We have also shown that psychopathic offenders appear to use alternative neural systems to process affective stimuli. These findings support and extend previous lesion-based observations in psychopaths and provide in vivo visualization of the neural processes that may underlie the affective anomalies that clinicians have described in criminal psychopaths*. » (p. 683)
Müller et al. (2003) "Abnormalities in emotion processing within cortical and subcortical regions in criminal psychopaths: evidence from a functional magnetic resonance	- **6 criminels psychopathes** (PCL-R : 34-40) admis dans un hôpital psychiatrique de haute sécurité en Allemagne	Identifier les réseaux neuronaux associés au traitement émotionnel chez des criminels psychopathes	IRMf (ROI) Stimuli émotionnels à valence positive (e.g. photographie d'un couple heureux) ou négative (photographie d'un visage menaçant)	1/ En réaction aux stimuli négatifs : **hyperactivité** de l'hémisphère droit au niveau du cortex préfrontal et de l'amygdale, **hypoactivité** de l'hémisphère droit au niveau du cortex cingulaire subgénual et du gyrus temporal, hypoactivité de l'hémisphère gauche au niveau du	« *These findings underline the hypotheses that psychopathy is neurobiologically reflected by dysregulation and disturbed functional connectivity of emotion-related brain regions. These findings may be interpreted within a framework including prefrontal*

imaging study using pictures with emotional content"	- **6 sujets témoins** (d'âge equivalent, PCL-R ≤ 10) Critères d'exclusion : troubles neuropsychiatriques et autres troubles de la personnalité, antécédents de traumatisme crânien, dépression, consommation de substances durant les six derniers mois			cortex cingulaire dorsal et du gyrus parahippocampal 2/ En réaction aux stimuli positifs : hyperactivité de l'hémisphère gauche dans les régions orbitofrontales, réduction de l'activité dans le lobe frontal et temporal médian	*regions that provide top-down control to and regulate bottom-up signals from limbic areas*. » (p. 152)
Raine et al. (2004) "Hippocampal structural asymmetry in unsuccessful psychopaths"	- **16 "unsuccessful psychopaths"** : criminels psychopathes (PCL-R : 23-40) ayant été condamnés pour leur crime - **12 "successful psychopaths"** : criminels psychopathes (PCL-R : 23-31) ayant échappé à des poursuites judiciaires[125] **23 sujets témoins** (âge équivalent, PCL-R : 2-14) Critère d'exclusion : antécédents d'épilepsie	Déterminer si les asymétries fonctionnelles de l'hippocampe observées chez les criminels sont caractéristiques des "unsuccessful psychopaths"	IRM	**Forte asymétrie structurelle** de l'hippocampe antérieur chez les « unsuccessful psychopaths » (par rapport aux deux autres catégories) : **réduction significative du volume** de l'hippocampe gauche	« *Atypical anterior hippocampal asymmetries in unsuccessful psychopaths may reflect an underlying neurodevelopmental abnormality that disrupts hippocampal-prefrontal circuitry, resulting in affect dysregulation, poor contextual fear conditioning, and insensitivity to cues predicting capture*. » (p. 186) « *[O]ur results provide further support for the notion that successful and unsuccessful psychopaths are distinct subgroups with different autonomic, cognitive, and neuroanatomic deficits, provide initial findings from brain imaging research on these two subgroups,*

[125] Bien que l'étude de Raine et ses collaborateurs (2004) n'en fasse pas mention, nous pouvons préciser que la distinction entre les « unsuccessful psychopaths » et les « successful psychopaths » remonte à la fin des années 1970 dans un article publié par la psychologue américaine Cathy Spatz Widom (1977). Cet article propose d'isoler les caractéristiques psychologiques, sociales et démographiques des « successful psychopaths » (ou « nonincarcerated psychopaths »), c'est-à-dire les psychopathes ayant commis un acte immoral mais qui ont réussi à échapper à des poursuites judiciaires.

					and more broadly support prior imaging, EEG, and neurocognitive research implicating the hippocampus–mesial temporal cortex in antisocial groups. » (p. 190)
Kiehl et al. (2004) "Temporal lobe abnormalities in semantic processing by criminal psychopaths as revealed by functional magnetic resonance imaging"	- **8 criminels psychopathes** (PCL-R ≥ 28) détenus dans une prison de haute sécurité en Colombie-Britannique (Canada) - **8 sujets témoins** (âge, statut socioéconomique des parents, niveau d'éducation et QI équivalents) Critères d'exclusion : antécédents de traumatisme crânien ou trouble psychotique, toxicomanie	Etudier l'architecture neurale sous-tendant le traitement lexico-sémantique des criminels psychopathes	IRMf *Lexical decision task* : alternance de mots concrets ou abstraits	1/ Entre les criminels psychopathes et le groupe contrôle : similitudes dans le pattern d'activation associé au traitement des mots concrets et abstraits 2/ Chez les criminels psychopathes : **déficits d'activité** dans le gyrus temporal et le cortex environnant pendant le traitement des mots abstraits, et faible performance comportementale (temps de réaction plus lent)	« *These data may have implications for the treatment and management of psychopaths. If part of the psychopathic syndrome is due to failing to appropriately integrate semantically abstract information because of a deficit or limitation in processing the abstract representations of stimuli, then training and treatment protocols that specifically address these deficits may be better suited for effectively treating the condition. Moreover, perhaps treatment and management of psychopathic individuals would be improved if these programs emphasized concepts in more concrete forms.* » (p. 38)
Birbaumer et al. (2005) "Deficient fear conditioning in psychopathy: a functional magnetic resonance imaging study"	- **10 criminels psychopathes** (PCL-R : 15-30) - **10 sujets témoins** (âge et niveau d'éducation équivalents, PCL-R ≤ 2)	Etudier les réactions cérébrales lors d'un conditionnement à la peur chez des criminels psychopathes par rapport à un groupe contrôle	IRMf (ROI) *Aversive differential delay conditioning*	Chez les criminels psychopathes, contrairement aux sujets témoins : 1/ Pas d'activité significative dans le circuit préfronto-limbique (amygdale, cortex orbitofrontal, insula, cortex cingulaire antérieur)	« *This inability to emotionally relate neutral and biologically significant events rather than a lack of response to biologically relevant stimuli in general may be at the core of psychopathy. The absence of emotional associative ability ("cold emotion") may be crucial for psychopathic behavior and should*

	Critères d'exclusion : antécédents de troubles mentaux, traumatisme crânien, consommation de médicaments psychotropes			2/ Faible conductance cutanée et valence émotionnelle	*be targeted by behavioral and somatic interventions*. » (p. 805)
Yang et al. (2005) "Volume reduction in prefrontal gray matter in unsuccessful criminal psychopaths"	- **16 "unsuccessful psychopaths"** : criminels psychopathes ayant été condamnés pour leur crime (PCL–R : 23-40) - **13 "successful psychopaths"** : criminels psychopathes ayant échappé à des poursuites judiciaires (PCL–R : 23-31) - **23 sujets témoins** (PCL–R ≤ 14) Aucun critère d'exclusion	Déterminer si les anomalies structurales au niveau du cortex préfrontal sont spécifiques aux "unsuccessful psychopaths"	IRM	1/ **Corrélation négative** entre le **score de psychopathie** (PCL-R) et le **volume de matière grise** préfrontale 2/ **Réduction significative (22.3%) du volume de matière grise** du cortex prefrontal chez les "unsuccessful psychopaths"	« *These results demonstrating for the first time a prefrontal structural deficit in community psychopaths provide partial support for a prefrontal theory of psychopathy but highlight an important difference between successful and unsuccessful psychopaths.* » (p. 1103)
Yang et al. (2010) "Morphological alterations in the prefrontal cortex and the amygdala in unsuccessful psychopaths"	- **16 "unsuccessful psychopaths"** (PCL–R ≥ 23) - **10 "successful psychopaths"** (PCL–R ≥ 23) - **27 sujets témoins** (PCL–R ≤ 14) Aucun critère d'exclusion	Comparer les anomalies morphologiques au niveau des régions corticales (cortex préfrontal) et sous-corticales (amygdale) entre les "unsuccessful psychopaths" et les "successful psychopaths"	IRM (ROI)	Chez les "unsuccessful" : **réduction significative du volume de matière grise et de l'épaisseur corticale** au niveau du cortex frontal médian, cortex orbitofrontal et de l'amygdale	« *This study provides the first evidence of greater prefrontal and amygdala structural deficits in unsuccessful psychopaths, which may predispose them to poor behavioral control and impaired decision-making, thus making them more prone to convictions*. » (p. 546) « *[T]his study supports previous hypotheses that disturbances in a complex neural network including the prefrontal cortex and the amygdala may contribute to the underlying neuroanatomical*

					pathology of psychopathy and further suggests that such disturbances may be associated particularly with criminal arrests and convictions. » (p. 553)
Boccardi et al. (2010) "Abnormal hippocampal shape in offenders with psychopathy"	26 criminels violents et récidivistes admis dans un hôpital médico-légal psychiatrique dans l'attente de leur procès : - **12 criminels présentant un haut score de psychopathie** (PCL–R ≥ 30) - **14 criminels présentant un score moyen de psychopathie** (PCL–R ≤ 28.7) **25 sujets témoins** (âge equivalent) Critères d'exclusion : antécédents de psychose, traumatisme cérébral ou autre trouble de la personnalité	Comparer le volume et la morphologie de l'hippocampe entre un groupe contrôle et un groupe de criminels violents et récidivistes présentant différents degrés de psychopathie	IRM	1/ Chez les deux groupes de criminels : **élargissement anormal** des bords latéraux de l'hippocampe (droite et gauche) 2/ Chez les criminels présentant un haut score de psychopathie (par rapport aux deux autres groupes) : **dépression significative** le long de l'axe longitudinal de l'hippocampe	« *These results indicate that habitually violent offenders exhibit a specific abnormal hippocampal morphology, in the absence of total gray matter volume changes that may relate to different autonomic modulation and abnormal fear-conditioning.* » (pp. 438-439)
Harenski et al. (2010) "Aberrant neural processing of moral violations in criminal psychopaths"	64 criminels admis dans une prison nord-américaine à sécurité moyenne dont : - **16 criminels psychopathes** (PCL–R ≥ 30)	Déterminer si les systèmes neuronaux sous-tendant les décisions morales diffèrent entre les psychopathes et les non-psychopathes	IRMf (ROI) Tâche invitant les sujets à évaluer le caractère moralement acceptable d'une situation : présentation d'une série	1/ Chez les psychopathes, pendant l'évaluation des situations moralement répréhensibles : « **activité atypique** » dans certaines régions impliquées dans les décisions morales (cortex préfrontal	« *These results demonstrate neural abnormalities in moral picture processing in psychopaths, and indicate that psychopaths utilize different brain regions when making*

	- **16 criminels non-psychopathes** (PCL–R : 7-18) d'âge et QI équivalents aux criminels psychopathes - **32 criminels témoins** Critères d'exclusion : antécédents de crise épileptique, traumatisme crânien, trouble psychotique		de vignettes décrivant une situation moralement répréhensible (e.g. une agression physique), une situation dépourvue de contenu moral (e.g. une main mutilée) et une situation neutre (e.g. une conversation entre deux individus)	ventromédian et cortex temporal antérieur) 2/ Chez les non-psychopathes : corrélation positive entre l'activité de l'amygdale et la sévérité de la sanction attribuée à l'action immorale 3/ Chez les psychopathes : corrélation négative entre la sévérité de la sanction attribuée à l'action immorale et l'activité du cortex temporal postérieur	*moral decisions than do nonpsychopaths*. »
<u>Gregory et al. (2012)</u> "The antisocial brain : psychopathy matters"	**44 criminels auteurs de crimes violents** (meurtre, tentative de meurtre, viol, coups et blessures graves), dont : - 17 criminels souffrant d'un **trouble de la personnalité psychopathique** (ASPDP) - 27 criminels souffrant d'un **trouble de la personnalité antisociale** (ASPD-P) **22 sujets témoins** (âge et niveau d'éducation équivalents) Critères d'exclusion : antécédents de maladie neurologique, traumatisme	Isoler les différences de structure de matière grise entre les criminels violents souffrant de psychopathie, et les criminels violents souffrant seulement d'un trouble de la personnalité antisociale	IRM (VBM, ROI)	1/ Chez les criminels psychopathes (ASPDP) : **réduction significative du volume de matière grise** (bilatéralement) dans le cortex préfrontal rostral antérieur (aire 10 de Brodmann) et les lobes temporaux poles (aire 20/38 de Brodmann) 2/ Chez les criminels non-psychopathes (ASPD-P) : volume de matière grise comparable à celui des sujets témoins	« *Reduced grey matter volume within areas implicated in empathic processing, moral reasoning, and processing of prosocial emotions such as guilt and embarrassment may contribute to the profound abnormalities of social behavior observed in psychopathy. Evidence of robust structural brain differences between persistently violent men with and without psychopathy adds to the evidence that psychopathy represents a distinct phenotype. This knowledge may facilitate research into the etiology of persistent violent behavior.* »

	crânien, trouble bipolaire ou psychotique				
<u>Decety et al. (2013)</u> "Brain response to empathy-eliciting scenarios involving pain in incarcerated individuals with psychopathy"	80 criminels incarcérés dans une prison américaine de moyenne sécurité ; - **27 criminels présentant un haut score de psychopathie** (PCL-R ≥ 30) - **28 criminels présentant un score moyen de psychopathie** (PCL-R : 21-29) - **25 criminels présentant un bas score de psychopathie** (PCL-R ≤ 20)	Identifier les différences de pattern d'activité cérébrale entre des criminels psychopathes et non-psychopathes	IRMf (ROI) *Pain interactions task* : stimuli visuels décrivant un préjudice physique causé à autrui *Pain expressions task* : stimuli visuels décrivant des expressions faciales de douleur	Chez les criminels présentant un haut score de psychopathie : 1/ **Réduction de l'activité** du cortex prefrontal ventromédian, cortex orbitofrontal latéral, et substance grise périaqueducale 2/ **Augmentation de l'activité de l'insula**, corrélée positivement au facteur 1 (relatif aux caractéristiques affectives et interpersonnelles) et au facteur 2 (relatif au style de vie et aux tendances antisociales) de la PCL-R	« *In response to pain and distress cues expressed by others, individuals with psychopathy exhibit deficits in the ventromedial prefrontal cortex and orbitofrontal cortex regardless of stimulus type and display selective impairment in processing facial cues of distress in regions associated with cognitive mentalizing. A better understanding of the neural responses to empathy-eliciting stimuli in psychopathy is necessary to inform intervention programs*. »
<u>Leutgeb et al. (2015)</u> "Brain abnormalities in high-risk violent offenders and their association with psychopathic traits and criminal recidivism"	- **40 criminels violents présentant un haut score de psychopathie** (PCL-R ≥ 20), détenus dans une prison de haute sécurité à Graz (Autriche) - **37 sujets témoins** d'âge équivalent Critères d'exclusion : criminels condamnés pour des crimes sexuels, antécédents de dépression, trouble bipolaire,	Comparer le volume de matière grise d'un échantillon de criminels violents présentant un important risque de récidive avec des sujets issus de la population générale	IRM (VBM, ROI)	1/ **Réduction du volume de matière grise** dans le cortex préfrontal 2/ **Augmentation du volume de matière grise** dans le cervelet et les ganglions de la base 3/ **Corrélation négative** entre le **volume de matière grise** du cortex préfrontal et de l'amygdale d'une part, et le **score de psychopathie** et	« *The current investigation revealed that in violent offenders deviations in grey matter volume of the prefrontal cortex as well as areas involved in the motor component of impulse control (cerebellum, basal ganglia, SMA) are differentially related to psychopathic traits and the risk for violence recidivism. The results might be valuable for improving existing risk assessment tools*. » « *It remains unclear, if the abnormalities are causal for criminality, and moreover, if they*

	psychose, trouble du déficit de l'attention avec hyperactivité, maladie neurologique ou organique, antécédents d'alcoolisme ou toxicomanie			le **risque de récidive violente** d'autre part 4/ **Corrélation positive** entre le **volume de matière grise** du cervelet, des ganglions de la base et de l'aire motrice supplémentaire d'une part, et le **score de psychopathie** et le **risque de récidive violente** d'autre part	*are reversible […] Science is still far from establishing neurobiological criteria as a contribution to final decisions in the court. However, the integration of objective neurobiological criteria, and an understanding of the interaction of biological and psychological factors could provide a higher quality of risk assessment. In addition, understanding neurobiological correlates of mental disorders could be helpful for improving therapeutic interventions. Lastly, a neurobiological evaluation of treatment success in forensic psychiatry would be of great value for the decision if patients should be discharged, or not*. » (p. 209-210)
<u>Hosking et al. (2017)</u> "Disrupted prefrontal regulation of striatal subjective value signals in psychopathy"	**49 criminels psychopathes** (20-45 ans) incarcérés dans deux prisons de moyenne sécurité (Wisconsin) Critères d'exclusion : antécédents de trouble psychotique ou bipolaire, traumatisme crânien, consommation de médicaments psychotropes	Identifier les mécanismes neurobiologiques sous-tendant les décisions inadaptées chez des psychopathes incarcérés	IRMf *Delay discounting task* : choix entre une petite récompense immédiate ou une grosse récompense tardive	1/ **Dérégulation du circuit cortico-striatal** : connectivité plus faible entre le striatum et le cortex médian 2/ **Corrélation positive** entre la dérégulation du circuit cortico-striatal et le **nombre de condamnations criminelles**	« *Finally, weaker cortico-striatal regulation predicted more frequent criminal convictions. These data suggest that cortico-striatal circuit dysregulation drives maladaptive decision making in psychopathy, supporting the notion that reward system dysfunction comprises an important neurobiological risk factor.* » (p. 221)

3.3. La « vision neurocriminologique »

3.3.1. L'effacement du social

Le premier constat qui ressort de l'analyse de la littérature est celui d'un programme de recherche extrêmement cohérent sur le plan théorique. La neurocriminologie s'est dotée d'un appareil conceptuel et méthodologique qui répond à ce que Mallory Fallin et ses collaborateurs (2018) appellent la « vision neurocriminologique ». Cette notion est entendue dans le sens du concept de « vision professionnelle » proposé par le linguiste et anthropologue Charles Goodwin (1994). Elle définit une manière de concevoir et d'interpréter les phénomènes empiriques qui garantit une vision cohérente et unifiée du champ disciplinaire. Comme le précisent Fallin et coll. (2018) cette « vision neurocriminologique » comprend donc « une combinaison de choix de recherche, de stratégies rhétoriques et de tendances interprétatives » qui structurent la manière dont les neurocriminologues étudient le cerveau des criminels. Les auteurs relèvent par exemple quelques principes méthodologiques fondamentaux, notamment la démarche comparative consistant à isoler les différences neurobiologiques entre un groupe de criminels et un groupe contrôle, ou encore la construction d'un paradigme expérimental qui inclue la réalisation d'une tâche permettant de mesurer un comportement associé à la criminalité (exemple : impulsivité, agressivité, insensibilité morale ou affective).

Mais surtout, cette « vision neurocriminologique » s'accompagne d'un « neurodiscours » qui pervertit véritablement le pouvoir explicatif des résultats. Elle implique la mise en place d'une double stratégie qui consiste d'une part, à redéfinir le problème de la criminalité en termes neurobiologiques, et d'autre part, à construire un modèle explicatif qui exclue systématiquement toute interprétation alternative (en l'occurrence, sociologique). Ce modèle explicatif répond à une stratégie d'envergure car il ne s'agit pas seulement de négliger l'influence des facteurs socio-environnementaux dans le développement du comportement criminel, mais bien de les *effacer* au profit d'une explication exclusivement neurobiologique : le modèle neurobiologique de la criminalité est construit de telle sorte à invalider la pertinence des explications sociologiques.

À cet égard, Fallin et coll. (2018) ne manquent pas de relever toute l'ambivalence du discours neurocriminologique, qui se défend de négliger le social, tout en proposant un modèle entièrement construit autour du niveau neurobiologique. Depuis la fin des années 1990, la neurocriminologie s'efforce d'étudier les interactions entre les facteurs socio-environnementaux et neurobiologiques, dans une dynamique de recherche qui consiste à identifier les « bases biosociales de la violence » (Raine et al. 1997b).[126] Comme le montrent Fallin et coll., ces études (dites intégratives) sont gouvernées par une démarche circulaire car les facteurs socio-environnementaux sont intégrés dans un modèle explicatif qui se prononce toujours en faveur du niveau neurobiologique :

> « *The result is an account of crime that transforms a multitude of social, environmental, and behavioral processes into neurological dysfunction.* » (Fallin et al. 2018).

Autrement dit, il s'agit de minimiser l'influence causale des facteurs socio-environnementaux, en les reléguant systématiquement à un niveau de description neurobiologique. La stratégie consiste par exemple à réduire les facteurs de risque de criminalité (par exemple, les antécédents de toxicomanie, d'abus physiques et/ou sexuels pendant l'enfance, etc.) à leurs effets neurologiques, c'est-à-dire à les associer à la survenue de lésions cérébrales : si le premier évènement joue un rôle causal dans le développement d'un comportement criminel, c'est parce qu'il laisserait une « empreinte » dans le cerveau de l'individu. Ainsi, même en intégrant des facteurs socio-environnementaux dans le modèle proposé, celui-ci garantit au cerveau son rôle privilégié dans l'explication du comportement.[127]

Il faut bien ici comprendre que cette dynamique de recherche s'inscrit au cœur même de la stratégie déployée par la neurocriminologie pour assurer sa propre légitimité scientifique. Elle est portée par des figures emblématiques du « programme fort », lequel vise ni plus ni moins qu'à expliquer la criminalité par des « prédispositions neurobiologiques ». Adrian Raine, qui a publié la première étude utilisant la TEP sur une

[126] Cette voie de recherche ne relève pas du champ de la neurocriminologie à proprement parler, mais de ce qui est communément appelé la « criminologie biosociale » (Eichelberger and Barnes 2015). Pour des revues de littérature scientifique de ce domaine de recherche, voir notamment Raine (2002) ; Scarpa and Raine (2007) ; Patrick (2008) ; Chen et al. (2015).

[127] Nous reviendrons en détail sur les fondements de cette stratégie dans le prochain chapitre. Pour l'heure, notre objectif se limite à l'analyse de la démarche globale adoptée par la neurocriminologie contemporaine.

population criminelle, est l'un des principaux acteurs de ce double discours. Depuis le début des années 1990, il s'applique à rechercher dans le cerveau des criminels ce qu'il nomme lui-même « les prédispositions à la violence » (Raine et al. 1997a, 1998a, 1998b), tout en cosignant régulièrement des études intégratives visant à identifier les « modèles biosociaux » de la criminalité (Brennan and Raine 1997 ; Rudo-Hutt et al. 2011 ; Rocque et al. 2012 ; Chen et al. 2015).[128]

Ces études sont souvent invoquées dans la littérature en neurocriminologie, vraisemblablement dans un souci d'autolégitimation et pour anticiper les critiques de cérébro-centrisme. Il est en effet certain que la neurocriminologie cherche un moyen d'étendre sa reconnaissance institutionnelle en multipliant les initiatives interdisciplinaires.[129] En mobilisant les ressources conceptuelles et méthodologiques des sciences sociales, elle ne cherche d'ailleurs pas tant à s'implanter dans le champ, mais bien plutôt à renforcer son statut d'autorité scientifique (Fallin et al. 2018) : l'essentiel de sa stratégie consiste, rappelons-le, à invalider la pertinence des explications sociologiques dans l'étude du phénomène criminel. Elle est donc gouvernée par une logique éliminativiste, qui n'appelle pas seulement à redéfinir le problème de la criminalité dans le langage des neurosciences, mais à substituer ses déterminants sociologiques par des explications neurobiologiques. En clair, la neurocriminologie ne se contente pas de défendre un modèle neurobiologique du comportement criminel ; elle cherche à faire du cerveau l'objet d'étude privilégié pour étudier les déterminants sociaux de la criminalité.

À partir d'une analyse de la littérature publiée entre 1999 et 2016, Fallin et coll. (2018) proposent d'expliquer comment cette « vision neurocriminologique » parvient à effacer le social dans l'étude du phénomène criminel. Les auteurs identifient trois stratégies qui régissent selon eux toutes les étapes de la recherche empirique, de la conceptualisation à la construction du modèle expérimental, jusqu'à l'interprétation des données : 1/ la sélection de variables de contrôle favorables à une interprétation exclusivement neurobiologique des phénomènes étudiés ; 2/ la transformation de

[128] Adrian Raine est le neurocriminologue le plus cité dans la littérature. Une analyse quantitative effectuée par le sociologue Julien Larregue (2019) montre que le nombre de citations de ses articles a été multiplié par dix entre 1994 et 2015 (voir Annexe n°8).

[129] Il est par exemple fréquent que les revues de littérature publiées par les neurocriminologues mobilisent les études consacrées aux bases génétiques du comportement criminel, en complément de leur approche neurobiologique (e.g. Nordstrom et al. 2011).

comportements sociaux complexes en un petit nombre de tâches expérimentales extrêmement simples ; et 3/ la production d'inférences interprétatives qui excluent les données susceptibles de compromettre la « vision neurocriminologique ». Il paraît important de présenter les spécificités et les limites de ces stratégies, dans la mesure où elles nous renseignent sur la logique qui gouverne la construction des études menées sur le cerveau des criminels. Ces trois stratégies contribuent à promouvoir l'agenda de la recherche en neurocriminologie, en favorisant un modèle explicatif entièrement construit autour du niveau de description neurobiologique.

3.3.2. Trois stratégies éliminativistes

La première stratégie se traduit par l'exclusion de facteurs démographiques et socio-économiques (âge, genre, niveau d'éducation, catégorie socio-professionnelle, etc.) dans la collecte ou l'interprétation des données. À l'exception des recherches intégratives que nous avons citées précédemment, les études ne tiennent pas compte des facteurs sociaux de la criminalité et il est rare que les résultats de la neuroimagerie soient mis en relation avec les caractéristiques sociodémographiques ou socioéconomiques des individus. Comme en atteste le tableau de synthèse de notre revue de la littérature, les neurocriminologues préfèrent intégrer des variables neurologiques comme les antécédents de traumatisme crânien ou de maladie mentale : l'objectif annoncé est de limiter les biais liés à des facteurs confondants, c'est-à-dire d'éviter que les anomalies cérébrales observées chez les criminels ne soient attribuées aux antécédents de troubles mentaux ou organiques.

Or comme le montrent Fallin et coll. (2018), ce choix méthodologique répond surtout à la volonté de contrôler les variables qui pourraient compromettre la « vision neurocriminologique ». Ces variables sont sélectionnées de manière à neutraliser le pouvoir explicatif des facteurs socio-environnementaux de la criminalité. Plutôt que d'interroger la dynamique de l'interaction avec les anomalies cérébrales observées, les neurocriminologues cherchent à montrer que le développement de la criminalité est uniquement lié à des facteurs neurobiologiques. Il en résulte un modèle explicatif exclusif de toute autre interprétation possible, qui attribue au cerveau un rôle premier dans la manifestation du comportement criminel.

La deuxième stratégie mobilisée par les neurocriminologues consiste à opérationnaliser un comportement complexe (par exemple, le contrôle de soi) au travers d'un petit nombre de tâches expérimentales. Comme nous le préciserons plus loin dans l'analyse des stratégies de recherche privilégiées par les neurocriminologues, le paradigme expérimental est globalement similaire entre les études. Il comprend la réalisation d'une tâche permettant de mesurer un état psychologique associé à la manifestation d'un comportement criminel (impulsivité, agressivité, insensibilité morale ou affective, etc.). Les tâches expérimentales impliquent par exemple le traitement de stimuli émotionnels (e.g. Müller et al. 2003), la capacité à contrôler son comportement (e.g. Habermeyer et al. 2003), ou encore la formation d'un jugement moral (e.g. Harenski et al. 2010). Les résultats de la neurocriminologie s'étendent alors sur deux niveaux d'analyse : le niveau comportemental (c'est-à-dire la performance des individus dans la réalisation de la tâche) et le niveau cérébral (correspondant aux données issues de la neuroimagerie anatomique ou fonctionnelle).

D'après les neurocriminologues, la réduction des performances dans la réalisation de ces tâches nous renseigne sur les capacités fonctionnelles qui font défaut chez les criminels : les déficits observés sur le plan cognitif, affectif ou social sont directement corrélés à la manifestation d'un comportement impulsif, violent ou antisocial. En clair, la construction du paradigme expérimental repose sur l'idée que les tâches expérimentales permettent de mesurer les capacités fonctionnelles des individus (exemple : le contrôle de soi), mais aussi que les performances individuelles sont équivalentes aux comportements réellement observés (tels qu'ils sont exercés dans l'environnement naturel). De ce point de vue, il paraît essentiel de questionner la pertinence et la portée explicative des tâches utilisées par les neurocriminologues : le problème de l'opérationnalisation pose la question de savoir si ces tâches expérimentales permettent de rendre compte d'un comportement aussi complexe que le contrôle de soi ou la formation d'un jugement moral.

En examinant les modèles expérimentaux les plus couramment utilisés dans les études, Fallin et coll. montrent que les neurocriminologues conceptualisent les comportements comme des objets universels, dépouillés de leur caractère contingent ou situationnel. Bien que certaines études s'efforcent de maximiser la validité écologique des tâches expérimentales en immergeant les participants dans des scénarios très complexes, le paradigme expérimental sous-spécifie le comportement tel qu'il est exercé dans l'environnement naturel. Un exemple en est donné par l'utilisation fréquente de la

Go/NoGo task pour mesurer le contrôle de soi des individus. Si cette tâche cognitive est pertinente pour évaluer la capacité à inhiber une action motrice, elle n'est en revanche pas suffisante pour rendre compte de la dimension prospective du contrôle de soi, à savoir la capacité à élaborer un plan d'action, à répondre à des raisons d'agir ou encore à s'identifier à la volonté formée.

Nous laissons ici de côté la question de savoir sous quelles conditions épistémologiques un comportement aussi complexe que le contrôle de soi pourrait être accessible à l'investigation neuroscientifique. Retenons néanmoins qu'il est inexact de considérer que la performance à la *Go/NoGo task* nous renseigne sur la capacité des sujets à envisager les conséquences de leurs actions sur le long terme, ou à contrôler leur comportement conformément à des règles morales. Compte tenu du caractère artificiel (décontextualisé et sous-spécifié) du phénomène psychologique étudié (la capacité à inhiber une action motrice), on peut aisément contester la validité écologique de ce modèle pour rendre compte du contrôle de soi tel qu'il est exercé dans l'environnement naturel.

De manière générale, l'exclusion de variables sociales ou contextuelles soulève des problèmes conceptuels et empiriques qui rendent difficile l'extrapolation des résultats aux conditions réelles d'exécution des comportements. Pourtant, les neurocriminologues font fi de ces limites et trouvent dans ces modèles expérimentaux un moyen de justifier encore davantage la pertinence d'une approche neurobiologique du comportement criminel. Comme le souligne Fallin et ses coll., la simplification de ces modèles garantit la possibilité de contrôler les variables qui pourraient discréditer la « vision neurocriminologique » : en opérationnalisant un comportement aussi complexe que le contrôle de soi ou la formation d'un jugement moral, les neurocriminologues parviennent à détourner les facteurs sociaux ou contextuels au profit d'une explication (exclusivement) neurobiologique des phénomènes étudiés. Ici encore, la stratégie ne relève pas d'un simple réductionnisme méthodologique consistant à décomposer un comportement en un petit nombre de tâches expérimentales ; il s'agit d'invalider la pertinence d'une analyse contextuelle ou systémique, en effaçant l'influence des facteurs sociaux de la criminalité.

La dernière stratégie identifiée par Fallin et ses coll. concerne non plus la construction du modèle expérimental, mais la manière dont les données empiriques sont interprétées. Elle consiste à négliger voire discréditer les résultats qui entrent en

contradiction avec la « vision neurocriminologique ». Cette stratégie s'exprime en particulier dans un souci constant de privilégier les résultats issus de la neuroimagerie au détriment des données comportementales, relatives à la performance des individus pendant la réalisation des tâches cognitives. Comme le relèvent Fallin et ses coll., la littérature en neurocriminologie fait état de résultats contradictoires selon que l'analyse est menée sur le plan comportemental ou cérébral. Dans la plupart des études, les données neurobiologiques révèlent des différences importantes entre les criminels et le groupe contrôle (notamment en termes de niveaux d'activité cérébrale), alors qu'aucune différence n'est observée sur le plan comportemental : environ un tiers des articles analysés par Fallin et ses coll. ne rapporte pas de différence statistiquement significative entre les performances cognitives des criminels et des sujets issus de la population générale.

Or, ces données comportementales sont minimisées, voire délibérément occultées dans l'interprétation des résultats. Comme en atteste le tableau de synthèse de notre revue de la littérature, les résultats sont axés sur le recensement des anomalies cérébrales et mentionnent rarement les données concernant les performances cognitives des individus. Comme le montrent Fallin et ses coll. (2018), les neurocriminologues mobilisent diverses stratégies rhétoriques pour contester la fiabilité des données comportementales et montrer qu'elles sont moins rigoureuses sur le plan méthodologique. Cette logique interprétative est particulièrement manifeste dans leur réticence à envisager la possibilité que le groupe de criminels et le groupe contrôle ne présentent pas de différence significative sur le plan neurobiologique : même si aucune différence n'est observée sur le plan comportemental, les neurocriminologues persistent à rechercher des différences neurobiologiques entre les deux groupes. Pour Fallin et ses coll. (2018), ce biais interprétatif révèle une stratégie systématisée, dont l'objectif est de montrer que le comportement criminel est uniquement lié à une différence dans la conformation cérébrale des individus :

> « *Attending to how researchers deal with these "inconvenient facts" reveals the extent to which neurocriminology aims to reinforce a neurological account of crime by using interpretive and rhetorical strategies to reify neurological differences between deviant and normal populations. In short, when presented with conflicting evidence, neurocriminologists privilege the neuroscientific ones in a manner that reflects their field-specific interests.* » (Fallin et al. 2018)

Nous pouvons également supposer que ce biais interprétatif s'explique par la volonté de défendre l'apport de la neuroimagerie pour révéler des déficiences fonctionnelles imperceptibles sur le plan comportemental. Comme nous le montrerons plus loin (section 3.4), les données neurobiologiques sont parfois interprétées comme le témoin d'une discordance possible entre l'environnement socio-familial du sujet et son niveau de dangerosité. Le pendant de cette stratégie consiste à favoriser l'usage de la neuroimagerie pour repérer les stigmates d'une dangerosité insaisissable par les outils de prédiction traditionnellement utilisés dans l'évaluation du risque de récidive. Il s'agit à la fois de faire de la neuroimagerie l'outil privilégié pour évaluer la dangerosité des individus, et d'invalider la pertinence d'une analyse prédictive centrée sur les facteurs contextuels ou historiques du sujet.

Ce dernier point nous donne l'occasion de préciser un aspect fondamental qui n'est pas abordé dans l'article de Fallin et coll. (2018), et qui est difficilement perceptible si l'on s'en tient au discours scientifique en lui-même : il concerne la nature des finalités recherchées et la question de leur relation avec le modèle explicatif défendu par la neurocriminologie. Si l'analyse proposée par Fallin et coll. (2018) révèle toute la circularité de la démarche adoptée par les neurocriminologues pour défendre un modèle neurobiologique de la criminalité, elle reste en revanche aveugle aux finalités qu'ils poursuivent. Il semble pourtant que l'ambition de ce programme de recherche ne se limite pas à une simple redéfinition neurobiologique de la criminalité (y compris de ses déterminants sociaux). Comme nous le montrerons tout au long de ce chapitre, cette stratégie constitue la prémisse d'un projet normatif qui vise à définir la neuroimagerie comme l'outil privilégié pour évaluer la dangerosité d'un individu. Il s'agit non seulement d'exclure toute interprétation alternative à l'explication neurobiologique, mais surtout de garantir la possibilité d'utiliser ces données hors de leur cadre d'élaboration et d'interprétation, et en l'occurrence, dans le système pénal.

3.3.3. Quel modèle pour quelle finalité ?

La question des finalités recherchées par la neurocriminologie est primordiale si l'on s'intéresse au pouvoir explicatif de ce programme de recherche. En s'interrogeant

sur la valeur épistémique de ses résultats, on pose dans le même temps la question de savoir si ces derniers sont à la hauteur de ses ambitions, mais aussi si les finalités poursuivies n'introduisent pas un biais dans la construction des modèles expérimentaux. Ce dernier point fait l'objet d'une discussion approfondie dans le champ de la philosophie régionale des sciences (Potochnick 2017 ; Chirimuuta 2018). Il s'inscrit dans le cadre d'une réflexion plus générale concernant les contraintes épistémologiques propres aux stratégies de modélisation d'un objet complexe, et à leurs conséquences sur la nature de l'explication proposée.

Dans son ouvrage consacré au problème de la modélisation scientifique (plus justement nommée « *idealization* »), Angela Potochnick (2017) interroge la relation entre les finalités poursuivies par un programme de recherche et la manière dont les modèles sont conçus pour étudier le phénomène d'intérêt. Elle montre que la nature des finalités détermine non seulement le type de modèle à construire, mais aussi le type d'explication qui en résulte. En outre, la recherche de l'explication n'est pas toujours première dans la démarche empirique, car elle peut être supplantée par une logique prédictive (voire interventionniste), qui dispense les chercheurs d'identifier les mécanismes impliqués dans la réalisation du phénomène. En clair, certains programmes de recherche ne visent pas tant à *expliquer* le phénomène qu'à le prédire ; le modèle expérimental est alors conçu de manière à garantir cette finalité, alors même qu'il ne donne pas accès au mécanisme impliqué.

S'il n'est pas de notre propos ici (du moins à ce stade) d'expliciter davantage cette analyse, précisons néanmoins qu'elle constitue le fil conducteur de notre réflexion. Le problème soulevé par Angela Potochnick (2017) est pertinent à plus d'un titre, dans la mesure où il fournit une grille de lecture supplémentaire pour analyser le pouvoir explicatif de la recherche en neurocriminologie. Si l'ambition de la neurocriminologie contemporaine est bien d'*expliquer* les comportements criminels par des « prédispositions neurobiologiques », il se pose la question de savoir si les données issues de la neuroimagerie donnent accès aux mécanismes impliqués dans la réalisation des comportements. Ce premier niveau d'analyse concerne en définitive la question de la nature de l'explication neuroscientifique. À cela s'ajoute le problème de l'opérationnalisation d'un objet complexe, qui apparaît encore plus nettement si les recherches s'inscrivent dans une logique *prédictive* : si l'ambition de la neurocriminologie est de construire des modèles prédictifs d'un comportement violent ou agressif, on peut par exemple se demander si la tâche expérimentale demandée aux

participants représente le comportement sous tous ces aspects. Dans ce cas, la question posée est à la fois empirique et conceptuelle, car elle oblige à préciser les critères de définition du comportement que l'on cherche à isoler.

Ces questions épistémologiques très générales nous donnent l'occasion d'insister sur la nécessité d'identifier la nature et la légitimité scientifique des finalités poursuivies par la neurocriminologie : la question de son pouvoir explicatif nous invite à préciser ses ambitions, mais aussi à déterminer si elles sont empiriquement justifiées. Ainsi, dans la continuité des analyses proposées par Potochnick (2017) et Fallin et coll. (2018), nous proposons de clarifier les ambitions de la neurocriminologie contemporaine, en explicitant la logique des stratégies de recherche privilégiées dans les études. Notre démarche diffère de celle adoptée par Fallin et coll. (2018), car elle met l'accent sur la relation entre le modèle explicatif proposé par les neurocriminologues et les finalités qu'ils poursuivent. Nous défendons également une approche empirique qui ne se limite pas à l'analyse du discours en tant que tel, mais vise plutôt à dégager les présupposés sous-jacents aux choix méthodologiques opérés.

À ce titre, il paraît important de préciser les raisons pour lesquelles l'analyse du discours neurocriminologique ne suffit pas à identifier la nature des finalités recherchées. La première tient à l'ambivalence du discours lui-même, qui entretient un certain flou quant aux ambitions annoncées, mais qui laisse régulièrement apparaître des conclusions hâtives et imprécises. Comme en atteste le tableau de synthèse de notre revue de la littérature, les études font preuve d'une relative prudence vis-à-vis des objectifs annoncés : la plupart se restreignent à la question de savoir si les criminels sont « associés à » ou « caractérisés par » des anomalies structurelles et/ou fonctionnelles. Aucune étude ne formule explicitement l'ambition d'utiliser la neuroimagerie pour estimer la dangerosité d'un individu. Pourtant, les conclusions exposées révèlent bien souvent une extrapolation des données, laquelle surestime largement le pouvoir explicatif de la neuroimagerie. Généralement, la question de son usage judiciaire – lorsqu'elle est posée − est brièvement mentionnée dans les dernières lignes de la conclusion, ce qui pose évidemment la question de savoir si elle est justifiée sur le plan empirique.

Par ailleurs, et de manière surprenante, ces publications ne font aucune mention des premières études EEG menées sur les populations criminelles. Les chercheurs s'efforcent bien souvent de défendre le caractère inédit de leur approche et n'ont vraisemblablement pas connaissance de la profondeur historique de leur programme de

recherche. Ce constat nous amène donc à la question de la continuité (théorique et empirique) entre ces deux projets : dans la mesure où la neurocriminologie contemporaine fait fi de son héritage historique, on peut légitimement se demander si elle répond aux mêmes finalités que la neurocriminologie des années 1940-1980.

3.3.4. Un « cerveau-type » pour chaque type de criminalité

Un premier élément de réponse peut être apporté par l'analyse des différentes formes que recouvre l'objet de recherche de la neurocriminologie contemporaine. Si l'on prend soin d'examiner les catégories de criminalité étudiées dans ces travaux, on constate que son objet d'étude s'est démultiplié depuis les premières recherches basées sur l'EEG. Alors que la neurocriminologie des années 1940-1980 s'est construite autour d'une figure emblématique de la dangerosité criminelle (celle du criminel récidiviste, agressif, et qui commet des crimes violents sans raison), la neurocriminologie contemporaine a quelque peu élargi son objet de recherche en s'intéressant à de nouvelles formes d'infractions, parmi lesquelles la violence conjugale (Bueso-Izquierdo et al. 2016 ; Verdejo-Román et al. 2019) ou les crimes et délits sexuels commis sur mineurs (e.g. Schiltz et al. 2007 ; Cantor et al. 2008 ; Poeppl et al. 2013 ; Massau et al. 2017).

La démultiplication de son objet d'étude révèle un changement de paradigme dans la manière d'étudier le phénomène criminel. Si l'avènement de ces nouvelles figures de la dangerosité est bien évidemment lié au contexte historique, social et juridique de ces recherches, il semble aussi que les techniques de neuroimagerie aient favorisé une approche résolument typologique, centrée sur l'analyse différenciée des comportements criminels : tout au long de son développement, la neurocriminologie contemporaine s'institue comme un programme de naturalisation des différents types de criminalité. L'approche localisatrice qui gouverne l'usage de la neuroimagerie a en effet introduit une différence fondamentale par rapport aux études basées sur l'EEG. Elle a encouragé une démarche consistant à isoler les caractéristiques neurobiologiques propres à chaque type de criminalité. L'objectif n'est donc plus de caractériser sur le plan neurophysiologique une figure multiforme (et finalement abstraite) de la dangerosité, mais de définir autant de « profils neurobiologiques » qu'il existe de comportements criminels. Autrement dit, il ne s'agit plus d'opposer un « cerveau normal » à un « cerveau anormal » pour

circonscrire les propriétés neurobiologiques de l'individu dangereux, mais de construire un « cerveau-type » pour chaque type de criminalité.

Cette stratégie de recherche présente des conséquences importantes en ce qui concerne la nature de l'explication fournie par la neurocriminologie contemporaine. Comme nous l'avons précisé dans le chapitre précédent, la neurocriminologie des années 1940-1980 échoue à identifier les marqueurs propres à chacune des figures de la dangerosité, dans la mesure où elle ne fait que les réduire à une seule et même catégorie neurophysiologique : l'excès d'anomalies EEG est décrit comme la condition de possibilité de toutes les formes de dangerosité. Dans la neurocriminologie contemporaine, la caractérisation neurobiologique des criminels bénéficie d'une multitude d'approches (notamment lésionnelle) qui permettent de *localiser* les anomalies structurelles et fonctionnelles associées à différentes formes de criminalité. En clair, alors que les premières études EEG proposaient de fonder le critère de dangerosité sur le degré d'anomalies cérébrales, la définition neurobiologique de la dangerosité répond ici à une démarche localisatrice : la caractérisation neurobiologique des criminels ne repose pas sur le degré d'anomalies, mais sur la localisation de ces anomalies.[130]

Sans entrer dans une discussion épistémologique générale concernant les limites techniques, méthodologiques et théoriques de la démarche localisatrice (e.g. Uttal 2001), il paraît important d'interroger l'apport de cette stratégie pour informer la question de la dangerosité d'un individu, par rapport à l'approche neurophysiologique des années 1940-1980. Dans la mesure où le paradigme de la neurocriminologie contemporaine repose sur la conjonction d'une démarche localisatrice *et* typologique, il se pose la question de savoir si ses principes méthodologiques diffèrent fondamentalement des stratégies mobilisées dans les études EEG. Plus encore, dans la perspective qui est la nôtre d'interroger la relation entre la nature des finalités recherchées et le modèle explicatif proposé, on peut se demander *ce qui justifie* cette démarche, à la fois sur le plan théorique et empirique.

Sur ce point, la question que nous allons chercher à résoudre est double. D'une part, il s'agit de montrer en quoi la démarche typologique et localisatrice répond à

[130] Comme en atteste le tableau de synthèse de notre revue de la littérature, une étude IRM (Aigner et al. 2000) fait exception à la règle, puisqu'elle propose d'évaluer la fréquence des anomalies associées à différentes formes de criminalité (« crimes très violents » ; « crimes peu ou pas violents » ; « crimes de nature sexuelle »).

l'agenda de la recherche en neurocriminologie. La construction d'un modèle neurobiologique propre à chaque type de criminalité n'est-elle pas dictée par une logique prédictive ? D'autre part, il s'agit de déterminer si les résultats issus de la neuroimagerie garantissent une explication pertinente pour isoler les caractéristiques propres aux différentes formes de criminalité. En clair, cette double exigence nous invite à analyser toutes les étapes de la construction des études en neuroimagerie : des présupposés sous-jacents à la construction du modèle expérimental jusqu'à l'interprétation des données, en passant par les choix méthodologiques opérés en amont, notamment en ce qui concerne l'objet de recherche privilégié (c'est-à-dire les catégories de criminalités retenues) dans les études.

Une dernière remarque s'impose avant d'entrer dans le détail de notre analyse. Comme nous l'avons précisé plus haut, notre revue de la littérature distingue trois types d'études selon les catégories de criminalité examinées : 22 études menées sur des criminels violents (généralement condamnés pour homicide volontaire), 15 études menées sur des criminels sexuels (poursuivis pour la plupart pour des actes de pédophilie) et 25 études menées sur des criminels psychopathes. Nous avons choisi d'isoler les principes méthodologiques qui régissent ces trois voies de recherche, principalement pour trois raisons.

D'abord, il paraît important d'insister sur les études menées sur les criminels violents, dans la mesure où elles ont contribué à définir les principales ambitions et stratégies de recherche du programme : la neurocriminologie contemporaine s'est essentiellement formée autour de la figure du criminel violent. À ce titre, une attention particulière sera portée à une série d'études publiées par Adrian Raine (1994, 1997a, 1998a, 1998b), car elles ont posé les bases théoriques et méthodologiques de l'étude neurobiologique du phénomène criminel. Ensuite, il est important de préciser que les paradigmes expérimentaux diffèrent selon les catégories criminelles étudiées. Les études menées sur les auteurs d'infractions sexuelles et les criminels psychopathes font appel à des stratégies de recherche qui leur sont propres, lesquelles répondent à des objectifs variés qu'il est nécessaire d'expliciter : si ces recherches sont gouvernées par la même logique que les études menées sur les criminels violents, elles vont progressivement redéfinir le champ d'intervention de la neurocriminologie, en favorisant une conception psychopathologique du comportement criminel. Enfin, dans la mesure où l'ensemble de ces études poursuivent une visée typologique – consistant à isoler un « cerveau type »

pour chaque type de criminalité – il nous parait important de déplacer notre analyse sur le terrain empirique, afin d'évaluer la légitimité scientifique de cette ambition. En clair, la distinction des études selon leur objet de recherche nous donne l'occasion de comparer la nature des anomalies associées à chacune des catégories criminelles examinées : les données empiriques permettent-elles réellement de dégager un « profil neurobiologique » spécifique à chaque type de criminalité ?

3.4. La problématique du crime violent

3.4.1. Une classification dichotomique

Si l'on retrouve fréquemment la figure du criminel violent décrite dans les premières études EEG, c'est parce qu'elle apparaît comme le véritable fil rouge des recherches menées en neuroimagerie. En parcourant la littérature publiée depuis le début des années 1990, on constate que l'immense majorité des études a pour objectif d'isoler les spécificités cérébrales des individus condamnés pour des crimes violents : homicides volontaires, viols, voies de fait graves, etc. La neurocriminologie contemporaine s'est d'ailleurs construite autour de cette ambition, puisque les premières publications assument la volonté de localiser les déficits fonctionnels responsables d'une « prédisposition à la violence » (Raine et al. 1997a, 1998a, 1998b).

Adrian Raine, figure emblématique de la neurocriminologie contemporaine, a largement contribué à la construction de ce projet, au travers d'une série d'études utilisant la TEP sur des criminels condamnés pour meurtre (Raine et al. 1994, 1997a, 1998a, 1998b). Au total, ces études ont porté sur une quarantaine de criminels ayant plaidé l'irresponsabilité pour cause psychiatrique (« *not guilty by reason of insanity* », NGRI) ou l'inaptitude à subir le procès (« *incompetent to stand trial* », IST).[131] Elles marquent

[131] Comme nous l'avons précisé dans le tableau de synthèse, il est important de noter que tous ces criminels n'ont pas bénéficié d'une irresponsabilité pour cause psychiatrique car certains (leur proportion n'est pas mentionnée dans l'étude) ont été condamnés à des peines de prison.

une étape cruciale dans l'histoire de la neurocriminologie, car ce sont les premières études utilisant la neuroimagerie fonctionnelle sur une population criminelle.

Comme nous l'avons précisé plus haut, ces études ne font aucune mention des recherches EEG publiées depuis le début des années 1940. Elles s'inscrivent plutôt dans la continuité des études neurologiques démontrant un lien entre les dysfonctionnements du cortex préfrontal et la manifestation d'un comportement violent (Weiger and Bear 1988 ; Bear 1989). Les recherches menées par Adrian Raine ont aussi été inspirées par une série d'études utilisant la TEP ou la tomodensitométrie (CT-scan) sur des populations psychiatriques connues pour des actes de violence, en particulier pour des agressions sexuelles (Volkow and Tancredi 1987 ; Garnett et al. 1988 ; Graber et al. 1982 ; Hendricks et al 1988 ; voir aussi Langevin et al. 1990). Comme le souligne Raine lui-même dans sa première étude publiée en 1994 :

> « *These studies are clearly important in illustrating the potential relevance of brain imaging to criminal behavior, but there have been no previous brain-imaging studies of serious violence. Murder is a violent criminal act of major public concern, and pleas of not guilty by reason of insanity (NGRI) in murder cases require critically important decisions to be made by forensic psychiatrists and psychologists*. » (Raine 1994, p. 366).

Si l'ambition d'utiliser les données issues de la neuroimagerie pour répondre au problème de la responsabilité pénale n'est pas clairement annoncée dans cet article, la question de l'apport de la neuroimagerie pour évaluer la dangerosité des individus semble en revanche constituer la problématique centrale des recherches menées sur le cerveau des criminels violents. L'ambition d'isoler les caractéristiques neurobiologiques de l'individu dangereux bénéficie d'ailleurs d'une stratégie comparative que l'on décelait déjà dans les premières publications reposant sur l'EEG. Elle consiste à assigner chaque individu dans la catégorie des « crimes violents » et des « crimes non violents », afin de caractériser leurs différences anatomiques et/ou fonctionnelles.

À cet égard, il est important de noter que les critères de classification sont généralement basés sur les données relatives au dossier judiciaire des sujets, c'est-à-dire en fonction des différents motifs de condamnation. La sélection des criminels suppose d'isoler de l'échantillon les individus « violents ou très violents » (généralement condamnés pour homicide volontaire ou viol) et les individus « peu ou pas violents »

(condamnés dans la plupart des cas pour vol, fraude ou trafic de drogue).[132] Il est rare que les études utilisent des outils actuariels d'évaluation de la violence, comme le font les recherches publiées dans le champ psychiatrique ou psychopathologique. Les quelques études adoptant une approche actuarielle utilisent des instruments d'auto-évaluation de l'impulsivité, de l'agressivité ou de l'anxiété comme le *Barratt's Impulsiveness Scale* (BIS, Patton et al. 1995), le *Reactive-Proactive Aggression Questionnaire* (RPQ, Raine et al. 2006), ou encore la *Welsh Anxiety scale (*Welsh, 1956).[133]

En se détournant ainsi de la stratégie d'évaluation actuarielle – pourtant largement utilisée en clinique – la neurocriminologie contemporaine propose une classification dichotomique arbitraire qui repose sur la même logique que les premières publications des années 1940. Les études menées sur les criminels violents ont adopté un choix méthodologique qui n'est pas sans rappeler la volonté d'opposer sur le plan neurophysiologique les figures antagonistes du « criminel dangereux » et du « criminel normal ». Pour rappel, les publications des années 1940-1980 étaient motivées par un partage constant entre deux catégories de criminels : le criminel dit « normal » présentant un EEG « normal » et « stable », et le criminel « dangereux » présentant un degré d'anomalies EEG proportionnel à son instabilité comportementale. Cette démarche comparative était favorisée par l'observation d'une proportion d'anomalies similaire entre la première catégorie (c'est-à-dire les criminels ne présentant pas un grand danger pour la société) et les sujets issus de la population générale. C'est ainsi que ces derniers se sont vus peu à peu supplantés par la constitution d'un nouveau groupe contrôle : celui constitué par les criminels dits « normaux », généralement stables sur le plan psychologique, le plus souvent condamnés pour homicide involontaire ou légitime-défense, et ayant commis un crime isolé.

La neurocriminologie contemporaine s'est dotée d'une stratégie complémentaire qui permet de renforcer la validité statistique de ses hypothèses. Contrairement aux études EEG, les études de neuroimagerie font généralement appel à des sujets issus de la

[132] Les enregistrements IRM sont effectués sans que les expérimentateurs n'aient connaissance de ces informations, afin de ne pas introduire un biais dans l'acquisition ou l'interprétation des données. Il est en effet souvent précisé dans les études que les neuroradiologues sont « aveugles » aux antécédents judiciaires ou au diagnostic clinique des individus.

[133] Des tests de personnalité comme le *Temperament and Character Inventory* (TCI, Cloninger et al. 1994) et des tests de Q.I. sont parfois réalisés en complément, ainsi que des mesures de la psychopathie comme la *Psychopathy Checklist-Revised* (PCL-R), dont nous parlerons plus loin.

population générale. Le cas échéant, elles adoptent la même stratégie que dans les publications des années 1940-1980, à savoir comparer deux groupes de criminels représentant deux niveaux de dangerosité (exemple : « criminels violents » vs. « criminels non-violents »). D'autres études sont fondées sur une double comparaison dans laquelle le groupe contrôle est représenté à la fois par des sujets issus de la population générale, et par des criminels condamnés pour des infractions non violentes (généralement des individus condamnés pour des crimes contre la propriété). Dans ce cas, il s'agit non seulement d'isoler les spécificités cérébrales des criminels par rapport aux individus n'ayant jamais fait l'objet d'une condamnation, mais aussi de comparer les données neurobiologiques entre deux figures antagonistes de la criminalité.

Ainsi, en l'absence de sujets témoins issus de la population générale, la comparaison s'établit entre deux catégories mutuellement exclusives. Elle peut par exemple s'effectuer sur la base des antécédents familiaux des individus ou des circonstances de leur crime : entre des criminels ayant souffert de carences psychosociales et des criminels ayant bénéficié d'un environnement socio-familial stable (Raine et al. 1998a) ; entre des « criminels prédateurs » ayant commis leur crime de sang-froid et de manière préméditée, et des « criminels affectifs » ayant commis leur crime de manière impulsive et dans un état de grande confusion émotionnelle (Raine et al. 1998b) ; ou encore entre des criminels psychopathes ayant été condamnés pour leur crime (les « *unsuccessful psychopaths* ») et d'autres ayant échappé à des poursuites judiciaires (les « *successful psychopaths* ») (Raine et al. 2004 ; Yang et al. 2005, 2010).[134]

S'il est bien évidemment tentant de voir dans cette double catégorisation une opposition *de nature* entre un « criminel normal » et un « criminel dangereux », c'est parce qu'elle apparaît en filigrane dans tous les travaux publiés par Adrian Raine. Parmi la demi-douzaine d'études menées par Raine que nous avons retenues dans notre revue de la littérature, celle menée sur les antécédents familiaux des sujets (Raine et al. 1998a) mérite une attention particulière. La raison est double : d'une part, elle figure parmi les rares études à intégrer les facteurs socio-environnementaux dans son analyse, s'inscrivant

[134] Bien que l'étude de Raine et ses collaborateurs (2004) n'en fasse pas mention, nous pouvons préciser que la distinction entre les « *unsuccessful psychopaths* » et les « *successful psychopaths* » remonte à la fin des années 1970 dans un article publié par la psychologue américaine Cathy Spatz Widom (1977). Cet article proposait d'isoler les caractéristiques psychologiques, sociales et démographiques des « *successful psychopaths* » (ou « *nonincarcerated psychopaths* »), c'est-à-dire les psychopathes ayant commis un acte immoral mais qui ont réussi à échapper à des poursuites judiciaires.

dans la dynamique de recherche interdisciplinaire amorcée à la fin des années 1990, et consacrée aux « bases biosociales de la violence ». D'autre part, elle a le mérite de mettre à jour la logique des stratégies de recherche employées dans la neurocriminologie contemporaine. Elle marque en effet l'une des premières tentatives visant à reconnaître la dangerosité d'un individu selon ses caractéristiques neurobiologiques.

3.4.2. Un criminel « biologiquement anormal »

L'étude de Raine et coll. (1998a) constitue la troisième de la série de publications parues entre 1994 et 1998, et utilisant la TEP sur des criminels condamnés pour meurtre (Raine et al. 1994, 1997a, 1998a, 1998b). Elle a porté sur une quarantaine de criminels répartis en un groupe d'individus ayant souffert de carences psychosociales (abus physiques et/ou sexuels pendant l'enfance, négligence parentale, antécédents familiaux de criminalité, etc.) et un groupe ayant bénéficié d'un environnement socio-familial stable.[135] Les résultats montrent que ces derniers présentent des déficits fonctionnels (en particulier un « hypométabolisme préfrontal ») nettement plus importants que les criminels ayant souffert de carences psychosociales. Après avoir montré que ces différences ne pouvaient être attribuées à des antécédents de lésions cérébrales ou de schizophrénie, ou à leur capacité à exécuter la tâche demandée, les auteurs de l'étude ont estimé que ces anomalies pouvaient constituer une « prédisposition à la violence » chez les individus n'ayant pas souffert de carence psychosociale :

> « *These results suggest that murderers lacking psychosocial deficits are characterized by prefrontal deficits. It is argued that among violent offenders without deprived home backgrounds, the "social push" to violence is minimized, and consequently, brain abnormalities provide a relatively stronger predisposition to violence in this group.* » (Raine et al. 1998a, p. 1).

Il est important de s'interroger sur le sens des arguments mobilisés par Adrian Raine et ses collaborateurs pour justifier la pertinence de leur modèle explicatif. Dans la

[135] Pour plus de précisions sur cette étude, nous invitons le lecteur à consulter le tableau de synthèse de notre revue de la littérature, présenté dans la section 3.2.

mesure où les seules différences perceptibles entre les deux groupes concernent les antécédents familiaux et le degré d'anomalies cérébrales des individus, les chercheurs font jouer à ces derniers deux rôles mutuellement exclusifs dans le développement de la criminalité. Il existerait d'un côté une criminalité émanant des carences psychosociales subies pendant l'enfance ; de l'autre, une criminalité résultant de déficits fonctionnels au niveau du cortex préfrontal. Autrement dit de ce point de vue, les anomalies préfrontales compensent elles-mêmes l'absence de déterminants socio-familiaux de la criminalité : elles sont interprétées comme la seule explication causale d'une criminalité qui ne répond pas à des facteurs contextuels ou historiques (antécédents familiaux et histoire personnelle de l'individu).

Ces résultats marquent une étape importante dans le projet typologique adopté par la neurocriminologie, car ils dessinent une opposition de nature entre deux catégories de criminels aux dynamiques socio-familiales totalement divergentes. Adrian Raine et ses collaborateurs opposent sur le plan neurobiologique deux figures antagonistes de la criminalité, dont l'une représente une dangerosité d'autant plus marquée qu'elle est imprévisible et imperceptible sur le plan socio-familial. Plus encore, ils voient dans ces résultats un argument pour déconstruire l'idée selon laquelle « les criminels ayant bénéficié d'un environnement familial normal » sont « biologiquement normaux » :

> « *One implication of these findings, in conjunction with the previous literature, is that they challenge the assumption often made in forensic settings that violent offenders from normal home backgrounds are biologically "normal", whereas those offenders exposed to abnormal rearing conditions are more likely to be biologically dysfunctional*. » (Raine et al. 1998a, p. 7)

En somme, la criminalité induite par des circonstances sociales défavorables s'oppose à la dangerosité des individus « biologiquement anormaux » ayant commis des crimes violents malgré un environnement familial parfaitement stable. Les chercheurs proposent de renverser le rapport à la normalité/anormalité entre le criminel, son cerveau, et son environnement familial, afin d'élaborer une définition (exclusivement) neurobiologique de la dangerosité criminelle. Plus précisément, ils invitent à voir les données neurobiologiques comme le témoin d'une discordance possible entre la qualité de l'environnement familial et la dangerosité de l'individu. En l'occurrence, certains individus ayant bénéficié d'une dynamique socio-familiale stable se révèlent aussi violents que ceux ayant souffert de carences psychosociales ; mais surtout, ces criminels

se distinguent des seconds en ce qu'ils sont caractérisés par des anomalies préfrontales nettement plus importantes. Ils incarnent alors la figure de l'individu dangereux qui échappe à un déterminisme social pour entrer dans le registre de « l'anomalie cérébrale » : du point de vue des neurocriminologues, un « cerveau anormal » associé à un environnement familial normal fait du criminel un individu *naturellement* dangereux, car doté d'une dangerosité *intrinsèque*. À l'inverse, la figure du « criminel normal » se dessine au travers d'une description neurobiologique plus neutre, qui fait peser à l'environnement son entière responsabilité dans la manifestation des comportements violents : ici, le « criminel normal » ne se définit pas par son degré d'anomalies cérébrales, mais seulement par son environnement criminogène.

Cette double équation n'est pas sans rappeler le modèle explicatif proposé par le psychiatre britannique Denis Williams (1969) pour justifier le caractère inné de certaines formes de criminalité. Pour rappel, l'étude de Williams proposait un modèle « endogène » et « exogène », correspondant respectivement à la catégorie des « criminels dangereux » (récidivistes, agressifs et impulsifs) et des « criminels normaux » (ayant commis un crime isolé). En l'absence de données relatives à l'environnement socio-familial des sujets, les seuls critères de distinction entre ces deux modèles étaient fondés sur le degré de leurs anomalies EEG et les circonstances de leur crime : l'excès d'anomalies EEG chez les « criminels dangereux » était associé à une origine « endogène », tandis que l'EEG « normal et stable » des « criminels normaux » était attribué à une origine « exogène ».

Bien que les facteurs socio-environnementaux soient pris en compte dans l'étude de Raine et coll. (1998a), son modèle explicatif repose sur le même principe puisqu'il consiste à fonder le critère de dangerosité sur des données neurobiologiques. Mais la stratégie employée par Raine et ses collaborateurs s'avère encore plus efficace au regard des finalités recherchées, précisément parce qu'elle tient compte de l'environnement socio-familial des sujets. Il ne s'agit pas ici simplement d'associer la nature de l'anomalie cérébrale (par exemple, un EEG « instable » ou « paroxystique ») avec la personnalité du criminel ou la nature du crime commis (un individu « instable » ou un comportement « paroxystique »). En distinguant les criminels selon leur environnement socio-familial, il s'agit bien plutôt d'isoler un indice de la dangerosité aveugle aux critères statistiques classiques basés sur les antécédents familiaux des sujets. Autrement dit, la stratégie adoptée par Raine offre l'avantage de faire apparaître un indice neurobiologique d'une dangerosité insaisissable par les outils de prédiction (statiques ou dynamiques)

traditionnellement utilisés dans l'évaluation du risque de récidive. Elle vise à la fois à défendre l'apport de la neuroimagerie pour estimer le niveau de violence d'un individu, et à invalider la pertinence d'une analyse prédictive centrée sur son histoire personnelle ou sur ses antécédents familiaux.

De là apparaît la fonction centrale de la démarche typologique : en isolant les caractéristiques neurobiologiques du criminel violent, la neuroimagerie se voit créditée d'une fonction prescriptive, autorisant ainsi son usage pour repérer les signes d'une dangerosité imperceptible sur le plan socio-familial. C'est en révélant une caractéristique inaccessible à l'observation et à l'analyse statistique que Raine et coll. (1998a) parviennent à faire de la neuroimagerie l'outil privilégié pour évaluer la dangerosité d'un individu.

Les conclusions avancées par cette étude sont paradigmatiques du projet typologique consistant à isoler les spécificités cérébrales des individus représentant une grande menace pour la société. Il s'agit globalement de comparer les données neurobiologiques entre deux figures antagonistes de la criminalité, afin d'isoler les caractéristiques distinctives de l'individu représentant le plus haut degré de dangerosité. Cette stratégie a d'ailleurs été appliquée dans la dernière étude de la série publiée par Raine et coll. (1998b).

Cette étude propose « d'évaluer les différences en termes de fonctionnement cortical et sous-cortical entre des meurtriers affectifs et prédateurs » : elle vise à caractériser les différences d'activité cérébrale entre les individus ayant commis leur crime dans un état de détresse émotionnelle, et ceux condamnés pour un crime prémédité et perpétré de sang-froid.[136] Les auteurs ont fait l'hypothèse que les premiers étaient caractérisés par une hyperactivité corticale et des déficits fonctionnels au niveau du cortex préfrontal, tandis que les seconds « montreraient un fonctionnement cérébral relativement normal en comparaison » (Raine et al. 1998b, p. 321). Ici encore, on voit clairement s'esquisser une approche catégorielle visant à justifier une opposition *de nature* entre un « criminel normal » et un « criminel dangereux ». Si le « criminel dangereux » apparaît comme la figure emblématique de l'individu impulsif, instable et entièrement soumis à

[136] Comme nous le précisons dans le tableau de synthèse de notre revue de la littérature, cette classification correspond à des critères de psychiatrie légale définis par Meloy (1988 ; 1997) et Dodge (1991).

ses états émotionnels, celle du « criminel normal » caractérise au contraire le sujet disposant d'un contrôle de soi absolu :

> « *Results support the hypothesis that emotional, unplanned impulsive murderers are less able to regulate and control aggressive impulses generated from subcortical structures due to deficient prefrontal regulation. It is hypothesized that excessive subcortical activity predisposes to aggressive behaviour, but that while predatory murderers have sufficiently good prefrontal functioning to regulate these aggressive impulses, the affective murderers lack such prefrontal control over emotion regulation*. » (p. 319)

De manière générale, les déficits structurels ou fonctionnels sont souvent interprétés comme un « facteur de risque » de criminalité (Aigner et al. 2000 ; Raine et al. 2001 ; Hosking et al. 2017) ou le signe d'une « prédisposition » neurobiologique aux comportements violents ou agressifs (Raine et al. 1994, 1997a, 2001 ; Schiltz et al. 2013 ; Poeppl et al. 2015 ; Chen et al. 2016). Tout au long des études menées sur les criminels violents, on voit ainsi apparaître un appareillage conceptuel et méthodologique dicté par une logique prédictive, qui s'exprime au travers d'une ambition doublement explicative et prescriptive. Il s'agit non seulement de construire un modèle neurobiologique de la dangerosité criminelle, mais aussi de garantir la possibilité d'utiliser la neuroimagerie sur des cas individuels pour *reconnaître* les signes de cette dangerosité.

Il n'est pas difficile de comprendre qu'en cherchant à révéler les « prédispositions neurobiologiques » aux comportements violents ou agressifs, les neurocriminologues voient dans la neuroimagerie un moyen de répondre à la problématique de la récidive. Comme nous venons de le voir, cette ambition prédictive bénéficie d'un ensemble de stratégies qui dictent à la fois la définition de l'objet d'étude (c'est-à-dire la sélection d'une catégorie de criminalité) et les choix méthodologiques privilégiés. D'une part, il paraît désormais évident que les figures de la dangerosité décrites par la neurocriminologie révèlent une grande partie des finalités qu'elle poursuit : les comportements violents et la récidive criminelle constituent le dénominateur commun de l'ensemble des études publiées depuis le début des années 1990. D'autre part, si l'on prête attention aux choix méthodologiques employés dans ces recherches, on constate qu'ils ne font que conforter la « vision neurocriminologique », telle que l'ont définie Fallin et coll. (2018) : les stratégies mobilisées sont elles-mêmes l'expression d'une

logique prédictive, puisque les données neurobiologiques sont collectées et interprétées de telle sorte à favoriser une explication (exclusivement) neurobiologique de la criminalité.

L'étude publiée par Adrian Raine et ses collaborateurs (1998a) est exemplaire à plus d'un titre, dans la mesure où elle prend le contre-pied de la plupart des recherches publiées dans le champ − en intégrant des variables liées aux antécédents familiaux des sujets − tout en détournant l'interprétation des données au profit d'un modèle centré sur une explication neurobiologique. Elle parvient à justifier l'existence d'une « prédisposition neurobiologique » à la violence, en révélant les apparentes discordances entre la qualité de l'environnement familial et la dangerosité des individus. Elle décrit un criminel « biologiquement anormal » malgré un environnement familial parfaitement stable, qui incarne finalement la figure prototypique du criminel entièrement soumis à ses instincts violents, qui échappe aux déterminants sociaux pour obéir à un autre déterminisme : celui de l'anomalie cérébrale.

L'analyse du champ de la neurocriminologie contemporaine serait incomplète si nous n'évoquions pas toutes les formes que recouvre son objet de recherche, et qui contribuent à nourrir encore davantage son ambition prédictive. Ces deux dernières décennies ont en effet vu apparaître de nouvelles figures de la dangerosité en raison de leur inclination à la violence, à l'impulsivité et à la récidive : les auteurs d'infractions sexuelles, et en particulier les pédocriminels (e.g. Schiltz et al. 2007 ; Sartorius et al. 2008 ; Poeppl et al. 2013 ; Kneer et al. 2019), et les criminels souffrant d'un trouble de la personnalité psychopathique (e.g. Kiehl et al. 2001, 2004 ; Yang et al. 2005, 2010 ; Leutgeb et al. 2015 ; Hosking et al. 2017).

3.5. Les nouvelles figures de la dangerosité

3.5.1. Le criminel psychopathe

Il est intéressant de constater que la neurocriminologie a conservé un vif intérêt pour le criminel psychopathe, qui constituait l'objet de recherche privilégié des premières

études EEG menées dans les années 1940. À ce jour, les études portant sur des criminels souffrant d'un trouble de la personnalité psychopathique concentrent près de la moitié des recherches publiées dans le domaine de la neurocriminologie.[137] Ici encore, il paraît difficile d'expliquer la résurgence de cette figure sans y voir l'expression des finalités recherchées par les neurocriminologues. À première vue, elle semble avoir été favorisée par les recherches démontrant une forte prévalence de la psychopathie dans les populations carcérales (e.g. Hare et Jutai 1983) : si la neurocriminologie a concentré son attention sur le criminel psychopathe, c'est vraisemblablement parce qu'il cristallise la figure dominante (et paradigmatique) de la dangerosité.

Pour autant, il importe de nuancer cette hypothèse car les ambitions annoncées dans les publications sont relativement confuses. Si l'on s'en tient au corps des articles, il est difficile de savoir si ces recherches répondent à la volonté d'isoler les caractéristiques neurobiologiques du « criminel dangereux », ou si elles se livrent à une étude neurobiologique de la psychopathie en tant que telle. La plupart affichent la volonté d'identifier « les bases neurales de la psychopathie », et plus précisément d'isoler les régions cérébrales associéss au traitement émotionnel (Kiehl et al. 2001 ; Müller et al. 2003) ou lexico-sémantique (Kiehl et al. 2004), à la formation d'un jugement moral (Harenski et al. 2010), ou encore au contrôle de soi (Hosking et al. 2017).

Dans un sens, la recherche de modèles neurobiologiques du contrôle de soi, des jugements moraux ou du traitement émotionnel fournit un indice sur les finalités poursuivies par les chercheurs, dans la mesure où l'impulsivité et l'insensibilité morale ou affective sont considérées comme des facteurs de risque de récidive (e.g. Hare 1991). Comme le remarque le neuroscientifique Russell Poldrack dans une revue de la littérature consacrée à la « neuroprédiction » des comportements violents :

> « *Psychopathy is a particularly useful model for understanding the neurobiology of violence because it encompasses both affective-interpersonal symptoms thought to underlie goal-directed violence and impulsive-antisocial symptoms linked to reactive violence*. » (Poldrack et al. 2018, p. 116).

[137] Sur les 62 études que nous avons retenues dans notre revue de la littérature, 17 ont été menées sur des criminels souffrant de psychopathie. Nous avons délibérément réduit cet échantillon, afin de limiter au mieux les éventuels biais de sélection en faveur des criminels présentant des troubles de la personnalité, qui pourraient introduire une confusion sur la relation établie entre le comportement criminel et la présence d'anomalies cérébrales.

Effectivement, si les études menées sur les criminels psychopathes concentrent la majeure partie de la littérature en neurocriminologie, c'est en raison du lien étroit établi entre la psychopathie et la récidive violente (e.g. Hart, Kropp, and Hare 1988). Cette association a fait naître une ambition majeure chez les neurocriminologues : utiliser la neuroimagerie comme un instrument de mesure du risque de récidive, en complément des outils diagnostics habituellement utilisés comme l'échelle de psychopathie (révisée) de Robert Hare (Psychopathy Checklist-Revised, PCL-R, Hare 1991, 2003).[138]

Toutes les études menées sur les criminels psychopathes utilisent la PCL-R, quel que soit l'outil de neuroimagerie utilisé (IRM anatomique ou fonctionnelle). Cette échelle en trois points (0-1-2) comprend 20 items répartis en deux facteurs principaux : un facteur relatif aux composantes affectives, interpersonnelles et narcissiques (facteur 1), et un facteur décrivant le style de vie et la tendance antisociale des individus (facteur 2).[139] Les scores varient de 0 à 40 et sont attribués à chaque criminel sur la base d'un entretien clinique, mais aussi en fonction des antécédents comportementaux figurant dans le dossier judiciaire des sujets.

La plupart des études proposent de comparer les données IRM entre des criminels souffrant de psychopathie (score $\geq$ 30) et des sujets témoins issus de la population générale (e.g. Kiehl et al. 2004). D'autres études cherchent plutôt à comparer les données recueillies chez des sujets témoins avec celles de criminels présentant différents scores de psychopathie (e.g. Decety et al. 2013). Dans ce cas, l'objectif annoncé est d'identifier les différences anatomiques et fonctionnelles entre les criminels psychopathes et les criminels non-psychopathes. Enfin, certaines études proposent d'identifier les anomalies cérébrales associées à chacun des facteurs de la PCL-R (Thijssen and Kiehl 2017 ; Deming et al. 2018). Il s'agit ici non plus de corréler la nature ou le degré des anomalies

[138] Effectivement, si la PCL-R a été initialement conçue comme un outil de diagnostic de la psychopathie, de nombreuses études ont démontré son utilité pour évaluer le risque de récidive violente (e.g. Hart et al. 1994).

[139] Plus précisément, le facteur 1 comprend les aspects suivants : loquacité/charme superficiel, grande estime de soi, mensonge pathologique, duperie/manipulation, absence de culpabilité ou de remords, affect superficiel, manque d'empathie, incapacité à assumer la responsabilité de ses actions. Le facteur 2 inclut : le besoin constant de stimulation, la tendance au parasitisme, l'incapacité à poursuivre des objectifs sur le long terme, l'impulsivité, l'irresponsabilité, le faible contrôle de soi, les troubles du comportement pendant l'enfance, la délinquance juvénile, la versatilité criminelle (Hare 1991, 2003).

avec le score total de la PCL-R, mais d'isoler les anomalies fonctionnelles associées à chaque dimension (affective et comportementale) de la psychopathie.

Globalement, les résultats montrent que les criminels psychopathes présentent des déficits structurels et fonctionnels au niveau des régions impliquées dans la régulation émotionnelle ou le contrôle de soi, mais surtout que le score de psychopathie est proportionnel au degré d'anomalies observées. Les études utilisant l'IRM anatomique montrent que les criminels présentant un haut score de psychopathie sont associés à une réduction significative du volume de matière grise au niveau du cortex préfrontal (Yang et al. 2005, 2010 ; Gregory et al. 2012 ; Decety et al. 2013 ; Leutgeb et al. 2015). Les résultats des études de neuroimagerie fonctionnelle sont plus hétérogènes, puisqu'ils rapportent des déficits d'activité dans un certain nombre de régions corticales et sous-corticales (e.g. Kiehl et al. 2001), mais aussi des asymétries fonctionnelles entre les deux hémisphères (Müller et al. 2003).[140]

Comme nous l'avons précisé plus haut, la majorité des études menées sur les criminels psychopathes répondent à la volonté d'étudier « les bases neurales de la psychopathie » dans l'objectif, à terme, de définir des critères neurobiologiques du risque de récidive violente. Certaines stratégies de recherche répondent d'ailleurs directement à cette ambition prédictive. L'une de ces stratégies consiste à corréler le degré d'anomalies observées avec le score de psychopathie et le risque de récidive violente des individus, ou encore avec le nombre de leurs condamnations antérieures. Une étude IRM menée sur 40 criminels violents a par exemple montré que le volume de matière grise du cortex préfrontal et de l'amygdale était inversement proportionnel au score de psychopathie et au risque de récidive des individus (Leutgeb et al. 2005). En clair, la réduction du volume de matière grise était d'autant plus importante que le risque de récidive était élevé. Une autre étude utilisant cette fois-ci l'IRMf a montré une corrélation positive entre un dysfonctionnement dans le circuit cortico-striatal et le nombre de condamnations criminelles des individus (Hosking et al. 2017). Dans la mesure où cette anomalie était

[140] L'hétérogénéité des résultats peut s'expliquer par le fait que les études n'utilisent pas toujours le même paradigme expérimental. L'enregistrement de l'activité cérébrale peut s'effectuer pendant la réalisation d'une tâche impliquant des stimuli émotionnels (e.g. Müller et al. 2003), la formation d'un jugement moral (e.g. Harenski et al. 2010), ou encore la capacité à contrôler son comportement (e.g. Hosking et al. 2017).

révélée par la réalisation d'une tâche impliquant un contrôle comportemental, elle a été perçue comme un « facteur neurobiologique du risque de récidive » :

> « *Finally, weaker cortico-striatal regulation predicted more frequent criminal convictions. These data suggest that cortico-striatal circuit dysregulation drives maladaptive decision making in psychopathy, supporting the notion that reward system dysfunction comprises an important neurobiological risk factor.* » (Hosking et al. 2017, p. 221).

Les conclusions de Leutgeb et ses collaborateurs (2005) sont tout aussi optimistes quant à la possibilité d'utiliser la neuroimagerie comme un outil d'évaluation du risque de récidive. Après avoir constaté que les réductions du volume de matière grise (corticale et sous-corticale) étaient associées à un risque de récidive violente nettement plus élevé, les chercheurs ont estimé que ces résultats « pouvaient être précieux pour améliorer les outils actuels d'évaluation du risque » (Leutgeb et al. 2005). Plus encore, tout en s'interrogeant sur la question de savoir « si les anomalies sont causalement liées à la criminalité, et surtout, si elles sont réversibles », ils suggèrent que les données neurobiologiques pourraient également jouer un rôle dans les mesures de réhabilitation, ainsi que pour informer les questions relatives à la détention préventive des individus :

> « *Science is still far from establishing neurobiological criteria as a contribution to final decisions in the court. However, the integration of objective neurobiological criteria, and an understanding of the interaction of biological and psychological factors could provide a higher quality of risk assessment. In addition, understanding neurobiological correlates of mental disorders could be helpful for improving therapeutic interventions. Lastly, a neurobiological evaluation of treatment success in forensic psychiatry would be of great value for the decision if patients should be discharged, or not*. » (Leutgeb et al. 2005).

3.5.2. Le pédocriminel

L'intérêt porté aux pédocriminels (et plus généralement, aux délinquants et criminels sexuels) remonte à la fin des années 1980, sous l'impulsion d'une série d'études utilisant la tomodensitométrie (CT-scan) sur des individus condamnés pour agression

sexuelle ou pédophilie (Graber et al 1982 ; Hucker et al. 1986, 1988 ; Langevin et al. 1988, 1989a, 1989b ; Hendricks et al. 1988 ; Wright et al. 1990).[141] Les résultats de certaines de ces études ont été résumés par Adrian Raine (1993) dans un tableau que nous présentons en annexe (Annexe n°9). Comme il le souligne lui-même, ces études présentent un certain nombre de limites liées notamment à la taille de l'échantillon et à la grande variabilité des résultats obtenus. Certaines études n'ont rapporté aucune différence significative entre les scans cérébraux des délinquants/criminels sexuels et des sujets témoins poursuivis pour des délits non-sexuels (Langevin et al. 1989a, 1989b). D'autres études s'accordent à montrer que les criminels sexuels sont associés à des anomalies structurelles au niveau du cortex temporal (Hucker et al. 1986, 1988 ; Wright et al. 1990 ; Langevin et al. 1990).

Malgré une grande hétérogénéité des résultats, ces études ont favorisé la multiplication des études IRM sur la pédophilie à partir des années 2000. Le paradigme expérimental est globalement similaire entre les études. La plupart reposent sur l'utilisation de l'IRMf et sont réalisées en présence d'un groupe contrôle (sujets sains). Il s'agit de comparer les enregistrements de l'activité cérébrale entre les deux groupes de sujets pendant la présentation de stimuli visuels sexuellement explicites (photographies d'adultes et d'enfants nus ou partiellement dénudés), puis pendant la présentation d'images neutres (photographies d'adultes et d'enfants habillés et/ou images non figuratives). Les données IRMf recueillies en réponse aux photographies d'enfants sont ensuite comparées aux données issues des deux autres catégories de stimuli (photographies d'adultes et images neutres). La présentation de stimuli visuels sexuellement explicites constitue le paradigme expérimental le plus fréquemment utilisé. Précisons néanmoins que certaines études incluent des tâches impliquant un contrôle attentionnel comme la *Go/NoGo task* (e.g. Habermeyer et al. 2013).

Comme l'ont relevé Mohnke et al. (2014), les études IRMf portant sur la pédophilie répondent généralement à trois exigences : identifier dans le groupe de pédophiles les régions préférentiellement activées pendant la présentation des stimuli sexuellement explicites ; isoler les différences d'activité cérébrale entre les pédophiles et le groupe contrôle pendant la présentation des différents stimuli ; et déterminer s'il existe

[141] Précisons que la grande majorité de ces études sont canadiennes et on été réalisé en combinant la tomodensitométrie avec des tests neurosychologiques. Pour une revue de la littérature de ces études, voir Joyal et al. (2007).

une spécificité fonctionnelle dans le traitement des stimuli entre les deux groupes (enfants pour les pédophiles, adultes pour les sujets sains). Globalement, les données collectées chez les pédophiles rapportent des déficits fonctionnels et structurels, notamment au niveau du cortex frontal et des régions sous-corticales. Les travaux sont cependant discordants concernant la localisation précise de ces anomalies, puisqu'elle démontre une grande variabilité selon les études : cortex préfrontal dorsolatéral et hypothalamus (Walter et al. 2007), amygdale (Sartorius et al. 2008 ; Ristow et al. 2016), ou encore cortex cingulaire et/ou insulaire (Poeppl et al. 2011 ; Ponseti et al. 2018).

Ces études répondent à une volonté d'utiliser la neuroimagerie fonctionnelle pour préciser les critères de diagnostic et de classification de la pédophilie (Poeppl et al. 2011 ; Ponseti et al. 2012, 2018), mais aussi pour isoler les anomalies propres à certaines caractéristiques phénotypiques. Des études ont en effet cherché à associer la nature des anomalies observées (structurelles ou fonctionnelles) avec l'histoire personnelle des individus et/ou les caractéristiques des infractions commises : antécédents de maltraitance ou d'abus sexuels pendant l'enfance (Walter et al. 2007), nombre de condamnations (Schiltz et al. 2007), ou encore âge des victimes (Poeppl et al. 2013). Cette dernière étude a par exemple montré qu'il existait chez les pédophiles une corrélation directe entre l'âge de leurs victimes et le volume de matière grise au niveau du cortex orbito-frontal et du gyrus angulaire, suggérant que « la préférence pour les enfants de bas âge est associée à une perte de matière grise » au niveau de ces régions (Poeppl et al. 2013, p. 3).

Sans entrer davantage dans le détail de ces études, il convient de résumer brièvement les principes communs sur lesquels s'appuient les recherches menées sur le cerveau des criminels psychopathes et des pédocriminels. Bien que les paradigmes expérimentaux diffèrent selon les catégories criminelles étudiées par les neurocriminologues, ces études obéissent à la même logique. On comprend bien que l'objectif n'est pas tant de préciser les critères de diagnostic de la psychopathie ou de la pédophilie, que d'isoler les caractéristiques neurobiologiques de criminels reconnus comme étant associés à des comportements typiques tels que l'impulsivité et l'insensibilité morale ou affective, et considérés comme des indicateurs du risque de récidive. Autrement dit, l'intérêt porté aux criminels psychopathes ou aux pédocriminels ne tient pas seulement au fait qu'ils incarnent la figure emblématique de la dangerosité ; il s'explique surtout par la possibilité d'opérationnaliser leurs comportements sur le plan neuroscientifique, afin d'isoler des marqueurs neurobiologiques du risque de récidive.

Ces deux figures de la criminalité apparaissent comme un objet d'étude privilégié pour isoler les indicateurs prédictifs des comportements impulsifs, violents, ou révélateurs d'une grande insensibilité morale ou affective. En ce sens, cette stratégie peut être mise en parallèle avec la logique qui sous-tendait les études EEG menées sur les sujets épileptiques, lesquels offraient aux neurocriminologues la possibilité d'identifier les indicateurs neurophysiologiques d'un comportement imprévisible et violent. Les stratégies de recherche mobilisées depuis les années 1990 relèvent du même raisonnement prototypique : les neurocriminologues cherchent à révéler les indicateurs prédictifs d'un comportement dangereux, tout en se dispensant de rechercher les mécanismes neurobiologiques impliqués.

Précisons à cet égard que si la neurocriminologie contemporaine a bénéficié d'une plus grande rigueur méthodologique – par la construction de paradigmes expérimentaux permettant d'isoler les phénomènes psychologiques d'intérêt −, force est de constater qu'elle ne parvient pas à isoler les caractéristiques neurobiologiques distinctives des catégories criminelles étudiées. Bien que les techniques de neuroimagerie offrent une résolution spatiale nettement plus fine que l'EEG, elles n'ont finalement pas tant nourri le champ de la neurocriminologie, dans la mesure où la localisation des anomalies témoigne d'une grande variabilité selon les études : le statut épistémologique de la relation établie entre les anomalies cérébrales et le comportement criminel est limité par le manque de spécificité fonctionnelle des régions considérées. En outre, une grande partie des études menées sur les criminels psychopathes et les pédocriminels reposent sur la neuroimagerie anatomique (IRMa), qui ne nécessite pas d'identifier une tâche cognitive donnée pour mesurer l'activité cérébrale des individus : ces résultats ne nous disent finalement rien sur les comportements cibles (l'impulsivité, l'insensibilité morale ou affecive, etc.) ; tout au plus permettent-ils de recenser les irrégularités et les déficits structurels qui distinguent les criminels de la population générale.

Enfin, il est intéressant de noter la résurgence d'une stratégie particulièrement représentative des études EEG menées dans les années 1940-1980 : celle consistant à associer les données neurobiologiques avec un ensemble de caractéristiques relatives aux traits psychologiques des individus, à la nature des crimes commis, ou au nombre de condamnations pénales. Les résultats ne sont d'ailleurs pas plus probants ou explicites sur le plan épistémologique puisque l'appréciation de la dangerosité se mesure, ici encore, par l'excès d'anomalies cérébrales : à défaut d'isoler des « profils neurobiologiques » propres à chaque type de criminalité, le seul consensus qui se dégage de ces études est

l'association établie entre le degré d'anomalies et le niveau de dangerosité de l'individu. Malgré tous les efforts déployés par les neurocriminologues pour fonder les catégories criminelles sur des catégories neurobiologiques spécifiques, ils tendent seulement à faire de l'excès d'anomalies cérébrales la condition de possibilité de toutes les formes de dangerosité. À ce titre, le lien de proportionnalité établi entre le niveau de déficit structurel et le score de psychopathie (Leutgeb et al. 2005) ou l'âge des victimes des pédocriminels (Poeppl et al. 2013) est particulièrement révélateur de l'ambition de fonder le critère de dangerosité sur le degré d'anomalies cérébrales : l'argument ne repose pas sur l'identification d'un mécanisme neurobiologique pleinement caractérisé, mais seulement sur l'observation d'une forte proportion d'anomalies cérébrales.

Ces limites épistémologiques nous amènent à une dernière question : si l'ambition de la neurocriminologie est d'isoler les marqueurs neurobiologiques du risque de récidive, dans quelle mesure les tâches cognitives utilisées dans les études de neuroimagerie fonctionnelle permettent-elles de rendre compte d'un comportement aussi complexe que le contrôle de soi ou la formation d'un jugement moral ? En clair, les modalités d'opérationnalisation des comportements impulsifs ou immoraux peuvent-elles apporter des informations utiles à la prédiction du risque de récidive ?

3.6. Le problème de l'opérationnalisation

Le problème de la portée prédictive, explicative ou normative des données issues de la neuroimagerie fonctionnelle est indissociable de la question de la pertinence des tâches cognitives utilisées dans les protocoles expérimentaux. Comme nous l'avons mentionné plus haut (section 3.3.), la construction du paradigme expérimental repose sur deux hypothèses : l'idée que les tâches cognitives permettent de mesurer les capacités fonctionnelles des individus d'une part, et d'autre part, que les performances individuelles sont équivalentes aux comportements tels qu'ils sont exercés dans un environnement social complexe. La question de la pertinence explicative des tâches utilisés peut alors être abordée sur un plan conceptuel et empirique, en interrogeant à la fois la validité conceptuelle du modèle – en examinant les critères de définition du comportement que l'on cherche à isoler – et la validité de la mesure, c'est-à-dire en déterminant si le modèle expérimental est pertinent pour mesurer ce comportement, et s'il le représente sous tous

ses aspects. Le problème de la pertinence et de la portée explicative des tâches utilisées nous invite également à analyser des contraintes épistémologiques qui pèsent sur la construction des modèles expérimentaux : la question est de savoir dans quelle mesure la « vision neurocriminologique » oriente les choix méthodologiques opérés dans la conceptualisation et l'opérationnalisation des comportements individuels.

À partir de ces trois niveaux d'analyse, on peut questionner les modalités d'opérationnalisation du contrôle de soi, telles qu'elles sont mises en œuvre dans la recherche en neurocriminologie. Ce concept occupe une place centrale dans les investigations empiriques, car les neurocriminologues semblent s'inspirer d'une théorie proposée par le sociologue Travis Hirschi et le criminologue Michael Gottfredson dans un ouvrage intitulé *A General Theory of Crime* (1990). Bien qu'elle ne soit jamais mentionnée dans la littérature en neurocriminologie, cette théorie a certainement eu une influence sur les choix méthodologiques opérés par les neurocriminologues car elle défend l'idée selon laquelle la perte du contrôle de soi constitue la source principale de tous les comportements criminels : selon Hirschi et Gottfredson, la capacité à contrôler son comportement constitue le critère décisif pour évaluer la propension d'un individu à commettre un crime. La perte du contrôle de soi recouvre selon eux quatre aspects : la recherche d'une gratification immédiate et concrète, la tendance à s'engager dans des comportements à risque, l'incapacité à envisager et maintenir un objectif sur le long terme, et l'égocentrisme (Gottfredson and Hirschi 1990, p. 89). Compte tenu des implications que pourraient présenter les modèles neuroscientifiques du contrôle de soi dans l'évaluation du risque de récidive, il paraît crucial d'évaluer la valeur épistémique des protocoles expérimentaux conçus par les neurocriminologues, tant sur le plan méthodologique que conceptuel.

Le problème épistémologique posé par l'opérationnalisation du contrôle de soi se situe sur deux plans. Le premier concerne la nature du modèle utilisé par les neurocriminologues. La plupart des études ont recours à la *Go/NoGo task,* qui permet d'évaluer la capacité d'un individu à inhiber une action motrice.[142] Rares sont les études qui utilisent des tâches cognitives permettant de mesurer la dimension prospective du contrôle de soi comme la *Delay Discounting Task,* qui invite le sujet à faire le choix entre

[142] Le participant est invité à effectuer une action motrice (par exemple, appuyer sur un bouton) en réponse à un certain stimulus (condition *Go*), et à inhiber cette action en réponse à un autre stimulus (condition *NoGo*).

une petite récompense immédiate ou une grosse récompense tardive,[143] ou la *Iowa gambling task*, qui permet d'évaluer la capacité à évaluer la balance bénéfice/risque de ses actions. Si ces paradigmes expérimentaux posent eux-mêmes un certain nombre de difficultés épistémologiques, on peut considérer qu'ils sont nettement plus pertinents du point de vue du comportement cible. Le fait de privilégier la *Go/NoGo task* au détriment des autres tâches cognitives nous renseigne d'ailleurs sur la manière dont les neurocriminologues conceptualisent l'impulsivité criminelle : le modèle ne s'intéresse qu'à la cause proximale de l'action (l'incapacité à inhiber une « pulsion » motrice) et ne rend pas compte de la dynamique en amont de son initiation. De fait, le modèle sous-spécifie le comportement tel qu'il est exercé dans l'environnement naturel, car les neurocriminologues font l'impasse sur les questions relatives à la définition de leur objet expérimental : ils échouent à isoler les critères de définition du concept qu'ils cherchent à opérationnaliser, réduisant le contrôle de soi à un « construit hypothétique », dépouillé de son caractère prospectif, dynamique et situationnel.

Le second problème concerne les conclusions tirées par les neurocriminologues du point de vue de l'applicabilité pratique de leurs résultats. Il semble que le caractère artificiel (décontextualisé et sous-spécifié) du modèle introduise un biais définitionnel dans l'interprétation des données, de par la nature des finalités recherchées : les anomalies cérébrales associées à la réalisation des tâches cognitives sont interprétées comme des marqueurs neurobiologiques des comportements impulsifs (et par conséquent du risque de récidive), alors même que les conditions d'opérationnalisation de ces comportements effacent tous les aspects normatifs qui leur sont inhérents. Dans la mesure où les dimensions contextuelle, prospective, morale et socio-affective du contrôle de soi sont exclues des conditions expérimentales, elles ne peuvent être réintroduites *a posteriori*, pour informer la capacité de l'individu à envisager les conséquences de son acte sur le long terme, ou à contrôler son comportement conformément à des raisons morales.

Finalement, on pourrait se satisfaire de ce modèle s'il avait pour seule ambition d'isoler les régions cérébrales impliquées dans l'inhibition d'une action motrice. Cependant, dès lors que ces résultats engagent une fonction prescriptive, la fiabilité du modèle se heurte au problème de l'inférence inverse. L'inférence inverse est un raisonnement qui consiste à inférer l'exercice d'un comportement ou d'un processus

[143] Cette tâche cognitive s'inspire du *Delay of Gratification Test* (ou communément appelé « test du marshmallow ») elaboré par Walter Mischel and Ebbe B. Ebbesen (1970).

psychologique à partir d'un état cérébral donné. Il a été très largement discuté dans le domaine des neurosciences et de la philosophie des neurosciences (e.g. Poldrack 2006, 2011 ; Machery 2014). Comme l'a montré notamment Russel Poldrack, l'inférence inverse peut servir d'outil heuristique pour préciser le statut de la relation entre les états cérébraux et les processus mentaux, mais elle est limitée par la spécificité fonctionnelle des régions considérées : une telle inférence est valide si et seulement si la région R est activée uniquement lorsque la fonction F est engagée. Compte tenu de la grande variabilité des régions cérébrales isolées par les neurocriminologues – et de la grande variabilité des protocoles expérimentaux eux-mêmes −, la portée prédictive, explicative ou normative des résultats s'en trouve affectée : le manque de spécificité fonctionnelle des régions isolées ne peut pas garantir la solidité épistémologique de la relation entre les anomalies cérébrales et la manifestation d'un comportement impulsif.

On pourrait d'ailleurs invoquer les mêmes arguments en ce qui concerne les modalités d'opérationnalisation du jugement moral, qui sous-tend la plupart des études menées sur les criminels psychopathes. Le paradigme expérimental repose généralement sur la présentation d'un dilemme sacrificiel inspiré du « dilemme du tramway », initialement formulé par Philippa Foot (1967). Ici encore, il se pose la question de savoir 1/ si ce modèle est pertinent vis-à-vis du concept psychologique que l'on cherche à étudier (l'incapacité à évaluer la valeur morale de son acte) ; 2/ si les données recueillies pendant la résolution de ces dilemmes garantissent une explication générale des jugements moraux ; et 3/ si elles permettent de tirer des conclusions normatives utiles à l'évaluation du risque de récidive de l'individu. Si l'on admet que les jugements moraux (au même titre que le contrôle de soi) peuvent faire l'objet d'une investigation neuroscientifique, il reste à s'assurer que le modèle expérimental est pertinent vis-à-vis des finalités poursuivies par les neurocriminologues. Or, la possibilité d'utiliser ces modèles pour évaluer le risque de récidive paraît doublement injustifiée : non seulement parce que les conditions de validité conceptuelle et de validité de la mesure ne sont pas remplies, mais aussi parce que le manque de spécificité fonctionnelle des régions isolées limite considérablement la portée explicative et normative des données.

La question de l'apport potentiel des données d'imagerie cérébrale dans l'évaluation du risque de récidive a été abordée sous l'angle du problème de l'opérationnalisation d'un comportement socialement complexe. Il aurait été également judicieux de mentionner les limites techniques, méthodologiques et théoriques de la

neuroimagerie fonctionnelle, qui fragilisent encore davantage le statut épistémologique de la relation entre les anomalies cérébrales et les comportements impulsifs ou immoraux. Sans entrer dans le détail de cette discussion, il convient de préciser que l'IRMf ne mesure pas *directement* l'activité cérébrale mais seulement une conséquence métabolique de cette activité (l'augmentation de la concentration en sang oxygéné). Son principe de fonctionnement repose sur une chaîne inférentielle qui relie l'activité neuronale et le signal BOLD (*blood-oxygen-level dependent*), à savoir le signal qui détecte la suroxygénisation du sang au niveau des régions mobilisées pendant la réalisation de la tâche cogntive. En clair, la valeur épistémique des données issues de l'IRMf repose sur l'hypothèse selon laquelle une information sur la concentration en sang oxygéné dans une région donnée nous renseigne sur l'implication de cette région dans la réalisation de la tâche cognitive (Forest 2014).

Le pouvoir explicatif des données issues de l'IRMf dépend donc de la « distance inférentielle » entre les différents niveaux d'organisation qui séparent le signal BOLD et l'activité cognitive (Roskies 2007). Comme l'a montré Denis Forest, l'une des manières de résoudre ce problème consiste à multiplier les moyens d'investigations empiriques en diversifiant les méthodologies et les outils d'analyse. Dès lors, l'implication d'une région donnée dans l'exercice d'une fonction psychologique peut être acceptée comme valide en fonction de sa cohérence avec la connaissance d'arrière-plan obtenue par d'autres moyens d'investigation : la valeur épistémique de la relation entre l'état cérébral et l'état psychologique est justifiée si elle fait partie « d'un ensemble de propositions qui se soutiennent et se confirment mutuellement sur la base d'enquêtes indépendantes et instructives » (Forest 2014, p. 57). Par exemple en ce qui concerne la neurocriminologie, la question de savoir si cette exigence est satisfaite nous engagerait à comparer les résultats d'imagerie cérébrale avec les résultats obtenus dans les études privilégiant l'approche lésionnelle ou neuropsychologique.

Si nous avons insisté sur les contraintes épistémologiques qui pèsent sur la construction des modèles expérimentaux, c'est parce qu'elles soulignent la nécessité de questionner la nature de l'explication fournie par la neurocriminologie au regard des finalités qu'elle poursuit. De toute évidence, l'efficacité prédictive des données restera limitée à une possibilité statistique tant que les neurocriminologues ne se seront pas accordés sur un protocole pertinent vis-à-vis du phénomène psychologique qu'ils cherchent à étudier, et surtout tant qu'ils n'auront pas mis en évidence une spécificité fonctionnelle entre les régions isolées et l'exercice du phénomène psychologique. La

question de l'efficacité explicative des données n'a quant à elle pas été totalement résolue car elle ne peut être dissociée de la « vision neurocriminologique » telle que nous l'avons définie au début de ce chapitre. Il semble en effet que l'ambition de tirer des conclusions normatives sur la seule base des données neurobiologiques ne repose finalement pas sur une stratégie explicative au sens d'une explication *mécaniste,* qui permettrait d'accéder à la dynamique fonctionnelle de la relation entre les anomalies cérébrales et les comportements impulsifs, violents ou agressifs. Bien que les neurocriminologues cherchent à s'implanter sur le terrain des neurosciences cognitives, ils s'interdisent de préciser le statut épistémologique de la relation entre leurs données empiriques et les catégories psychologiques qu'ils cherchent à naturaliser. Leur logique éliminativiste vis-à-vis des autres stratégies explicatives révèle une volonté hégémonique de déployer leur champ d'action sur le terrain judiciaire, en se dispensant d'une recherche laborieuse (et véritablement neuroscientifique) qui consisterait à identifier les mécanismes impliqués dans l'altération des fonctions cognitives des individus.

CHAPITRE 4

CERVEAU « MONSTRUEUX » OU CERVEAU « MALADE » ?

Si le chapitre précédent s'est attaché à analyser le cadre méthodologique des recherches en neurocriminologie contemporaine, nous allons ici concentrer notre attention sur les concepts mobilisés dans l'interprétation des données empiriques. La manière dont le cerveau des criminels est représenté à l'intérieur du champ nous invite à identifier les conditions épistémologiques pour lesquelles les données neurobiologiques pourraient affecter les questions relatives à la responsabilité et à la prise en charge des criminels. Contrairement à ce que l'on pourrait attendre du projet neurocriminologique, les données neurobiologiques recueillies chez les criminels n'ont pas vocation à servir la stratégie rhétorique du « *my brain made me do it defense* », c'est-à-dire la recherche d'une preuve de l'irresponsabilité pour cause psychiatrique. Les neurocriminologues ne s'intéressent pas à la question de la responsabilité des criminels présentant d'importantes anomalies cérébrales car leur projet semble être plutôt tourné vers une logique médicale : tout au long du développement de la neurocriminologie, on voit s'enraciner un discours invitant à traiter le problème de la criminalité comme un problème médical.

La question de la médicalisation du crime recoupe certains enjeux propres au champ du droit pénal et de la sociologie ou de la philosophie de la psychiatrie. Elle a été largement discutée en France à partir des années 1970, notamment sur la base des travaux menés par Michel Foucault (1972, 1975) et Robert Castel (1976). Globalement, ces enjeux s'articulent autour du problème de la psychiatrisation du droit pénal et de la confusion entretenue entre le crime, les troubles mentaux et les troubles de la personnalité. Encore aujourd'hui, les nouvelles stratégies préventives se sont accompagnées d'une médicalisation accrue de la déviance, l'approche médicale du crime étant perçue comme une réponse possible à l'impératif de protection sociale (Mbanzoulou et al. 2008 ; Danet

2008 ; Kaluszynski 2008).[144] Depuis les années 1980 aux Etats-Unis, les mesures pénales en matière de lutte contre la récidive contribuent à brouiller les frontières entre la criminalité, la psychiatrie et la psychopathologie, en imposant une « tolérance zéro » pour les criminels souffrant de troubles sévères de la personnalité : la notion de dangerosité est confondue dans une indistinction croissante entre les maladies mentales et les troubles de la personnalité.

Si l'approche médicale du crime et ses implications sur les notions fondamentales du droit pénal ont fait l'objet d'une importante littérature, la dimension neurobiologisante de cette stratégie a été peu traitée.[145] Pourtant depuis le milieu des années 2000, une nouvelle dynamique de médicalisation du crime a émergé en parallèle des recherches utilisant la neuroimagerie sur les populations criminelles, dont la logique s'est implantée à l'intérieur même du champ de la neurocriminologie. Certains neurocriminologues comme Adrian Raine assument la volonté d'inscrire le criminel dans un modèle psychopathologique, en promouvant la nécessité de redéfinir le problème de la criminalité comme un problème de santé publique (Raine 1993, 2014). Bien que cette dynamique de recherche s'inspire des données statistiques rapportant une fréquence élevée de troubles psychiatriques dans les populations carcérales (e.g. Schiltz et al. 2013), elle ne répond pas à une stratégie de *psychiatrisation* du droit pénal. Elle est gouvernée par sa propre logique, puisqu'elle est pensée de manière à favoriser une extension du champ d'intervention de la neurocriminologie.

Dans ce chapitre, nous proposons d'analyser les fondements empiriques de cette stratégie de pathologisation et de médicalisation du criminel. Au-delà des enjeux proprement éthiques d'une telle approche, cette ambition pose un problème épistémologique majeur car les données issues de la neuroimagerie fournissent un type

[144] Ces réflexions s'inscrivent dans le cadre d'un vif débat concernant le principe de la rétention de sûreté, introduit dans le droit pénal français en février 2008 par la loi relative à la déclaration d'irresponsabilité pénale pour cause de trouble mental. Précisons qu'une procédure similaire avait déjà été adoptée en Grande-Bretagne en 1999, sous l'influence des travaux menés par Charles Murray et Richard J. Herrnstein dans *The Bell Curve* (1994).

[145] Elle a été abordée sur le plan théorique par le sociologue Julien Larregue (2019, 2020), dans le cadre d'une reconstruction historique de la criminologie biosociale. Selon lui, cette dynamique de recherche (qui comprend rappelons-le, le champ de la neurocriminologie) s'inscrit dans une logique de médicalisation du contrôle social qui prend sa source dans la criminologie et la psychiatrie du XIX^{e} siècle.

d'explication qui se prête facilement à une possibilité d'intervention : elles offrent la possibilité d'identifier un marqueur neurobiologique sur lequel on peut agir directement grâce aux techniques de neurostimulation comme la stimulation cérébrale profonde (*Deep Brain Stimulation* ou DBS).[146] Si l'on veut évaluer la légitimité scientifique de cette ambition, il convient donc de déterminer d'une part, si la conception psychopathologique du criminel est justifiée sur le plan empirique, et d'autre part, si la stratégie neuro-interventionniste est efficace du point de vue des finalités pratiques et épistémiques de la neurocriminologie.

Nous examinerons en premier lieu le statut épistémologique des catégories conceptuelles mobilisées dans la littérature en neurocriminologie contemporaine. Nous verrons que la caractérisation des anomalies cérébrales des criminels relève d'une double qualification qui nous engage à questionner la portée explicative de ces données au regard du problème de la responsabilité pénale et de la prise en charge des criminels (section 4.1). Nous montrerons dans un deuxième temps que l'interprétation des données neurobiologiques est entièrement dictée par la nature des finalités poursuivies par les chercheurs : la stratégie de médicalisation du crime s'inscrit au cœur même de la démarche neurocriminologique, dans le sens où elle introduit un biais de confirmation en faveur d'une conception psychopathologique du criminel (section 4.2).

Dans la troisième section, nous décrirons les différentes voies de recherche que recouvrent ce projet neuro-interventionniste : des enjeux proprement théoriques de ce qui est appelé la « neuroaugmentation morale », aux investigations empiriques visant à utiliser la stimulation cérébrale profonde pour corriger le cerveau des criminels. Au-delà des enjeux liés à l'acceptabilité morale ou à la faisabilité pratique de cette stratégie neuro-interventionniste, nous nous intéresserons surtout à la question de son apport épistémique dans l'explication neurobiologique du comportement criminel (section 4.4) : plutôt que de s'attarder sur les enjeux éthiques de la médicalisation du criminel, nous porterons notre analyse sur le terrain épistémologique en examinant l'argument selon lequel les

[146] La stimulation cérébrale profonde est une technologie neurochirurgicale développée au milieu des années 1980. Elle consiste à moduler l'activité des populations neuronales par le biais d'électrodes connectées à un boîtier de contrôle, dont la fréquence est ajustable et réversible. Initialement développée pour le traitement de troubles neurologiques comme la maladie de Parkinson, son application a été étendue à certains troubles psychiatriques ou neuropsychiatriques résistants aux traitements conventionnels, comme le syndrome de Gilles de la Tourette ou le trouble obsessionnel compulsif (e.g. Holtzheimer and Mayberg 2011).

techniques de neurostimulation permettent de confirmer l'influence causale des anomalies cérébrales dans la manifestation du comportement criminel. En somme, nous examinerons successivement la question de la pertinence et de la légitimité scientifique de la stratégie neuro-interventionniste, puis la question de son efficacité explicative : si la démarche localisatrice adoptée au début des années 1990 échoue à isoler les mécanismes neurobiologiques impliqués dans la manifestation du comportement criminel, l'usage des techniques de neurostimulation permet-il de statuer sur la responsabilité causale de telle ou telle région cérébrale ?

Précisons que notre analyse sera conduite selon une approche empirique qui n'a pas pour vocation première de nourrir les réflexions philosophiques (ou juridiques) sur la médicalisation du crime. Il s'agit surtout de prendre pour objet les arguments mobilisés à l'intérieur du champ de la neurocriminologie, en les mettant en perspective avec la « vision neurocriminologique » telle que nous l'avons décrite dans le chapitre précédent. Notre réflexion progressera ainsi en trois étapes. Nous procèderons d'abord à une analyse interne au projet avant de sortir du champ de la neurocriminologie basée sur la neuroimagerie, pour examiner les nouvelles perspectives de recherche qui se déploient depuis le milieu des années 2000. Enfin, pour répondre à la question de l'apport épistémique de la stratégie neuro-interventionniste, nous situerons notre analyse sur le plan de la philosophie des neurosciences, en prenant comme point de départ quelques réflexions épistémologiques générales sur la nature de l'explication neuroscientifique.

4.1. Une terminologie médicale

4.1.1. Un cerveau défaillant, déficient, irrégulier

En parcourant la littérature publiée depuis le début des années 1990, on constate que la caractérisation du cerveau des criminels use d'une terminologie relativement nuancée, mais qui renvoie essentiellement à trois états. Quelles que soient les catégories criminelles examinées, les anomalies observées se traduisent principalement par des déficits, des atrophies et des asymétries. Autrement dit, la description des données neurobiologiques est toujours corrélative d'un dysfonctionnement, d'une insuffisance ou d'une irrégularité : réduction du métabolisme du glucose, réduction du volume de matière grise, réduction de la connectivité fonctionnelle, atrophies, asymétries structurelles ou fonctionnelles, etc. Parmi les concepts les plus employés dans la littérature, on peut citer : « *deficit* » ; « *deficiencies* » ; « *reduction* » ; « *dysfunction* » ; « *dysregulation* » ; « *impairment* » ; « *alterations* » ; « *atrophy* » ; « *asymmetry* » ; « *aberrant* » ; « *abnormal activation* » ; « *abnormal brain structure* » ; « *abnormal shape* », etc.

Avant d'expliciter le statut épistémologique de ces catégories conceptuelles, une précision importante doit être apportée en ce qui concerne les concepts exprimant l'état d'irrégularité. La caractérisation du cerveau des criminels n'est pas toujours associée à des déficiences, car elle fait souvent mention d'un excès d'activité dans certaines régions cérébrales, en particulier dans les structures sous-corticales comme l'amygdale (e.g. Raine et al. 1998b). Pour autant, il est intéressant de noter que ces anomalies ne sont pas considérées pour elles-mêmes, car elles sont systématiquement mises en relation avec les déficits observées dans les autres régions cérébrales. L'excès d'activité dans une région donnée n'est décrit qu'en référence à une insuffisance observée dans une autre région, de sorte que ces anomalies sont finalement reléguées à un état de « déséquilibre cérébral ».

Cette notion de « déséquilibre » est régulièrement évoquée dans les études utilisant la neuroimagerie fonctionnelle, en particulier sous les termes de « déséquilibre inter-hémisphérique » ou de « déséquilibre fronto-limbique » (Hecht 2011 ; Angus et al. 2016 ; Siep et al. 2018). Il s'agit globalement de montrer l'existence d'une dérégulation des systèmes cognitifs et émotionnels, à partir des irrégularités ou des asymétries fonctionnelles observées entre les deux hémisphères (e.g. Raine et al. 1998b, 2001 ;

Müller et al. 2003) ou entre les régions corticales et sous-corticales (e.g. Raine et al. 1997a ; Yang et al. 2010 ; Poeppl et al. 2011). Dans ce dernier cas, il s'agit surtout de mettre en relation les altérations du circuit neuronal reliant le cortex préfrontal à l'amygdale d'une part, avec les incapacités fonctionnelles des individus d'autre part. Adrian Raine défend par exemple un lien entre l'impulsivité de certains criminels, et le déséquilibre induit par l'« hyperactivité sous-corticale » et l'« hypoactivité préfrontale » :

> « *Results support the hypothesis that emotional, unplanned impulsive murderers are less able to regulate and control aggressive impulses generated from subcortical structures due to deficient prefrontal regulation. It is hypothesized that excessive subcortical activity predisposes to aggressive behaviour, but that while predatory murderers have sufficiently good prefrontal functioning to regulate these aggressive impulses, the affective murderers lack such prefrontal control over emotion regulation.* » (Raine et al. 1998b, p. 319).

Dans les études de neuroimagerie anatomique, la notion de « déséquilibre » se rapporte davantage à un état de disproportion ou d' irrégularité des structures cérébrales : les études font très souvent mention d'« atrophies » (e.g. Schiltz et al. 2007 ; Tiihonen et al. 2008 ; Ly et al. 2012), d'« asymétries » (e.g. Raine et al. 2004 ; Boccardi et al. 2010), d' « altérations » ou d' « aberrations » morphologiques (e.g. Yang et al. 2010 ; Ermer et al. 2012 ; Korponay et al. 2017). Ces anomalies structurelles ne renvoient ici plus tant à une désorganisation ou à un désordre cérébral, qu'à un état de difformité ou de défaut de développement. Les irrégularités observées apparaissent comme le résultat d'un agrégat de déficits et d'atrophies inégalement répartis dans le cerveau, qui confèrent à ce dernier un aspect globalement disproportionné.

Enfin, si l'on tient compte de l'ensemble des études (anatomiques et fonctionnelles) publiées depuis le début des années 1990, les références à la difformité se rapportent également à un excès d'anomalies cérébrales. Rappelons en effet que tout au long du développement de la neurocriminologie, le critère de dangerosité de l'individu est fixé par son degré d'anomalies cérébrales : les criminels récidivistes et qui commettent des crimes violents ou sans raison sont systématiquement associés au plus haut degré d'anomalies (Raine et al. 1998b). Dans ce cas, la caractérisation neurobiologique des criminels répond moins à une différence qualitative qu'à une différence de degré, puisque les références à la difformité apparaissent comme un critère de mesure (quantitative) de

la dangerosité de l'individu. Autrement dit, les irrégularités ou les disproportions observées chez les criminels participent à la construction d'une typologie indexée sur leur niveau de dangerosité. Il ne s'agit pas seulement de mettre en relation une anomalie morphologique avec une incapacité fonctionnelle donnée, mais de faire de l'excès d'anomalies le stigmate caractéristique de la dangerosité de l'individu.

4.1.2. Une double qualification

Globalement depuis le début des années 1990, l'interprétation des données neurobiologiques recueillies chez les criminels fait état de dysfonctionnements, d'insuffisances et d'irrégularités. Ainsi, si la caractérisation du cerveau des criminels dépend en grande partie de l'outil de neuroimagerie (anatomique ou fonctionnelle) utilisé, la terminologie employée relève essentiellement du champ lexical de la défaillance, de la déficience ou de la difformité. Il paraît important de préciser le statut épistémologique de ces trois catégories conceptuelles, car elles nous renseignent sur la manière dont le criminel est représenté à l'intérieur du champ de la neurocriminologie. De toute évidence, les deux premières catégories favorisent la représentation d'un « cerveau défectueux », caractérisé non pas tant par ses différences, mais surtout par ses manques et ses faiblesses. Elles ne décrivent pas une simple figure de l'altérité correspondant à un écart par rapport à une norme préétablie ; elles traduisent une « infériorité biologique » qui se décline sur le registre de l'infirmité et de la défectuosité.

Les anomalies observées sont généralement mises en relation avec les incapacités fonctionnelles des individus (notamment leur incapacité à contrôler leur comportement), mais aussi avec les caractéristiques relatives aux infractions commises (nature de la condamnation, fréquence de récidive, etc.). Ce modèle explicatif vise à montrer que le comportement criminel est entièrement expliqué par la présence d'anomalies cérébrales : il consiste à assimiler les catégories juridiques, psychologiques et neurobiologiques, dans l'objectif de relier le crime à son auteur. Plus précisément, il s'agit d'associer les insuffisances observées sur le plan neurobiologique à une inaptitude ou un défaut qui pourrait faciliter l'expression d'un comportement criminel. Il en résulte un modèle psychopathologique qui qualifie l'individu sous les mêmes traits que son cerveau. La caractérisation du cerveau du criminel opère ainsi un glissement conceptuel en faveur

d'une conception pathologique du sujet lui-même : de l'observation d'une « anomalie » traduisant un écart à la moyenne, à l'identification d'une « anormalité » qui relève du domaine de la pathologie. D'un concept empirique décrivant les données neurobiologiques recueillies chez les criminels, l'interprétation des résultats finit par introduire un concept normatif axé sur l'individu lui-même, et qui recouvre une qualification médicale.

Les références à la difformité viennent compléter ce tableau clinique en accentuant encore davantage les écarts par rapport à la norme biologique qui caractérise le non-criminel. Elles marquent une extrême variation qui dépasse largement le domaine de l' « anormalité biologique » pour entrer dans celui de la « monstruosité » : les irrégularités, les aberrations et les disproportions observées chez le criminel expriment une altérité radicale qui l'oppose en tous points au sujet issu de la population générale, et dont on peut se demander si elle relève véritablement du registre de la pathologie. Ces anomalies étendent le champ de l' « anormal » au domaine du monstrueux, au travers d'une terminologie qui ne s'inscrit pas à proprement parler dans un discours médical. Elles traduisent un écart qualitatif (morphologique), qui ne tient plus seulement à une différence de degré mais à une différence de *nature*. Les irrégularités et les asymétries apparaissent finalement comme la somme de toutes les déviations, les stigmates d'une aberration qui complète le caractère proprement pathologique des défaillances et des déficiences.

En somme, la terminologie employée pour caractériser le cerveau des criminels révèle la nature équivoque d'un individu porteur à la fois d'une infirmité et d'une difformité. Elle inscrit le criminel dans le domaine confus de la pathologie et de la monstruosité, en le reléguant alternativement à la catégorie du défaut, de l'incomplétude et de l'aberration. Or, si l'on poursuit cette analyse à la lumière des problématiques relevant du champ pénal, cette double qualification présente des implications majeures sur la notion de responsabilité et la question de la prise en charge de la criminalité. Dans un sens, le recours à une terminologie médicale pour décrire le cerveau des criminels pourrait contribuer à les disqualifier sur le plan juridique, en favorisant une explication pathologique (et par conséquent une conception déresponsabilisante) de la criminalité : le criminel n'appartiendrait plus à la sphère juridique, mais bien plutôt à la sphère médicale. En revanche, si l'on tient compte de sa dimension tératologique, les questions d'ordre juridique apparaissent sous un autre angle. Alors que les catégories de la défaillance et de la déficience appellent une prise en charge médicale, les références à la

difformité semblent recentrer le problème autour de la notion de dangerosité. Le « criminel monstrueux » réapparaît comme une catégorie juridique à part entière, c'est-à-dire comme un sujet susceptible de répondre à la question de la responsabilité pénale ; mais surtout, les références à la difformité réintroduisent une logique centrée sur la mesure quantitative du danger qu'il représente. Elles réaffirment sa dangerosité en marquant le passage d'une différence de degré à une différence de nature, et en consolidant l'association établie entre le degré d'anomalies cérébrales et le niveau de dangerosité de l'individu.

Ainsi, en questionnant le statut épistémologique des catégories conceptuelles mobilisées pour décrire le cerveau des criminels, on constate que leur caractérisation ne conduit pas seulement à les opposer à la norme biologique qui caractérise le non-criminel ; elle invite à une reconfiguration des problématiques liées à la responsabilité et à la dangerosité de l'individu. Comme nous venons de le voir, les catégories de l'infirmité et de la difformité semblent répondre à deux logiques différentes : une logique médicale et thérapeutique − qui dissout le statut juridique du criminel − et une logique centrée sur la question de sa dangerosité, qui garantit l'imputabilité du crime à son auteur. Le passage de la catégorie de l' « anormalité » à la catégorie de la « difformité » révèle un glissement épistémologique dans lequel le critère de dangerosité se substitue à la question de la responsabilité : de la représentation d'une « dangerosité pathologique » à une dangerosité perçue comme un état permanent, qui certifie dans le même temps l'imputabilité et l'incorrigibilité de l'individu.

4.1.3. La question de la validité empirique de l'interprétation pathologique

Compte tenu de l'importance de ces enjeux vis-à-vis des notions fondamentales du droit pénal, il convient de préciser notre analyse sur un plan proprement empirique. Si l'on veut prendre la mesure de la portée des résultats issus de la neurocriminologie, il devient nécessaire d'interroger le pouvoir explicatif (réel) des anomalies cérébrales, notamment en ce qui concerne leur relation avec le comportement criminel. Par ailleurs, le recours à une terminologie médicale pour décrire le cerveau des criminels nous invite à préciser les conditions pour lesquelles les données empiriques pourraient effectivement affecter les questions relatives à leur responsabilité et à leur prise en charge.

Notre analyse conceptuelle laisse alors apparaître deux problèmes épistémologiques fondamentaux. Le premier concerne le glissement conceptuel opéré dans l'interprétation des données, entre ce qui relève de l'« anomalie » (concept empirique) et de l' « anormalité » (concept normatif). On peut en effet se demander si la caractérisation du cerveau des criminels répond à un modèle véritablement descriptif. Dans la mesure où la terminologie employée favorise une reconfiguration des problématiques liées à la responsabilité et à la dangerosité de l'individu, il se pose la question de savoir si elle est entièrement justifiée sur le plan empirique, c'est-à-dire si elle est compatible avec les données issues de la neuroimagerie. Le problème apparaît encore plus nettement si l'on se réfère à ce que nous avons décrit plus haut à propos de la « vision neurocriminologique ». Si l'on admet que celle-ci structure toutes les étapes de la construction des études, on peut s'attendre à ce que l'interprétation des données soit directement subordonnée aux finalités poursuivies par la neurocriminologie. La question est donc de savoir si la terminologie employée dans l'interprétation des données a un statut véritablement descriptif, ou si elle répond à un modèle normatif visant à promouvoir les finalités recherchées par les neurocriminologues.

Le second problème, qui s'inscrit dans la continuité du précédent, concerne la portée explicative de ces anomalies et la nature de la relation établie avec la manifestation du comportement criminel. Cette question doit d'abord être examinée du point de vue de la recherche en neurocriminologie, telle qu'elle est posée et traitée dans la littérature. Avant de s'interroger sur le statut épistémologique de ces données, il convient en effet d'expliciter l'argument justifiant l'ambition explicative de la neurocriminologie, c'est-à-dire l'idée selon laquelle le comportement criminel est entièrement expliqué par la présence d'anomalies cérébrales. Comme nous l'avons précisé plus haut, les anomalies observées sont généralement mises en relation avec les incapacités psychologiques des individus, afin de démontrer leur responsabilité causale dans le développement d'un comportement criminel. Il s'agit d'associer l'altération d'une fonction cognitive (par exemple, l'incapacité à inhiber une action motrice) à une anomalie cérébrale caractérisée et localisée (un dysfonctionnement des régions impliquées dans le contrôle de soi). L'argument général consiste alors à attribuer à chacune des anomalies observées un rôle causal dans la manifestation d'un comportement violent, agressif ou impulsif. Dans l'objectif qui est le nôtre d'évaluer le pouvoir explicatif des anomalies cérébrales, il paraît essentiel de questionner la validité de cet argument, c'est-à-dire de définir le statut épistémologique de la relation entre les anomalies cérébrales, l'altération des capacités

fonctionnelles des individus, et la manifestation du comportement criminel. Dans cette perspective, il ne s'agit pas seulement de déterminer si ces anomalies rendent effectivement impossible l'exercice d'une fonction, mais surtout de savoir si elles fournissent une *explication causale* des incapacités fonctionnelles des individus.

Finalement, ces deux problèmes nous invitent à interroger la validité du modèle interprétatif qui convertit l'anomalie en pathologie et, plus généralement, à questionner la nature de la relation établie entre les anomalies cérébrales et la manifestation du comportement criminel. Existe-t-il un « mécanisme pathologique » qui permettrait d'expliquer les incapacités fonctionnelles des individus, et par conséquent de rendre compte de la manifestation d'un comportement impulsif, violent ou agressif ? Les données empiriques justifient-elles une conception pathologique du criminel, et permettent-elles de statuer sur la responsabilité causale des anomalies cérébrales dans la manifestation du comportement criminel ?

4.2. Un modèle normatif et non descriptif

4.2.1. Le « *language of disorder* »

Le premier problème épistémologique que nous proposons de résoudre concerne la question de savoir si la terminologie employée pour décrire le cerveau des criminels est entièrement justifiée sur le plan empirique. Cette question a été discutée par la philosophe Marga Reimer (2008) en ce qui concerne la manière dont la psychopathie est caractérisée dans les recherches en psychologie et neurosciences cognitives. Bien que cette analyse ne porte pas à proprement parler sur la littérature en neurocriminologie, il nous parait important de la mentionner ici car elle fait état d'une stratégie similaire à celle adoptée par les neurocriminologues.

En axant son analyse sur les travaux publiés par le neuropsychologue américain James Blair (e.g. Blair et al. 2005), Marga Reimer montre que la terminologie employée pour caractériser la psychopathie relève de ce qu'elle nomme « *the language of disorder* » : l'interprétation des données fait apparaître des signes de défaillance ou de déficience, à la fois sur le plan neurologique, cognitif et émotionnel. Pourtant, si la

psychopathie est décrite comme un « désordre biologique », elle apparaît davantage comme une « pathologie normative » dans le sens où les concepts mobilisés pour caractériser le trouble relèvent d'un discours normatif incompatible avec les données empiriques. D'après Reimer, l'interprétation des données s'exprime dans un « langage métaphorique » chargé de présupposés, et qui vise à montrer que la psychopathie résulte d'un processus pathologique. Le « *language of disorder* » n'est donc pas tant le résultat d'une investigation empirique, que l'expression d'une stratégie de pathologisation du trouble. Cette terminologie pourrait d'ailleurs aussi bien être remplacée par une description plus neutre, dans des termes qui n'évoquent pas une « déficience » mais simplement une « différence ».

Comme le souligne Reimer, le « *language of disorder* » est largement inspiré par le présupposé d'une relation directe entre les troubles du comportement et les troubles du fonctionnement cérébral. Dans le champ des neurosciences cognitives, les cas cliniques comme celui de Phineas Gage ont encouragé les inférences causales entre les états cérébraux et les états comportementaux. Plus précisément, ils ont favorisé l'idée selon laquelle « *healthy brains underlie pro-social behavior* » et « *unhealthy brains underlie antisocial behavior* » (Reimer 2008, p. 193) : les causes ressemblent à leurs effets, ce qui invite les chercheurs à confondre les désordres du cerveau et du comportement dans un processus de même nature. De ce point de vue, la nature des anomalies cérébrales reflèterait la nature des anomalies observées sur le plan comportemental. Ce présupposé débouche alors sur un modèle interprétatif qui défend l'idée selon laquelle les causes d'un état pathologique sont elles-mêmes pathologiques : si l'on part du présupposé qu'un trouble comportemental est de nature pathologique, les données neurobiologiques sont *de facto* reléguées à un état pathologique. Autrement dit, la pathologisation d'un trouble du comportement ou de la personnalité peut faciliter les biais en faveur d'une pathologisation du fonctionnement cérébral. Or d'après Reimer, c'est précisément ce qui est à l'œuvre dans la manière dont la psychopathie est caractérisée dans la littérature en psychologie et neurosciences cognitives. Le « *language of disorder* » qui structure l'interprétation des données empiriques est fondé sur le présupposé selon lequel la psychopathie résulte d'un processus pathologique ; il en ressort une pétition de principe qui exclut toute interprétation alternative à l'explication pathologique.

Ce raisonnement circulaire est d'autant plus problématique que la manière dont la psychopathie est caractérisée dans la littérature en psychologie ou neurosciences cognitives présente des implications majeures sur la question de la responsabilité des

psychopathes ayant commis un acte immoral. Elle peut en effet introduire un biais dans l'attribution de leur responsabilité morale, mais également vis-à-vis des questions relatives à leur prise en charge. D'après Reimer, le « *language of disorder* » favoriserait une atténuation de la responsabilité des individus, et constituerait un argument empirique en faveur d'une prise en charge médicale.[147] À l'inverse, une description plus neutre évoquant une « différence » et non plus une « déficience » conserverait leur entière responsabilité, et inviterait à se détourner des perspectives de traitement de la psychopathie.

Sur ce point, il convient de préciser les limites de cette analyse en ce qui concerne les implications du « *language of disorder* » sur la notion de responsabilité morale. L'argument selon lequel il favoriserait une déresponsabilisation de l'individu est discutable (ou du moins insuffisant) si l'on admet une conception « capacitarienne » de la responsabilité (Vincent 2008 ; Turner 2010).[148] Selon cette conception, l'attribution de la responsabilité morale (mais aussi pénale) repose sur l'intégrité de certaines capacités psychologiques, comme la capacité à contrôler son comportement ou la capacité à former un jugement moral. De ce point de vue, le « *language of disorder* » n'a finalement que peu d'influence sur les jugements moraux portant sur la responsabilité des individus. Il n'affecte pas nécessairement la responsabilité de l'individu, sauf si l'on peut démontrer un lien avec l'altération des capacités requises pour être moralement responsable de ses actions. Autrement dit, il ne suffit pas de révéler la présence d'un dysfonctionnement, mais de montrer que ce dysfonctionnement affecte les capacités psychologiques nécessaires à l'agentivité morale. Comme le souligne la philosophe Nicole Vincent (2008) :

> « *Whether someone's responsibility is reduced or not depends not on whether their neurological condition is a disorder per se, but rather that it depends on how that neurological condition affects the mental capacities which are required for moral agency — i.e. that capacity and not disorder is what determines responsibility.* » (Vincent 2008, p. 202).

[147] Sur la question de la responsabilité (morale ou pénale) des psychopathes, voir notamment Levy (2007), McGeer (2008), Vargas (2010), Malatesti et McMillan (2010), Shoemaker (2011).

[148] Rappelons que la conception « capacitarienne » (ou « volitionniste ») de la responsabilité est héritée de la « *capacity-responsability theory* » de Herbert Hart (1968).

Ce dernier point nous donne l'occasion d'insister sur la nécessité de préciser la nature de la relation entre les incapacités fonctionnelles des individus et les dysfonctionnements neurobiologiques observés. Cette question empirique est fondamentale si l'on veut évaluer le pouvoir explicatif de la neurocriminologie, car elle nous renseigne sur la responsabilité causale des anomalies cérébrales dans la manifestation d'un comportement criminel. La question est d'autant plus importante si l'on admet une conception capacitarienne de la responsabilité, car ces anomalies pourraient affecter les capacités psychologiques nécessaires à l'agentivité morale et ainsi justifier une atténuation de la responsabilité de l'individu. En clair, la validité empirique de l'interprétation pathologique est essentiellement liée au statut épistémologique de la relation entre les anomalies cérébrales et l'altération des capacités psychologiques des individus. C'est à ce niveau d'analyse que doit se situer le débat sur la question de savoir si la conception psychopathologique du criminel est justifiée au regard des données neurobiologiques.

4.2.2. Un soupçon de raisonnement *ad hoc*

La principale contrainte épistémologique réside dans le fait que les données issues de la neuroimagerie ne donnent pas accès aux mécanismes impliqués dans l'altération des fonctions cognitives des individus. La localisation des anomalies cérébrales ne permet pas d'identifier les mécanismes susceptibles d'*expliquer* les incapacités psychologiques des criminels ; tout au plus permet-elle de recenser les déficits structurels et fonctionnels qui les distinguent de la population générale. L'association établie entre les anomalies cérébrales et les incapacités fonctionnelles des individus est donc fragilisée par l'absence de relation logique entre les données neurobiologiques et l'explication pathologique : rien ne garantit la validité de l'inférence selon laquelle le comportement criminel résulte d'un mécanisme pathologique.

La démarche localisatrice adoptée par la neurocriminologie laisse plutôt apparaître un biais de confirmation en faveur d'une conception pathologique du comportement criminel. La localisation des anomalies cérébrales consiste en effet moins à isoler les structures responsables des incapacités fonctionnelles des individus, qu'à *objectiver* l'existence d'un processus pathologique à l'origine de leur comportement.

Autrement dit, cette démarche localisatrice introduit un soupçon d'argument *ad hoc* dans lequel la conception (*a priori*) pathologique du criminel s'impose comme la seule grille de lecture des données neurobiologiques : l'interprétation de ces données est conçue de telle sorte à confirmer le statut pathologique des anomalies, et en conséquence, à apporter la preuve de l'origine pathologique du comportement criminel.

Dans la continuité de cette analyse, on peut s'interroger sur les raisons pour lesquelles le modèle explicatif exclut toute interprétation alternative à l'explication pathologique. Marga Reimer a montré que cette pétition de principe présentait des implications majeures sur la question de la responsabilité et de la prise en charge des psychopathes ayant commis un acte immoral. En invoquant le « *language of disorder* » pour caractériser la psychopathie, les chercheurs en psychologie ou neurosciences cognitives finiraient par encourager (à leur insu) le développement de stratégies thérapeutiques visant à corriger les dysfonctionnements neurologiques des individus.

À l'inverse, on peut également faire l'hypothèse que la caractérisation de la psychopathie est elle-même biaisée par les finalités poursuivies par les chercheurs. Si le « *language of disorder* » peut servir d'argument empirique pour justifier le recours à un traitement médical, il est aussi possible qu'il soit lui-même dicté par une logique thérapeutique. L'interprétation des données ne viserait pas seulement à apporter la preuve de l'origine pathologique de la psychopathie, mais à caractériser celle-ci de telle sorte à justifier la pertinence d'un traitement médical. De ce point de vue, toute la circularité de la démarche reposerait sur une stratégie de médicalisation du trouble, dont le caractère pathologique correspondrait davantage à une « construction normative » qu'à une réalité empirique.

Les récentes contributions consacrées à l'étude neurobiologique de la psychopathie vont justement dans le sens de cette stratégie, puisque la plupart plaident en faveur d'une approche interventionniste basée sur le traitement pharmacologique ou neurotechnologique des individus (Glannon 2014a ; Canavero 2014 ; Umbach et al. 2015 ; Wonderly 2016 ; Hosking et al. 2017). Le « *language of disorder* » a joué un rôle déterminant dans la définition et la mise en place de cette stratégie, car il sert d'argument empirique pour justifier la pertinence d'une approche centrée sur la neurostimulation des régions fonctionnellement déficientes : c'est à partir des recherches neurobiologiques menées par James Blair (e.g. 2003) ou Adrian Raine (e.g. Glenn, Raine, and Schug 2009) qu'a émergé la possibilité de rehausser le niveau d'activité cérébrale des individus pour

corriger leur comportement. Ce projet neurointerventionniste est notamment défendu par le philosophe Walter Glannon (2014a), qui voit dans les dispositifs neurotechnologiques ou pharmacologiques une éventuelle « alternative à l'incarcération » pour les criminels souffrant de psychopathie :

> « *It is possible that psychopharmacology or stimulation of psychopaths' brains could modulate the dysfunctional neural pathways underlying their disordered personality and behavior. This could promote pro-social behavior and reduce the probability of violent aggression.* » (Glannon 2014a, p. 44).

De là apparaît l'un des principaux enjeux posés par la neurocriminologie contemporaine. Si le « *language of disorder* » est régi par une visée thérapeutique, on peut supposer que la caractérisation du cerveau des criminels répond elle aussi à une logique médicale. Cette hypothèse est plausible si l'on en croit la multiplication des études consacrées à ce qui est communément appelé le « *moral neuroenhancement* » (neuroaugmentation morale) ou la « neurocorrection » des comportements violents ou agressifs (e.g. Young et al. 2010 ; Choy, Raine, and Hamilton 2018b). Comme nous le montrerons à la fin de ce chapitre, ces études garantissent à la neurocriminologie le moyen d'étendre son champ d'intervention, en encourageant l'idée selon laquelle il serait possible de traiter le problème de la criminalité en corrigeant le cerveau des criminels. En l'occurrence, la stimulation cérébrale profonde bénéficie d'une attention particulière depuis le milieu des années 2000, car elle nourrit la promesse d'une « réhabilitation par le traitement » (Greely 2008).

Du reste, l'hypothèse selon laquelle les résultats de la neurocriminologie répondent à une logique médicale mérite d'être démontrée sur un plan plus structurel, en se référant à la « vision neurocriminologique » telle que nous l'avons définie dans le chapitre précédent. Pour rappel, cette notion correspond au cadre théorique et méthodologique structurant l'étude neurobiologique des criminels, et qui est conçu de telle sorte à promouvoir l'agenda de la recherche en neurocriminologie. Si nous avons jusqu'ici insisté sur le projet typologique visant à fonder le critère de dangerosité sur le degré d'anomalies cérébrales, il convient désormais de mettre en perspective cette stratégie avec la logique interventionniste qui se déploie depuis le milieu des années 2000. L'objectif qui est le nôtre n'est donc pas seulement de mettre en évidence une continuité théorique entre les deux projets, mais de montrer que la recherche en neurocriminologie

est entièrement dictée par une logique interventionniste : celle-ci préside à la construction du paradigme expérimental et se déploie tout au long de l'interprétation des données.

Cette analyse nous permettra finalement de montrer en quoi la caractérisation du cerveau des criminels est intrinsèquement liée à la nature des finalités recherchées. Le biais de confirmation en faveur d'une pathologisation du criminel répond entièrement à la « vision neurocriminologique », puisque l'essentiel de la stratégie consiste à définir le phénomène criminel comme un problème médical susceptible d'être traité par des outils neuroscientifiques. Tout au long de l'interprétation des données, on voit s'opérer la construction d'un « cerveau anormal » entièrement conçu comme un « cerveau à corriger » : un cerveau « déficient » ou « défectueux », mais surtout *rectifiable*. Le criminel apparaît comme la cible d'une intervention neurobiologique spécifique ; une sorte de construction neuroessentialiste de « l'individu à corriger », cette figure de « l'anormal » décrite par Michel Foucault comme l'objet d'une technologie de transformation et de normalisation.[149]

4.2.3. La construction d'un « cerveau à corriger »

La stratégie consistant à intervenir sur le cerveau des criminels pour corriger leurs comportements apparaît comme une composante essentielle de la démarche typologique adoptée depuis le début des années 1990. Si l'on prête attention aux conclusions avancées par les neurocriminologues, on constate que l'analyse des résultats s'accompagne souvent d'une réflexion sur la possibilité de corriger les anomalies associées aux comportements violents, agressifs ou impulsifs. De nombreuses études établissent un lien direct entre les résultats issus de la neuroimagerie et les perspectives de traitement des individus, à l'instar des conclusions établies par Aigner et ses collaborateurs (2000), qui ont observé une proportion d'anomalies quatre fois plus importante chez les criminels violents que chez les criminels condamnés pour des infractions non violentes (65,5% contre 16,6%) :

> « *[t]he basis for diagnosis of such dramatic forms of behavioral disorders should therefore include modern brain imaging techniques such as magnetic resonance*

[149] Michel Foucault, *Les Anormaux*. Cours au Collège de France, 1974-1975, Paris, Gallimard-Seuil, 1999.

imaging in order to allow patients appropriate treatment and legal protection. » (Aigner et al. 2000, p. 63).

De nombreux autres exemples illustrent cette volonté d'utiliser la neuroimagerie pour orienter et adapter la prise en charge des criminels selon les données issues de la neuroimagerie, comme en témoignent ces quelques extraits issus des études que nous avons retenues dans notre revue de la littérature :

« *Should the brainimaging findings of fronto-temporal abnormalities in violent offenders be further replicated and shown to remain stable over time, as well as being closely related to violent criminal behavior, research on functional consequences and possible treatment strategies would be of great interest to society*. » (Soderstrom et al. 2000, p. 40).

« *The findings of this exploratory study thus indicate that the comparison of subgroups of offenders may be important to gain more information about specific deficits and underlying brain mechanisms, which might then also be relevant for selecting adequate treatments in forensic facilities*. » (Bertsch et al. 2013, p. 604).

« *This work has important implications for the development of biomarkers for youth on a high-risk trajectory of antisocial behavior and it also has important implications for developing novel interventions to help reduce the enormous fiscal and emotional toll homicidal behavior has on society*. » (Cope et al. 2014, p. 806)

Du point de vue des neurocriminologues, les anomalies cérébrales des criminels apparaissent comme un modèle neurobiologique pertinent pour justifier la mise en place d'une stratégie neuro-interventionniste. Ces anomalies favoriseraient l'idée selon laquelle la dangerosité criminologique est liée à une pathologie organique, et qu'elle relève par conséquent d'un traitement biomédical. Il faut bien ici comprendre que l'assimilation du crime au domaine de la pathologie repose sur la localisation d'un dysfonctionnement organique qui exclut toute cause psychiatrique : les anomalies cérébrales plaideraient en faveur d'une *neuropathologisation* (et non pas d'une psychiatrisation) du crime. La nuance est importante car la logique qui sous-tend cette qualification est intrinsèquement liée au mode d'intervention privilégié par les neurocriminologues. La caractérisation du cerveau des criminels vise à inscrire l'individu dans le registre pathologique, tout en excluant la maladie mentale du modèle explicatif. Il ne s'agit pas de montrer que le

criminel relève d'un traitement psychiatrique, mais bien de le définir comme une catégorie biomédicale susceptible de faire l'objet d'une intervention directe, neurobiologique et spécifique.

Cette stratégie passe par la construction d'un « cerveau à corriger », c'est-à-dire d'un « cerveau anormal » qui garantit les critères de sa propre normalisation : la caractérisation du cerveau des criminels définit elle-même les modalités permettant de neutraliser les conditions d'émergence d'un comportement dangereux. Rappelons en effet que les concepts mobilisés dans la littérature relèvent de la catégorie du défaut ou de l'incomplétude, et non pas d'une simple altérité biologique. La caractérisation du cerveau des criminels ne consiste pas seulement à recenser les différences neurobiologiques par rapport à une norme prédéfinie ; elle porte en elle-même les conditions de possibilités d'une normalisation. Il s'agit de rechercher l'irrégularité à laquelle s'oppose une norme préétablie, mais qui garantit néanmoins une possibilité de correction.

De ce point de vue, on comprend mieux les ressorts de la démarche typologique adoptée par la neurocriminologie pour fonder les catégories juridiques sur des catégories neurobiologiques. Cette démarche permet de fixer les termes d'une dangerosité accessible à une intervention neuroscientifique. Il s'agit dans le même temps de repérer les éléments démontrant la dangerosité de l'individu, tout en précisant les critères d'une correction possible. La démarche localisatrice a également joué un rôle déterminant dans la définition et la mise en place de cette stratégie, son avantage étant qu'elle dispense les chercheurs d'identifier les mécanismes impliqués dans la manifestation du comportement criminel : la caractérisation et la localisation des anomalies cérébrales suffisent à définir les modalités d'une intervention neurobiologique. Ainsi du point de vue des neurocriminologues, il suffirait de rehausser le niveau d'activité des régions fonctionnellement déficientes, pour corriger les anomalies observées sur le plan comportemental.

Cette stratégie interventionniste pose deux problèmes épistémologiques fondamentaux, dont les solutions sont mutuellement exclusives. D'une part, elle fait apparaître un problème majeur en ce qui concerne le pouvoir explicatif de la neurocriminologie : celui d'un programme de recherche entièrement construit autour d'une logique interventionniste, et qui ne fournit aucune explication des phénomènes étudiés. Les données empiriques sont interprétées au travers d'une grille de lecture dont les termes sont définis conformément à une stratégie de pathologisation et de

médicalisation du crime. Le modèle interprétatif vise à définir les modalités d'une mesure neuro-corrective, sans parvenir à fournir une *explication* causale du comportement criminel. Ce modèle efface dans le même temps les questions relatives à la responsabilité des criminels, puisqu'il dissout définitivement leur statut juridique. En d'autres termes, la caractérisation du cerveau des criminels participe à la transformation d'une catégorie juridique en une catégorie biomédicale susceptible de faire l'objet d'une intervention neurobiologique spécifique ; elle fait donc du cerveau du criminel un objet naturel « rectifiable », dont le statut juridique est révoqué au profit d'une approche strictement médicale.

D'un autre côté, on peut s'interroger sur l'apport épistémologique (potentiel) de la stratégie neuro-interventionniste. Si l'on peut aisément contester la validité épistémique d'un programme de recherche dicté par une logique médicale, on peut aussi admettre que la stratégie neuro-interventionniste présente elle-même un certain rôle épistémique. Les récentes recherches consacrées à la « neurocorrection » du comportement criminel s'efforcent justement de défendre la pertinence de cette approche pour étudier les relations de causalité entre les états cérébraux d'un côté, et les états psychologiques ou comportementaux de l'autre. En intervenant directement sur le cerveau des criminels, cette stratégie permettrait de statuer sur le rôle causal de chacune des anomalies cérébrales, et ainsi d'isoler les régions ayant une responsabilité dans la manifestation d'un comportement violent ou agressif. Il est d'ailleurs intéressant de constater que la plupart de ces études répondent au double objectif de « réduire la violence ou l'agressivité » des individus, et « d'identifier les mécanismes qui sous-tendent » ces comportements (e.g. Choy, Raine, and Hamilton 2018b). En clair, ces recherches se défendent de centrer leur ambition sur les perspectives de traitement des criminels ; elles cherchent à acquérir une forme de légitimité scientifique dans le champ des neurosciences cognitives, en contribuant à l'identification des mécanismes impliqués dans le contrôle de soi, la régulation émotionnelle ou encore la formation d'un jugement moral.

À cet égard, il est tentant de considérer cette démarche comme la finalité ultime de la recherche en neurocriminologie. Les neurocriminologues ne chercheraient plus à recenser les anomalies cérébrales associées au comportement criminel, mais à tester les liens de causalité entre la stimulation ou l'inhibition d'une région donnée d'une part, et ses effets sur le plan comportemental d'autre part. De ce point de vue, la stratégie neuro-interventionniste permettrait de combler les limites de la démarche localisatrice pour répondre à l'ambition explicative des neurocriminologues. Si la localisation des

anomalies n'apporte en elle-même aucune explication des phénomènes étudiés, la stimulation ou l'inhibition d'une région donnée pourrait contribuer à statuer sur la nature de la relation (causale ou non) entre un état cérébral donné et son effet psychologique ou comportemental.

Cette question reste en suspens, car elle exige une analyse approfondie des résultats et du cadre méthodologique des recherches consacrées à la « neurocorrection » du comportement criminel. La question du pouvoir explicatif de la stratégie neuro-interventionniste nous invite ainsi à sortir du champ de la neurocriminologie basée sur la neuroimagerie, pour analyser les nouveaux domaines d'application utilisant les techniques de neurostimulation comme la stimulation cérébrale profonde (*Deep Brain Stimulation* ou DBS) ou la stimulation magnétique transcrânienne (*Transcranial Magnetic Stimulation* ou TMS). Depuis que leur application a été étendue au traitement de certains troubles psychiatriques ou neuropsychiatriques, ces techniques ont favorisé la multiplication des recherches consacrées à l'amélioration des capacités cognitives ou morales des individus (e.g. Douglas 2013). Leur usage dans le champ criminologique s'inscrit donc dans le cadre de ce qui est appelé le « *neuroenhancement* » du contrôle de soi ou des jugements moraux. Il inclut diverses tentatives visant à modifier les états cérébraux des individus, dans l'objectif de renforcer les capacités psychologiques nécessaires à une agentivité responsable.

Cette dynamique de recherche regroupe un large éventail d'études dont le statut disciplinaire est encore indéterminé. Parmi les contributions relatives au « *neuroenhancement* », il est en effet important de distinguer les apports théoriques qui s'inscrivent dans le champ de la philosophie morale ou de la bioéthique, et les recherches expérimentales basées sur l'utilisation de techniques de neurostimulation, qui relèvent du champ de la neurocriminologie et du neurodroit. Naturellement, ces deux voies de recherche se distinguent par le type d'approche adopté à l'égard du « *neuroenhancement* » : philosophique pour la première, neuroscientifique pour la seconde. Elles font néanmoins preuve d'une certaine perméabilité car les contributions philosophiques s'inscrivent généralement dans une ambition naturalisatrice. Celles-ci répondent en effet moins à une réflexion critique ou épistémologique, qu'à un programme de naturalisation de la morale, lequel cherche à défendre la pertinence de l'approche interventionniste pour répondre à des questions éthiques. Ces réflexions appartiennent en ce sens davantage au domaine de la neuroéthique (ou aux « neurosciences de

l'éthique ») – c'est-à-dire à un projet d'intégration des outils neuroscientifiques dans le champ éthique – que de la philosophie morale ou de l'éthique en tant que telle.[150]

Dans la section suivante, nous proposons de préciser les perspectives et les apports épistémologiques de ces deux voies de recherche, dans l'objectif de caractériser les différentes formes que recouvre le projet neuro-interventionniste défendu depuis le milieu des années 2000. La première voie de recherche correspond à ce qui est communément appelé le « *moral neuroenhancement* » ou neuroaugmentation morale. Elle est fondée sur une conception capacitarienne de la responsabilité, car elle défend l'idée selon laquelle une amélioration des capacités morales des individus conduirait à un renforcement de leur responsabilité (e.g. Vincent 2013). Globalement, cet axe de recherche s'articule autour de la question de savoir sous quelles conditions il pourrait être acceptable d'intervenir sur le cerveau d'un individu pour modifier son agentivité morale. Autrement dit, il s'agit de définir les conditions pour lesquelles la stratégie neuro-interventionniste permettrait d'améliorer les capacités psychologiques requises pour être moralement responsable. Comme nous le verrons plus loin, ce débat théorique entretient une certaine confusion avec le champ de la neurocriminologie, car certains philosophes conçoivent la neuroaugmentation morale comme « un idéal de réhabilitation criminelle » (e.g. Glannon 2014a).

Par ailleurs, nous verrons que ces propositions théoriques ont progressivement investi le champ de la neurocriminologie et du neurodroit, en se poursuivant au travers d'une série d'investigations empiriques consacrées aux effets de la neurostimulation sur la formation d'un jugement moral (Young et al. 2010 ; Tassi et al. 2012 ; Jeurissen 2014 ; Ye et al. 2015 ; Sellaro et al. 2015 ; Buckholtz et al. 2015 ; Ginther 2016 ; Darby and Pascual-Leone 2017). Une brève mention de ces études nous permettra de sortir du débat proprement théorique portant sur la « neuroaugmentation morale », pour analyser les perspectives de recherche empirique telles qu'elles sont envisagées depuis le milieu des années 2000, dans le domaine de la neurocriminologie et du neurodroit.

Enfin, parmi les contributions relatives au « *neuroenhancement* », la deuxième voie de recherche que nous proposons de décrire correspond à ce que nous avons appelé

[150] La distinction entre « les neurosciences de l'éthique » et « l'éthique des neurosciences » a été introduite par la philosophe Adina Roskies (2002). Elle a également été discutée par Eric Racine (2008), qui a décrit les différentes voies de recherche que recouvre le champ de la neuroéthique, notamment les études menées en neuro-ingénierie et en neuropsychopharmacologie.

la « neurocorrection » du comportement criminel. Bien qu'elle ne réponde pas explicitement à cette appellation, elle relève bel et bien du champ de la neurocriminologie, car elle s'inscrit dans le prolongement des recherches relatives aux anomalies cérébrales des criminels. La stratégie invoquée consiste à utiliser les techniques de neurostimulation (notamment la DBS) pour corriger les dysfonctionnements associés à la manifestation d'un comportement agressif, violent ou impulsif. En somme, cette voie de recherche prend pour cible le « cerveau à corriger » tel qu'il a été décrit tout au long des études publiées depuis le début des années 1990. Elle défend l'idée selon laquelle les techniques de neurostimulation permettraient de corriger les anomalies associées au comportement criminel, et ainsi prévenir le risque de récidive (Franzini et al. 2005, 2013 ; Hortensius, Schutter, and Harmon-Jones 2012 ; Dambacher et al. 2015 ; Riva et al. 2015 ; Choy, Raine, and Hamilton 2018b).

4.3. Le « *neuroenhancement* » comme idéal de réhabilitation ?

4.3.1. La « neuroaugmentation » morale

Sur un plan purement théorique, la neuroaugmentation morale peut se définir comme « [t]out changement dans un agent moral A, effectué ou facilité de façon significative par l'application d'une neurotechnologie, et pour laquelle le résultat attendu est que A soit un meilleur agent moral » (Earp, Douglas, and Savulescu 2017, p. 168). Cette voie de recherche regroupe un large éventail d'études utilisant des approches pharmacologiques ou neurotechnologiques pour améliorer les capacités morales des individus. Elle inclut le développement de projets neuro-interventionnistes consistant à faciliter les comportements pro-sociaux comme l'altruisme et l'empathie, ou visant à supprimer les biais implicites inhérents aux jugements moraux (e.g. Douglas 2008).

La question de la neuroaugmentation morale occupe une place de plus en plus grande dans les débats contemporains en neuroéthique, en particulier en ce qui concerne ses implications sur les notions de contrôle de soi, d'agentivité, de libre arbitre et d'identité personnelle (Schechtman 2010 ; Lipsman and Glannon, 2013 ; Glannon 2014b ; Vincent 2014 ; Roskies 2015 ; Faucher 2018). Les discussions portent

essentiellement sur la notion d'agentivité morale, et la question de savoir si la neuroaugmentation affecte les capacités requises pour être tenu moralement responsable de ses actions (voir en particulier Roskies 2015). Le problème du libre arbitre pèse finalement peu dans les débats, dans la mesure où les théories proposées reposent sur une conception compatibiliste centrée sur les notions d'agentivité et de contrôle de soi.

Plus généralement, les débats s'articulent autour de la question de savoir si la neuroaugmentation morale est théoriquement souhaitable, ou du moins s'il existe des raisons éthiques de s'y opposer (Persson and Savulescu 2008 ; Harris 2011 ; Pacholczyk 2011 ; Shook 2012 ; Douglas 2015 ; Earp, Douglas, and Savulescu 2017). Certains philosophes comme Ingmar Persson et Julian Savulescu encouragent les recherches biomédicales dans ce domaine, en invoquant « un impératif urgent d'augmenter le caractère moral de l'humanité » (Persson and Savulescu 2008, p. 162 ; voir aussi Persson and Savulescu 2013). Si les propositions restent souvent spéculatives, la plupart des philosophes ayant contribué au débat y voient une démarche souhaitable voire nécessaire, à condition qu'elle respecte le consentement libre et éclairé de l'individu (voir en particulier DeGrazia 2014 ; Douglas 2015).

L'une des propositions consiste par exemple à utiliser les technologies biomédicales pour réduire les biais implicites de nature émotionnelle (Douglas 2008, 2013). Partant de l'idée que les émotions peuvent affecter la formation d'un jugement moral, le philosophe Thomas Douglas propose d'atténuer ce qu'il nomme les « *counter-moral emotions* », c'est-à-dire les émotions qui interfèrent avec le raisonnement moral et les comportements pro-sociaux :[151] l'amélioration morale passerait donc par « une réduction du degré des émotions ressenties par l'agent » (Douglas, 2008, p. 231).[152] Cette

[151] Sans entrer dans le détail de cette discussion, précisons que le modèle dualiste opposant raison et émotion fait figure de paradigme dans le domaine de la neurobiologie des jugements moraux, depuis la publication d'une série d'études proposant d'utiliser les techniques de neuroimagerie pour trancher le débat entre l'utilitarisme et le déontologisme (e.g. Greene et al. 2001, 2004) : le conflit entre ces deux types de jugements moraux s'expliquerait par une dissociation entre les systèmes cognitif et émotionnel. Ce modèle dualiste est pourtant contestable sur le plan méthodologique et conceptuel (Penavayre, Brun, et Boraud 2019), mais aussi du point de vue de la biologie évolutionniste : le développement des structures cérébrales montre que les processus de décision ne découlent pas d'un conflit mais d'une collaboration dynamique entre les deux systèmes (Boraud et al. 2018).

[152] Parmi les « *counter-moral emotions* », Thomas Douglas cite par exemple les pulsions agressives ou violentes et l'aversion vis-à-vis de certains groupes raciaux. Les enjeux philosophiques relatifs à cette proposition ont notamment été discutés par John Harris (2011).

proposition a été largement discutée par John Harris (2011), lequel voit dans l'amélioration morale une notion « fondamentalement incomprise » qui repose sur une conception erronée de la morale. Il lui parait vain de centrer le problème sur les « *counter-moral emotions* » plutôt que sur l'amélioration des capacités cognitives elles-mêmes, car la neuroaugmentation morale doit selon lui essentiellement reposer sur l'amélioration de la connaissance et de la compréhension des règles morales : « *moral enhancement must, in large part, consist of cognitive enhancement* ». (Harris 2011, p. 106).

Dans le cadre du débat concernant le lien entre le « *neuroenhancement* » et l'agentivité morale, des réflexions plus générales ont été menées autour de la question de la responsabilité morale des individus ayant fait l'objet d'une neurointervention (Vincent 2011, 2013 ; Bublitz and Merkel 2013 ; Klaming and Haselager 2013 ; Shaw 2014 ; Koi et al. 2017 ; voir aussi Bigenwald and Chambon 2019). Ces discussions s'inscrivent dans la continuité du débat relatif à la responsabilité d'un agent manipulé par un tiers et ayant commis un acte immoral (Fischer and Ravizza 1998). La perspective d'une neuroaugmentation morale a suscité un renouvellement du problème, car elle interroge le lien entre les capacités psychologiques et la responsabilité de l'agent. Elle introduit en effet l'idée selon laquelle l'amélioration des capacités de l'individu (par exemple, sa capacité à contrôler son comportement) conduirait à un renforcement de sa responsabilité (Vincent 2011, 2013, 2014 ; Shaw 2014). Comme le résume la philosophe Nicole Vincent (2013) : « *hypercapacity does result in hyperresponsibility* » (Vincent 2013, p. 309).

C'est dans ce cadre théorique qu'ont émergé les réflexions autour de la possibilité d'utiliser des techniques d'interventions neuroscientifiques sur des criminels, pour faciliter leur réhabilitation (Shaw 2014 ; Vincent 2014). D'après les philosophes Elizabeth Shaw et Nicole Vincent, la conception capacitarienne de la responsabilité justifie la mise en place d'une stratégie neurointerventionniste, car elle encourage l'idée selon laquelle une amélioration des capacités psychologiques des individus permettrait de favoriser les comportements moralement responsables. Les techniques de neurostimulation permettraient ainsi de développer ou renforcer leur agentivité responsable, et par conséquent de prévenir les comportements violents, agressifs ou impulsifs. Comme le souligne Elizabeth Shaw (2014) : « *[a]dvances in neuroscience might make it possible to develop techniques for directly altering offenders' brains, in order to make offenders more responsible and law-abiding*. ».

Cette ambition s'est exprimée sur un plan purement théorique, à partir d'une distinction conceptuelle proposée par Nicole Vincent (2009), entre ce qu'elle nomme la « *virtue responsibility* » et la « *capacity responsibility* » (voir aussi Vincent 2011 ; Shaw 2014). L'amélioration de la « *virtue responsibility* » consisterait à inculquer à l'individu des valeurs morales, de telle sorte à développer sa conscience morale. L'amélioration de la « *capacity responsibility* » viserait quant à elle à renforcer les capacités psychologiques nécessaires à l'exercice d'une agentivité responsable, comme la capacité à contrôler son comportement ou à former un jugement moral. Sans entrer dans le détail de cette distinction, retenons que cette seconde notion définit les conditions pour lesquelles il pourrait être acceptable d'intervenir sur le cerveau d'un individu pour modifier son agentivité morale. Autrement dit pour Nicole Vincent et Elizabeth Shaw, la perspective d'une neuroaugmentation morale ne peut être envisagée que si l'on définit la notion de responsabilité comme une « *capacity responsibility* ». Il serait théoriquement (et éthiquement) possible d'intervenir sur le cerveau des individus pour renforcer leur responsabilité, à condition que l'on s'en tienne à améliorer les capacités nécessaires à leur agentivité morale. En l'occurrence, la stratégie consisterait à stimuler les régions cérébrales impliquées dans le contrôle de soi ou la capacité à former un jugement moral, leur permettant ainsi d'orienter leurs conduites conformément à un ensemble de normes morales ou juridiques.

De manière intéressante, ce débat théorique a été redéfini dans un registre plus empirique, au travers d'une série d'études consacrées aux effets de la neurostimulation sur la formation d'un jugement moral (Young et al. 2010 ; Tassi et al. 2012 ; Jeurissen 2014 ; Ye et al. 2015 ; Sellaro et al. 2015 ; Buckholtz et al. 2015 ; Ginther 2016 ; Darby and Pascual-Leone 2017). Ces études reposent sur l'utilisation des techniques de neurostimulation comme la DBS ou la TMS, et sont généralement menées sur des sujets issus de la population générale. Bien qu'elles ne comportent aucune référence aux contributions philosophiques présentées plus haut, la plupart répondent aux propositions théoriques relatives à la neuroaugmentation morale. Dans un sens en effet, les enjeux proprement théoriques de la neuroaugmentation ont fini par s'implanter sur le terrain de l'investigation empirique, les techniques de neurostimulation étant perçues comme un moyen de tester les liens de causalité entre la modification des états cérébraux d'une part, et leurs effets sur la formation des jugements moraux d'autre part. Une précision néanmoins : contrairement aux contributions philosophiques que nous venons de décrire,

peu d'études empiriques évoquent explicitement l'ambition d'améliorer les capacités morales des individus. Les objectifs annoncés sont relativement flous, et le statut disciplinaire de ces études est tout aussi confus. Dans la mesure où elles s'inscrivent dans une dynamique complémentaire à la stratégie de « neurocorrection » du comportement criminel, on peut les situer à l'intersection du champ de la neuroéthique d'une part, et de la neurocriminologie ou du neurodroit d'autre part.

Si l'on s'en tient aux objectifs annoncés, ces recherches relèvent du champ de la neuroéthique car elles proposent d'utiliser les techniques de neurostimulation (DBS ou TMS) pour étudier le rôle des processus cognitifs et émotionnels dans la formation d'un jugement moral. Plus précisément, elles visent à étudier le rôle des émotions ou des croyances dans le raisonnement moral, en modulant l'activité cérébrale des participants pendant la résolution d'un dilemme moral. Le paradigme expérimental consiste à perturber l'activité d'une ou plusieurs région(s) cérébrale(s) impliquée(s) dans la formation d'un jugement moral (exemple : cortex préfrontal droit, cortex préfrontal dorso-latéral, jonction temporo-pariétale, etc.) ; les individus sont alors chargés d'évaluer la responsabilité d'un agent ayant commis un acte immoral, avant et pendant la stimulation des régions cérébrales. La finalité de ces recherches consiste à analyser les relations de causalité entre la perturbation d'une région donnée et ses effets sur la résolution des dilemmes moraux, dans l'objectif d'identifier les mécanismes neurobiologiques impliqués dans l'évaluation de la responsabilité d'un agent ayant commis un acte immoral.

Derrière cet objectif général apparaît une logique qui ne relève pas à proprement parler du champ de la neuroéthique. Certaines études ont cherché à interpréter leurs résultats à la lumière des problématiques relevant du droit pénal, notamment en ce qui concerne la qualité et la rationalité des décisions de justice (e.g. Buckholtz et al. 2015 ; Ginther 2016 ; Bellucci et al. 2017). À ce titre, ces études s'inscrivent explicitement dans le champ du neurodroit, puisqu'elles défendent l'idée selon laquelle les outils neuroscientifiques permettraient de résoudre les problèmes relatifs à l'appréciation de la responsabilité et à la détermination de la peine. Grâce à l'utilisation combinée de l'IRMf et de la TMS, Joshua Buckholtz et ses collaborateurs (2015) ont par exemple isolé deux types de réseaux neuronaux correspondant d'une part, à l'évaluation de la responsabilité d'un agent ayant commis un acte immoral, et d'autre part, à la détermination de la sanction appropriée (voir aussi Buckholtz et al. 2008 ; Buckholtz and Marois 2012 ; Bellucci et al. 2017). Si les auteurs de ces études font preuve d'une certaine prudence

quant à la finalité qui sous-tend cette démarche catégorielle, leurs résultats ont par ailleurs suscité une réflexion théorique sur la possibilité d'utiliser les outils neuroscientifiques pour évaluer la capacité des jurés à reconnaître la culpabilité d'un prévenu (Ginther et al. 2016). La stratégie (à ce stade, essentiellement spéculative) consisterait à mesurer l'intégrité des réseaux neuronaux qui sous-tendent l'évaluation de la responsabilité d'un individu et la détermination de la sanction appropriée.

Parmi les recherches que nous avons citées plus haut, l'étude publiée par Young et ses collaborateurs (2010) a posé les bases d'une voie de recherche complémentaire, consacrée à la « neurobiologie de la théorie de l'esprit » et à ses implications dans le champ pénal.[153] Cette étude a montré qu'une perturbation de l'activité de la jonction temporo-pariétale altérait la capacité des sujets à inférer les états mentaux d'un agent ayant commis un acte immoral. Les sujets avaient alors tendance à déresponsabiliser l'agent, n'étant plus capables de reconnaître ses intentions et ses motivations et *de facto*, de former un jugement moral. Ces résultats ont permis d'une part, de confirmer le lien entre la théorie de l'esprit et la formulation d'un jugement moral, et d'autre part, de montrer que la jonction temporo-pariétale joue bien un rôle dans la capacité à inférer les états mentaux d'autrui. De ce point de vue, la stratégie neuro-interventionniste nous renseigne sur l'influence causale exercée par cette région sur la capacité à évaluer la responsabilité morale d'un individu : sa perturbation fonctionnelle apparaît comme une condition suffisante pour altérer la capacité à former un jugement moral. Ces résultats ont alors nourri les réflexions sur la possibilité d'utiliser l'IRMf pour mesurer l'intégrité de cette région, et ainsi évaluer la capacité d'un individu à apprécier la valeur morale de ses actions et de celles d'autrui (Ginther et al. 2016).

Globalement, les études consacrées aux effets de la neurostimulation sur la formation d'un jugement moral entretiennent un certain flou quant à la nature des finalités recherchées. Il est difficile de savoir si ces recherches visent à promouvoir la neuroaugmentation morale dans une perspective de réhabilitation criminelle, à favoriser le développement de stratégies visant à supprimer les biais émotionnels inhérents à la

[153] La théorie de l'esprit désigne la capacité à inférer les états mentaux d'autrui (les intentions, les croyances, les émotions, etc.). Depuis la découverte du système miroir à la fin des années 1980 (Rizzolatti et al. 1988), elle fait l'objet de nombreuses recherches en neurosciences cognitives (e.g. Fletcher et al. 1995 ; Gallagher and Frith 2003 ; Eddy et al. 2016).

formation d'un jugement moral, ou plus généralement, à améliorer la capacité à statuer sur la responsabilité morale d'un individu. Du reste, quelle que soit la cible privilégiée par la stratégie du « *neuroenhancement* », l'utilisation combinée de la neuroimagerie et des techniques de neurostimulation est mise au service d'une approche catégorielle (normative), visant à repenser le système pénal et la gestion de la criminalité à la lumière des avancées neuroscientifiques. Elle défend l'apport des outils neuroscientifiques pour compenser les limites des principes fondamentaux du droit pénal, comme l'évaluation de la responsabilité ou la prévention du risque de récidive.

L'ambition d'utiliser les techniques de neurostimulation comme une stratégie de réhabilitation s'est développée en parallèle de ces recherches, au travers d'une série d'investigations empiriques visant à rehausser le niveau d'activité cérébrale des criminels. Ces études répondent à la stratégie de pathologisation du crime amorcée dès les premières publications relatives aux anomalies cérébrales des criminels. Le « *language of disorder* » qui gouverne l'interprétation des données neurobiologiques sert d'argument empirique pour justifier la mise en place d'une stratégie neuro-interventionniste : il permet de définir les modalités permettant de neutraliser les conditions d'émergence d'un comportement (potentiellement) dangereux.

4.3.2. La « neurocorrection » du comportement criminel

Depuis le milieu des années 2000 et l'extension de son champ d'application à certains troubles psychiatriques ou neuropsychiatriques, la stimulation cérébrale profonde (DBS) bénéficie d'une attention croissante dans le domaine criminologique. Le juriste américain Henry Greely a mené l'une des premières réflexions concernant la possibilité d'utiliser cette technique « pour prévenir ou réduire le risque de récidive criminelle » (Greely 2008). Il paraît important de préciser le sens de son argument, car il nous renseigne sur la manière dont les résultats de la neurocriminologie ont été perçus dans le champ du droit pénal. Cet article est d'ailleurs paradigmatique du discours consistant à attribuer aux outils neuroscientifiques un rôle prescriptif pour repenser le système pénal américain. D'après Greely, les neurosciences « pourraient nous apporter de nouvelles options en matière de condamnation en offrant la promesse d'une réhabilitation efficace par le biais d'un traitement » (Greely 2008, p. 104). Le caractère réversible et ajustable

de la DBS en fait un outil potentiellement efficace pour prévenir le risque de récidive, en offrant la possibilité de moduler l'activité cérébrale des criminels présentant un défaut ou un excès d'activité dans certaines régions cérébrales :

> « *DBS [Deep Brain Stimulation] provides a plausible method for stimulating greater activity in regions where diminished function is thought to play a role in crime. It could also provide an adjustable, intermittent, and reversible method of inhibiting activity in areas where over-activity may contribute to criminal activity and thus serve some of the functions of neurosurgery, but without its permanence.* » (Greely 2008, p. 114).

Il est important d'insister sur le fait que la stratégie neuro-interventionniste repose essentiellement sur le « *language of disorder* », tel qu'il a été défendu tout au long des études menées sur le cerveau des criminels : la présence de déficits ou de dysfonctionnements sert d'argument pour justifier le recours à la technique du « *neuroenhancement* ». Plus généralement, l'argument repose sur l'idée selon laquelle il suffirait de modifier le cerveau du criminel pour modifier son comportement. D'après Greely, l'efficacité de la stratégie neuro-interventionniste est garantie par le fait qu'un changement au niveau neurobiologique induit nécessairement un changement au niveau comportemental.

Pour cette raison, il lui semble préférable de favoriser le développement de cette stratégie au détriment des programmes traditionnels de réhabilitation, ces derniers opérant en définitive autant sur le plan comportemental que cérébral : « [l]a plupart des sanctions pénales fonctionnent en affectant le cerveau du criminel. Tout type de réhabilitation n'agit qu'en modifiant le comportement du criminel, ce qui ne fonctionne qu'en modifiant son cerveau. » (Greely 2008, p. 133). Henry Greely ne voit en effet « aucune différence qualitative entre le fait d'agir *directement* pour modifier le cerveau d'un criminel – par des médicaments, un traitement chirurgical, la stimulation cérébrale profonde ou des vaccins, si l'innocuité et l'efficacité ont été démontrés – et le fait d'agir *indirectement* – par la peine, la réhabilitation, la théorie cognitive, ou les conditions de libération conditionnelle, à supposer que les résultats soient les mêmes » (Greely 2008, p. 134).

De ce point de vue, l'usage des techniques de neurostimulation dans le champ criminologique repose sur un double argument. Le premier est empirique ; il bénéficie du modèle neurobiologique défendu depuis le début des années 1990, qui décrit l'individu dangereux comme un « cerveau à corriger ». En reléguant le cerveau des criminels à la

catégorie du défaut ou de l'incomplétude, les neurocriminologues ont fini par encourager les initiatives visant à rehausser le niveau d'activité cérébrale des individus, contribuant à renouveler la question de la réhabilitation dans la communauté des juristes. Le deuxième argument repose sur un présupposé neuroessentialiste. Il défend l'idée selon laquelle il suffirait d'agir sur le cerveau des criminels pour corriger leurs comportements, et ainsi prévenir leur risque de récidive. Autrement dit, le « *neuroenhancement* » est perçu comme une méthode efficace de réhabilitation, car il permettrait d'améliorer dans le même temps les états psychologiques et cérébraux de l'individu. Notons également que Greely évoque un argument de type utilitariste pour justifier la pertinence de l'approche neuro-interventionniste. Cette approche s'est en effet formée autour de la problématique de la récidive et de l'impératif de protection sociale ; elle s'inscrit dans le cadre d'une prévention médicalisée du crime qui consiste à neutraliser (littéralement) le risque de récidive.

Ces arguments sont partagés par la plupart des neurobiologistes ayant étudié la question du « *neuroenhancement* » dans le champ criminologique. Globalement, ces recherches visent à utiliser les techniques de neurostimulation (la DBS ou la TMS) pour corriger les comportements impulsifs ou agressifs (Franzini et al. 2005, 2013 ; Hortensius, Schutter, and Harmon-Jones 2012 ; Dambacher et al. 2015 ; Riva et al. 2015 ; Choy, Raine, and Hamilton 2018b).[154] Elles assurent à la neurocriminologie le moyen d'étendre son champ d'intervention, en encourageant l'idée selon laquelle il serait possible de traiter le problème de la criminalité en intervenant directement sur le cerveau des criminels.

Parmi les recherches que nous citons ci-dessus, la plupart sont des études longitudinales menées sur des volontaires présentant des hauts niveaux d'agressivité ou de violence, et dans certains cas sur des sujets sains issus de la population générale. La taille de l'échantillon est très variable puisqu'elle oscille entre une demi-douzaine (e.g. Franzini et al. 2013) et plusieurs dizaines d'individus (Choy, Raine, and Hamilton 2018b). L'objectif annoncé est d'étudier les effets sur le long terme (parfois sur plusieurs années) d'une stimulation ou d'une inhibition des régions cérébrales associées aux comportements violents ou agressifs, en mesurant les conséquences de cette intervention

[154] Pour une revue de la littérature des études utilisant des dispositifs pharmacologiques ou neurotechnologiques sur des populations criminelles, voir Choy, Focquaert, et Raine (2018a) et Chew et al. (2018).

sur le plan comportemental. L'amélioration du comportement est évaluée grâce à un paradigme expérimental permettant de mesurer l'agressivité des sujets (exemple : le *Taylor Aggression Paradigm*, Taylor 1967), ou à partir d'un questionnaire distribué dans l'environnement socio-professionnel ou socio-familial des individus. Après avoir évalué leurs niveaux d'agressivité ou de violence, les participants sont généralement répartis en deux groupes : un groupe traité par stimulation cérébrale, et un groupe contrôle bénéficiant d'une intervention placebo (« *sham stimulation* »).

Les résultats de ces études montrent que la neurostimulation produit une amélioration rapide et significative des comportements individuels, avec une disparition complète des manifestations d'agressivité ou de violence. Franziska Dambacher et ses collaborateurs (2015) ont par exemple constaté que la stimulation du cortex préfrontal dorsolatéral droit (DLPFC) diminuait nettement l'agressivité proactive des hommes, par rapport aux niveaux d'agressivité mesurés avant stimulation.[155] La stimulation de cette même région a donné des résultats similaires dans l'étude menée par Olivia Choy, Adrian Raine et Roy Hamilton (2018b), en permettant une réduction de l'agressivité physique et sexuelle, ainsi qu'une évolution de la conscience morale des 39 participants du groupe expérimental. Par ailleurs, l'étude de Paolo Riva et ses collaborateurs (2015) a montré qu'une stimulation du cortex préfrontal ventrolatéral droit (VLPFC) réduisait de manière significative l'agressivité résultant d'une situation d'exclusion sociale.

Sans entrer dans le détail des paradigmes expérimentaux utilisés pour mesurer l'agressivité des sujets, il convient de préciser que les régions ciblées diffèrent selon les études : cortex préfrontal dorsolatéral (Dambacher et al. 2015, Choy, Raine, and Hamilton 2018b), cortex préfrontal ventrolatéral (Riva et al. 2015), ou encore hypothalamus dorso-médian (Franzini et al. 2005, 2013). Il est surtout important de noter que la sélection des régions d'intérêt repose parfois sur les résultats d'une seule étude de neuroimagerie, ce qui pose la question de la robustesse de la relation entre le comportement agressif et les régions d'intérêt : on peut aisément contester la validité empirique de la stratégie neuro-interventionniste si les études sur lesquelles elle repose ne garantissent pas la spécificité fonctionnelle de ces régions.

[155] Précisons que cette étude n'a pas observé de différence statistiquement significative chez les femmes. Cela peut s'expliquer par le fait que la plupart des études de neuroimagerie (sur lesquelles s'appuie la sélection des régions d'intérêt) sont effectuées sur des hommes. Or, des études ont montré que les hommes et les femmes ne présentaient pas les mêmes patterns d'activité cérébrale au repos (e.g. Zhang and Li 2012), ce qui a des conséquences sur la collecte des données pendant la réalisation de la tâche cognitive.

Les neurocriminologues contournent cette limite en cherchant à justifier la pertinence explicative de la stratégie neuro-interventionniste : même si cette stratégie n'est pas entièrement fondée sur le plan empirique, son efficacité explicative est garantie par le fait qu'elle permet de tester les relations de causalité entre une région donnée et ses effets sur le plan comportemental. Autrement dit, les neurocriminologues cherchent à défendre l'apport de la stratégie neuro-interventionniste pour isoler les régions ayant une responsabilité causale dans la manifestation des comportements violents ou agressifs. Cette stratégie leur donne ainsi l'occasion d'assurer leur propre légitimité scientifique, en contribuant à l'identification des mécanismes neurobiologiques impliqués dans le contrôle de soi ou la régulation émotionnelle. Pour autant, nous ne saurions insister davantage sur le fait que si les neurocriminologues cherchent à s'implanter dans le champ des neurosciences cognitives, la finalité recherchée est bien de recourir aux techniques de neurostimulation pour neutraliser le risque de récidive, puisqu'ils voient dans ces résultats les fondements possibles d'une prévention médicalisée du crime :

> « *Ultimately, such findings could help to examine whether the brain can be a direct target for potential supportive interventions in clinical settings dealing with overly aggressive patients and/or violent offenders*. » (Dambacher et al. 2015).

> « *Results provide experimental evidence that increasing activity in the prefrontal cortex can reduce intentions to commit aggression and enhance perceptions of moral judgment. Findings shed light on the biological underpinnings of aggression and theoretically have the potential to inform future interventions for aggression and violence*. » (Choy, Raine, and Hamilton 2018b).

4.3.3. Quels enjeux éthiques ?

Les débats concernant l'usage des techniques de neurostimulation sur les populations criminelles tournent essentiellement autour d'une analyse d'ordre éthique (Bomann-Larsen 2013 ; Douglas 2014 ; Fuss et al. 2015 ; Craig 2016 ; Pugh et Douglas 2016 ; Petersen 2017 ; Birks and Buyx 2018 ; Vallentyne 2018 ; Barn 2019).[156] Les

[156] Pour une réflexion plus générale sur les enjeux éthiques de la neurostimulation, voir Bell et al. (2009) ; Clausen (2010) ; Schermer (2011) ; Grant et al. (2014).

principaux enjeux ont été recensés dans un ouvrage collectif dirigé par les philosophes Thomas Douglas et David Birks (2018). Loin de trancher le débat concernant la pertinence ou la faisabilité d'une approche médicale du crime, cet ouvrage a le mérite d'amorcer une discussion éthique sur ce qu'ils nomment les « *crime-preventing neurointerventions* » (CPN) à savoir les interventions « qui exercent un effet physique, chimique ou biologique sur le cerveau dans l'objectif de diminuer la probabilité de commettre certaines formes de criminalité. » (Douglas and Birks 2018, p. 2).

Tout en préservant un juste équilibre entre les arguments favorables et défavorables, les auteurs proposent d'identifier les problèmes éthiques fondamentaux qui limitent cette stratégie neuro-interventionniste, notamment en ce qui concerne la question du consentement et de l'intégrité physique de l'individu, du libre arbitre et de l'identité personnelle (Shaw 2018 ; Vallentyne 2018 ; Bennet 2018 ; Pugh 2018 ; voir aussi Craig 2016). Ils proposent également d'interroger ce qui justifie cette approche sur le plan juridique (Forsberg 2018 ; Matravers 2018), et de déterminer les conditions de son acceptabilité morale : la question est de savoir sous quelles conditions il pourrait être moralement acceptable d'utiliser les CPN dans le cadre d'une procédure préventive, et si elles peuvent être envisagées comme une alternative à l'incarcération ou aux méthodes traditionnelles de réhabilitation (Douglas 2018 ; Bullock 2018 ; Pugh 2018 ; Lippert-Rasmussen 2018 ; voir aussi Birks and Buyx 2018 ; Koi, Uusitalo, and Tuominen 2018 ; Barn 2019).[157]

Sur ce point, les arguments favorables défendent l'idée selon laquelle la stratégie neuro-interventionniste est nettement moins restrictive pour les libertés fondamentales, mais aussi « plus efficace et plus humaine » pour réduire le risque de récidive (Birks and Douglas 2018). Certains philosophes préconisent ainsi les CPN comme une alternative à l'incarcération, ou pour réduire le temps de détention (Douglas 2018 ; voir aussi Douglas 2014 ; Ryberg 2012). Les principales réserves concernent l'atteinte à l'intégrité physique et mentale des sujets : bien que les effets de certaines CPN (notamment la stimulation cérébrale profonde) soient réversibles, elles peuvent s'avérer bien plus préjudiciable que l'incarcération, dans la mesure où elles agissent directement sur les états cérébraux du

[157] Plus précisément, il s'agit de déterminer s'il existe une différence fondamentale entre les « interventions directes », qui proposent de corriger le comportement criminel en agissant directement sur le cerveau de l'individu, et les « interventions indirectes », qui visent à modifier ses états mentaux selon une approche holistique (voir en particulier Lippert-Rasmussen 2018).

sujet (Shaw 2018 ; voir aussi Birks and Buyx 2018). Par ailleurs, la question du consentement est jugée fallacieuse, dans la mesure où les CPN s'imposent comme la seule alternative à une mesure nettement plus coercitive (l'incarcération) : en pratique, l'obtention du consentement libre et éclairé serait impossible à appliquer (du moins illusoire), puisque le criminel n'aurait pas d'autre alternative. Pour certains philosophes, il n'y aurait d'ailleurs pas de raison d'exiger le consentement avant la neuro-intervention, si elle n'est pas demandée en cas d'incarcération (Douglas 2014, 2016 ; Koi, Uusitalo, and Tuominen 2018).

Une autre préoccupation principale est évoquée dans l'ouvrage collectif dirigé par Thomas Douglas et David Birks (2018). Elle concerne la généralisation d'une approche médicalisée du crime à l'ensemble des populations criminelles, y compris aux criminels ne présentant aucun trouble mental ou organique. Effectivement, les études consacrées à la « neurocorrection » du comportement criminel proposent de faire de l'individu dangereux la cible d'une intervention jusqu'ici réservée aux délinquants sexuels ou aux criminels atteints d'un trouble neurologique ou psychiatrique.[158] La stratégie neuro-interventionniste est définie comme un traitement *thérapeutique* pour corriger les anomalies cérébrales des criminels, et ainsi améliorer leur contrôle de soi ou leur conscience morale. De ce point de vue, toutes les formes de criminalités associées à la présence d'anomalies cérébrales pourraient appeler une intervention curative, indépendamment du profil clinique de l'individu.

Or, comme le soulignent Christopher Chew et coll. (2018) dans leur revue de la littérature consacrée aux « *crime-preventing neurointerventions* », il est essentiel de distinguer les stratégies thérapeutiques visant à traiter une pathologie, des stratégies mobilisées pour corriger un comportement déviant ou améliorer les capacités cognitives ou morales de l'individu. Cette distinction s'intègre dans une réflexion plus générale sur l'élargissement du champ d'intervention des techniques de « *neuroenhancement* » (e.g. Douglas 2013). La stratégie neuro-interventionniste préconisée par les neurocriminologues s'inscrit en effet dans la continuité du basculement observé depuis l'extension du champ d'application des techniques de neurostimulation, entre un usage visant à traiter des troubles neurologiques, psychiatriques ou neuropsychiatriques, et un usage visant à améliorer les capacités cognitives ou morales de sujets sains.

[158] Pour une revue de la littérature des stratégies biomédicales utilisées pour réduire la récidive des auteurs d'infractions sexuelles, voir Alschuler (2003) et Forsberg (2018).

En somme, l'objet d'intervention privilégié par les neurocriminologues s'étend dans un enchevêtrement confus entre le domaine de la criminalité, de la psychopathologie, et de la simple altération cognitive ou morale. D'une part, la perspective d'améliorer les capacités cognitives ou morales des criminels questionne les limites entre la « normalité » et l' « anormalité », puisqu'elle invite à définir les critères objectifs permettant de justifier la pertinence d'une telle stratégie sur des sujets sains : la question est de savoir sous quelles conditions il pourrait être légitime de restaurer les capacités cognitives ou morales de sujets ne présentant aucun trouble mental ou organique. D'autre part, la généralisation de cette stratégie à l'ensemble des populations criminelles nourrit encore davantage la confusion entre le crime, la maladie mentale et les troubles de la personnalité ou du comportement, en brouillant les frontières entre la « dangerosité psychiatrique » (celle du criminel porteur d'une maladie mentale) et la « dangerosité criminologique » (l'individu dangereux qui n'appartient pas au registre pathologique). Du point de vue des neurocriminologues, ces deux catégories se confondent dans une présomption de « dangerosité pathologique » qu'il serait nécessaire de corriger, au même titre qu'un trouble neurologique ou neuropsychiatrique. En clair, tous les criminels présentant des anomalies cérébrales seraient susceptibles de faire l'objet d'une prise en charge thérapeutique, malgré l'absence de pathologie (mentale ou organique) caractérisée.

Les questions d'ordre épistémologique sont relativement absentes du débat concernant la mise en place d'une stratégie neuro-interventionniste sur les populations criminelles. Elles ne sont pas traitées dans l'ouvrage collectif dirigé par Thomas Douglas et David Birks (2018), lesquels insistent largement sur les enjeux éthiques posés par les stratégies médicamenteuses ou neurotechnologiques. En outre, en raison de la pauvreté de la littérature consacrée à la « neurocorrection » du comportement criminel, l'usage des techniques de neurostimulation n'a pas fait l'objet d'une analyse spécifique, et les réflexions sur la pertinence de cette stratégie concernent généralement les auteurs d'infractions sexuelles. Peu d'attention été porté à la question de sa généralisation à l'ensemble des populations criminelles, et les auteurs ne font pas mention des études de neuroimagerie, qui constituent pourtant le fondement empirique de la stratégie neuro-interventionniste.

Ce constat nous renvoie à une question fondamentale qui n'a jusqu'ici pas été abordée dans la littérature critique consacrée à la « neurocorrection » du comportement

criminel. Compte tenu du caractère parfois spéculatif des propositions formulées par les neurocriminologues, on peut s'interroger sur la validité empirique de la stratégie neuro-interventionniste : pour l'heure, la question n'est pas tant de savoir si elle est (éthiquement) souhaitable, mais surtout si elle est justifiée sur le plan empirique. Il est en effet crucial d'envisager la réflexion épistémologique comme un préalable aux questions d'ordre éthique ou pragmatique. Avant de questionner la faisabilité ou la valeur éthique d'une prévention médicalisée du crime, il convient d'abord de s'interroger sur la pertinence et le bien-fondé d'une telle approche.

Cette question a été abordée dans la section précédente (4.2) portant sur le statut (descriptif, explicatif ou normatif) de l'interprétation pathologique défendue par les neurocriminologues. Comme nous l'avons indiqué, la stratégie neuro-interventionniste prend pour cible le « cerveau à corriger » tel qu'il a été décrit tout au long des études utilisant la neuroimagerie. En remontant à la source de l'argument, nous avons montré que le modèle explicatif relevait d'une pétition de principe qui exclut toute interprétation alternative à l'explication pathologique. L'interprétation des données est entièrement dictée par une stratégie de pathologisation et de médicalisation du crime, ce qui limite considérablement le pouvoir explicatif de la neurocriminologie. Cette analyse nous a ainsi permis de contester les arguments empiriques qui plaident en faveur d'une conception psychopathologique du criminel : les données neurobiologiques recueillies depuis le début des années 1990 ne justifient pas le développement d'une stratégie neuro-interventionniste consistant à traiter le problème de la criminalité en corrigeant le cerveau des criminels.

Une question corolaire est néanmoins restée en suspens. Elle concerne l'apport épistémique de la stratégie neuro-interventionniste, pour étudier les relations de causalité entre les anomalies cérébrales et la manifestation d'un comportement criminel. Contrairement à la neuroimagerie qui permet de révéler des corrélations entre un phénomène psychologique donné et l'état cérébral correspondant, les techniques de neurostimulation offrent l'avantage de tester les relations de causalité entre les deux variables (e.g. Krakauer et al. 2017). La question est donc de savoir si les études consacrées à la « neurocorrection » du comportement criminel répondent à l'ambition explicative de la neurocriminologie, c'est-à-dire si elles permettent de montrer que le comportement criminel est entièrement expliqué par la présence d'anomalies cérébrales. Si l'on admet que la stratégie neuro-interventionniste constitue la finalité ultime de la recherche en neurocriminologie, il est en effet nécessaire de questionner son apport

épistémologique par rapport à la démarche localisatrice adoptée au début des années 1990.

Ce problème pose en définitive la question du pouvoir explicatif des interventions effectuées sur le cerveau des criminels. À titre d'exemple, la relation établie entre une stimulation du cortex préfrontal et l'inhibition de l'agressivité du sujet permet-elle d'affirmer que les déficits d'activité au niveau de cette région jouent un rôle causal et explicatif dans la manifestation du comportement agressif ? Plus généralement, le problème épistémologique que nous allons chercher à résoudre se décline sur plusieurs niveaux : 1/ La stratégie neuro-interventionniste permet-elle d'isoler les régions cérébrales ayant une responsabilité causale dans le développement d'un comportement violent ou agressif ? ; 2/ La stratégie neuro-interventionniste permet-elle d'identifier les mécanismes neurobiologiques impliqués dans le développement d'un comportement violent ou agressif ? 3/ La stratégie neuro-interventionniste permet-elle de fournir une explication causale complète du comportement violent ou agressif ?

Ces questions méritent d'être posées à l'aune des réflexions épistémologiques très générales concernant la nature de l'explication neuroscientifique. Dans la section suivante, nous poserons les bases du modèle mécaniste de l'explication scientifique, avant de présenter brièvement les stratégies de décomposition et de localisation qui lui sont propres : ces deux stratégies de recherche sont utilisées dans le champ des neurosciences cognitives, pour identifier les mécanismes impliqués dans la réalisation d'un comportement (Bechtel et Richardson 1993). La question sous-jacente sera ici de savoir si la stratégie neuro-interventionniste mobilisée par les neurocriminologues répond à un modèle d'explication mécaniste, c'est-à-dire si elle permet d'accéder au fonctionnement interne du mécanisme par lequel la cause produit son effet : la relation causale établie entre les déficits d'activité cérébrale et les incapacités fonctionnelles des individus permet-elle comprendre comment le mécanisme produit le comportement lorsque celui-ci fonctionne normalement ?

Nous analyserons ensuite les conditions pour lesquelles la stratégie neuro-interventionniste peut être efficace sur le plan explicatif, même si elle ne donne pas accès aux mécanismes impliqués dans la réalisation du phénomène. Le modèle interventionniste de James Woodward (2003, 2008) nous permettra d'une part, d'analyser les conditions de vérification de la relation causale inférée grâce à l'intervention, et d'autre part, de déterminer si elle est réellement informative sur le plan explicatif. Nous

verrons à cet égard que la faiblesse épistémique de la stratégie employée par les neurocriminologues ne tient pas seulement au statut de la relation entre les états cérébraux et comportementaux, mais aussi aux contraintes épistémologiques propres à l'opérationnalisation d'un comportement complexe.

4.4. La question du pouvoir explicatif de la stratégie neuro-interventionniste

4.4.1. Le modèle mécaniste de l'explication

Il revient à William Bechtel et Robert Richardson, dans leur ouvrage majeur *Discovering Complexity* (1993) d'avoir exposé l'un des plus influents modèles de l'explication scientifique, à savoir le modèle mécaniste de l'explication des phénomènes biologiques. Le modèle mécaniste a été proposé comme une alternative au modèle déductif-nomologique (DN) de l'explication scientifique. D'après ce dernier, un phénomène est expliqué lorsque la proposition qui le décrit peut être déduite d'un énoncé conjoint de lois et de conditions initiales (Hempel and Oppenheim 1948). L'explication présente alors la structure logique d'un argument déductif, constitué par un *explanandum* (le phénomène que l'on cherche à expliquer) et un *explanans* (ce qui explique le phénomène). Le phénomène décrit par la proposition particulière est expliqué par la théorie s'il peut en être déduit logiquement. Autrement dit, une bonne explication est une déduction valide de l'énoncé décrivant l'*explandum,* à partir de l'énoncé décrivant l'*explanans*.

S'il existe un important débat concernant le type d'explication le plus opérant en neurosciences, il est généralement admis qu'elles font appel à des explications de type mécaniste (Bechtel et Richardson 1993 ; Bechtel 2008 ; Craver 2007 ; Glennan 1996, 2002, 2005, 2008 ; Machamer, Darden et Craver 2000 ; Craver and Darden 2013).[159] Les

[159] D'autres types d'explication ont été proposés, notamment l'explication dite dynamique (Chemero and Silberstein 2008), sémantique (Chalmers and Jackson 2001) ou encore topologique (Huneman 2010 ; Kostic 2018).

philosophes mécanistes contestent la pertinence du modèle DN pour rendre compte de l'explication en neurosciences, notamment parce que les propositions neuroscientifiques peuvent difficilement prétendre au statut de loi (Bechtel et Richardson 1993). De manière générale, les propositions en sciences biologiques sont plutôt conçues comme des hypothèses que l'on peut tester empiriquement, dont l'objectif est de comprendre *comment* tel ou tel phénomène est engendré (Salmon 1984). En outre, l'explication en neurosciences répond à une définition fonctionnelle, c'est-à-dire qu'elle vise à identifier les mécanismes impliqués dans la réalisation d'une fonction. L'explication neuroscientifique n'est donc pas nomologique, mais systémique : elle vise à rendre compte du comportement d'un système complexe, en identifiant les fonctions exécutées par ses parties, et les mécanismes impliqués dans leur réalisation.

Pour la suite de notre analyse, il est important de préciser brièvement ce que recouvrent les concepts de « mécanisme » et d' « explication mécaniste ». Selon la définition proposée par William Bechtel et Robert Richardson, une explication est mécaniste lorsqu' « elle propose de rendre compte du comportement d'un système à partir des fonctions exécutées par ses parties et des interactions entre ses parties » (Bechtel et Richardson 1993, p. 17). Un mécanisme se définit alors comme « une structure accomplissant une fonction en vertu des parties qui la composent, des opérations qui la composent, et de leur organisation. Le fonctionnement concerté du mécanisme est responsable d'un ou de plusieurs phénomènes » (Bechtel 2008, p. 13). D'autres définitions du concept de mécanisme ont été proposées par les philosophes Peter Machamer, Lindley Darden, Carl F. Craver (2000) ou encore Stuart Glennan (2002) :

> « *Mechanisms are entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions.* » (Machamer, Darden et Craver 2000, p. 3).

> « *A mechanism for a behavior is a complex system that produces that behavior by the interaction of a number of parts, where the interactions between parts can be characterized by direct, invariant, change-relating generalizations*. » (Glennan 2002, p. 344)

Comme le precise Stuart Glennan (2008), ces définitions partagent globalement quatre caractéristiques : 1/ les mécanismes produisent des phénomènes ou des comportements ; 2/ ils sont composés de parties qui forment une structure et qui

interagissent les unes avec les autres pour produire un comportement donné ; 3/ les interactions entre ces parties sont stables et régulières, mais elles peuvent être altérées dans des circonstances particulières (par exemple sous l'effet d'une intervention sur le mécanisme en question) ; 4/ les mécanismes sont hiérarchiquement organisés, c'est-à-dire que leurs parties et les interactions entre ces parties s'opèrent sur différents niveaux d'organisation.[160] Il est surtout important de retenir qu'un mécanisme n'est pas défini par la somme des parties qui le composent, mais par les interactions entre ces parties et la manière dont elles sont organisées. Comme le souligne Glennan :

> « *Mechanisms are individuated not simply by what parts they have, but by how those parts are organized. A heap of parts does not make a mechanism. Rather the characteristic spatial, temporal, and functional organization of the parts explains the behavior of the mechanism*. » (Glennan 2008, p. 378).

A cet égard, Stuart Glennan (2005) distingue deux niveaux de description du modèle mécaniste : la caractérisation du comportement produit par le mécanisme (la description dite comportementale) et la caractérisation du fonctionnement de ce dernier (la description dite mécaniste). Le niveau de description comportemental nous renseigne sur *ce que fait* le mécanisme, tandis que la description proprement mécaniste répond à la question de savoir *comment* le mécanisme produit le comportement. Cette distinction rend plus explicite la définition proposée par Bechtel et Richardson (1993), selon laquelle l'explication mécaniste consiste à « rendre compte du comportement d'un système à partir des fonctions exécutées par ses parties et des interactions entre ses parties » : elle vise à identifier *quel* mécanisme produit le comportement, et surtout *par quel moyen* ce dernier est réalisé. La description du fonctionnement interne du mécanisme impliqué dans la réalisation du comportement (c'est-à-dire la description proprement mécaniste) apparaît alors comme une explication causale de ce comportement.[161] Elle permet d'identifier la séquence causale des évènements qui ont une responsabilité dans sa réalisation. Précisons enfin qu'il existe plusieurs relations possibles entre les descriptions comportementales et mécanistes, car un même comportement peut être produit par

[160] Cette conception est très proche de l'explication *multi-niveaux* proposée par Carl F. Craver (2009), selon laquelle les mécanismes et leur réalisation s'opèrent sur différents niveaux de description.

[161] Comme le précise Glennan, le niveau de description comportemental est analogue à l'*explanandum*, tandis que le niveau de description mécaniste joue le rôle de l'*explanans* : le premier permet de caractériser le comportement qui requiert une explication, et le second fournit l'explication.

différents mécanismes : il peut être réalisé par différents composants ou par différentes modalités d'interactions entre ces composants (Glennan 2005).

L'une des manières d'évaluer l'adéquation entre ces deux niveaux de description consiste à isoler chacun des composants du système considéré, afin de déterminer leur responsabilité causale dans la réalisation du comportement. Cette heuristique correspond à ce que Bechtel et Richardson ont appelé la stratégie de décomposition et de localisation, qui est propre à l'explication mécaniste, et qui s'applique en particulier dans le domaine des neurosciences cognitives (Bechtel and Richardson 1993 ; voir aussi Machamer, Darden, and Craver 2000 ; Bechtel and Zawidski 2002 ; Forest 2008). Elle définit de manière complémentaire la décomposition d'un système complexe en ses divers composants, et un projet localisateur qui assigne à chacun d'eux une responsabilité causale spécifique.[162]

La stratégie de décomposition caractérise les composants du système par leurs relations fonctionnelles : c'est cette décomposition fonctionnelle qui est constitutive du mécanisme. Elle permet d'isoler l'ensemble des interactions entre les composants du système, afin de comprendre comment il fonctionne. La stratégie localisatrice permet d'attribuer à chacun des composants isolés une localisation spécifique, laquelle peut correspondre à des parties individualisées du cerveau. Cette localisation peut elle-même guider la stratégie de décomposition, en assignant un rôle fonctionnel aux composants isolés. Comme le précisent Bechtel et Richardson (1993), la complémentarité de ces deux stratégies permet de réviser continuellement les hypothèses initiales, de manière à obtenir la description la plus complète possible du phénomène d'intérêt :

> « L'une des forces de la décomposition et de la localisation comme stratégie scientifique est qu'elle facilite une représentation toujours plus réaliste du domaine exploré, même lorsque la représentation initiale est considérablement déformée : les échecs de la localisation peuvent être aussi éclairants que ses succès » (Bechtel et Richardson 1993, p. 8).

[162] Nous aurons l'occasion de revenir sur cette double stratégie dans le prochain chapitre, car elle permet de rendre compte des différences (et des similitudes) entre le projet localisateur de la neurocriminologie d'une part, et de l'organologie de Franz-Joseph Gall d'autre part.

Cette double stratégie permet de guider la recherche des mécanismes neurobiologiques impliqués dans un comportement. Elle s'applique dans le domaine des neurosciences cognitives, lorsqu'il s'agit de comprendre le fonctionnement d'un phénomène complexe et d'identifier les mécanismes impliqués dans sa réalisation. Pour autant, l'explication mécaniste s'avère être une tâche difficile lorsque l'on s'intéresse à des fonctions cognitives fortement intégrées, car la dynamique neuronale est dans ce cas loin d'être linéaire.[163] La complexité des réseaux d'interaction fait qu'il est très difficile d'isoler les rôles fonctionnels associés à chacun des composants du système considéré : les stratégies de décomposition et de localisation peuvent échouer à isoler les mécanismes causalement responsables du comportement.

Dans ce cas, la recherche des mécanismes neurobiologiques exige des moyens d'investigation supplémentaires. La méthode d'inférence indirecte permet d'étudier les propriétés d'un mécanisme lorsque ce dernier n'est pas immédiatement accessible (Glennan 2005). Elle consiste à déduire la structure ou le fonctionnement interne du mécanisme, en examinant son comportement externe dans diverses conditions expérimentales : en observant les effets sur le plan comportemental d'une intervention ciblée sur l'une des causes du phénomène considéré, on peut identifier le mécanisme par lequel la cause produit son effet. Selon certains philosophes mécanistes, une altération du mécanisme peut fournir des informations pertinentes sur son mode de fonctionnement. Comme le souligne Glennan (2005) : « *[o]bserving when and how a mechanism breaks can tell one a lot about how a mechanism works. One tests the models by seeing if they break in ways similar to the actual mechanism*. » (Glennan 2005, p. 459).

De ce point de vue, la stratégie interventionniste peut s'avérer convaincante sur le plan explicatif, car elle permet de déterminer si une modification du mécanisme produit des effets sur le phénomène d'intérêt : si tel est le cas, cela signifie que le mécanisme est causalement efficace dans la réalisation du phénomène. Pourtant, il se pose la question de savoir si ce type d'intervention fournit une explication causale complète du phénomène d'intérêt. On peut en effet mettre en doute la portée explicative de cette stratégie, notamment parce que les circonstances expérimentales de l'intervention peuvent générer un *hiatus* entre le modèle comportemental (c'est-à-dire le comportement tel qu'il est

[163] Comme nous aurons l'occasion de le préciser dans le chapitre suivant (section 5.2.), le caractère dynamique des interactions entre les réseaux de neurones constitue une source majeure des difficultés posées par les stratégies de décomposition et de localisation.

opérationnalisé) et le phénomène tel qu'il est mis en jeu dans l'environnement naturel (Potochnick 2017). Si la stratégie interventionniste offre l'avantage de tester l'efficacité causale du mécanisme, il se pose la question de savoir si elle permet de comprendre comment le mécanisme produit le comportement lorsque celui-ci fonctionne normalement.

Le modèle interventionniste de l'explication proposé par James Woodward (2003) permet d'évaluer l'efficacité explicative de l'intervention, même si elle ne donne pas accès au fonctionnement interne du mécanisme : d'après ce modèle, il n'est pas nécessaire d'identifier les mécanismes impliqués dans la réalisation du phénomène pour que l'intervention soit efficace sur le plan explicatif. Les conditions pour lesquelles l'intervention garantit une explication pertinente (du moins l'explication la plus complète possible) du phénomène d'intérêt se résument dans ce que Woodward (2003) appelle « la généralisation invariante ». De ce point de vue, la question du pouvoir explicatif de la stratégie interventionniste nécessite d'évaluer le degré de stabilité de la relation qui unit l'*explanandum e*t l'*explanans* : il s'agit de déterminer les conditions pour lesquelles un changement opéré au niveau cérébral (par exemple, une neurostimulation d'une région cérébrale impliquée dans la manifestation d'un comportement agressif) produit systématiquement un changement sur le plan comportemental (en l'occurrence une inhibition des comportements agressifs).

4.4.2. Le modèle interventionniste de James Woodward

De manière générale, lorsque l'on s'interroge sur la finalité ou l'efficacité des stratégies interventionnistes dans le domaine des sciences biologiques ou médicales, il est d'usage de distinguer deux types d'interventionnisme : un interventionnisme médical, consistant à manipuler expérimentalement un état biologique dans l'objectif de rétablir une fonction ou de corriger un dysfonctionnement, et un interventionnisme à visée explicative, qui a pour fonction de fournir une explication causale d'un phénomène, en testant les relations de causalité qui l'unissent à un ou plusieurs état(s) biologique(s).

Cette deuxième stratégie correspond à ce que James Woodward (2003) nomme « *the manipulationist or interventionist account of explanation and causation* ». D'après Woodward, la stratégie interventionniste peut être conçue comme un outil heuristique

pour étudier les relations de causalité entre deux variables : elle permet de vérifier si une variable X exerce une influence causale sur une variable Y. Si en intervenant sur X, on produit un changement stable et régulier sur la valeur de Y, alors on peut affirmer que les deux variables sont liés par une relation causale.[164] En outre, le changement sur la valeur de Y ne doit pas être un simple épiphénomène de l'intervention, c'est-à-dire qu'il ne doit se produire qu'en vertu du changement de la valeur de X (Woodward 2003, 2008 ; voir aussi Craver 2007). La stratégie interventionniste permet ainsi de générer des explications causales pertinentes, en identifiant les conditions pour lesquelles un changement sur la valeur de X produit un changement sur la valeur de Y.

Précisons que la pertinence explicative de la stratégie interventionniste ne repose pas sur la formulation d'une loi stipulant les conditions nécessaires et suffisantes pour que la relation causale soit vérifiée : l'inférence causale ne repose pas sur une relation nomologique entre les deux variables, comme c'est le cas dans le modèle déductif-nomologique de l'explication. Woodward se réfère plutôt à la notion d' « invariance » (ou de « stabilité »), c'est-à-dire à l'ensemble des conditions pour lesquelles un changement dans l'*explanandum* produit systématiquement un changement dans *l'explanans* (Woodward 2003, 2008). Les relations causales nécessitent un certain degré d'invariance pour qu'elles puissent être pertinentes sur le plan explicatif. Ce degré d'invariance est évalué de la manière suivante :

> « Si la relation causale entre X et Y résiste à une vaste gamme d'interventions I_1, I_2…, I_n sur X, et si elle persiste dans plusieurs contextes C_1, C_2…, C_n, alors on peut affirmer que le degré d'invariance de cette relation causale est élevé. »

La notion d'invariance décrit l'ensemble des conditions de vérification et de généralisation de la relation causale. Elle permet de mesurer l'étendue de la dépendance causale qui unit les deux variables, de manière à garantir l'explication la plus complète possible (Potochnick 2017). Il existe en effet des cas où la relation causale est vérifiée pour un ensemble très restreint d'interventions possibles, ou dans des circonstances très précises. Inversement, elle peut rester inchangée quelles que soient les conditions, ce qui renforce la stabilité et la robustesse de la relation entre les deux variables.

---

[164] Les types d'intervention évoqués par Woodward comprennent par exemple les ablations neurochirurgicales, la répression transcriptionnelle de gènes spécifiques, ou les techniques de neurostimulation comme la stimulation magnétique transcrânienne.

D'après Woodward, l'une des manières de tester et vérifier l'inférence causale est de répondre à des questions du type « *what-if-things-had-been-different* ». Ces questions visent à identifier les conditions pour lesquelles l'*explanans* aurait été différent, si l'intervention affectait par exemple une autre variable, ou si les conditions expérimentales ou environnementales étaient différentes. Elles permettent d'isoler le facteur le plus pertinent sur le plan explicatif, c'est-à-dire celui qui satisfait toutes les conditions de l'inférence causale, quelles que soient les variations du milieu extérieur. Pour illustrer le problème de l'invariance, Woodward (2008) cite l'exemple d'une inférence causale défendue notamment par Richard Dawkins dans *The Extended Phenotype* (1982) :

> « *[C]onsider a gene Z with the following characteristics : when subjects possess a defective form they are unable to learn to read, when the possess the normal form (and when many other conditions, both genetic and environmental are satisfied, including exposure to a culture in which there is writing, availability of appropriate instruction etc.) they do learn to read.* » (Woodward 2008, p. 222).

L'affirmation selon laquelle « la version fonctionnelle du gène Z est la cause de la capacité d'apprentissage de la lecture » n'est que partiellement vérifiée, car la relation qui unit le gène et la capacité d'apprentissage est sensible à des changements qui s'opèrent à un autre niveau que le niveau génétique. Comme le souligne Woodward, une modification du gène Z produira incontestablement des effets sur l'apprentissage de la lecture, mais seulement si toutes les conditions nécessaires à l'acquisition de cette capacité sont présentes. Le cas échéant, on ne pourra pas être certain que les effets observés sur le plan comportemental sont bien le résultat d'une modification du gène, et qu'ils ne sont pas seulement liés à des circonstances socio-environnementales particulières.

Les questions de type contrefactuelles (« *what-if-things-had-been-different* ») sont alors très utiles pour vérifier si le gène Z constitue un facteur pertinent pour expliquer la capacité d'apprentissage de la lecture, et exclure la possibilité qu'une autre cause intervienne dans le modèle explicatif : la question n'est pas tant de savoir si ce niveau d'explication (en l'occurrence génétique) est plus pertinent qu'un autre, mais si la relation causale inférée grâce à l'intervention est une généralisation *invariante*. On peut par exemple se demander si les effets sur le plan comportemental resteraient inchangés quel que soit le niveau d'éducation de l'individu. Ces effets pourraient tout aussi bien résulter

de la double influence causale de l'intervention et d'un défaut d'éducation, auquel cas la pertinence explicative du gène Z s'en trouverait affaiblie.

Ainsi, dans la mesure où la relation qui unit le gène et la capacité d'apprentissage de la lecture est sensible à des facteurs ou à des conditions qui ne relèvent pas de l'intervention en elle-même, on peut en conclure que l'étendue de la dépendance causale est trop instable pour justifier une généralisation explicative. Autrement dit, si la stratégie interventionniste permet de tester et vérifier la relation de causalité, le degré d'invariance de cette relation est insuffisant pour justifier une inférence *explicative* : cette relation causale est une généralisation instable, car elle n'est vérifiée que dans des conditions particulières.

D'un point de vue général, le degré de stabilité de la relation entre l'*explanandum* et l'*explanans* permet à la fois d'opérer des manipulations causalement efficaces, et de vérifier si l'intervention est informative sur le plan explicatif. En clair, le degré d'invariance de la relation est un critère suffisant pour justifier la pertinence de la stratégie interventionniste (c'est-à-dire fournir des informations utiles à la manipulation), mais aussi pour mesurer la portée explicative de cette stratégie. Comme le résume Woodward (2008, p. 149) :

> « *Other things being equal, explanations that appeal to more stable causal relationships both provide more extensive possibilities for manipulation and control and often at least seem to provide better or deeper explanations – for one thing, such generalizations will often contribute information about more extensive patterns of dependence linking explanans and explanandum variables (than what is contributed by less stable generalizations).* »

Une dernière remarque doit être mentionnée avant d'évaluer la portée explicative de la stratégie neuro-interventionniste, telle qu'elle est mobilisée dans les études consacrées à la « neurocorrection » du comportement criminel. Comme nous l'avons précisé plus haut, les questions de type contrefactuelles ne visent pas à montrer que le facteur isolé est plus pertinent qu'un autre niveau de description ; mais il convient d'ajouter que selon Woodward, une généralisation invariante assure *de facto* l'explication la plus complète possible. Le degré d'invariance de la relation est donc un outil heuristique puissant pour montrer en quoi certains facteurs ont une pertinence causale et explicative plus grande qu'un autre niveau de description (Woodward 2008).

4.4.3. Un interventionnisme explicatif ou médical ?

Cette brève présentation du modèle mécaniste (Bechtel and Richardson 1993 ; Glennan 2005, 2008) et du modèle interventionniste de l'explication (Woodward 2003, 2008) apporte des éléments précieux pour répondre à la question du pouvoir explicatif de la stratégie neuro-interventionniste propre au champ de la neurocriminologie. Pour rappel, la question est de savoir si les études consacrées à la « neurocorrection » du comportement criminel répondent à l'ambition explicative de la neurocriminologie, c'est-à-dire si elles permettent de montrer que le comportement criminel est entièrement expliqué par la présence d'anomalies cérébrales.

Rappelons d'ailleurs que nous distinguons la question de la légitimité ou de la pertinence de la stratégie neuro-interventionniste d'une part, et la question de son *efficacité* d'autre part. Il ne s'agit pas ici de questionner la légitimité scientifique de cette stratégie (cette question ayant été traitée dans la section 4.2.) mais de déterminer son efficacité sur le plan explicatif, mais aussi sur le plan prédictif et curatif. En outre, les finalités recherchées par les neurocriminologues s'expriment à la fois sur le terrain de la pratique médicale et des neurosciences cognitives : l'usage des techniques de neurostimulation vise à la fois à corriger un dysfonctionnement, mais aussi à nourrir les réflexions propres au champ des neurosciences cognitives.[165] A ce titre, les neurocriminologues s'inscrivent dans le cadre de la méthodologie classique des neurosciences, en cherchant à identifier les mécanismes impliqués dans le contrôle de soi, la régulation émotionnelle ou encore la formation d'un jugement moral. La question que nous posons ici revient donc à préciser le statut épistémologique de la stratégie neuro-interventionniste : obéit-elle à une visée purement médicale, ou permet-elle de fournir une explication neurobiologique des comportements ?

[165] Comme nous l'avons précisé plus haut, les études consacrées à la « neurocorrection » du comportement criminel répondent au double objectif de « réduire la violence ou l'agressivité » des individus, et « d'identifier les mécanismes qui sous-tendent » ces comportements (e.g. Choy, Raine, and Hamilton 2018b).

On peut difficilement mettre en doute l'efficacité de la stratégie neuro-interventionniste sur le plan curatif. Si l'on s'en tient aux résultats des études consacrées à la « neuro-correction » du comportement criminel, cette stratégie produit incontestablement des effets sur le comportement cible : la stimulation des régions fonctionnellement déficientes empêche la manifestation de comportements violents ou agressifs. Bien que cette stratégie ne soit pas entièrement justifiée sur le plan empirique (notamment en raison du manque de spécificité fonctionnelle des régions ciblées), elle est *opérationnelle* si l'on se réfère à un interventionnisme à visée médicale. Le problème posé par cette stratégie réside précisément dans le fait qu'elle est efficace du point de vue des finalités *pratiques* de la neurocriminologie, mais absolument stérile sur le plan épistémique.

Trois arguments peuvent être avancés. Premièrement, telle qu'elle est mobilisée par les neurocriminologues, la stratégie neuro-interventionniste n'apporte aucune information sur le mode de fonctionnement du phénomène, car elle ne donne pas accès aux mécanismes neurobiologiques impliqués dans le rétablissement des fonctions cognitives. La relation de causalité est bien avérée (dans le sens où la cause produit son effet), mais elle ne permet pas de montrer *quel type* de mécanisme régit cette relation, et encore moins *par quel moyen* elle est physiquement réalisée : elle n'apporte aucune information sur la séquence causale des évènements qui produisent les effets sur le plan comportemental. D'ailleurs, si l'on reprend la distinction proposée par Glennan (2005) entre les deux niveaux de description de l'explication mécaniste, on peut sans nul doute affirmer que ces interventions nous renseignent uniquement sur le niveau comportemental de l'explication. Elles permettent seulement de déterminer *ce que fait* le mécanisme, sans que l'on ait la moindre idée des modalités de son expression, ou des interactions qui ont lieu entre ses différents composants.

Sur ce point, il convient de rappeler que les régions ciblées par les neurocriminologues diffèrent selon les études. On pourrait y voir un effort de systématisation consistant à identifier l'ensemble des mécanismes neurobiologiques impliqués dans la manifestation des comportements violents ou agressifs : en élargissant la gamme des interventions possibles, les neurocriminologues se donnent les moyens de tester toutes les configurations possibles entre un comportement donné et l'ensemble des états cérébraux qui le sous-tendent. Pourtant, force est de constater que ces interventions ne visent qu'à tester et vérifier l'efficacité causale de quelques-unes des relations établies grâce aux techniques de neuroimagerie. Elles ont le statut de *confirmation* de la relation

qui unit les comportements violents ou agressifs à certaines régions cérébrales, mais elles n'apportent aucune information supplémentaire sur les interactions possibles entre les différents réseaux de neurones. Autrement dit, elles permettent *a minima* de statuer sur la responsabilité causale de telle ou telle région cérébrale dans la manifestation des comportements violents ou agressifs, mais elles sont bien loin de fournir une explication causale complète de ces comportements.

En ce sens, il convient de préciser que l'absence de mécanisme (ou plutôt l'échec de l'explication mécaniste) n'affecte pas l'efficacité prédictive de la stratégie interventionniste. Les données issues de la neuroimagerie fournissent elles-mêmes des informations pertinentes pour prédire les effets de la neurostimulation sur le plan comportemental ; mais le pouvoir prédictif des interventions est limité par le manque de spécificité de la relation entre structure et fonction. Dans la mesure où elles ne permettent pas d'identifier le(s) mécanisme(s) par le(s)quel(s) la cause produit son effet, il est impossible d'identifier les conditions pour lesquelles la cause produira systématiquement son effet. En clair, il est possible de prédire quel type de comportement sera affecté si l'on intervient sur l'un de ses corrélats neurobiologiques, mais l'efficacité prédictive de ces interventions est limitée à une possibilité statistique.

Deuxièmement, si l'on s'appuie sur la théorie interventionniste de l'explication défendue par Woodward (2003, 2008), on peut dire que les interventions effectuées sur le cerveau des criminels ne vont pas dans le sens d'une généralisation explicative : la relation établie entre les états cérébraux et les états comportementaux ne remplit pas toutes les conditions pour justifier la portée explicative de la stratégie interventionniste. Comme nous l'avons précisé dans la description des études consacrées à la « neurocorrection » du comportement criminel, la sélection des régions d'intérêt repose parfois sur les résultats d'une seule étude de neuroimagerie. Du point de vue des neurocriminologues, la corrélation établie entre les déficits d'activité dans une région donnée et les incapacités fonctionnelles des individus est perçue comme un critère suffisant pour justifier l'usage des techniques de neurostimulation pour rétablir les fonctions cognitives des individus.

Or, le manque de spécificité fonctionnelle des régions ciblées ne permet pas de mesurer l'étendue de la dépendance causale qui unit les états cérébraux et comportementaux : il ne permet pas de garantir l'invariance de la relation entre le comportement (agressif ou violent) et l'état cérébral qui lui est associé. Il est alors

impossible de déterminer si cette région cérébrale est plus pertinente qu'une autre sur le plan explicatif, ni de s'assurer que la disparition du comportement agressif ne résulte pas d'une autre cause que celle qui a fait l'objet d'une intervention. Il est même très probable qu'en raison du manque de spécificité fonctionnelle des régions ciblées, l'intervention neutralise un autre effet que le comportement cible, accentuant encore davantage l'instabilité de la relation entre les deux phénomènes.

On peut donc contester à la fois la pertinence et la portée explicative de la stratégie interventionniste : non seulement les données issues de la neuroimagerie sont insuffisantes (car essentiellement corrélationnelles) pour garantir l'invariance de la relation entre les deux phénomènes, mais l'intervention elle-même ne remplit pas toutes les conditions de vérification de la relation causale. Celle-ci n'est vérifiée que pour un ensemble très restreint d'interventions possibles, ce qui la rend très sensible à des facteurs ou à des conditions qui ne relèvent pas de l'intervention en elle-même. A titre d'exemple, il se pose la question de savoir si la relation causale resterait inchangée quel que soit le paradigme expérimental utilisé pour mesurer l'agressivité des sujets. Il ne s'agit pas ici de contester le fait que les neurocriminologues utilisent le même paradigme avant et après l'intervention – ce choix méthodologique est pertinent pour mesurer les effets de l'intervention sur le plan comportemental – mais de s'interroger sur la sensibilité de la relation causale dans d'autres situations expérimentales. Ici encore, les questions de type contrefactuelles peuvent être utiles pour vérifier si la relation causale ne s'opère pas à un autre niveau que le niveau neurobiologique, ou exclure la possibilité qu'une autre cause intervienne en amont.

Ce dernier point nous amène à un troisième argument, qui n'est pas tant lié aux conditions de vérification de la relation causale qu'aux contraintes épistémologiques propres à l'opérationnalisation d'un comportement complexe. De toute évidence, la stratégie neuro-interventionniste est confrontée au même problème que la stratégie localisatrice reposant sur les techniques de neuroimagerie : sa portée explicative est fragilisée par l'existence d'un biais définitionnel entre les critères de définition (normatifs) du comportement cible et les modalités de son opérationnalisation. La mise en place de cette stratégie nécessite la construction d'un modèle qui réduit toute la complexité de la relation entre les états cérébraux et comportementaux, pour la transposer dans un objet opérationnalisable et manipulable : les comportements violents ou agressifs

sont réduits à des « construits hypothétiques » sur lesquels on peut intervenir par le biais de leurs marqueurs neurobiologiques.

De fait, l'opérationnalisation de ces comportements efface la pertinence explicative du contexte socio-environnemental (qui leur est pourtant inhérent) ; elle exclut les aspects normatifs de leur définition pour les figer dans une seule et même description neurobiologique. Le problème qui subsiste ici n'est pas seulement celui de l'inadéquation entre la description neurobiologique et la description comportementale, mais celui de la traduction empirique d'un concept éminemment normatif. Naturellement, le rôle causal joué par le contexte socio-environnemental ne peut s'appliquer dans les conditions expérimentales, en raison des contraintes qui pèsent sur l'opérationnalisation d'un objet aussi complexe que le comportement violent ou agressif ; pourtant, le modèle explicatif qui en résulte réintègre la dimension normative de ces comportements afin de les relier à une explication causale : la stratégie neuro-interventionniste exclut la dimension normative de la définition comportementale, alors même que son modèle explicatif réintroduit tous les aspects normatifs du comportement.

Pour résumer cette section, on peut retenir que les études consacrées à la « neurocorrection » du comportement criminel ne répondent pas à l'ambition explicative de la neurocriminologie. S'il est permis de douter de l'efficacité prédictive et curative de cette stratégie – en raison principalement du manque de spécificité fonctionnelle des régions considérées −, on ne peut en revanche que rejeter la question de son pouvoir explicatif. La stratégie neuro-interventionniste permet bel et bien d'isoler les régions cérébrales ayant une responsabilité causale dans la manifestation d'un comportement violent ou agressif, puisqu'elle offre la possibilité d'étudier les relations de causalité entre les états cérébraux et leurs effets sur le plan comportemental. Pour autant, elle ne permet ni d'accéder au fonctionnement interne du mécanisme par lequel la cause produit son effet, ni de fournir une explication causale complète du comportement criminel. Les études consacrées au « *neuroenhancement* » ou à la « neurocorrection » du comportement criminel correspondent à un interventionnisme médical qui n'a aucune portée explicative, et qui est donc peu enclin à nourrir les réflexions propres au champ des neurosciences cognitives.

L'objectif de ce chapitre était d'analyser la légitimité scientifique et la pertinence épistémologique de la stratégie de pathologisation et de médicalisation du criminel, telle

qu'elle est envisagée dans le projet neurocriminologique. Nous avons montré que cette stratégie n'était pas justifiée au regard des données empiriques, mais surtout qu'elle n'apporte aucune information sur le statut épistémologique de la relation entre les anomalies cérébrales et la manifestation du comportement criminel : l'usage des techniques de neuroimagerie ne permet pas d'identifier les mécanismes susceptibles d'expliquer les incapacités psychologiques des individus ; tout au plus permet-il d'objectiver l'existence d'une irrégularité ou d'un déficit d'activité, c'est-à-dire d'un état cérébral largement indéfini, mais suffisamment saillant pour faire l'objet d'une intervention directe, neurobiologique et spécifique. Cette analyse nous a ainsi permis de rendre compte de manière plus explicite la « vision neurocriminologique » qui structure les études menées depuis le début des années 1940 : elle montre comment la nature des finalités recherchées peut introduire un biais systématique dans l'interprétation des données. Tout au long du développement de la neurocriminologie, on voit s'opérer la construction d'un projet qui déploie son champ d'action sur le terrain médical, sans même parvenir à isoler des mécanismes pathologiques tangibles, pleinement caractérisés, et qui pourraient contribuer à démontrer la pertinence d'une explication neurobiologique du comportement criminel.

Cette stratégie se révèle très préoccupante pour le maintien de la confiance vis-à-vis des neurosciences elles-mêmes, car elle masque le pouvoir explicatif réel de ses outils. Elle contribue à fragiliser le discours neuroscientifique en brouillant les frontières entre la fonction descriptive, explicative et normative des données issues de la neuroimagerie. Les outils neuroscientifiques sont sollicités bien au-delà de leur champ d'investigation, pour répondre à une logique de protection sociale et justifier dans le même temps l'intérêt exclusif porté au cerveau dans l'étude du phénomène criminel. Les investigations empiriques consacrées à la « neurocorrection » des comportements violents ou agressifs marquent en effet un tournant dans la stratégie de légitimation de la neurocriminologie et la tentation hégémonique qui l'anime. Elles ont favorisé un déplacement des enjeux liés à l'apport des outils neuroscientifiques dans le domaine de la criminologie : de la catégorisation ou la prédiction des comportements criminels, jusqu'à un usage interventionniste qui fait des techniques de neurostimulation un outil préventif et curatif idéal pour résoudre le problème de la récidive.

À cet égard, il peut être intéressant de mettre en perspective cette stratégie avec trois présupposés définis par Eric Racine et ses collaborateurs (2005), qui caractérisent selon eux l'exportation des données neuroscientifiques hors de leur cadre d'élaboration :

le neuroessentialisme, le neuroréalisme et la neuropolitique.[166] Le premier, rappelons-le, se résume dans l'affirmation « je suis mon cerveau » ou « mon cerveau définit ce que je suis » (Roskies 2002). Il définit l'idée qu'en étudiant le cerveau de l'individu, on étudie l'individu lui-même, dans toute sa complexité et ses dimensions (subjectives, intersubjectives, socio-écononomiques, etc.). Le neuroréalisme s'articule autour du discours de la preuve et de la croyance selon laquelle les images IRMf sont des « photographies du cerveau » (Dumit 1999 ; Roskies 2007). Il justifie l'ambition de rechercher dans le cerveau la preuve objective et infaillible d'un phénomène intrinsèquement subjectif, éclipsant par la même occasion l'immense complexité d'acquisition et d'interprétation des données issues de la neuroimagerie (Forest 2014). Enfin, la neuropolitique désigne toute tentative visant à utiliser les neurosciences pour promouvoir un agenda politique ou personnel.[167] Elle est particulièrement manifeste dans les neurodisciplines, lorsqu'il s'agit de construire des catégories normatives susceptibles de résoudre des problèmes sociétaux, comme par exemple les inégalités en matière d'apprentissage dans le cas de la neuroéducation ou, en l'occurrence ici, le problème de l'insécurité et la nécessité de renforcer les mesures préventives pour lutter contre la récidive.

De toute évidence, le neuroessentialisme constitue le point de départ des recherches en neurocriminologie : si le criminel *est* (uniquement) son cerveau, alors il parait légitime de privilégier le cerveau *seul* dans l'étude des causes du crime, ou dans la prédiction et la correction du comportement criminel. Sur le plan purement théorique, le neuroessentialisme permet effectivement de garantir la légitimité et la pertinence des finalités recherchées par les neurocriminologues : si le cerveau contient toutes les

[166] Ces trois présupposés ont été formulés dans l'objectif de rendre compte de la manière dont la couverture médiatique des neurosciences contribue à masquer le pouvoir explicatif réel de ses outils. En jouant sur la confusion entre ce que les neurosciences peuvent *décrire* ou *expliquer*, ces présupposés ont finalement contribué à légitimer socialement leur usage hors du cadre médical ou scientifique. Comme nous aurons l'occasion de le préciser plus loin, le neuroessentialisme, le neuroréalisme et la neuropolitique sont en effet inhérents à la plupart des neurodisciplines. Il est important de préciser qu'ils ne sont pas constitutifs des neurosciences en elles-mêmes, mais bien de leur l'exportation hors de leur cadre théorique et méthodologique.

[167] La neuropolitique telle que l'ont défini Racine et al. (2005) ne doit pas être confondue avec le concept de « neuro-politique » (ou « *neuropolitics* »), qui correspond à un domaine d'application des neurosciences, dont la finalité est d'étudier les corrélats neuronaux des diverses opinions politiques (e.g. Connolly 2002 ; Oxley et al. 2008).

conditions matérielles nécessaires au développement de la criminalité, alors il suffirait d'intervenir directement sur le cerveau des criminels pour corriger leurs comportements. C'est pourquoi les neurocriminologues ont cherché, dès les premières publications dans les années 1940, à construire une typologie des irrégularités associées aux comportements violents ou agressifs, et pouvant faire l'objet d'une intervention spécifique. Leur stratégie consiste finalement à internaliser des processus pathologiques, c'est-à-dire à utiliser les outils neuroscientifiques (EEG, IRMa, IRMf, TEP) pour *objectiver* l'existence d'une entité pathologique parfaitement individualisée dans le cerveau des criminels. De là apparaît l'argument empirique pour répondre à l'agenda de la recherche en neurocriminologie, et promouvoir l'usage des techniques de neurostimulation comme une alternative à l'incarcération ou aux méthodes traditionnelles de réhabilitation.

Dans la perspective historique qui a servi de point de départ à notre analyse, il pourrait sembler judicieux de questionner les affinités et les différences entre la stratégie neuro-interventionniste de la neurocriminologie et les interventions psychochirurgicales de la première moitié du XX^{e} siècle. On pourrait en effet interpréter ce projet neuro-interventionniste comme un retour à la lobotomie préfrontale d'Egas Moniz (1874-1955), de Walter Freeman (1895-1972) et James W. Watts (1904-1994), ou encore aux expériences menées sur l'électrostimulation par José Delgado (1915-2001).[168] Cette question ne saurait être tranchée sans une enquête historique et épistémologique des conditions d'émergence de ces recherches, du contexte scientifique dans lequel elles s'inscrivent, et sans négliger les considérations proprement éthiques sur le caractère invasif et irréversible de ces techniques (voir à ce sujet Fangerau et al. 2011). Retenons que la résurgence de ces stratégies interventionnistes tient en grande partie au rôle prépondérant attribué au cerveau dans l'explication des comportements déviants. Les interventions psychochirurgicales ont été largement inspirées par le cas clinique de Phineas Gage et les premières découvertes physiologiques sur le rôle exercé par le cortex frontal dans la régulation émotionnelle (Missa 2013).

[168] La première lobotomie préfrontale a été pratiquée à Lisbonne en 1935 par le neurologue portugais Egas Moniz. Aux Etats-Unis, les neurologues Walter Freeman et James W. Watts ont employé massivement cette technique tout au long des années 1940-1950. Malgré le prix Nobel de physiologie et de médecine attribué à Egas Moniz en 1949, qui récompensait « la valeur thérapeutique de la leucotomie dans certaines psychoses », la psychochirurgie a été largement discréditée avec la découverte des premiers neuroleptiques comme la chlorpromazine, au début des années 1950 (Tierney 2000 ; Missa 2006).

On peut également situer ces interventions dans la continuité des études phrénologiques et criminologiques du XIX^{e} siècle, dans la mesure où elles ont posé les principes fondamentaux de la localisation cérébrale des instincts violents (Rafter 2008 ; Pustilnik 2009). C'est pourquoi il convient désormais d'expliciter chacun des termes de la filiation entre l'organologie de Franz-Joseph Gall, l'anthropologie criminelle de Cesare Lombroso, et le projet neurocriminologique qui se déploie depuis les années 1940. L'intérêt exclusif porté au cerveau dans l'étude, la prédiction ou la correction du comportement criminel nous invite à remonter aux sources de cet héritage, en reconstituant la cadre théorique et méthodologique qui structure le discours neurobiologique sur le crime.

CHAPITRE 5

RECONSTRUCTION HISTORIQUE DE LA CONCEPTION (NEURO)BIOLOGISANTE DU CRIMINEL

Peu de recherches ont exploré la question de la filiation entre la neurocriminologie et la phrénologie ou la criminologie du XIX^e^ siècle, aussi bien dans le champ de l'épistémologie et de l'histoire des sciences, que sur le terrain de la criminologie elle-même. Les quelques études publiées à ce jour montrent qu'elle s'apparente à une « nouvelle phrénologie » (Markowitsch 2008 ; Dartigues 2017), ou encore qu'elle réactive l'idéal d'une prévention médicalisée du crime (Larregue 2019). Citons également les deux seules études ayant mis en perspective les thèses de Lombroso avec les recherches consacrées à la neurobiologie de la violence (Pustilnik 2009) ou aux bases neurogénétiques de la criminalité (Sirgiovanni 2017). Ces recherches présentent des limites en raison du fait qu'elles ne reposent pas sur une analyse explicite des présupposés, des méthodes ou des concepts mobilisés dans la littérature. Elles relèvent en effet davantage d'une critique *a priori* de l'approche biologisante du phénomène criminel que d'un examen approfondi du cadre conceptuel et méthodologique dans lequel les recherches sont menées : l'analyse proprement empirique est délaissée au profit d'une enquête discursive centrée sur les prétentions de la neurocriminologie à investir le champ pénal.[169] Le discours neurocriminologique est d'ailleurs souvent réduit à son incarnation emblématique, le neuroscientifique Adrian Raine, dont les ambitions et les stratégies de recherche ne sont pas toujours représentatives du champ.

[169] Ces analyses entretiennent une confusion entre la neurocriminologie et le neurodroit, car elles ne rendent pas compte des spécificités méthodologiques et des finalités propres aux deux champs. Comme nous l'avons mentionné précédemment, la neurocriminologie s'est d'abord développée dans un *continuum* de réflexions criminologiques et judiciaires, avant de se structurer autour de l'ambition de fonder l'appréciation de la dangerosité sur des critères neurobiologiques : elle ne s'intéresse pas à la question de la responsabilité des criminels présentant d'importantes anomalies cérébrales. Le neurodroit étudie quant à lui l'apport des neurosciences pour informer les questions relatives à l'évaluation de la responsabilité pénale. Il répond clairement à la volonté d'intégrer les neurosciences dans le système pénal, notamment pour combler les insuffisances de l'expertise psychiatrique.

En outre, l'analyse critique de la littérature fait fi de la profondeur historique du programme de recherche puisqu'elle ne fait aucune mention des publications des années 1940-1980. La naissance du champ disciplinaire est généralement située au début des années 1990, date de publication des premières études utilisant la neuroimagerie anatomique ou fonctionnelle ; c'est pourquoi la plupart des critiques se sont articulées autour de la réactualisation du principe organologique de Franz-Joseph Gall. Rappelons que si cette filiation semble justifiée par l'adoption d'une démarche localisatrice, elle n'a jamais été réellement démontrée sur le plan épistémique. Compte tenu des implications majeures qu'elle revêt sur le pouvoir explicatif de la neurocriminologie, et étant donné la grande complexité de l'organologie de Gall, l'argument mérite une analyse approfondie des présupposés, des concepts et des méthodes gouvernant les deux approches. Il est en effet remarquable que leurs affinités ne se limitent pas à une simple démarche localisatrice basée sur une conception modulaire du fonctionnement cérébral. La neurocriminologie repose sur un certain nombre de principes hérités de l'organologie de Gall, qu'il est nécessaire d'expliciter. Reste à démontrer ce qui différencie l'approche phrénologique consistant à localiser l' « organe du meurtre » − condition matérielle nécessaire au développement de l' « instinct criminel » − du projet neurocriminologique qui propose de recenser les anomalies cérébrales impliquées dans la manifestation d'un comportement impulsif, violent ou agressif. Il convient en effet de déterminer si la stratégie adoptée par la neurocriminologie tire un bénéfice explicatif plus grand qu'une instanciation directe des instincts individuels dans des régions individualisées du cerveau.

Par ailleurs, l'analyse critique de la neurocriminologie a jusqu'ici négligé l'autre acteur de cette continuité historique, Cesare Lombroso, dont les thèses sont pourtant bien plus problématiques du point de vue épistémologique. L'analyse de la littérature nous a déjà permis d'entrevoir quelques éléments attestant cette filiation, notamment la démarche typologique et classificatoire qui préside à l'étude du cerveau des criminels depuis le début des années 1940, ou encore la stratégie de pathologisation du crime qui se déploie tout au long du développement de la neurocriminologie. Cette filiation mérite une analyse attentive des présupposés sur lesquels elle repose, car ces derniers nous renseignent sur le pouvoir explicatif de ce programme de recherche.

Si l'on veut prendre la mesure de cet héritage historique, l'examen comparatif des différentes thèses doit répondre à deux exigences : nous proposons d'une part, de reconstruire le paradigme de la neurocriminologie à partir de l'analyse conceptuelle et

méthodologique de la littérature publiée depuis le début des années 1940, et d'autre part, de montrer en quoi les thèses de Franz-Joseph Gall et Cesare Lombroso sont réinvesties à l'intérieur du champ de la neurocriminologie.

La première exigence nous engage à proposer un cadre théorique unifié, susceptible de rendre compte de l'ensemble des résultats publiés dans le domaine. Si nous avons jusqu'ici centré notre analyse sur les ambitions du projet, les stratégies de recherche et les concepts employés dans la littérature, il convient désormais d'interroger les présupposés qui sous-tendent la construction des études publiées depuis le début des années 1940. Nous avons montré que l'interprétation des données répondait à un modèle normatif visant à promouvoir les finalités recherchées par les neurocriminologues. La circularité de la démarche résulte d'un présupposé causaliste et d'un biais de confirmation en faveur d'une conception pathologique du comportement criminel : l'interprétation pathologique et l'inférence causale s'imposent comme la seule grille de lecture des données neurobiologiques. Dans la continuité de cette analyse, nous montrerons que toutes les hypothèses avancées pour rendre compte des anomalies observées reposent sur un postulat innéiste, et sont conçues de telle sorte à justifier la prééminence du cerveau dans l'explication du comportement criminel.

La deuxième exigence est d'isoler les grands principes théoriques et méthodologiques qui attestent d'une filiation entre le projet neurocriminologique et les thèses issues de la phrénologie et de la criminologie du XIX^{e} siècle. Nous voudrions à cet égard montrer que cette continuité historique se déploie au-delà d'une simple entreprise de naturalisation du crime, et qu'elle ne peut se réduire à un débat théorique sur la réintroduction d'un déterminisme biologique dans l'explication de la criminalité. Notre analyse ne se limitera d'ailleurs pas à un examen comparatif des méthodes ou des résultats issus des différentes approches, puisqu'elle vise à montrer que les thèses de Gall et Lombroso imprègnent toutes les étapes de la construction des études : des stratégies de recherche privilégiées par les neurocriminologues jusqu'aux concepts mobilisés dans l'interprétation des données, en passant par les présupposés qui sous-tendent chacun des choix méthodologiques opérés.

Dans cette perspective, il paraît important de rappeler les différents modèles interprétatifs défendus depuis le début des années 1940 : de l'hypothèse d'une « immaturité corticale » à celle d'un « déséquilibre cérébral », l'interprétation à donner aux anomalies cérébrales des criminels a beaucoup évolué en fonction des outils utilisés

et des données disponibles. La première étape de notre analyse propose donc de reconstituer le cadre conceptuel et méthodologique commun à l'ensemble des hypothèses explicatives avancées par les neurocriminologues (section 5.1). Cela nous permettra d'une part, de dégager le principe théorique fondamental dans lequel se déploie le paradigme de la neurocriminologie, et d'autre part, d'isoler un certain nombre de présupposés hérités des études phrénologiques et criminologiques du XIX^{e} siècle.

Nous montrerons dans un deuxième temps que ce paradigme est gouverné par l'adoption d'un postulat innéiste, fruit d'un double héritage phrénologique et lombrosien (section 5.2). Après avoir rappelé les différentes formes que recouvre l'innéisme criminel dans la phrénologie et l'anthropologie du XIX^{e} siècle, nous procèderons à un examen attentif du projet localisateur de la neurocriminologie contemporaine. L'objectif est de déterminer si cette stratégie localisatrice découle d'une conception innéiste du comportement criminel, au même titre que la localisation phrénologique du « penchant au meurtre ». S'il existe de bonnes raisons de voir la neurocriminologie contemporaine comme une simple réactualisation du principe organologique, nous verrons que le problème n'est pas tant lié à la démarche localisatrice en elle-même, mais à la confiance épistémique que l'on peut accorder aux données issues de la neuroimagerie, et aux conditions d'opérationnalisation des processus psychologiques.

La question de la filiation entre la neurocriminologie et l'anthropologie lombrosienne fera l'objet d'une analyse plus détaillée, qui entend dépasser les enjeux purement méthodologiques concernant la démarche typologique et classificatoire adoptée par les neurocriminologues. Bien qu'elle se garde de toute référence à l'anthropologie criminelle du XIX^{e} siècle, la neurocriminologie s'est construite autour d'un cadre conceptuel et méthodologique qui réactualise la théorie du criminel-né à la lumière des données neurobiologiques. Cette analyse nous permettra d'invalider le pouvoir explicatif de la neurocriminologie : à défaut d'isoler les mécanismes neurobiologiques responsables des incapacités fonctionnelles des individus, le paradigme de la neurocriminologie est condamné à confirmer des présupposés hérités de l'anthropologie criminelle italienne, notamment l'idée que le criminel est porteur d'une disposition intrinsèque et irrésistible.

Par ailleurs, et dans un souci d'anticiper les objections qui pourraient être formulées à l'encontre de ce « procès de l'innéité », nous montrerons l'échec des tentatives visant à étudier la dynamique des interactions entre les facteurs socio-environnementaux et les facteurs neurobiologiques (section 5.3). Tous les efforts déployés par la neurocriminologie pour devancer les accusations de cérébro-centrisme ou

d'innéisme ne font que réactualiser les grands principes de l'organologie de Gall et de la typologie criminelle de Lombroso. Le paradigme de la neurocriminologie ne laisse aucune marge de manœuvre pour interpréter les données à la lumière des caractéristiques socio-économiques ou de l'histoire personnelle des individus, ce qui ne fait que renforcer l'existence d'un biais de confirmation en faveur d'une conception innéiste du comportement criminel.

Enfin, nous tenterons de prendre de la hauteur sur la question de l'innéisme criminel, en analysant les deux principales thèses qui sous-tendent cette analyse : le cérébro-centrisme, qui favorise l'idée que le cerveau *seul* suffit à expliquer l'ensemble des comportements individuels, et l'internalisme des processus psychologiques, à savoir la thèse selon laquelle les conditions d'émergence du comportement criminel sont internalisées dans des états cérébraux spécifiques (section 5.4). L'examen critique de ces thèses nous donnera l'occasion de sortir du cadre de référence de notre analyse, pour montrer en quoi la neurocriminologie illustre de manière paradigmatique certaines critiques formulées à l'égard des neurodisciplines. Plusieurs enjeux épistémologiques seront discutés tout au long de cette section, notamment la question de savoir si l'on peut donner un sens à l'affirmation selon laquelle le criminel (ou l'individu en général) est « d'essence cérébrale », ou s'il est possible de tirer des conclusions normatives à partir des seules données neurobiologiques. En prenant pour objet certaines applications envisagées dans le champ du neurodroit, nous proposerons quelques arguments pour invalider les tentatives visant à fonder des catégories normatives (relatives, par exemple, à la question de la responsabilité ou de la dangerosité d'un individu) sur des critères neurobiologiques.

5.1. Le paradigme de la recherche en neurocriminologie

5.1.1. Cinq hypothèses explicatives

La question de savoir comment interpréter les anomalies cérébrales des criminels a fait l'objet d'un long débat, tout au long de la seconde moitié du XX^{e} siècle (Williams 1941 ; Hill 1944 ; Volavka 1990 ; Pillmann et al. 1999 ; Lindberg et al. 2005 ; Patrick 2008 ; Loeber and Pardini 2008). Comme nous l'avons signalé dans le second chapitre (section 2.1), les premières publications ont évoqué l'hypothèse d'une « immaturité corticale » (Hill and Watterson 1942) ou d'un « défaut de développement du système nerveux central » (Hill 1944). Rappelons qu'il s'agit d'une hypothèse purement théorique, dont la première mention apparaît dans une étude menée sur le cerveau de psychopathes admis dans une institution psychiatrique. C'est en observant des similitudes entre les données EEG de jeunes enfants et de psychopathes « à prédominance agressive » que les psychiatres Denis Hill et Donald Watterson (1942) ont suggéré l'idée qu'un défaut de maturation cérébrale pouvait être à l'origine de ces anomalies. Cette hypothèse s'est progressivement imposée dans le champ de la neurocriminologie, puisqu'elle a été invoquée pour expliquer l'origine des anomalies EEG des criminels jusqu'à la fin des années 1960 (Silverman 1943 ; Hill 1944, 1955 ; Stafford-Clark et al. 1951 ; Williams 1969 ; voir aussi Milstein 1988).

Il convient surtout de rappeler que ce modèle interprétatif s'est constitué selon un principe d'identité psychophysiologique, à partir du recoupement des données EEG avec un ensemble de critères d'appréciation relatifs aux traits psychologiques des individus (impulsivité, irritabilité, agressivité, etc.). Les neurocriminologues ont fait le constat que les criminels et les psychopathes présentaient des caractéristiques neurophysiologiques et psychologiques que l'on rencontre habituellement chez les enfants. L'interprétation des données EEG s'est alors structurée autour du principe selon lequel les criminels et les

psychopathes sont « immatures », tant sur le plan comportemental que cérébral : leur « immaturité corticale » reflèterait leur « immaturité comportementale ».[170]

À partir des années 1950, les études conjuguant les mesures EEG avec des tests de personnalité ont encouragé la formulation d'un autre principe d'identité psychophysiologique : celui d'une association entre « l'instabilité de la personnalité » et « l'instabilité de l'EEG » (e.g. Levy and Kennard 1953). Ce principe de correspondance est issu d'une démarche typologique consistant à opposer deux grandes catégories neurophysiologiques : un EEG « normal » et « régulier », que l'on observe généralement chez les criminels ayant commis un crime isolé et/ou présentant une structure de la personnalité relativement stable, et un EEG « très anormal », « instable » et « irrégulier » chez les criminels violents et/ou récidivistes. De ce point de vue, le tracé EEG révèle une irrégularité proportionnelle à l'instabilité émotionnelle et comportementale des individus : « l'instabilité de l'EEG » apparaît comme un critère d'évaluation quantitative de la dangerosité de l'individu, le degré d'irrégularité du tracé EEG étant proportionnel au niveau de violence ou d'impulsivité des sujets.

Ce principe finira par s'imposer comme le modèle privilégié pour expliquer le cas particulier des crimes violents et sans motif apparent, tout au long du développement de la neurocriminologie. Les neurocriminologues s'efforceront de démontrer un lien entre la manifestation d'un « comportement paroxystique » et l'observation d'un « EEG paroxystique », en instaurant un principe d'équivalence en termes de propriétés entre la nature du crime et la nature de l'anomalie observée (e.g. Hill 1952 ; Howard 1984).[171]

En somme, depuis le début des années 1940, la caractérisation neurophysiologique des criminels est marquée par le souci d'accorder la forme et la nature du tracé EEG avec les traits de personnalité de l'individu (immaturité comportementale ou morale, impulsivité, instabilité) ou les caractéristiques relatives aux infractions commises (nature du crime, fréquence de récidive, etc.). L'interprétation des données EEG recueillies chez les criminels a ainsi contribué à instaurer un double principe d'identité psychophysiologique : entre l'« immaturité comportementale » et l'« immaturité

[170] Daniel Silverman (1944), qui a publié la première étude neurophysiologique sur des psychopathes ayant fait l'objet d'une condamnation criminelle, résumait ce parallélisme psychophysiologique ainsi : « *the psychopath's brain, as well as his personality, is immature* » (Silverman 1944, p. 41).

[171] Comme nous l'avons vu dans le premier chapitre, la formulation de ce principe a été favorisée par l'hypothèse d'un lien entre l'épilepsie et la manifestation d'un comportement criminel.

corticale » d'une part, et entre l'« instabilité de de la personnalité » et l'« instabilité de l'EEG » d'autre part. Ce dernier principe s'est réaffirmé au travers de l'association établie entre l' « EEG paroxystique » et le « comportement paroxystique » des criminels les plus violents. Il a permis de fixer les termes d'une dangerosité identifiable sur le plan neurophysiologique, car décrite comme le stigmate d'un comportement aux propriétés équivalentes.

Depuis le développement de la neuroimagerie anatomique et fonctionnelle, la typologie criminelle a bénéficié d'une stratégie localisatrice consistant à recenser les anomalies cérébrales associées aux incapacités fonctionnelles des individus. Cette stratégie propose d'évaluer le volume ou le niveau d'activité des structures cérébrales impliquées dans diverses aptitudes cognitives, émotionnelles ou sociales (notamment la capacité à contrôler son comportement). Pour les neurocriminologues, il ne s'agit plus simplement d'établir une taxonomie des relations psychophysiologiques, mais de défendre un lien de causalité entre les incapacités fonctionnelles des individus et les dysfonctionnements dans les régions qui leur sont associées. Cette stratégie vise à répondre à l'ambition explicative de la neurocriminologie : l'objectif n'est pas seulement d'isoler les caractéristiques neurobiologiques propres à chaque type de criminalité, mais surtout de montrer que le comportement criminel est entièrement *expliqué* par la présence d'anomalies cérébrales.

De manière générale, ce modèle explicatif consiste à mettre en relation les déficits neurologiques avec les déficits observés sur le plan cognitif, affectif ou social (e.g. Raine et al. 1998a ; Aigner 2000). Il propose par exemple d'associer la réduction du volume ou de l'activité des régions préfrontales avec une réduction des aptitudes cognitives ou émotionnelles des criminels : altération du contrôle de soi, défaut de régulation émotionnelle, incapacité à former un jugement moral, etc. En clair, l'interprétation des données neurobiologiques consiste à instaurer un lien de causalité entre une insuffisance dans une région donnée et une déficience majeure sur le plan psychologique et social : elle s'est structurée autour d'un modèle consistant à expliquer les « déficits psychosociaux » des individus par des « déficits structurels ou fonctionnels » parfaitement localisés dans leur cerveau.

L'interprétation des données neurobiologiques recueillies chez les criminels s'est également articulée autour de la notion de « déséquilibre cérébral », de « déséquilibre inter-hémisphérique » ou de « déséquilibre fronto-limbique » (Raine et al. 1998b ; Hecht

2011 ; Angus et al. 2016 ; Siep et al. 2018). Ce modèle explicatif consiste à montrer l'existence d'une dérégulation des systèmes cognitifs et émotionnels, à partir des irrégularités ou des asymétries observées entre les deux hémisphères, ou entre les régions corticales et sous-corticales. Les neurocriminologues évoquent par exemple un lien entre l'impulsivité de certains criminels et le déséquilibre induit par une « hyperactivité sous-corticale » et une « hypoactivité préfrontale » (e.g. Raine et al. 1998b) : l'excès d'activité dans les régions limbiques n'est plus modulé par le cortex préfrontal, ce qui entraîne une altération du contrôle émotionnel, attentionnel ou motivationnel. Précisons que l'hypothèse d'un « déséquilibre cérébral » s'est également imposée dans les études EEG, les neurocriminologues évoquant un lien entre les « asymétries hémisphériques » et la manifestation d'un comportement agressif. Ces asymétries se traduisent généralement par une latéralisation fonctionnelle – par exemple, une hyperactivité de l'hémisphère gauche et une hypoactivité de l'hémisphère droit – pendant le traitement de stimuli aversifs (e.g. Keune et al. 2012, 2018).

De manière générale, il est intéressant de constater que ces irrégularités ou ces asymétries sont toujours attribuées à une dérégulation des systèmes de contrôle cognitif ou émotionnel. D'après les neurocriminologues, l'instabilité émotionnelle ou comportementale des criminels résulterait d'une désorganisation générale de leur cerveau. Autrement dit, l'interprétation des données vise à démontrer un lien entre ce « déséquilibre cérébral » et le « déséquilibre psychologique » des individus.

Enfin, si l'on tient compte de l'ensemble des études utilisant l'EEG et la neuroimagerie (anatomique ou fonctionnelle) sur les populations criminelles, on voit apparaître un autre principe d'identité psychophysiologique, fondé non plus sur la nature des anomalies mais sur leur degré de sévérité. Quel que soit l'outil d'exploration cérébrale utilisé, l'excès d'anomalies cérébrales est décrit comme un indicateur neurobiologique de la dangerosité de l'individu. Depuis les premières publications relatives à la psychopathie et aux troubles du comportement pendant l'enfance, les neurocriminologues s'accordent à montrer que les individus les plus dangereux sont associés au plus haut degré d'anomalies cérébrales. L'association établie entre le niveau de dangerosité et le degré d'anomalies cérébrales constitue d'ailleurs l'un des résultats les plus significatifs du domaine de la neurocriminologie. L'immense majorité des études EEG montre que la proportion d'anomalies est corrélée à la sévérité du crime commis, la fréquence de récidive, ou encore au niveau de violence, d'impulsivité ou d'agressivité des sujets (e.g.

Hill et Watterson 1942 ; Stafford-Clark et Taylor 1949, 1951 ; Levy et Kennard 1953 ; Williams 1969 ; Howard et al. 1984 ; Convit et al. 1991 ; Pillmann et al. 1999 ; Aigner et al. 2000 ; Lindberg et al. 2005). Dans les études de neuroimagerie, le critère de dangerosité de l'individu est fixé non seulement par le degré d'anomalies, mais aussi par leur localisation dans le cerveau. Les anomalies préfrontales apparaitraient comme un critère déterminant dans l'appréciation de la dangerosité du sujet, parce qu'elles affectent les régions impliquées dans le contrôle de soi ou la régulation émotionnelle.

5.1.2. Le criminel ressemble à son cerveau

Si l'on résume les différents modèles interprétatifs que nous venons d'évoquer, on voit apparaître un principe méthodologique commun à l'ensemble des études menées sur le cerveau des criminels. Les conclusions de ces études reposent systématiquement sur le recoupement des données issues de l'EEG ou de la neuroimagerie, avec un ensemble de critères relatifs aux traits de personnalité des individus ou à la nature des crimes commis. Plus encore, tous les modèles interprétatifs défendus depuis les années 1940 s'appliquent à confondre les désordres du cerveau et du comportement dans un processus de même nature. Quel que soit l'outil d'exploration cérébrale utilisé, l'interprétation des données est dictée par le souci de démontrer une équivalence en termes de propriétés entre les anomalies observées sur le plan neurologique et comportemental. Elle vise à instaurer un principe de correspondance entre un état psychologique évocateur d'un comportement potentiellement dangereux (immaturité, instabilité, impulsivité, etc.) et un état cérébral décrivant le même type de propriété.

Ce principe peut être résumé en six propositions, chacune définie par un rapport d'identité entre une catégorie psychologique ou comportementale donnée et une catégorie neurobiologique spécifiant la nature ou le degré des anomalies observées. Ces six propositions découlent directement des choix méthodologiques opérés pour définir les critères d'appréciation de la dangerosité criminelle. Elles sont le résultat des efforts conjoints de plusieurs centaines d'études menées dans l'objectif de dresser le « profil neuropsychologique » de l'individu dangereux :

- son « immaturité comportementale » correspond à une « immaturité corticale » (e.g. Hill and Watterson 1942) ;

- son « instabilité psychologique » est associée à un « EEG instable » (e.g. Levy and Kennard 1953) ;
- son « comportement paroxystique » est associé à un « EEG paroxystique » (e.g. Hill 1952) ;
- ses « déficits psychosociaux » sont corrélés à des « déficits structurels et fonctionnels » (e.g. Raine et al. 1998a) ;
- son « déséquilibre psychologique » est associé à un « déséquilibre cérébral » (e.g. Raine et al. 1998b) ;
- son « niveau de dangerosité » est proportionnel à son « degré d'anomalies cérébrales » (e.g. Stafford-Clark and Taylor 1949 ; Aigner et al. 2000).

En somme, depuis le début des années 1940 jusqu'aux récentes recherches utilisant la neuroimagerie anatomique ou fonctionnelle, l'interprétation des données tend à montrer que le criminel ressemble à son cerveau : « immature », « instable », « paroxystique », « déficient », « déséquilibré », et d'autant plus dangereux qu'il est porteur d'un grand nombre d'anomalies. Plus encore, dans la mesure où l'interprétation des données est corrélée à la fois aux traits de personnalité de l'individu et aux caractéristiques relatives aux infractions commises, on peut supposer que ce principe s'est formé dans l'objectif de relier le crime à son auteur. En assimilant les catégories juridiques, psychologiques et neurobiologiques, les neurocriminologues cherchent à établir une équivalence nomologique entre le criminel, son cerveau, et son crime : il s'agit de montrer que le criminel ressemble à la fois à son cerveau et à son crime.

Il se pose la question de savoir si ce principe est issu d'une véritable démonstration empirique, ou s'il est l'expression d'un présupposé inhérent à l'approche biologique de la criminalité. De ce point de vue, la relation entre les états psychologiques et cérébraux relèverait davantage d'une identification théorique visant à circonscrire une typologie des relations psychophysiologiques, que d'une identité empirique résultant d'une analyse approfondie du point de vue fonctionnel. Autrement dit, on peut se demander si l'association établie entre le crime, le criminel, et son cerveau n'est pas l'expression d'un présupposé internaliste (ou innéiste), qui consiste à montrer que la dangerosité du criminel s'exprime au travers d'une potentialité cérébrale préexistante.

Sur ce point, deux éléments de réponse peuvent être avancés. D'abord, il convient de préciser que les six propositions que nous avons isolées précédemment définissent

l'unique grille de lecture des données recueillies chez les criminels. Chaque état psychologique est lié à sa description neurobiologique dans un rapport d'équivalence descriptive qui exclut toute interprétation alternative. La stratégie consiste à confondre les états psychologiques et cérébraux dans un processus de même nature, afin de justifier l'inférence causale entre les deux états : il s'agit de rendre équivalent les deux niveaux de description, afin de montrer que les propriétés de niveau supérieur (par exemple, un défaut de maturité comportementale ou morale) héritent des propriétés de niveau inférieur (un défaut de développement du cortex cérébral).

L'interprétation des données révèle en effet une coextensivité des concepts décrivant chacun des états psychologiques et cérébraux : la force et la spécificité de leur relation sont assurées par le fait qu'ils dénotent le même type de propriété. La valeur extensive des propriétés engage ainsi un rapport nomologique entre les deux niveaux de description, ce qui permet d'assigner à chaque état psychologique une seule et même explication neurobiologique : l' « immaturité comportementale » des individus serait entièrement expliquée par leur « immaturité corticale », leur « instabilité psychologique » par l' « instabilité de leur tracé EEG », etc. Cette stratégie s'avère particulièrement efficace vis-à-vis de l'ambition explicative de la neurocriminologie, puisqu'elle dispense les neurocriminologues de rechercher l'ensemble des mécanismes étiologiques. Chaque modèle interprétatif est conçu de telle sorte à exclure la possibilité qu'un état psychologique soit associé à une autre configuration cérébrale, alors même qu'aucune relation de causalité n'est démontrée.

Par ailleurs, la spécificité de la relation établie entre les états psychologiques et cérébraux introduit un soupçon d'argument *ad hoc* quant au principe selon lequel le criminel ressemble à son cerveau. En cherchant à confondre les catégories juridiques, psychologiques et neurobiologiques, la neurocriminologie vise à doubler le crime dans le cerveau du criminel. La stratégie consiste à faire apparaître les motifs neurologiques et comportementaux qui relient le crime à son auteur, dans l'objectif d'identifier dans son cerveau *ce qui le rend* criminel. De ce point de vue, le rapport d'équivalence épistémique qui unit les états psychologiques et cérébraux répond à l'ambition de rechercher l'anomalie cérébrale qui préfigure une irrégularité comportementale donnée. Il vise à montrer que la dangerosité de l'individu s'exprime au travers d'une potentialité préexistante, d'un état cérébral qui détermine lui-même les conditions d'émergence d'un comportement impulsif, violent ou agressif.

De toute évidence, cette stratégie se fait l'écho d'une conception préformationniste du développement du comportement criminel, car elle suggère que les conditions de sa manifestation sont le produit de leur propre déploiement interne.[172] Les comportements violents, agressifs ou impulsifs résulteraient d'une « auto-construction » endogène suivant un schéma prédéterminé émergeant des états cérébraux vers les états psychologiques du sujet : de l' « immaturité corticale » vers l' « immaturité comportementale », de l' « instabilité de l'EEG » à l' « instabilité psychologique », etc. Ainsi de ce point de vue, les différents modèles interprétatifs défendus depuis les années 1940 constituent autant d'éléments démontrant la dangerosité intrinsèque des criminels. Cette dangerosité serait en quelque sorte « déjà-là » dans leur cerveau, puisqu'elle se manifesterait par le biais d'un processus préexistant.

En somme, le principe selon lequel le criminel ressemble à son cerveau répond à la mise en œuvre d'une stratégie heuristique pour guider la recherche des corrélats neurophysiologiques ou neurobiologiques du comportement criminel. Chacune des anomalies cérébrales est recherchée en tant que « preuve » d'une insuffisance fonctionnelle ou d'une irrégularité comportementale donnée. Or, la faiblesse épistémique de cette stratégie réside dans le fait qu'elle contraint la construction du modèle explicatif au prix d'une pétition de principe : les neurocriminologues s'efforcent de montrer que le criminel (et plus précisément son cerveau) porte en lui les germes d'une criminalité latente, tout en garantissant un modèle exclusif de toute autre explication possible.

5.1.3. Les prémisses de la filiation

Ici apparaît clairement l'une des sources de la filiation entre la neurocriminologie, l'organologie de Franz-Joseph Gall, et l'anthropologie criminelle de Cesare Lombroso.

[172] Dans la mesure où les neurocriminologues ne font pas référence à la biologie du développement, le concept de « préformationnisme » que nous employons ici ne se rapporte pas à l'ontogénèse au sens de l'*embryogénèse*, mais plutôt au sens de la psychologie du développement. Le concept d'internalisme pourrait sembler plus approprié, mais il ne rendrait pas compte de la dynamique « auto-constructive » qui sous-tend l'exercice des phénomènes psychologiques : du point de vue du paradigme de la neurocriminologie, ces derniers ne sont pas seulement *intrinsèques* au cerveau du criminel ; ils préexistent dans des états cérébraux spécifiques, avant leur manifestation.

Le principe selon lequel le criminel ressemble à son cerveau et à son crime est solidaire d'un postulat innéiste qui, bien qu'il ne soit jamais explicité dans la littérature, se décline dans les différents modèles explicatifs défendus par la neurocriminologie. L'ambivalence de ce programme de recherche réside précisément dans le fait que, contrairement à la phrénologie et à la criminologie du XIX^{e} siècle, la neurocriminologie se défend de fonder l'étude du cerveau des criminels sur un postulat innéiste : la spécificité de sa démarche est de faire passer ce présupposé pour le résultat d'une démonstration empirique. Les neurocriminologues prétendent démontrer la dangerosité intrinsèque des individus, alors même que toutes les données recueillies sont interprétées selon un modèle préformationniste du comportement criminel.

Avant de reconstituer dans le détail la circularité de cette démarche, il convient de préciser que l'héritage phrénologique et lombrosien est loin de se limiter à l'adoption d'un postulat innéiste. Les six propositions que nous avons isolées précédemment révèlent les fondements d'une filiation conceptuelle et méthodologique qui imprègne chacun des modèles interprétatifs défendus par les neurocriminologues. Il est en effet remarquable que ces derniers mobilisent les mêmes inférences interprétatives, alors que les études ne comportent aucune référence à la phrénologie ou à l'anthropologie criminelle italienne. Depuis le début des années 1940, l'interprétation des données recueillies chez les criminels réinvestit sous différentes formes les thèses principales de Gall et Lombroso – en l'occurrence, l'hypothèse du déséquilibre cérébral, mais aussi la théorie de l'arrêt infantile et la théorie de l'épilepsie – alors qu'elles reposent sur des données et des méthodes d'investigation essentiellement différentes (du moins incomparables).

L'hypothèse de l' « immaturité corticale » réactualise le présupposé lombrosien selon lequel les criminels sont bloqués dans un « état d'enfance prolongée ». Elle renouvelle le postulat fondateur de la théorie du criminel-né, à savoir que le développement incomplet des « organes cérébraux » du criminel constitue le stigmate de son « infériorité biologique et morale » (sa paresse, son oisiveté, sa cruauté...). Chez Lombroso comme dans la neurocriminologie basée sur l'EEG, l'interprétation des données s'articule autour de l'idée que « l'immaturité » des criminels s'exprime aussi bien sur le plan cérébral que comportemental ou moral. À partir du constat selon lequel les anomalies recensées chez les criminels se rencontrent habituellement chez les enfants, les neurocriminologues ont cherché à interpréter le défaut de maturation du cortex cérébral comme un indicateur prédictif d'un comportement potentiellement dangereux.

L'hypothèse d'une relation entre l'épilepsie et le comportement criminel est fondée sur le même raisonnement que l'hypothèse de l'« épilepsie larvée », concept forgé par Benedict Augustin Morel en 1860, et réhabilité par Lombroso dans sa théorie du criminel-né. Le fondement empirique de l'hypothèse de l'épilepsie – rappelons-le, l'association établie entre la manifestation d'un « comportement paroxystique » et l'observation d'un « EEG paroxystique » – marque d'ailleurs un retour à la dynamique de la « force irrésistible », du « penchant naturel » ou de l' « instinct criminel », notions qui ont traversé toute la phrénologie et la criminologie du XIX^{e} siècle. Nous montrerons que cette dynamique de l'instinct est absolument centrale dans le programme explicatif de la neurocriminologie, car elle permet de réaffirmer la dangerosité du criminel présentant d'importantes anomalies cérébrales. L'excès d'anomalies cérébrales ferait de lui un individu *naturellement* dangereux, car doté d'un « instinct criminel » maintenu dans un état latent, et qui n'attend qu'à s'exprimer.

Précisons également que la neurocriminologie contemporaine n'échappe pas à cet héritage historique, puisqu'elle partage avec la phrénologie un socle méthodologique et conceptuel qui découle directement de sa stratégie localisatrice. Cette parenté s'exprime notamment au travers de l'hypothèse du « déséquilibre cérébral », notion héritée de l'organologie de Franz-Joseph Gall et de la typologie criminelle établie par Alexandre Lacassagne (1843-1924).[173] Telle qu'elle est mobilisée dans la littérature en neurocriminologie, cette notion favorise la représentation d'un cerveau plus ou moins vulnérable aux facteurs socio-environnementaux. Elle renouvelle l'idée proprement phrénologique selon laquelle les circonstances défavorables peuvent être contrebalancées par l'activité de certains « organes cérébraux ». Par ailleurs, nous montrerons que cette notion de « déséquilibre », couplée à la démarche typologique et localisatrice de la neurocriminologie, réactualise le présupposé selon lequel différents crimes répondent à différentes topographies cérébrales : les techniques de neuroimagerie sont mises au service d'une typologie classificatoire consistant à caractériser et classer les criminels

[173] Rappelons que Lacassagne, avant de fonder son « école du milieu social », a été l'un des premiers défenseurs des thèses lombrosiennes en France. Il a lui-même proposé une typologie criminelle basée sur le développement relatif des différentes parties du cerveau, qui comprenait trois catégories : les criminels de pensée (les criminels « frontaux »), les criminels d'acte, impulsifs (les criminels « pariétaux »), et les criminels de sentiments ou d'instincts (les criminels « occipitaux ») (Lacassagne, préface à E. Laurent 1908, pp. I-XVIII).

selon le développement relatif de leurs structures cérébrales, c'est-à-dire en comparant le volume ou le niveau d'activité des différentes régions de leur cerveau.

Enfin, un dernier aspect de cette filiation mérite d'être évoqué ici. L'association établie entre le degré d'anomalies cérébrales et le niveau de dangerosité de l'individu opère la construction anthropologique d'un criminel d'autant plus *dangereux* que son cerveau est *monstrueux*. Elle contribue à établir une gradation des anomalies indexée sur le niveau de dangerosité des individus, et qui n'est pas sans rappeler la hiérarchisation des « types criminels » de l'école anthropologique de Lombroso. Plus encore, elle réintroduit la figure du « monstre physique et moral » incarnée par le criminel-né, criminel contre-nature, réputé incorrigible et donc à exclure de la société. Une différence notable doit être néanmoins mentionnée ici, au regard de ce que nous avons montré dans le chapitre précédent : les finalités recherchées par la neurocriminologie marquent une rupture avec l'ambition de la criminologie italienne du XIX^{e} siècle, puisque la stratégie de pathologisation du crime n'appelle pas une politique d'exclusion, mais une prise en charge directe, neurobiologique et spécifique.

Ainsi s'esquissent les contours d'une filiation dont l'étendue nous engage à procéder à un examen attentif des présupposés sur lesquels elle repose. Comme nous l'avons mentionné plus haut, peu d'études ont procédé à une analyse historique du champ de la neurocriminologie. Bien que ne relevant pas d'une démarche proprement historique, la plus instructive est certainement celle publiée par la juriste américaine Amanda C. Pustilnik (2009). Elle propose d'évaluer la légitimité scientifique de la neurocriminologie contemporaine en comparant sa démarche avec trois moments dans l'histoire de la biologisation du phénomène criminel : les études phrénologiques effectuées sur les criminels au début du XIX^{e} siècle, les thèses issues de la criminologie positiviste de Cesare Lombroso, et les interventions neurochirurgicales reposant sur la lobotomie préfrontale, au milieu du XX^{e} siècle.

Cette analyse mérite d'être brièvement mentionnée ici car elle met en évidence les principes théoriques et méthodologiques qui fragilisent l'ambition d'utiliser la neuroimagerie comme un instrument de connaissance et de prédiction du comportement criminel. D'après Amanda C. Pustilnik, les recherches publiées dans le domaine de la neurocriminologie reposent essentiellement sur trois présupposés : 1/ l'idée que le cerveau constitue le niveau de description le plus pertinent pour comprendre ou expliquer le comportement criminel ; 2/ l'idée que ce dernier résulte de dysfonctionnements

parfaitement localisés dans le cerveau ; et 3/ que les criminels sont « biologiquement différents » (voire qu'ils relèvent de la pathologie) en raison des anomalies cérébrales qui les caractérisent.

Amanda C. Pustilnik propose de mettre en perspective ces présupposés avec trois problèmes épistémologiques inhérents aux études phrénologiques et criminologiques du XIX^{e} siècle : le problème des niveaux d'explication, le problème de la localisation, et le problème de l'altérité. Selon elle, ces obstacles épistémologiques sont essentiellement liés à une approche neuro-réductionniste du comportement criminel : le niveau de description neurobiologique est jugé le plus pertinent pour expliquer la manifestation d'un comportement violent, impulsif ou agressif. Il en résulte un double présupposé, favorisé par le constat selon lequel les criminels présentent un grand nombre d'anomalies cérébrales : l'idée que la criminalité relève du domaine de la pathologie, et qu'elle réside dans des régions parfaitement localisées et individualisées du cerveau.

Dans la continuité de cette analyse, nous proposons de circonscrire la question de la filiation à deux grands principes méthodologiques : la démarche typologique et classificatoire qui préside à l'étude du cerveau des criminels depuis les premières recherches utilisant l'EEG, et l'approche localisatrice permise par l'usage de la neuroimagerie anatomique et fonctionnelle depuis le début des années 1990. À la lumière de ce que nous avons décrit plus haut concernant la conception préformationniste du comportement criminel, nous montrerons que cet héritage ne réside pas tant dans une conception neuro-réductionniste du comportement criminel que dans l'adoption d'un postulat innéiste. Bien que l'innéisme criminel s'exprime dans des contextes socio-historiques spécifiques, elle est réactualisée à la faveur de certains principes méthodologiques très problématiques sur le plan épistémologique, et qui sont pour la plupart hérités des thèses de Cesare Lombroso. Ici encore, la spécificité de notre démarche est d'expliciter les présupposés, les concepts et les méthodes qui gouvernent l'étude neurobiologique des criminels. L'objectif est donc de mener une réflexion critique sur le double discours de la neurocriminologie, laquelle se défend de négliger l'influence des facteurs socio-environnementaux dans le développement de la criminalité, tout en cherchant à confirmer un modèle innéiste admis au préalable.

5.2. L'adoption d'un postulat innéiste

5.2.1. Trois conceptions de l'innéisme criminel

Dans la phrénologie et la criminologie italienne du XIX[e] siècle, l'innéisme criminel est un postulat qui guide la recherche des causes biologiques du crime. Chez Gall, elle prend la forme d'une *disposition intrinsèque* pouvant être modulée sous l'influence des circonstances socio-environnementales. L' « organe du meurtre », au même titre que l'ensemble des organes cérébraux, préexiste avant l'exercice de toute faculté. Il désigne la condition matérielle nécessaire à l'exercice de l'instinct criminel, lequel dépend du maintien de l'équilibre entre le degré d'activité des autres organes cérébraux d'une part, et l'influence des circonstances extérieures d'autre part : l' « organe du meurtre » ne prédispose pas nécessairement à la manifestation du penchant criminel, car son exercice peut être réprimé par des circonstances favorables comme l'éducation ou la morale.

Rappelons à ce titre que la prise en compte des facteurs socio-environnementaux n'est pas incompatible avec le postulat innéiste. C'est précisément parce que l'instinct criminel est inscrit dans les dispositions organiques des individus que l'influence environnementale peut s'exercer. Le cerveau joue bien un rôle premier dans la manifestation du penchant criminel car il possède toutes les conditions matérielles nécessaires à l'exercice des facultés. Autrement dit, le cérébro-centrisme et l'internalisme constituent le fondement théorique de la localisation cérébrale du « penchant au meurtre » : l'intérêt exclusif porté au cerveau dans l'étude du phénomène criminel s'accompagne d'une thèse internaliste, qui définit l'idée que les conditions d'exercice des facultés sont internalisées dans le cerveau.

Nous avons également vu que l'organologie de Gall était gouvernée par l'adoption d'un principe d'identité psychophysiologique. Ce principe a une valeur heuristique puisqu'il guide la stratégie de décomposition et de localisation des facultés (Bechtel and Richardson 1993). Il permet de faire le lien entre la pluralité des facultés (la décomposition de l'esprit) et la pluralité cérébrale (la division du cerveau), en associant chacune des facultés isolées à autant d'organes cérébraux qui leur sont spécifiques. La localisation du « penchant au meurtre » découle donc directement de ce principe : c'est

en posant un système de correspondance entre la pluralité des facultés et la pluralité cérébrale que Gall va chercher à ancrer les facultés dans les dispositions organiques de l'individu, et ainsi localiser l'organe cérébral responsable de l'altération du sens moral.

Chez Lombroso, l'innéisme criminel est conçu sur le modèle du *préformationnisme* : la criminalité résulte d'une « auto-construction » endogène. Tout au long de son anthropologie criminelle, Lombroso s'efforce de montrer que la criminalité est inhérente au corps du criminel, et qu'il est possible d'apprécier sa dangerosité en mesurant le degré de ses anomalies anatomiques ou morphologiques. Nous avons en effet montré que la recherche des causes biologiques du crime était entièrement dictée par le présupposé selon lequel le criminel ressemble à son crime : ses caractéristiques psychologiques ou comportementales (sa grande instabilité, ses traits infantiles, sa dégénérescence morale, etc.) doivent refléter la nature de ses anomalies anatomiques ou morphologiques (le déséquilibre des fonctions cérébrales, le développement incomplet de ses organes cérébraux, sa dégénérescence biologique, etc.). Autrement dit, Lombroso oriente sa recherche des causes du crime vers une confirmation empirique de sa thèse innéiste : il s'agit de faire apparaître les irrégularités anatomiques ou morphologiques qui relient le crime à son auteur, et qui attestent d'une pulsion intrinsèque et irrésistible.

Ce présupposé introduit une différence fondamentale par rapport à l'organologie de Gall et à sa localisation cérébrale du « penchant au meurtre ». L'origine du crime ne réside pas dans une région individualisée du cerveau ; elle est imputable à la monstruosité du criminel, à ses asymétries, ses disproportions et ses arrêts du développement. Le principe selon lequel le criminel ressemble à son crime a ainsi permis à Lombroso d'établir un lien de proportionnalité entre le degré des anomalies observées et le niveau de dangerosité de l'individu : si le crime est monstrueux (moralement), alors le criminel doit l'être tout autant (physiquement). La notion d'instinct est ici centrale, car elle permet non seulement d'ancrer le crime dans le corps du criminel, mais aussi de faire de la criminalité un état permanent et incurable : c'est autour de cette notion que s'articule le couplage entre la monstruosité physique et morale des criminels. Le crime est conçu comme un acte instinctif, c'est-à-dire comme l'expression d'une « force irrésistible », dont la cause biologique réside essentiellement dans le déséquilibre des fonctions cérébrales.

On retrouve ici une logique argumentative comparable à celle qui préside à l'étude neurobiologique du criminel depuis le début des années 1940. Pourtant, contrairement à

la phrénologie et à la criminologie du XIX^{e} siècle, la neurocriminologie ne repose pas explicitement sur un postulat innéiste. Tout au long de son développement, on voit plutôt s'opérer la mise en œuvre d'une stratégie visant à faire de l'innéisme criminel le résultat d'une démonstration empirique. Les neurocriminologues se sont dotés d'un ensemble d'outils conceptuels et méthodologiques pour justifier l'intérêt exclusif porté au cerveau dans l'étude du comportement criminel, et apporter la preuve que le criminel porte en lui les germes d'une dangerosité maintenue dans un état latent. La neurocriminologie renchérit ainsi sur le postulat central de la théorie du criminel-né – d'une relation entre le criminel et son crime – en cherchant à montrer que le criminel ressemble à la fois à son cerveau et à son crime. Elle poursuit l'entreprise d'essentialisation du crime amorcée par Lombroso, en défendant l'idée que les conditions d'émergence de la criminalité sont *internalisées* dans des états cérébraux spécifiques (l'immaturité corticale, l'instabilité de l'activité cérébrale, le déséquilibre fronto-limbique ou inter-hémisphérique, etc.).

Il est important d'insister sur le rôle joué par les résultats neuroscientifiques dans cette entreprise d'essentialisation du crime. Alors que la criminologie de Lombroso s'attache à révéler la « monstruosité biologique » du criminel « moralement monstrueux », la neurocriminologie s'inscrit dans une stratégie corolaire qui consiste à confondre le crime *et le cerveau* du criminel. Il ne s'agit pas seulement de rechercher les preuves d'une criminalité inhérente au criminel, mais de montrer que le criminel est « d'essence cérébrale ». Autrement dit, le modèle explicatif défendu par la neurocriminologie prend la forme d'un biais de confirmation en faveur du neuroessentialisme. Ce dernier a une valeur démonstrative, car il permet de confirmer la thèse selon laquelle le cerveau remplit à lui seul toutes les conditions d'émergence d'un comportement violent, impulsif ou agressif. En concevant les anomalies cérébrales comme autant de conditions matérielles inhérentes à l'organisation cérébrale du criminel, la neurocriminologie cherche à montrer que la connaissance du sujet criminel s'accomplit entièrement dans la connaissance de son cerveau, c'est-à-dire qu'il est possible de rendre compte de l'ensemble de ses manifestations comportementales par une partie de lui-même.

À première vue, cette thèse apparaît comme le produit d'un double héritage phrénologique et lombrosien. Elle réactualise le cérébro-centrisme et l'internalisme de Gall, en y ajoutant une dynamique auto-constructive qui conçoit le criminel comme la somme de ses déterminations internes. Le modèle explicatif de la neurocriminologie emprunte à l'organologie l'idée d'une disposition immanente et intrinsèque, et affirme

encore davantage la prééminence du cerveau dans l'explication du comportement criminel : il fait jouer au cerveau un rôle nettement plus actif que dans l'anthropologie criminelle de Lombroso, puisque le crime n'est plus seulement ancré dans le *corps* du criminel, mais dans des états cérébraux spécifiques.

Nous laissons provisoirement de côté la question de la validité épistémologique de cette thèse, car elle mérite une analyse attentive des arguments et des limites du cérébro-centrisme et de la conception internaliste des processus mentaux. Le problème n'est d'ailleurs pas tant lié au neuroessentialisme en lui-même, qu'à la tentation de voir ce dernier comme la seule grille de lecture des données empiriques : c'est cette pétition de principe que nous proposons d'analyser dans la suite de cette section, car elle réinvestit sous une nouvelle forme l'innéisme criminel telle qu'elle est défendue dans la phrénologie et la criminologie du XIX^{e} siècle.

Plusieurs questions seront discutées tout au long de cette analyse. En quoi le modèle explicatif de la neurocriminologie marque-t-il un retour à la dynamique de la « force irrésistible », du « penchant naturel » ou de l' « instinct criminel » ? Dans quelle mesure l'association établie entre le degré d'anomalies cérébrales et le niveau de dangerosité de l'individu réintroduit la figure lombrosienne du « monstre physique et moral » ? Comment expliquer le fait que les neurocriminologues mobilisent les mêmes inférences interprétatives que Gall ou Lombroso – notamment l'hypothèse de l'épilepsie, de « l'immaturité corticale » ou du « déséquilibre cérébral » − alors que les études ne comportent aucune référence à la phrénologie ou à l'anthropologie criminelle du XIX^{e} siècle ?

La première question que nous allons chercher à résoudre concerne le projet localisateur de la neurocriminologie contemporaine et l'accusation de « nouvelle phrénologie » dont elle fait l'objet (Markowitsch 2008 ; Dartigues 2017 ; Larregue 2019). Ces analyses s'inspirent largement de la critique formulée par William Uttal (2001), qui prend pour cible l'usage des techniques de neuroimagerie dans le champ des neurosciences cognitives. Précisons d'ores et déjà que notre approche ne s'inscrit pas dans la continuité de la thèse de William Uttal, dont la radicalité pourrait introduire une méprise sur les principes régissant le projet neurocriminologique. Sans contester le bien-fondé de l'expression de « nouvelle phrénologie » − du moins en ce qui concerne le projet neurocriminologique −, nous allons ici chercher à dépasser l'aspect strictement méthodologique de la filiation, pour questionner les présupposés qui sous-tendent le

projet localisateur de la neurocriminologie. Quelle différence y a-t-il entre d'une part, le principe d'identité psychoneurale qui structure l'interprétation des données recueillies depuis les années 1940, et d'autre part, le principe d'identité psychophysiologique énoncé par Gall pour faire le lien entre sa pluralité des facultés et sa pluralité cérébrale ? Dans les deux cas, ce principe est pensé comme un modèle heuristique pour ancrer le crime dans les dispositions organiques de l'individu : chez Gall, comme dans la neurocriminologie, il s'agit de décomposer un système complexe en un ensemble de catégories psychologiques, afin de les localiser (directement ou indirectement) dans des régions individualisées du cerveau. Nous montrerons que cette double stratégie n'a pas tout à fait le même statut épistémologique chez Gall et dans la neurocriminologie, mais que les modèles qui en résultent sont tout aussi limités sur le plan explicatif.

5.2.2. Le projet localisateur de la neurocriminologie contemporaine

Comme nous l'avons mentionné précédemment lorsque nous abordions la question du pouvoir explicatif de la stratégie neuro-interventionniste (section 4.4), le projet localisateur est un projet d'assignation de responsabilité causale (Bechtel et Richardson 1993). Il s'applique à attribuer des rôles fonctionnels à chacun des composants du système considéré dans l'objectif de fournir une explication mécaniste de ce dernier. Tel qu'il est mobilisé dans la neurocriminologie contemporaine (à partir des années 1990), le projet localisateur vise à mettre en relation 1/ les anomalies cérébrales observées chez les criminels, 2/ l'altération des capacités fonctionnelles des individus, et 3/ la manifestation du comportement criminel. Nous ne reviendrons pas sur la faiblesse épistémique de cette relation car notre objectif est ici de comprendre ce qui différencie cette démarche de la localisation phrénologique du « penchant au meurtre ». En outre, si l'analyse qui suit porte sur les études récentes utilisant les techniques de neuroimagerie, elle s'adresse plus globalement à la stratégie déployée depuis les années 1940 pour fonder des catégories psychologiques sur des catégories neurophysiologiques, et montrer que le criminel ressemble à son cerveau.

Si l'on s'en tient aux études publiées depuis les années 1990, on peut dire que la localisation des anomalies cérébrales des criminels suppose une décomposition fonctionnelle qui s'opère à la fois sur le plan cérébral et comportemental. Le cerveau du

criminel est divisé en un certain nombre d'*états* associés à une anomalie particulière (un déficit d'activité, une irrégularité de structure, une asymétrie fonctionnelle, etc.) et localisée dans une région individualisée (cortex préfrontal, amygdale, etc.). Le comportement criminel est lui-même conçu comme un système décomposable, dans la mesure où il est défini par l'association de plusieurs propriétés psychologiques (par exemple : l'impulsivité, l'absence de sens moral, etc.), qui traduisent autant d'incapacités opérationnalisables sur le plan expérimental (l'incapacité à inhiber une action motrice, l'incapacité à résoudre un dilemme moral, etc.). En définitive, le modèle explicatif de la neurocriminologie vise à faire correspondre les processus de haut niveau (les propriétés psychologiques ou comportementales) avec les processus de bas niveau (les propriétés cérébrales), de telle sorte à fournir une explication neurobiologique du comportement criminel. Il s'agit de rendre compte de l'ensemble des manifestations comportementales du criminel en les subdivisant en différents composants identifiables dans le cerveau : un état cérébral pour chaque incapacité fonctionnelle ; une incapacité fonctionnelle pour chaque propriété psychologique ou comportementale.

Contrairement au modèle de Gall qui repose sur une instanciation directe des facultés dans des régions individualisées du cerveau, cette stratégie correspond à la mise en œuvre d'une localisation complexe.[174] Selon la définition proposée par Bechtel et Richardson (1993), la localisation complexe ou indirecte (*complex localization* ou *localization multiply constrained*) nécessite de décomposer la capacité que l'on cherche à expliquer, c'est-à-dire d'isoler les sous-capacités qui contribuent à l'exercice de la capacité de rang supérieur. Cette stratégie ne repose pas sur une simple corrélation entre une structure et une fonction, car elle implique une détermination multiple de l'hypothèse localisatrice : le modèle doit satisfaire les multiples contraintes imposées sur les différents niveaux d'organisation, afin de fournir une explication mécaniste de la capacité générale (Forest 2008).

Précisons à cet égard que le projet explicatif de le neurocriminologie repose sur la localisation des processus qui affectent l'exercice des capacités psychologiques, et non pas sur la localisation des capacités elles-mêmes. La stratégie adoptée par la

[174] Comme le soulignent Bechtel et Richardson, le modèle de Gall correspond à une localisation simple car il repose sur une corrélation directe entre structure et fonction. À l'inverse, la localisation complexe renvoie à la stratégie adoptée par Carl Wernicke (1848-1905) : c'est en décomposant la capacité linguistique en un petit ensemble de tâches de rang inférieur qu'il parvient à localiser les centres du langage.

neurocriminologie ne relève pas d'une simple localisation cérébrale d'un « penchant au meurtre », puisqu'elle vise à isoler les états cérébraux impliqués dans l'altération des fonctions volitionnelles, cognitives ou morales. La différence entre ces deux modèles renvoie à la distinction proposée par Bechtel (2002) entre deux types de décomposition : la décomposition phénoménale et la décomposition mécaniste. La première correspond à la stratégie adoptée par Gall dans sa division de l'esprit en facultés distinctes et indépendantes : elle marque la première étape de son organologie (la psychologie phrénologique) qui lui permettra de construire sa psychophysiologie des facultés et d'isoler l'organe cérébral associé au penchant au meurtre.

Comme le précise Bechtel (2002, p. 231), « [u]ne décomposition phénoménale différencie les divers phénomènes produits par un système, et qui requièrent une explication ». Elle ne présente elle-même aucune valeur explicative dans la mesure où elle n'apporte pas d'information concernant la manière dont les facultés sont produites. Cette explication ne peut être apportée que par le biais d'une décomposition mécaniste, c'est-à-dire par l'identification des processus contribuant à l'exercice des capacités isolées initialement. La décomposition phénoménale peut donc être utile à la construction du modèle mécaniste, mais elle ne garantit pas elle-même une explication du phénomène considéré.

En gardant à l'esprit ces distinctions, résumons brièvement ce qui différencie le modèle de Gall de celui de la neurocriminologie. Le premier correspond à une localisation simple car il repose sur une instanciation directe des facultés. Rappelons que, selon Gall, la recherche des conditions matérielles des facultés revient à situer les facultés elles-mêmes. Il en est de même pour le penchant au meurtre, qui est directement localisé dans une région individualisée du cerveau, « au-dessus des oreilles, dans la ligne perpendiculaire de la colonne vertébrale. » (Gall, 1825, t.5, p. 454). La localisation du penchant au meurtre est elle-même issue d'une décomposition phénoménale, puisqu'elle est rendue possible par la division de l'esprit en facultés distinctes et indépendantes. Cette décomposition n'a aucune valeur explicative dans la mesure où elle repose sur une simple corrélation entre la structure (l'organe cérébral) et la fonction (la faculté correspondante). Selon Bechtel et Richardson (1993), le pouvoir explicatif de ce modèle est donc limité par le fait qu'il repose sur une localisation directe et « insuffisamment contrainte ». Pour autant, il est important de rappeler que l'exercice du penchant au meurtre dépend du maintien de l'équilibre entre le degré d'activité des autres organes cérébraux et l'influence des circonstances extérieures. Autrement dit, si le penchant au meurtre est associé à un

seul organe cérébral, son exercice implique une détermination multiple de la stratégie localisatrice. Comme le précise Denis Forest (2008, p. 32), « la localisation fonctionnelle proposée par Gall est sans doute insuffisamment contrainte (d'où son échec), mais elle ne l'est pas unilatéralement » : elle n'implique pas une décomposition mécaniste des facultés, mais elle est contrainte de manière multiple par l'organisation cérébrale de l'individu.

Le modèle explicatif de la neurocriminologie contemporaine repose sur une décomposition mécaniste et sur une localisation indirecte des processus psychologiques. Il vise à attribuer à chacune des incapacités fonctionnelles isolées lors de l'étape de décomposition un état cérébral spécifique et parfaitement individualisé dans le cerveau (par exemple, un déficit d'activité au niveau du cortex préfrontal). Contrairement à l'organologie de Gall, ce modèle localise plus d'une région responsable de l'altération des capacités : il suppose une localisation multiple des incapacités fonctionnelles, lesquelles sont associées à différentes propriétés psychologiques. La responsabilité causale n'est d'ailleurs pas tant imputée à une *région* individualisée du cerveau (comme chez Gall), qu'à un *état* cérébral particulier : un déficit d'activité, une irrégularité ou une asymétrie, etc. En clair, l'objectif de la neurocriminologie est de construire une cartographie des états cérébraux causalement responsables des incapacités fonctionnelles des individus, afin de montrer que le cerveau *seul* fournit une explication complète du comportement criminel.

5.2.3. Une « nouvelle phrénologie » ?

La question se pose de savoir si ces deux heuristiques répondent véritablement à l'ambition explicative de la neurocriminologie. La complémentarité des stratégies de décomposition et de localisation devrait logiquement répondre à la question de savoir *comment* les propriétés psychologiques des criminels sont physiquement réalisées dans leur cerveau, puisque ces deux stratégies imposent un certain nombre de contraintes sur la construction du modèle explicatif. Contrairement à Gall qui n'adopte pas le modèle mécaniste de l'explication, la neurocriminologie cherche à identifier les anomalies causalement responsables des incapacités fonctionnelles des individus. Pourtant,

plusieurs remarques nous amènent à tempérer le pouvoir explicatif des stratégies de recherche qui sous-tendent cette ambition.

En premier lieu, il est permis de douter de l'efficacité de la stratégie localisatrice, compte tenu du fait qu'elle porte sur des fonctions cognitives complexes et fortement intégrées. Le problème de la non-linéarité du système nerveux – c'est-à-dire le caractère dynamique des interactions entre les réseaux de neurones − constitue une source majeure des difficultés posées par les stratégies de décomposition et de localisation. William Bechtel (2002) a par exemple montré que ces stratégies pouvaient être difficilement appliquées aux systèmes dont l'activité émerge uniquement de l'interaction entre leurs divers composants. Elles sont globalement inefficaces si les composants du système entretiennent une relation de dépendance mutuelle et n'exercent eux-mêmes aucun rôle causal lorsqu'ils sont considérés isolément.[175]

Sans entrer dans le débat concernant la possibilité d'une explication mécaniste des capacités cognitives de haut niveau, on peut émettre des réserves quant à la scientificité d'une démarche consistant à *identifier* des propriétés psychologiques à des propriétés cérébrales spécifiques, sans même chercher à isoler les mécanismes impliqués. Rappelons que la localisation cérébrale des catégories psychologiques des criminels ne nous dit rien sur la responsabilité causale des états cérébraux isolés, car elle ne repose que sur des corrélations : le statut de la relation entre les états cérébraux et les états psychologiques interdit toute inférence causale, ce qui limite la possibilité d'isoler le rôle fonctionnel spécifique associé à chacun des états cérébraux. Il n'est d'ailleurs pas exclu que ces derniers soient le produit de l'interaction entre une multitude de composants, sans compter que les incapacités fonctionnelles qu'ils sous-tendent sont elles-mêmes loin de remplir toutes les conditions de manifestation d'un comportement criminel. En somme, le problème que nous soulevons ici concerne d'une part, la confiance épistémique que l'on peut accorder aux données issues de la neuroimagerie, et d'autre part, les conditions d'opérationnalisation des processus psychologiques. On ne saurait trop insister sur le fait que l'identification d'une catégorie psychologique donnée à un état cérébral spécifique échoue à expliquer la dynamique des interactions qui mènent à la réalisation d'un comportement aussi complexe que le comportement criminel : la description et la

[175] Comme le précise Bechtel (2002, p. 229) : « *[i]n normally operating systems the identity of the components is concealed by their smooth coordination in performing the system's overall activities. Often this smooth coordination involves non-linear interactions of large numbers of components.* »

localisation des structures physiques régulièrement associées aux phénomènes psychologiques et comportementaux ne suffisent pas à en fournir une explication mécaniste.

Par ailleurs, la conceptualisation des modèles expérimentaux introduit un problème majeur en ce qui concerne le pouvoir explicatif de la stratégie localisatrice. La principale difficulté épistémologique tient au fait que la recherche des corrélats neurobiologiques du comportement criminel est intrinsèquement dépendante de la formulation préalable d'une typologie psychologique : il existe de bonnes raisons de penser que la localisation des anomalies cérébrales des criminels n'est pas justifiée empiriquement dans la mesure où elle reste tributaire d'une décomposition phénoménale préalable. Le point de départ de ces recherches consiste en effet à isoler des catégories psychologiques contribuant au développement d'un comportement criminel (agressivité, impulsivité, instabilité, etc.), et qui requièrent une explication causale. On peut donc craindre que la localisation des anomalies cérébrales des criminels soit entièrement déterminée par les catégories psychologiques initialement isolées lors de l'étape de décomposition. Comme le précise Denis Forest (2008), la décomposition phénoménale d'un système invite à rechercher dans le cerveau le pendant anatomique de la capacité que l'on cherche à expliquer, sans pour autant justifier la pertinence causale de cette localisation. Cette stratégie ne fournit pas une base théorique satisfaisante pour la recherche d'un corrélat neurobiologique, car elle oblige « à concevoir l'architecture fonctionnelle du cerveau comme le reflet de la décomposition phénoménale telle qu'on l'a initialement effectuée. » (Forest 2008, p. 41). Mais surtout, elle restreint le champ des investigations empiriques dans la mesure où elle invite « à distribuer dans l'espace cérébral des rôles conçus de manière immuable » (*ibid*).

Le problème posé par cette démarche pourrait être partiellement résolu si elle s'inscrivait dans un cadre méthodologique permettant aux stratégies de décomposition et de localisation de s'informer mutuellement (voir à ce sujet Mundale and Bechtel 1996 ; Bechtel and Mundale 1999 ; Zawidski and Bechtel 2004). Effectivement, la valeur épistémique du modèle peut être sensiblement renforcée si ces deux stratégies maintiennent une relation d'interdépendance permettant de confirmer, réfuter ou réviser la caractérisation initiale des catégories psychologiques (Forest 2008). Ces stratégies deviennent alors très restrictives du point de vue de la construction du modèle explicatif, car leur dépendance réciproque introduit une contrainte supplémentaire qui permet de renforcer la robustesse du modèle : les catégories psychologiques imposent une contrainte

top-down sur la recherche des mécanismes neurobiologiques sous-jacents, ces derniers apportant en retour une contrainte *bottom-up* sur la caractérisation des catégories psychologiques. La décomposition fonctionnelle devient scientifiquement informée, dans le sens où elle peut faire l'objet d'une révision en fonction des données disponibles.[176]

Or, le projet explicatif de la neurocriminologie adopte visiblement une approche *top-down* qui maintient une stricte indépendance entre les stratégies de décomposition et de localisation : la décomposition phénoménale précède la localisation fonctionnelle et impose une adéquation inflexible entre les catégories psychologiques et neurobiologiques. De ce point de vue, on comprend mieux le rôle heuristique du postulat selon lequel le criminel ressemble à son cerveau. Il apparaît comme un modèle théorique solide pour guider et valider la stratégie *top-down* : l' « immaturité corticale » et « l'instabilité de l'activité cérébrale » sont recherchées en tant que marqueurs neurophysiologiques de « l'immaturité comportementale » et de « l'instabilité psychologique ». Dans la mesure où la contrainte *top-down* ne permet pas de réviser la caractérisation (pourtant arbitraire) des catégories psychologiques, elle oriente la recherche des processus neurobiologiques vers une confirmation empirique du principe d'identité psychoneurale : à l'indétermination des catégories psychologiques correspond une exigence de compatibilité (normative) avec les données disponibles, qui garantit un modèle exclusif de toute autre explication possible.

La faiblesse épistémique de cette stratégie tient donc au fait qu'elle contraint la construction du modèle explicatif selon la caractérisation initiale des catégories psychologiques, et introduit de fait un biais dans l'interprétation des données empiriques. La décomposition phénoménale « structure » le phénomène à expliquer dans le sens où elle limite de manière significative les possibilités d'investigation empirique. Dès lors, quelle différence existe-t-il entre le principe organologique de Gall et le modèle de la neurocriminologie ? Rappelons que Gall défend l'antériorité de sa psychologie des facultés sur son analyse anatomique : « [l]a connaissance des fonctions a toujours précédé celle des parties » (Gall, t.2, p. 21). Autrement dit, la décomposition précède la localisation selon une stratégie *top-down* qui contraint la division anatomique du cerveau

[176] D'après Bechtel (2002), les données issues de la neuroimagerie peuvent par exemple imposer une révision de la décomposition initiale, dans le cas où l'identification d'un état cérébral n'est pas compatible avec son rôle fonctionnel imputé au départ. En clair, si les données montrent qu'une région est (contrairement au modèle initial) inactive pendant la réalisation d'une tâche cognitive, elles peuvent conduire à une redéfinition des catégories psychologiques.

selon la caractérisation initiale des facultés. Dans la mesure où la division du cerveau dérive d'une psychologie des facultés déjà constituée, cette stratégie a des conséquences sur le statut épistémologique des arguments anatomiques. Comme le souligne Laurent Clauzade (2008a), elle introduit un soupçon d'argument *ad hoc*, car les arguments anatomiques ont le statut de confirmation de la psychologie des facultés : ils ne font que confirmer la décomposition fonctionnelle établie initialement.

Il semble que la démarche adoptée par les neurocriminologues soit tout aussi circulaire. Bien que leur modèle repose sur une localisation indirecte – puisqu'il ne s'agit pas de localiser le comportement criminel lui-même, mais les incapacités ou les déficits fonctionnels qui lui sont associées – les neurocriminologues réactualisent à leur insu la thèse de la pluralité des facultés et des organes. La localisation des processus psychologiques reste tributaire d'une décomposition phénoménale préalable, ce qui introduit un biais dans la collecte et l'interprétation des données. Chez Gall comme dans la neurocriminologie, il s'agit d'identifier une typologie psychologique à un complexe d'états cérébraux spécifiques, de sorte à garantir la validité du principe d'identité psychophysiologique (ou psychoneurale) : cette stratégie révèle toute la circularité d'une démarche condamnée à confirmer un modèle théorique admis au préalable, sans fournir aucune explication des phénomènes psychologiques étudiés.

Une dernière précision s'impose en ce qui concerne le statut épistémologique des catégories psychologiques isolées lors de l'étape de décomposition. Nous pourrions défendre l'idée que ces catégories ne sont rien d'autre que des « construits hypothétiques » dont la réalité empirique ne peut être démontrée (Uttal 2001). Pour autant, nous rejetons la thèse de l'inaccessibilité empirique des phénomènes mentaux, ou du rejet pur et simple du projet localisateur. Il existe de bonnes raisons de voir la neurocriminologie contemporaine comme une simple réactualisation du principe organologique de Gall, mais cela ne nous incline pas à renoncer à l'apport potentiel des techniques de neuroimagerie pour identifier les corrélats neurobiologiques des phénomènes mentaux. Le problème que nous soulevons est celui de la confiance épistémique que l'on peut accorder aux données issues de la neuroimagerie. Nous ne condamnons pas le projet localisateur dans son principe, mais seulement l'inférence interprétative qui consiste à faire du neuroessentialisme la seule grille de lecture des données empiriques. Comme nous le montrerons dans la dernière section de ce chapitre, le problème provient d'une erreur méréologique inhérente à la conception internaliste des processus psychologiques : si les anomalies cérébrales sont conçues comme autant de

conditions matérielles intrinsèques à l'organisation cérébrale des criminels, c'est parce que les neurocriminologues entretiennent une confusion entre les conditions d'exercice des capacités psychologiques et les capacités psychologiques elles-mêmes.

5.2.4. Un retour à la dynamique de l' « instinct criminel »

La filiation entre la neurocriminologie et l'anthropologie criminelle de Lombroso recouvre plusieurs aspects, que nous avons esquissés tout au long de notre analyse épistémologique : l'adoption d'une démarche typologique et classificatoire, la construction anthropologique d'un « cerveau criminel », la conception préformationniste et psychopathologique du comportement criminel, le postulat d'une « infériorité biologique », ou encore l'ambition d'apprécier la dangerosité d'un individu selon le degré de ses anomalies cérébrales. La filiation conceptuelle et méthodologique se révèle surtout dans l'effort déployé par les neurocriminologues pour démontrer un lien entre le criminel et son crime : cette démonstration est circulaire car les données sont interprétées selon une conception préformationniste du comportement criminel.

Nous n'insisterons pas sur la démarche typologique et classificatoire qui préside à l'étude du cerveau des criminels depuis le début des années 1940. De toute évidence, l'héritage lombrosien est avant tout dominé par l'ambition de fonder des catégories criminelles sur des catégories naturelles (neurobiologiques) spécifiques, dans l'objectif d'indexer le niveau de dangerosité des individus sur le degré de leurs anomalies cérébrales. En gardant à l'esprit les fondements théoriques et empiriques de cette ambition, nous allons ici chercher à comprendre pourquoi les neurocriminologues mobilisent les mêmes inférences interprétatives que Lombroso, alors qu'ils ne s'inscrivent pas explicitement dans la continuité de ses thèses.

Le cadre méthodologique qui préside à l'étude neurobiologique du criminel rappelle incontestablement la caractérisation anatomique des « types criminels » qui a pris forme au XIX^{e} siècle avec l'école criminelle italienne. Il repose sur l'idée qu'il existe un « cerveau type » correspondant à chaque type de criminalité, et qu'il est par conséquent possible de reconnaître la dangerosité d'un individu à partir de ses caractéristiques neurobiologiques. Plus encore, cette démarche typologique est mise au

service d'un modèle hiérarchique de la criminalité qui est entièrement imprégné par le paradigme de l'anthropologie criminelle italienne. L'héritage de Lombroso transparaît dans l'un des résultats les plus significatifs dans le domaine de la neurocriminologie : l'existence d'un lien de proportionnalité entre le degré d'anomalies cérébrales et le niveau de dangerosité de l'individu. L'observation d'une forte proportion d'anomalies chez les criminels les plus dangereux marque un retour à une monstruosité de nature, inscrite dans les prédispositions organiques de l'individu. On peut en effet voir ce résultat comme une réactualisation du constat dressé par Lombroso, qui confère au criminel-né « un maximum de monstruosité ». La figure du « monstre physique et moral » auquel est consacrée la majeure partie de l'ouvrage de Lombroso est réintroduite au travers de l'idée selon laquelle l'excès d'anomalies cérébrales est couplé à une immoralité absolue : celle du criminel violent, récidiviste, ou qui commet des crimes sans motif apparent.

Tel qu'il a été posé au XIX^{e} siècle, le problème de la monstruosité criminelle a introduit la question de la nature innée et pathologique de la criminalité. Lombroso a cherché à résoudre ce problème par le recours à la notion d' « instinct criminel » : si le crime est monstrueux, c'est précisément parce qu'il est instinctif. Cette notion a joué un rôle décisif dans la stratégie de pathologisation du criminel car elle a contribué à définir le crime comme l'expression d'une « force irrésistible ». Elle a aussi largement participé à l'ambition d'ancrer le crime dans la psychologie et l'anatomie du criminel, puisque cette essentialisation du crime s'est structurée autour de l'idée que le criminel est porteur d'un instinct intrinsèque et irrésistible. On peut dès lors se demander si le modèle préformationniste de la neurocriminologie relève du même raisonnement. S'il est désormais admis que la neurocriminologie et l'anthropologie du XIX^{e} siècle répondent à la même ambition d'établir un lien entre le criminel et son crime, il nous reste à montrer en quoi cette ambition défend l'innéité d'un « instinct criminel ».

D'abord, il convient de rappeler que les neurocriminologues échouent à identifier les mécanismes neurobiologiques impliqués dans l'altération des fonctions volitionnelles, cognitives ou morales. En s'interdisant de préciser le statut épistémologique de la relation entre les anomalies cérébrales et le comportement criminel, ils favorisent la représentation d'un individu porteur d'une dangerosité intrinsèque, hypothèse dont le seul fondement empirique repose sur l'observation d'une forte proportion d'anomalies cérébrales. L'argument ne repose pas sur la caractérisation d'une pathologie cérébrale en tant que telle, mais seulement sur l'observation d'une multitude d'anomalies structurelles et/ou fonctionnelles (atrophies, malformations, asymétries...) : il s'est formé à partir du constat

que ces anomalies évoquent un désordre pathologique, malgré l'absence de signes cliniques associés à une maladie psychiatrique ou neurologique. Pourtant, nous avons montré que cette interprétation correspondait moins à une réalité empirique qu'à une construction normative dictée par une finalité neuro-interventionniste. Les neurocriminologues cherchent à objectiver l'existence d'un processus pathologique à l'origine du comportement criminel, alors que les données n'apportent aucune information sur la séquence causale des évènements qui relient les états cérébraux et les états comportementaux : ils laissent suggérer l'existence d'une mécanique pathologique à l'origine d'une pulsion criminelle irrépressible, sans en fournir la moindre explication causale.

Comme chez Lombroso, la stratégie de pathologisation du crime est d'ailleurs nourrie par le présupposé selon lequel l'excès d'anomalies cérébrales traduit un défaut inné ou « constitutionnel ». Cette tendance interprétative traverse toute l'histoire de la neurocriminologie, depuis les premières publications relatives aux anomalies EEG des psychopathes (e.g. Silverman 1943) jusqu'aux récentes recherches basées sur la neuroimagerie (e.g. Raine et al. 1998a). L'excès d'anomalies est attribué à une criminalité « endogène », dans le seul objectif de promouvoir une conception innéiste et psychopathologique des crimes violents. On reconnaît ici un biais interprétatif récurrent chez Lombroso, dans sa tendance à réinterpréter les données à la lumière de ses présupposés, pour assurer la confirmation réciproque de ses thèses. Les modèles préformationniste et pathologique de la criminalité s'inscrivent dans un rapport de complémentarité mutuelle, pour garantir la pertinence des finalités poursuivies par les neurocriminologues.

Plus généralement, il semble que l'ambition de relier le criminel à son crime s'articule elle-même autour de la dynamique de l'instinct ou de la « force irrésistible ». Il n'est finalement pas surprenant que les différents modèles interprétatifs défendus depuis le début des années 1940 réintroduisent l'hypothèse de l'arrêt infantile ou l'hypothèse de l'épilepsie, telles qu'elles ont été formulées par Lombroso : si la neurocriminologie et l'anthropologie criminelle de Lombroso mobilisent les mêmes inférences interprétatives, c'est parce que le principe selon lequel le criminel ressemble à son crime repose globalement sur la même structure argumentative. Chacune des anomalies observées est mise en relation avec des critères d'appréciation relatifs aux traits psychologiques des criminels, selon un raisonnement analogique qui permet de les assimiler à un processus de même nature. La faiblesse épistémique de ce raisonnement réside précisément dans le

fait qu'il contraint la construction des hypothèses explicatives au prix d'une série de pétitions de principe : l'interprétation des données recueillies chez les criminels vise à montrer que leur dangerosité s'exprime au travers d'une potentialité préexistante, d'un état cérébral porteur d'une anomalie particulière, et qui préfigure une irrégularité comportementale donnée.

Précisons que la circularité du raisonnement est nettement plus évidente chez Lombroso que dans la neurocriminologie, car ses observations sont directement dictées par ses présupposés, notamment l'infériorité biologique des criminels, le caractère irrésistible de leur instinct, ou encore la présomption d'incorrigibilité. Lombroso cherche à confirmer ces présupposés en comparant les données recueillies chez les criminels, les hommes primitifs, les enfants et les épileptiques, et en recensant l'ensemble des analogies observées sur le plan biologique et comportemental. La théorie atavique du crime, la théorie de l'arrêt infantile et la théorie de l'épilepsie sont donc fondées sur un raisonnement par analogie qui consiste à relier chacune des anomalies physiques des criminels à autant d'irrégularités psychologiques caractéristiques des primitifs, des enfants et des épileptiques : les caractères régressifs des criminels évoquent la primitivité et la bestialité de leurs comportements, leurs arrêts de développement sont l'expression de leur infantilité morale, leur violence impulsive évoque une « pulsion irrésistible » aussi brutale et imprévisible qu'une crise épileptique, etc. En résumé :

- leur « dégénérescence biologique » correspond à leur « dégénérescence morale » ;
- le « développement incomplet » de leurs organes cérébraux reflète leurs « traits [psychologiques] infantiles » ;
- l' « instabilité » ou le « déséquilibre » de leurs fonctions cérébrales constitue le marqueur biologique de leur « pulsion irrésistible » ;
- leur « monstruosité physique » constitue l'indice visible de leur « monstruosité morale ».

On retrouve ici la même logique argumentative que celle qui a guidé l'interprétation des données neurobiologiques recueillies chez les criminels : le comportement criminel résulterait d'une « auto-construction » endogène, car les caractéristiques psychologiques des individus héritent des propriétés des états cérébraux correspondants. Sans revenir sur l'ensemble des arguments mobilisés dans les deux camps, on peut s'attarder sur les fondements théoriques (et empiriques) de l'hypothèse de

l'épilepsie. Cette hypothèse mérite une attention particulière, principalement pour deux raisons. Premièrement, chez Lombroso comme dans la neurocriminologie, elle constitue l'arrière-plan théorique de la stratégie de pathologisation du crime violent : elle cristallise l'ambition de faire du crime un acte *instinctif*, résultant d'une pulsion pathologique irrésistible. Elle est d'ailleurs paradigmatique de l'ambition d'établir un lien entre le criminel et son crime, puisqu'elle repose sur un principe d'équivalence en termes de propriétés entre les anomalies observées sur le plan anatomique et psychologique. Deuxièmement, telle qu'elle a été formulée dans la neurocriminologie des années 1940, cette hypothèse réactualise la problématique emblématique de l'anthropologie criminelle du XIX^{e} siècle : celle des crimes sans raison qui ne s'accompagnent d'aucun symptôme visible de folie. Nous verrons ainsi que, dans les deux cas, l'hypothèse de l'épilepsie a contribué à faire du crime violent un problème insoluble pour le droit pénal. Elle a favorisé un renouvellement des questions relatives à la prise en charge des criminels, en inscrivant la dynamique de l'instinct sur le double registre de la pathologie et de la tératologie.

5.2.5. Une force irrésistible mais corrigible

Dans la neurocriminologie des années 1940-1980, l'hypothèse d'une relation entre l'épilepsie et le comportement criminel s'est formée autour du constat que les criminels présentent des anomalies cérébrales caractéristiques du sujet épileptique. Elle repose sur l'observation d'une activité « paroxystique », malgré l'absence de symptômes convulsifs ou d'autres signes cliniques associés à l'épilepsie. La notion d' « épilepsie latente » a alors été introduite pour rendre compte des similitudes observées entre les criminels et les épileptiques, à la fois sur le plan neurophysiologique et comportemental. À bien des égards, cette notion entretient une certaine parenté avec le concept d' « épilepsie larvée » tel qu'il a été défini par Lombroso dans sa théorie du criminel-né. L'épileptique et le criminel-né sont confondus dans une même entité pathologique, dont le seul argument empirique repose sur l'observation d'une « instabilité » ou d'un « déséquilibre » des fonctions cérébrales : dans la neurocriminologie comme chez Lombroso, les anomalies cérébrales des criminels sont attribuées à une épilepsie

maintenue dans un état latent, c'est-à-dire à une condition pathologique qui pourrait favoriser l'expression d'une « pulsion irrésistible ».

Plus encore, cette parenté conceptuelle se fait l'écho d'une remarquable affinité méthodologique car, dans les deux cas, l'hypothèse de l'épilepsie repose sur une analyse comparative entre les anomalies observées sur le plan anatomique ou neurophysiologique d'une part, et psychologique ou comportemental d'autre part. L'association établie entre la manifestation d'un « comportement paroxystique » et l'observation d'un « EEG paroxystique » procède d'un raisonnement analogique qui n'est pas sans rappeler la relation instaurée par Lombroso entre l'acte criminel et l'accès épileptique. Ce principe de correspondance réactualise la double symétrie défendue à la fin du XIX^{e} siècle, entre la « soudaineté du crime » qui évoque une « décharge épileptique » et la « férocité inutile et exagérée » des épileptiques qui rappelle celle des criminels-nés : dans la neurocriminologie comme chez Lombroso, l'hypothèse de l'épilepsie s'est formée dans l'objectif de faire du crime violent un équivalent de la crise épileptique, et de l'épilepsie une véritable « caricature du crime ». On peut d'ailleurs voir une certaine résonance entre les deux extraits suivants, issus respectivement de l'une des premières études menées dans le champ de la neurocriminologie et de *L'Homme criminel* de Lombroso :

> « Tant du point de vue clinique que du point de vue électro-encéphalographique, de nombreuses preuves suggèrent une association entre certaines formes de psychopathie et l'épilepsie. La perturbation épisodique du comportement, la dangerosité explosive du psychopathe agressif qui n'a pas conscience des conséquences de son comportement et l'amnésie présumée, ont conduit de nombreux psychiatres à parler de personnalité épileptoïde. » (Hill and Watterson, 1942, p. 60)

> « Les accès de fureur épileptique, ou, pour mieux dire, d'équivalent psychique, par leur caractère d'instantanéité soudaine, par leur férocité inutile et exagérée, forment aux yeux de l'observateur attentif, dans une période aigüe et courte, ce que l'épileptique est normalement — une caricature du crime, de même que l'obtusion sensorielle, la stupidité momentanée qui suivent et accompagnent les accès ne sont que l'excès de l'état habituel et souvent congénital de l'épileptique. » (Lombroso, 1887, p. 426)

Il est intéressant de remarquer que cette hypothèse s'est déployée autour de la même problématique des crimes monstrueux. Chez Lombroso, l'instinct criminel s'inscrit dans le domaine confus du crime-folie : il répond au problème des crimes qui échappent aussi bien à l'assignation de raison qu'à l'assignation de folie. Dans la neurocriminologie, la dynamique de l'instinct se situe dans le champ indifférencié du crime et des troubles de la personnalité. Dans les deux cas, l'hypothèse de l'épilepsie a été avancée en vue d'expliquer les pulsions criminelles qui ne peuvent être imputées à la présence d'un trouble mental : les notions d' « épilepsie larvée » et d' « épilepsie latente » sont mobilisées pour rendre compte des anomalies qui ne peuvent être attribuées ni à la folie ni à l'épilepsie proprement dite. Cette condition pathologique contribue ainsi à résoudre le problème des crimes monstrueux, car elle permet de combler l'absence de trouble mental par « l'explication neurobiologique » (le déséquilibre des fonctions cérébrales chez Lombroso, l'activité cérébrale paroxystique dans la neurocriminologie) : ces anomalies sont perçues comme la source d'une mécanique irrépressible qui contribue elle-même à expliquer le crime inexplicable. De là apparaît la représentation d'un criminel entièrement soumis à ses instincts violents, porteur d'une dégénérescence intrinsèque qui le rend définitivement dangereux pour autrui.

L'hypothèse d'une relation entre l'épilepsie et le comportement criminel a donc joué un rôle décisif dans la stratégie de pathologisation du crime, aussi bien dans la neurocriminologie que dans la théorie du criminel-né de Lombroso. Elle répond au double objectif de démontrer l'origine « constitutionnelle » des anomalies cérébrales des criminels, et de promouvoir une conception psychopathologique des comportements violents, impulsifs ou agressifs. Cependant, une différence importante doit être soulignée en ce qui concerne la réponse apportée à la question de la prise en charge de la criminalité. De l'anthropologie lombrosienne à la neurocriminologie, l'hypothèse de l'épilepsie et la dynamique de l'instinct criminel répondent à la même ambition ; pourtant, la conception innéiste et pathologique de la criminalité n'a pas tout à fait les mêmes implications sur le plan social et juridique.

Rappelons que l'anthropologie criminelle de Lombroso ne favorise pas un renouvellement des questions relatives à la responsabilité pénale, car elle répond seulement à une exigence de protection sociale. La neurocriminologie adopte une position plus ambiguë face à la question de la responsabilité des criminels présentant d'importantes anomalies cérébrales. À partir des années 1960, l'hypothèse d'une relation entre l'épilepsie et le comportement criminel s'est progressivement diffusée dans le

champ pénal, au travers de ce que nous avons décrit sous le terme d'*epilepsy defense*. Cette disposition légale exceptionnelle a contribué à mener une réflexion inédite sur la possibilité d'utiliser l'EEG pour rechercher une cause d'irresponsabilité ou d'atténuation de la responsabilité pénale.[177] Pour autant, nous avons montré que la question des implications des données EEG sur la notion de responsabilité pénale était très rarement discutée à l'intérieur du champ de la neurocriminologie, à l'exception des premières études menées dans les années 1940 (Stafford-Clark et Taylor 1949). La reconstruction historique de ce programme de recherche nous a en effet permis de constater un basculement du point de vue des finalités poursuivies par les chercheurs, dès le début des années 1950. La question de la responsabilité a été très rapidement délaissée au profit de l'appréciation de la dangerosité, et les études se sont structurées autour de la possibilité d'en proposer une définition neurobiologique. La neurocriminologie et l'anthropologie lombrosienne obéissent donc à la même logique de protection sociale : il s'agit de faire de l'excès d'anomalies un critère d'appréciation de la dangerosité criminologique, afin de garantir la possibilité de repérer l'individu représentant une grande menace pour la société (le criminel violent, récidiviste, ou qui commet des crimes sans motif apparent).

La réponse apportée à la question de la prise en charge de l'individu dangereux n'a pourtant pas le même sens dans les deux cas. La neurocriminologie défend une médicalisation du crime en des termes qui garantissent une extension de son champ d'intervention : le crime est défini comme un problème médical susceptible d'être traité par des outils d'intervention neuroscientifique. On ne saurait trop insister sur le fait que cette stratégie a bénéficié d'une conception préformationniste et psychopathologique du comportement criminel, tout au long des études menées depuis les années 1940 : l' « instinct criminel » est corrigible en vertu du fait qu'il est intrinsèque au cerveau des criminels et qu'il relève du registre de la pathologie (de l'infirmité et de la défectuosité).

À l'inverse, Lombroso rejette toute possibilité de redressement de l'instinct criminel, précisément parce qu'il est inné et incontrôlable : le caractère intrinsèque et irrésistible de l'instinct rend le criminel-né insensible aux mesures correctives comme l'éducation ou la morale. Dans la pensée lombrosienne, l'instinct criminel est

[177] Pour rappel, l'*epilepsy defense* s'inscrit généralement dans le cadre de l'*insanity defense* et de la règle M'Naghten (Weinberg 1977). Elle fait l'objet de nombreuses controverses aux Etats-Unis, certains juristes ou psychiatres ayant noté un risque de biais liés à des facteurs confondants, notamment la présence de lésions cérébrales qui pourrait être attribuée à un autre trouble que l'épilepsie (e.g. Paul and Lange 1992).

intrinsèquement lié à une monstruosité de nature, ce qui atteste l'impuissance de la société dans la prise en charge du criminel-né. Lombroso fait de la monstruosité un qualificatif de dangerosité et d'incorrigibilité afin de justifier l'exclusion sociale du criminel-né. En caractérisant ce dernier sous les traits d'un monstre, il parvient à le disqualifier de son statut juridique, c'est-à-dire à rendre inintelligible la question de sa responsabilité pénale : le criminel ne relève ni d'une qualification juridique, ni d'une qualification médicale ; il devient un problème insoluble pour le droit pénal car sa dangerosité est conçue comme un état permanent et incurable.

Bien que les finalités poursuivies par les neurocriminologues marquent une rupture avec l'ambition de la criminologie lombrosienne, les deux projets défendent une thèse identique en ce qui concerne le rôle de la société dans l'émergence de la criminalité. Les neurocriminologues, comme Lombroso, cherchent à se dessaisir de la question des déterminants sociaux de la criminalité, non seulement pour asseoir leur autorité scientifique, mais aussi pour promouvoir leur propre agenda politique. L'essentialisation du crime rend légitime le discours consistant à disculper la société (les institutions sociales) de toute responsabilité dans le développement de la criminalité : si le crime est intrinsèque au cerveau des criminels, alors il n'y a pas besoin d'aller rechercher la cause de la criminalité en dehors de ce que peuvent nous dire les outils neuroscientifiques. On voit ainsi clairement émerger le projet idéologique qui sous-tend la recherche en neurocriminologie, et qu'elle partage avec l'anthropologie criminelle de Lombroso : démontrer l'impuissance des approches sociologiques pour expliquer le phénomène de la criminalité et garantir la supériorité des outils neuroscientifiques pour repérer les individus dangereux et renforcer la lutte contre la récidive. Les questions relatives à l'explication, la prévention ou la prise en charge de la criminalité apparaissent comme un problème exclusivement neurobiologique, dont la résolution ne peut s'opérer qu'à l'intérieur du champ de la neurocriminologie.

Ce dernier point nous invite à interroger la position adoptée par la neurocriminologie vis-à-vis des approches sociologiques du crime. Tout au long de cette section, nous avons montré l'existence d'un biais de confirmation en faveur d'une conception préformationniste du comportement criminel. Le principe selon lequel le criminel ressemble à son cerveau est moins le résultat d'une démonstration empirique que l'affirmation d'un postulat innéiste, fruit d'un double héritage phrénologique et lombrosien. Pourtant, nous avons déjà eu l'occasion de noter la grande prudence des

neurocriminologues pour récuser les accusations de cérébro-centrisme ou d'innéisme dont elle pourrait faire l'objet (Fallin et al. 2018). Les initiatives interdisciplinaires qui ont émergé à la fin des années 1990 témoignent d'un effort constant de devancer ou désamorcer ces accusations. Ainsi, dans un souci d'anticiper les objections éventuelles face aux critiques que nous venons de formuler, nous allons chercher à déconstruire toutes les tentatives visant à étudier les interactions entre les facteurs neurobiologiques et les facteurs socio-environnementaux : tous les efforts déployés par les neurocriminologues relèvent eux-mêmes d'un postulat innéiste, et sont conçus de sorte à invalider la pertinence des approches sociologiques du crime.

5.3. Les échecs du modèle intégratif

5.3.1. Un double discours à l'égard de l'influence environnementale

Bien que l'étude du cerveau des criminels soit gouvernée par l'adoption d'un postulat innéiste, il est intéressant de noter que la phrénologie et la criminologie du XIX^{e} siècle reconnaissent l'influence de l'environnement dans le développement de la criminalité. Gall défend l'influence mutuelle des facteurs organiques et socio-environnementaux dans l'expression du « penchant au meurtre », celle-ci pouvant être modulée en fonction des circonstances (favorables ou défavorables). Lombroso admet lui aussi que les circonstances sociales peuvent jouer un rôle dans l'expression des instincts violents, car il reconnaît l'existence des « criminels par passion » et des « criminels d'occasion ou d'habitude ». En somme, chez Gall, comme chez Lombroso, les circonstances socio-environnementales jouent un rôle *modulateur* : elles peuvent réprimer ou renforcer l'expression de l'instinct criminel, en interagissant avec les dispositions organiques de l'individu.[178]

[178] Il est important de rappeler que la prise en compte des facteurs socio-environnementaux n'est pas incompatible avec leur postulat innéiste : c'est précisément parce que l'instinct criminel est inscrit dans les dispositions organiques des individus que l'influence environnementale peut s'exercer.

La neurocriminologie tient un discours nettement plus ambigu à l'égard des facteurs socio-environnementaux de la criminalité. La plupart des études assument l'intérêt exclusif porté au cerveau dans l'étude du comportement criminel, puisqu'elles excluent délibérément les facteurs démographiques et socio-économiques (âge, genre, niveau d'éducation, catégorie socio-professionnelle, etc.) dans la collecte et l'interprétation des données. Par conséquent, les seules données qui sont prises en compte dans le modèle explicatif concernent les données neurophysiologiques (EEG) ou neurobiologiques (IRM, TEP), les données relatives au dossier judiciaire des sujets (type de condamnation, nombre de récidives, etc.), et, dans certains cas, les scores obtenus à l'échelle d'évaluation des comportements agressifs ou impulsifs.

De fait, en vertu des variables sélectionnées dans la construction des études, la référence au cerveau neutralise la pertinence des autres facteurs explicatifs (par exemple contextuels ou historiques) de la criminalité. Rares sont les études qui mènent une réflexion sur l'influence des facteurs socio-environnementaux dans l'apparition des anomalies cérébrales. Pour autant, nous avons cité quelques exemples dans lesquels les anomalies ne sont pas toutes attribuées à une cause neurobiologique : lorsque les causes socio-environnementales ne sont pas délibérément occultées dans la construction du modèle expérimental, elles sont invoquées *a posteriori* pour rendre compte des différences dans la fréquence des anomalies observées entre deux catégories criminelles et expliquer ainsi certaines formes de criminalité.

Pour rappel, l'interprétation des données s'opère souvent dans un partage dichotomique entre la catégorie des criminels dangereux − caractérisés par un excès d'anomalies − et la catégorie des criminels dits « normaux », présentant une faible proportion d'anomalies (Hill 1952 ; Williams, 1969 ; Raine et al. 1998a, 2004 ; Yang et al. 2005, 2010). Du point de vue des neurocriminologues, la différence entre les deux groupes s'expliquerait par l'influence causale de deux facteurs mutuellement exclusifs : il existerait une « criminalité endogène » dont la cause réside dans l'organisation cérébrale (défaillante, défectueuse ou difforme) de l'individu, et une « criminalité exogène » qui trouverait son explication dans les circonstances socio-environnementales. La comparaison des données recueillies chez différentes catégories criminelles est donc marquée par une tendance interprétative qui consiste à faire de l'excès d'anomalies un défaut « inné » ou « constitutionnel », et à reléguer les formes les moins sévères d'anomalies à une origine « exogène ». Autrement dit, l'influence causale des facteurs socio-environnementaux (l'existence même d'une « criminalité exogène »)

n'est reconnue que par défaut (*a posteriori*), lorsque certains crimes ne peuvent être attribués à une explication neurobiologique.[179]

De manière générale, la neurocriminologie contemporaine entretient un double discours à l'égard de l'influence des facteurs socio-environnementaux dans le développement du comportement criminel. Depuis la fin des années 1990, elle cherche à anticiper la critique du cérébro-centrisme, en multipliant les initiatives interdisciplinaires visant à rechercher les « bases biosociales de la violence » (e.g. Raine et al. 1997b). Or, nous avons montré dans le troisième chapitre comment les neurocriminologues parvenaient à détourner et à invalider la pertinence explicative des causes sociales du crime, en les reléguant systématiquement à un niveau de description neurobiologique. Rappelons que l'une des stratégies consiste à réduire les facteurs de risque de criminalité (par exemple, les antécédents d'abus sexuels, de toxicomanie, etc.) à leurs effets neurologiques : ce renversement de causalité leur permet de justifier l'idée que le comportement criminel est lié, en dernière instance, à des facteurs neurobiologiques. Plutôt que d'interroger la dynamique de l'interaction entre les anomalies cérébrales et les déterminants psychosociaux de la criminalité, les neurocriminologues cherchent à construire un modèle explicatif entièrement articulé autour du niveau neurobiologique. La dynamique de l'interaction entre les facteurs neurologiques et socio-environnementaux est elle-même formulée en termes neurobiologiques, réduite à un mécanisme figé et irréversible.

Cette logique éliminativiste constitue le fondement implicite des recherches menées dans le champ de la criminologie biosociale. Dans son ouvrage majeur *The psychopathology of crime* (1993), Adrian Raine défend notamment l'idée que ces recherches interdisciplinaires permettraient de résoudre les failles des explications sociologiques de la criminalité. L'hypothèse dite des « facteurs psychophysiologiques de protection » a par exemple été avancée pour expliquer le fait que des individus ne commettent pas de crime malgré un environnement psychosocial à risque (abus physiques et/ou sexuels pendant l'enfance, négligence parentale, antécédents familiaux de criminalité, etc.). D'après Raine, certains états neurophysiologiques comme un niveau

[179] Nous renvoyons le lecteur à l'analyse présentée dans la section 3.4. concernant l'étude menée par Adrian Raine (1998a), qui proposait d'isoler les différences neurobiologiques entre les criminels ayant souffert de carences psychosociales et les criminels ayant bénéficié d'un environnement socio-familial stable.

élevé d'excitation corticale, pourraient contrebalancer les circonstances défavorables, et minimiser le risque de développer un comportement criminel :

> « *[I]t is hypothesized that psychophysiological protective factors against the development of crime (e.g., high arousal) may help explain why some who are at risk for crime by virtue of possessing negative psychosocial factors do not develop such an outcome ; such a perspective could help potentiate findings from psychosocial research on crime.* » (Raine 2013 [1993], p. 190).

Il paraît inutile de préciser qu'il s'agit là d'un faux problème, conçu pour favoriser l'extension du champ d'intervention de la neurocriminologie. La portée explicative des données neurobiologiques s'étend au-delà même du domaine d'investigation de la criminologie (l'étude du comportement criminel) : elles sont mobilisées pour expliquer ce que Raine décrit comme des « faux positifs », c'est-à-dire l'absence de criminalité malgré un environnement criminogène.[180] Nous avions déjà cité une étude publiée par Raine et coll. (1998a) qui proposait d'interpréter les données neurobiologiques comme le témoin d'une discordance possible entre la qualité de l'environnement familial et la dangerosité de l'individu. D'après les auteurs de cette étude, les déficits localisés au niveau du cortex préfrontal permettaient d'expliquer les crimes violents commis par des individus ayant bénéficié d'un environnement socio-familial parfaitement stable. De ce point de vue, les anomalies cérébrales apparaissaient comme l'explication privilégiée des crimes ne pouvant être imputés à un environnement criminogène.

Avec l'hypothèse des « facteurs psychophysiologiques de protection », les données neurobiologiques sont mobilisées pour créer un problème qui, par définition, ne peut être résolu par les explications sociologiques traditionnelles. Il ne s'agit plus seulement de montrer que le cerveau garantit une explication complète du comportement criminel (quel que soit l'environnement socio-familial du sujet), mais d'invoquer des mécanismes neurobiologiques (hypothétiques) pour expliquer l'absence de criminalité. Précisons d'ailleurs que, selon Raine, cette hypothèse bénéficierait davantage au champ de la sociologie du crime qu'à la criminologie biosociale ou à la neurocriminologie, ce qui laisse peu de doute quant à sa volonté de s'implanter dans le champ des sciences

[180] Comme le souligne Raine : « *[m]any of these "false positives" may be explained by the fact that they possess, for example, higher arousal, which serves to protect them from developing criminal behavior.* » (Raine 1993, pp. 188-189).

sociales, et de remplacer les approches sociologiques traditionnelles par des explications neurobiologiques :

> « *This line of research, while involving biological risk factors, could paradoxically be of greater significance to socially based research on crime than psychophysiological research.* » (Raine 1997 [1993], p. 188).

5.3.2. L'éternel retour des thèses de Gall et Lombroso

Si l'on résume les différentes manières dont la neurocriminologie conçoit les interactions entre les facteurs neurobiologiques et socio-environnementaux, on voit apparaître quatre stratégies, qui répondent toutes à la volonté d'invalider la pertinence d'une approche sociale ou sociologique du crime : 1/ les causes socio-environnementales sont délibérément exclues dans la construction des études car les variables sont sélectionnées de sorte à favoriser une explication neurobiologique du comportement criminel ; 2/ les causes socio-environnementales sont invoquées par défaut, pour rendre compte des données qui rapportent une faible proportion d'anomalies cérébrales, et combler ainsi les limites du modèle neurobiologique de la criminalité ; 3/ les causes socio-environnementales sont reformulées en termes neurobiologiques, c'est-à-dire détournées au profit d'une explication neurobiologique de la criminalité ; 4/ les causes socio-environnementales sont tout simplement révoquées, désinvesties de toute pertinence pour expliquer le phénomène de la criminalité.

Il est difficile de ne pas voir dans ces stratégies une réactualisation des grands principes qui ont structuré la typologie criminelle au XIX^{e} siècle. D'abord, l'opposition établie entre la criminalité endogène et exogène rappelle la distinction proposée par Lombroso entre les criminels-nés (caractérisés par « un maximum de monstruosité ») et les criminels « d'occasion » ou « par passion », caractérisés par une plus faible proportion d'anomalies. Cette distinction renouvelle la typologie établie entre les criminels poussés au crime par leur organisation biologique, et les criminels qui y sont entrainés par la force des circonstances. La prégnance de la pensée lombrosienne est d'autant plus manifeste que cette distinction a pour seule fonction d'isoler les « criminels dangereux » des « criminels normaux ». Le fondement empirique de cette distinction est d'ailleurs largement contestable, précisément parce que l'origine du crime est inférée à partir des

seules données sur le degré des anomalies observées. Or, aucune donnée relative aux facteurs démographiques ou socio-économiques ne permet de retracer le contexte d'émergence du crime dans l'histoire personnelle de l'individu.

Comme chez Lombroso, cette typologie répond à la volonté d'expliquer toutes les formes de criminalité. Il s'agit non seulement de ranger chaque type de criminalité dans une catégorie naturelle (neurobiologique) spécifique, mais aussi d'expliquer les crimes qui ne peuvent être attribués à un excès d'anomalies cérébrales. Les neurocriminologues sont en effet contraints de céder à un élargissement théorique pour rendre compte des données rapportant une faible proportion d'anomalies cérébrales, qui contredisent ainsi le modèle innéiste de la criminalité. On retrouvait déjà chez Lombroso cet effort de systématisation, dont l'unique but était de combler les limites de sa théorie du criminel-né et d'anticiper les critiques qui pourraient souligner les incohérences de sa thèse.

Par ailleurs, les efforts déployés par les neurocriminologues pour tenir compte des facteurs socio-environnementaux réhabilitent les grands principes théoriques qui régissent l'organologie de Gall, en particulier sa thèse de la pluralité. Rappelons que Gall défend l'influence mutuelle des facteurs organiques et socio-environnementaux dans l'expression de l'« organe du meurtre ». Celui-ci ne prédispose pas nécessairement la manifestation du penchant car son exercice peut être modulé en fonction des circonstances socio-environnementales (favorables ou défavorables). Les circonstances favorables comme l'éducation ou la morale peuvent par exemple réprimer l'expression de l'« organe du meurtre » ; inversement, les circonstances défavorables peuvent être contrebalancées par l'activité de certains « organes cérébraux », en favorisant l'exercice des qualités morales ou intellectuelles.

L'hypothèse des « facteurs psychophysiologiques de protection » est fondée sur le même raisonnement. Du point de vue des neurocriminologues, il suffirait de posséder un état cérébral particulier (en l'occurrence, un niveau élevé d'excitation corticale) pour limiter les risques psycho-sociaux et s'affranchir d'un déterminisme social. Cette hypothèse n'est pas plus solide sur le plan épistémologique que la théorie du « déséquilibre cérébral » défendue par Gall, ou l'idée selon laquelle certains « organes cérébraux » peuvent contrebalancer les circonstances sociales défavorables, et neutraliser ainsi l'expression de l'instinct criminel. D'abord, elle favorise l'idée selon laquelle le cerveau possède toutes les conditions matérielles nécessaires à la manifestation ou à la répression d'un comportement criminel : le cerveau tient un rôle privilégié pour expliquer

tous les aspects de la criminalité, y compris lorsqu'elle ne se manifeste pas. Mais surtout, cette hypothèse réactualise l'idée que certains « organes cérébraux » peuvent exercer une influence modératrice sur les circonstances socio-environnementales défavorables, ce qui renforce encore davantage la prééminence du cerveau sur tous les autres facteurs explicatifs.

En outre, l'hypothèse des « facteurs psychophysiologiques de protection » ne correspond à aucune réalité empirique. Il s'agit d'une hypothèse purement théorique, dont l'argument repose sur un paralogisme. Il est fondé sur le constat selon lequel certains criminels ou psychopathes présentent une « sous-stimulation » du système nerveux central et autonome – c'est-à-dire une réduction de l'activité corticale, un ralentissement du rythme cardiaque, et une faible conductance cutanée – lorsqu'ils sont soumis à un protocole de conditionnement à la peur.[181] En clair, l'hypothèse des « facteurs psychophysiologiques de protection » repose sur le raisonnement suivant : dans la mesure où les criminels sont caractérisés par une « sous-stimulation » corticale, on peut supposer qu'un niveau élevé d'excitation corticale joue un rôle neuroprotecteur vis-à-vis de la criminalité.

Cette proposition conclut à partir de la négation d'un état cérébral, une inversion de la relation entre l'état cérébral et le comportement qui lui est associé : « si P équivaut à Q ; alors non-P équivaut à non-Q ». Autrement dit, l'hypothèse des « facteurs psychophysiologiques de protection » consiste à associer à toute occurrence d'un état négatif l'absence d'occurrence du comportement associé à cet état. Or, la validité de cette inférence est contestable au regard de la polysémie de la négation − qui peut traduire aussi bien une contradiction, une altérité ou une contrariété – et compte tenu du statut de la relation entre l'état cérébral et comportemental. On peut mobiliser un argument tiré de la neurocriminologie elle-même, pour soutenir que la « sous-stimulation » du système nerveux ne remplit pas toutes les conditions de manifestation du comportement criminel (ce dernier étant associé à une infinie variété d'anomalies cérébrales). La négation de cet état cérébral ne garantit pas nécessairement l'absence de criminalité car celle-ci peut se

[181] Plusieurs études longitudinales ont montré qu'une « sous-stimulation corticale » chez des enfants ou adolescents était associée à des condamnations criminelles plus fréquentes à l'âge adulte (Mednick et al. 1981 ; Raine et al. 1990 ; Niv et al. 2015, 2018). Dans la mesure où cet état cérébral (qui se traduisait en particulier par un excès d'ondes lentes dans les régions frontales) était présent bien avant la manifestation du comportement criminel, les chercheurs en ont déduit qu'il constituait un « indicateur prédictif », voire une « prédisposition psychophysiologique » du comportement criminel (Raine 1993).

traduire de multiples manières dans le cerveau. Le caractère contingent de la relation entre l'état cérébral et l'état comportemental interdit donc toute généralisation inductive : la négation du premier n'est pas un critère suffisant pour conclure à l'absence du second. Pour le dire encore autrement, montrer que tel comportement est observé en présence de tel état cérébral ne permet pas d'affirmer que la négation de ce dernier garantit l'absence du comportement.

Jusqu'ici, nous avons analysé la validité du projet neurocriminologique – et plus précisément la thèse de l'innéisme criminel − par le prisme de ses affinités avec les études phrénologiques et criminologiques du XIX^{e} siècle. Nous avons montré que cette filiation limitait considérablement le pouvoir explicatif de la neurocriminologie, ce qui nous a donné l'occasion de désamorcer le double discours mis en œuvre pour asseoir la légitimité scientifique de ses recherches. Il convient désormais de sortir du cadre théorique et méthodologique de la neurocriminologie, pour adopter une perspective plus générale sur la possibilité empirique d'une explication neurobiologique du comportement criminel.

En faisant abstraction de tous les biais inhérents à la construction des études menées depuis les années 1940, on peut très naïvement se demander si l'on peut donner un sens à l'ambition d'analyser le phénomène de la criminalité en étudiant uniquement le cerveau des criminels. Il se pose en effet la question de savoir si les raisons épistémologiques de l'échec de la neurocriminologie tiennent au cadre théorique et méthodologique dans lequel les recherches sont conduites (c'est-à-dire aux présupposés et aux stratégies de recherche mobilisés), ou au fait que les neurocriminologues portent un intérêt exclusif au cerveau dans l'étude du comportement criminel. Dans la mesure où les recherches en neurocriminologie adhèrent à un cérébro-centrisme de principe, il paraît difficile de contester leurs arguments sans se placer sur le terrain empirique. On peut cependant mener une réflexion plus générale sur le projet explicatif de ce programme de recherche : sans nécessairement renverser la chaîne de causalité en mobilisant d'autres stratégies explicatives (par exemple, sociologiques), il existe de bonnes raisons de considérer que le cerveau *seul* ne suffit pas à expliquer le comportement criminel.

5.4. Pourquoi le criminel *n'est pas* son cerveau

Les conclusions que nous avons tirées concernant le paradigme de la neurocriminologie nous invitent à interroger deux thèses : 1/ la thèse selon laquelle le comportement criminel est imputable au *seul* cerveau ; et 2/ la thèse selon laquelle les conditions d'émergence du comportement criminel sont internalisées dans des états cérébraux spécifiques. La première thèse est issue de ce qui est communément appelé le cérébro-centrisme, c'est-à-dire la tendance à se référer exclusivement au cerveau dans l'étude des comportements. La seconde correspond à une conception internaliste des processus psychologiques, à savoir l'idée que les conditions de l'activité mentale sont internes au cerveau. Le préformationnisme est tributaire de cette thèse, dans la mesure où il conçoit les phénomènes de haut niveau comme le produit de leur propre déploiement interne. De ce point de vue, les processus psychologiques résultent d'une « auto-construction » endogène et apparaissent comme la somme de leurs déterminations internes.

En s'affranchissant du cadre de référence de notre analyse épistémologique (la phrénologie et l'anthropologie criminelle du XIX^{e} siècle), nous allons ici chercher à déconstruire l'idée que le criminel est « d'essence cérébrale ». Le cérébro-centrisme et l'internalisme constituent le fondement théorique du projet explicatif de la neurocriminologie : ces deux thèses participent conjointement à la prétention d'expliquer les conditions d'émergence de la criminalité par certains types d'états cérébraux (l'immaturité corticale, l'instabilité de l'activité cérébrale, le déséquilibre fronto-limbique ou inter-hémisphérique, etc.). Il est donc crucial de questionner leur validité si l'on veut montrer pourquoi la référence au cerveau est insuffisante pour fournir une explication complète du comportement criminel.

Les critiques portées à l'encontre du cérébro-centrisme ou de l'internalisme relèvent de différentes écoles de pensée, dont nous ne donnerons qu'un aperçu très général. Globalement, elles défendent l'idée que le cerveau n'est pas une condition suffisante à l'exercice des processus cognitifs car ces derniers émergent de l'interaction entre le cerveau, le corps et l'environnement. Précisons d'ores et déjà que notre objectif n'est pas tant de contester la pertinence de l'approche cérébrocentrique – en montrant que le comportement criminel est situé « en dehors du cerveau » – mais plutôt de mettre en

doute l'idée que seuls les états internes jouent un rôle dans l'exercice des facultés psychologiques. Nous verrons en effet que la stratégie consistant à défendre une position « externaliste » vis-à-vis des phénomènes psychologiques ne suffit pas à déconstruire le projet explicatif de la neurocriminologie ; encore faut-il montrer pourquoi l'ambition d'identifier les conditions d'émergence du comportement criminel à des états cérébraux spécifiques n'a pas de sens.

Nous insisterons donc en particulier sur les faiblesses du modèle internaliste, dans sa dimension à la fois ontologique et épistémologique. Comme le souligne le philosophe externaliste Mark Rowlands (1999), la conception internaliste des processus mentaux se présente sous la forme de deux thèses : une thèse ontologique, selon laquelle « les états et processus mentaux sont localisés exclusivement dans le crâne des organismes doués de cognition » ; et une thèse épistémologique, selon laquelle « il est possible de comprendre la nature des états et processus mentaux en s'intéressant exclusivement à ce qui se produit dans le crâne des organismes doués de cognition » (Rowlands 1999, p. 8). Il est très important de distinguer ces deux thèses, car la critique de l'internalisme peut mobiliser différentes stratégies selon que l'on s'intéresse à la question de la nature des processus mentaux, ou à la manière dont ils sont étudiés et expliqués : il n'est pas nécessaire de questionner la thèse ontologique pour invalider la pertinence d'une conception internaliste des processus mentaux. En outre, la question qui nous intéresse est bien épistémologique. Il s'agit de déterminer si l'on peut donner un sens à l'affirmation selon laquelle la connaissance du sujet criminel (et de ses comportements) s'accomplit entièrement dans la connaissance de son cerveau.

Précisons que les critiques que nous allons exposer ci-après présentent un certain degré de généralité, c'est-à-dire qu'elles soulèvent des problèmes dont la portée dépasse le champ de la neurocriminologie. Cette analyse nous donnera ainsi l'occasion de montrer en quoi ce programme de recherche illustre de manière paradigmatique certaines critiques adressées aux neurosciences et plus précisément aux *domaines d'application* des neurosciences. Il convient en effet de lever une confusion éventuelle sur la cible de notre analyse critique : elle concerne bien les applications judiciaires des neurosciences (et plus généralement nous le verrons, toutes les neurodisciplines), et non pas les neurosciences en elles-mêmes. Si nous évoquerons des remarques et des objections issues du champ de la philosophie critique des neurosciences, nous restons très réservée quant à leur pertinence et à leur efficacité concernant le domaine neuroscientifique en tant que tel. À

notre sens, ces critiques soulèvent des problèmes métaphysiques ou épistémologiques qui ne sont pas inhérents à la méthodologie des neurosciences en elles-mêmes, mais bien plutôt à l'exportation de leurs outils ou de leurs résultats hors de leur cadre d'élaboration.

En outre, les critiques formulées à l'encontre du cérébro-centrisme ou de l'internalisme ne nous engageront pas à défendre l'idée que les processus mentaux sont inaccessibles à l'investigation neuroscientifique. Il ne s'agit pas pour nous de nier l'influence causale du cerveau dans l'exercice des facultés psychologiques, ni l'idée que ces derniers relèvent d'un traitement interne. Il s'agit plutôt d'une part, de contester le présupposé selon lequel *seuls* les états internes jouent un rôle dans l'exercice des facultés psychologiques, et d'autre part, de montrer pourquoi le niveau de description neurobiologique est insuffisant dans l'explication du comportement criminel : « le cerveau oui, mais pas le cerveau seul ».

5.4.1. Etudier le cerveau *seul* ne suffit pas

Les critiques dirigées contre l'intérêt exclusif porté au cerveau peuvent mobiliser plusieurs voies argumentatives, qui s'étendent de la philosophie pragmatiste ou externaliste jusqu'aux connaissances empiriques fournies par les sciences biologiques.[182] On peut par exemple s'appuyer sur des arguments tirés de la biologie de l'évolution ou de la biologie du développement et défendre une position « interactionniste », c'est-à-dire concevoir les phénomènes cognitifs comme le résultat d'une co-construction entre différents niveaux d'organisation (de la génétique à l'environnement physique et social de l'individu).[183] De ce point de vue, le cérébro-centrisme est difficilement compatible avec les recherches actuelles menées dans le domaine de la biologie, en particulier depuis l'émergence de l'épigénétique et la multiplication des études menées sur le système

[182] Pour une analyse détaillée des différentes critiques qui peuvent être formulées à l'encontre du cérébro-centrisme, nous renvoyons le lecteur au troisième chapitre de l'ouvrage *Neuroscepticisme* de Denis Forest (2014).

[183] Cette thèse s'oppose à l'internalisme dans la mesure où elle défend l'influence mutuelle de l'environnement et de l'expression des gènes dans le développement des organismes. On la retrouve sous diverses formes chez Jean Piaget (1967), Richard C. Lewontin (1994), ou encore chez la philosophe des sciences Susan Oyama (1985).

nerveux entérique : privilégier le cerveau *seul* dans l'étude des phénomènes cognitifs ne permet pas de rendre compte de la dynamique des interactions entre l'environnement et les processus physico-chimiques et biologiques, ni de comprendre le rôle joué par les réseaux neuronaux situés en dehors du cerveau.[184]

On peut également recourir à des expériences de pensée – notamment l'expérience dite de la « Terre-Jumelle » ou la théorie du « cerveau dans une cuve », proposées par Hilary Putnam − pour défendre un externalisme sémantique (externalisme de contenu), à savoir l'idée que les déterminants de la pensée ne sont pas intrinsèques au cerveau et ne font sens que si l'on prend en compte l'environnement du sujet.[185] À bien des égards, le cérébro-centrisme favorise la représentation d'un individu réduit à un « cerveau dans une cuve », isolé du reste de son corps et de son environnement. En privilégiant le cerveau seul dans l'étude des phénomènes mentaux, on prend ainsi le risque d'adhérer à un nouveau dualisme « cerveau - reste du sujet » : le cérébro-centrisme conduit en définitive à substituer la substance pensante et immatérielle de Descartes (la *res cogitans*) par une entité interne, matérielle (le cerveau).[186] Cette conception − qui relève d'ailleurs davantage du neuroessentialisme que du cérébro-centrisme en tant que tel − est problématique pour plusieurs raisons.

---

[184] La notion d'épigénétique est polysémique. Au sens où nous l'entendons ici, elle renvoie à « la science qui étudie le passage du génotype au phénotype » (Morange 2005). Elle s'oppose à l'internalisme (et en particulier au préformationnisme) car elle étudie l'influence de l'environnement sur l'expression des gènes : elle permet de comprendre que la seule connaissance de l'ADN ne suffit pas à expliquer comment les gènes fonctionnent. Par ailleurs, les études menées sur le système nerveux entérique (communément appelé « le second cerveau ») montrent que la cognition n'est pas concentrée dans le seul cerveau puisque certaines opérations neurocognitives se distribuent jusque dans le système digestif (Gershon 1999).

[185] L'argument de l'externalisme sémantique a été formulé dans le cadre d'une critique du fonctionnalisme computationnel, au milieu des années 1970. Comme nous le préciserons plus loin, l'argument de Putnam vise à contester l'idée que les états mentaux sont des états computationnels qui possèdent un contenu conceptuel. Il montre que les états intentionnels n'ont pas de référence intrinsèque, c'est-à-dire qu'ils ne font pas sens à l'intérieur de notre cerveau, indépendant de l'environnement du sujet et de son activité.

[186] La critique de ce que l'on appelle communément le « nouveau dualisme » n'est pas nouvelle. On la retrouve chez le philosophe pragmatiste John Dewey (1859-1952), qui affirmait dans *Democracy and Education* (1916) : « [t]rop souvent […] l'ancien dualisme de l'âme et du corps a été remplacé par celui du cerveau et du reste du corps. Mais le système nerveux n'est qu'un mécanisme particulier permettant de maintenir toutes les activités corporelles en un travail d'ensemble. » L'objection du « nouveau dualisme » a par la suite été développée par Max Bennett et Peter Hacker (2003), ou encore par W. Teed Rockwell (2005) et Alva Noë (2009), sous la forme de ce qu'ils ont appelé le « matérialisme cartésien ».

D'une part, parce qu'elle n'échappe pas au problème de l'erreur de catégorie formulé par Gilbert Ryle (1900-1976) dans *Le Concept d'esprit* (1949).[187] En considérant que le cerveau joue un rôle exclusif dans l'exercice des processus cognitifs, on se rend coupable d'un sophisme méréologique de division puisque l'on attribue à la partie (le cerveau) des propriétés qui relèvent d'un tout (l'individu dans son environnement) : c'est l'individu qui pense et agit, et non pas son cerveau.[188] Cette erreur renvoie à ce que les philosophes Max Bennett et Peter Hacker (2003) nomment « l'adhésion irréfléchie à une forme mutante de cartésianisme ». Si le cérébro-centrisme s'inscrit (par définition) dans la tradition moniste matérialiste, il ne se soustrait pas pour autant aux impasses du dualisme cartésien car les propriétés erronément attribuées à l'âme sont désormais attribuées, tout aussi erronément, au cerveau (Poirier et Payette 2007). [189]

D'autre part, le cérébro-centrisme favorise la représentation d'un individu réduit à ses déterminations internes, à des propriétés intrinsèques qui font sens à l'intérieur de son cerveau. Il peut ainsi conduire à une forme d'éliminativisme vis-à-vis des états mentaux car il tend à concevoir ces derniers comme des épiphénomènes, c'est-à-dire des évènements ne jouant aucun rôle dans la chaîne causale qui mène à la réalisation des comportements. De ce point de vue, l'idée selon laquelle le cerveau est une condition suffisante à l'exercice des processus cognitifs nous amènerait à céder définitivement à

[187] Gilbert Ryle illustre l'erreur de catégorie, erreur selon lui inhérente au dualisme cartésien, par cet exemple : un étudiant visite une université, avec sa bibliothèque, ses laboratoires, ses amphithéâtres, etc. avant de demander « mais où est l'université ? ». Cet étudiant fait une erreur méréologique en considérant que l'université appartient à la même catégorie que l'ensemble des éléments qui la constitue. S'il persiste à chercher l'université en visitant tous les éléments qui la constitue, il finirait d'ailleurs par croire que l'université est une entité immatérielle, indépendante de tous ses bâtiments. Pour Ryle, Descartes fait la même erreur en considérant que l'esprit existe indépendamment du corps car ces deux entités appartiennent à la même catégorie.

[188] Un sophisme méréologique désigne une erreur de raisonnement qui consiste à confondre le tout et la partie. Il peut s'agir d'un sophisme méréologique de division – lorsque l'on attribue à la partie une propriété qui relève du tout, ou d'un sophisme méréologique de composition – lorsque l'on attribue au tout une propriété de la partie.

[189] Précisons que l'argument de Bennett et Hacker s'inscrit dans le cadre d'une critique générale du projet des neurosciences. Les neuroscientifiques sont selon eux plongés dans une immense confusion méréologique car ils attribuent au cerveau (ou à des régions cérébrales) des propriétés qui ne peuvent s'appliquer qu'à des individus. Nous reviendrons sur cette thèse dans la prochaine section car si elle n'est selon nous pas inhérente aux neurosciences en tant que telles, elle est particulièrement pertinente en ce qui concerne les neurodisciplines comme la neurocriminologie ou le neurodroit.

une explication en termes de causes et non de raisons d'agir : elle conduirait à neutraliser la pertinence des états intentionnels pour expliquer les comportements, et à les éliminer au profit de ses déterminations internes. S'il n'est pas dans notre propos d'entrer dans le détail de cette discussion, on peut mettre en doute l'idée que les états intentionnels comme les raisons d'agir font sens à l'intérieur de notre cerveau. Sans pour autant défendre l'autonomie explicative des raisons d'agir vis-à-vis du niveau subpersonnel,[190] on peut contester l'idée qu'elles possèdent une référence intrinsèque dans la mesure où elles sont en grande partie déterminées par l'environnement physique et social de l'individu. En s'appuyant sur l'argument de l'externalisme sémantique de Hilary Putnam – l'idée que « les significations ne sont pas dans la tête » (Putnam 1975, p. 227) –, on peut effectivement montrer que les états intentionnels ne font sens qu'en vertu des relations que l'individu entretient avec son environnement.

Par ailleurs, la référence au seul cerveau s'oppose à un courant de pensée bien connu dans le champ de la philosophie critique des neurosciences, qui défend l'« extension corporelle » des processus cognitifs. La critique du cérébro-centrisme peut en effet puiser ses arguments dans la grande famille des théories dites de la « cognition incarnée » de l' « énactivisme » ou de l' « esprit étendu », que l'on retrouve sous diverses formes chez Francisco Varela (Varela, Thompson and Rosch 1991 ; Varela 1992), Mark Rowlands (1999), Andy Clark et David Chalmers (1998), ou encore chez Alva Noë (2009) et W. Teed Rockwell (2005).[191] Ces théories s'accordent à replacer le cerveau dans son contexte biologique et écologique, et à situer les processus cognitifs dans

[190] Ce problème renvoie à la distinction entre le niveau personnel (les explications par les raisons) et le niveau subpersonnel (les explications par les causes) (Dennett 1969). La relation entre ces deux niveaux d'explication fait l'objet de nombreux débats en philosophie de l'esprit (voir notamment Proust and Pacherie 2008). Globalement, la question est de savoir si les propriétés intentionnelles des états mentaux peuvent être réduites à des causes de nature physique, et en définitive, à des explications neurobiologiques. Certains philosophes prônent une forme de séparatisme entre les deux niveaux, c'est-à-dire qu'ils défendent l'autonomie explicative du niveau personnel vis-à-vis du niveau subpersonnel. De ce point de vue, les états intentionnels comme les raisons d'agir sont non seulement irréductibles, mais ils relèvent uniquement de la compétence des philosophes de l'esprit. On peut au contraire défendre l'idée que le niveau d'explication personnel n'est pas autonome et qu'il n'est pleinement satisfaisant que si l'on prend en compte le niveau subpersonnel.

[191] Le terme d'énaction (ou de cognition incarnée) recouvre plusieurs sens. Il a été introduit par Francisco Varela et définit de manière générale la théorie selon laquelle la cognition émerge de l'interaction de l'individu avec son environnement (Varela, Thompson and Rosch 1991).

l'environnement conçu au sens large. La cognition s'étend en dehors du cerveau et se distribue dans d'autres parties du système nerveux, dans le corps et l'environnement de l'individu. Pour reprendre un exemple cité par Alva Noë (2009), il n'y a donc pas plus de sens à situer les phénomènes mentaux dans le (seul) cerveau que de localiser la vision dans l'œil ou le processus de digestion dans l'estomac : ce dernier, comme le cerveau, n'est qu'un organe parmi d'autres, une partie d'un système comprenant de multiples niveaux d'organisation, qui interagissent entre eux pour accomplir leurs fonctions.

Alva Noë a consacré son ouvrage *Out of Our Heads* (2009) à la question de savoir « pourquoi vous n'êtes pas votre cerveau ». Sa thèse s'inscrit plus largement dans le cadre d'une critique de la théorie de l'identité du mental et du cérébral, de l'internalisme, et plus spécifiquement, de la possibilité empirique d'une explication neurobiologique de la conscience.[192] Selon lui, les phénomènes mentaux (et en particulier la conscience) désignent quelque chose *que nous faisons*, et non pas quelque chose *qui se produit en nous* (Noë 2004, 2009). L'idée selon laquelle la conscience est internalisée dans le cerveau n'est pas seulement dénuée de fondement empirique ; elle est en elle-même dépourvue de sens car « le cerveau n'a pas d'activité cognitive en propre » (Noë 2009, p. 24). La thèse de Noë prend ainsi la forme d'un rejet de principe contre toute assimilation des processus cognitifs à des états cérébraux, et contre toute tentative visant à localiser les corrélats neuronaux de la conscience.[193] Les neurosciences échouent selon lui à expliquer les propriétés fonctionnelles de la conscience (c'est-à-dire à identifier son rôle causal), car la seule référence au cerveau ne rend pas compte de la dynamique de l'interaction entre le reste du corps, l'activité sensorimotrice et l'environnement au sens large :

> « *The brain is not the locus of consciousness inside us because consciousness has no locus inside us. Consciousness isn't something that happens inside us: it is something that we do, actively, in our dynamic interaction with the world around us*. » (Noë, 2009, p. 24).

La question des limites topologiques que l'on peut assigner à la cognition fait l'objet d'un vif débat, de même que la fécondité des théories de l'énaction ou de l'esprit

[192] Pour une discussion de cette thèse, voir Forest (2012).

[193] Sur ce dernier point, voir en particulier Noë et Thompson (2004).

étendu. On peut en effet voir dans ces théories un anti-naturalisme de principe (ou du moins une certaine imperméabilité face aux connaissances neuroscientifiques) qui conduit, de fait, à minimiser le rôle joué par le cerveau dans l'exercice des processus cognitifs, et surtout à rendre caduque (ineffective) la recherche des fondements empiriques de leur extension corporelle. Selon Fred Adams et Ken Aizawa (2001), l'extension crânienne ou corporelle des processus cognitifs est une possibilité logique, mais elle n'a pas de réalité empirique. Sans entrer dans le détail de cette thèse (et de ses limites),[194] précisons qu'ils défendent ce qu'ils appellent un « intracrânialisme contingent » : à défaut d'identifier empiriquement les processus cognitifs qui ont lieu en dehors du corps, on est contraint d'admettre qu'ils se produisent à l'intérieur du cerveau.

Sans nécessairement souscrire à cette thèse, on peut considérer que les théories de l'énaction et de l'esprit étendu ne suffisent pas à invalider le modèle internaliste tel qu'il est défendu par la neurocriminologie. Tout au plus permettent-elles de récuser la thèse ontologique selon laquelle les processus psychologiques sont localisés exclusivement dans le cerveau des criminels. Mais qu'en est-il de la thèse épistémologique, à savoir l'idée qu'il est possible d'*expliquer* les conditions d'émergence de la criminalité en étudiant uniquement les états internes du criminel ? On ne saurait trop insister sur le fait qu'en privilégiant le cerveau seul, on s'interdit de saisir la complexité des relations qui l'unissent à l'environnement au sens large. Les critiques que nous avons adressées à l'encontre du cérébro-centrisme nous amènent ainsi à défendre une conception dite « pluraliste » de l'explication des phénomènes mentaux. De ce point de vue, la fécondité de l'explication nécessite d'articuler le niveau de description neurobiologique avec ce que nous savons *en dehors* du cerveau ; elle exige l'instauration d'un dialogue avec des disciplines connexes qui étudient les mêmes mécanismes, mais à des niveaux de description différents.[195]

Pour autant, on peut supposer que cette conception pluraliste de l'explication n'ait aucune prise sur le modèle défendu par les neurocriminologues. En concentrant notre critique autour du cérébro-centrisme, on fait l'impasse sur les arguments avancés en faveur du modèle internaliste de la criminalité et sur les erreurs conceptuelles et empiriques qui leur sont inhérentes : défendre l'idée que le comportement criminel est

[194] Denis Forest (2014) a bien montré que cette thèse reposait sur une conception erronée de la théorie de l'esprit étendu (qui est d'ailleurs définie par les auteurs sous le terme de « cognition transcrânienne ou transcorporelle »).

[195] Voir à ce sujet Faucher et Poirier (2001) ; Forest (2019).

situé en dehors du cerveau ne fait que contourner le problème. L'intérêt exclusif porté au cerveau dans l'étude du comportement criminel est contestable pour les raisons que nous avons mentionnées ci-dessus ; mais la stratégie consistant à défendre une position externaliste vis-à-vis des phénomènes psychologiques ne suffit pas à déconstruire les arguments empiriques mobilisés par les neurocriminologues. Car le projet neurocriminologique est bien celui d'une « neuroessentialisation » du crime, c'est-à-dire la recherche des preuves empiriques que le criminel est « d'essence cérébrale ». Il ne s'agit donc pas tant pour nous de contester la pertinence de l'approche cérébrocentrique que de rejeter l'identification (au sens d'une essentialisation) du sujet criminel à tout ou partie de son cerveau.

Il est désormais évident que le projet explicatif de la neurocriminologie souffre d'un biais de confirmation en faveur d'une conception innéiste du comportement criminel, qui amène à exclure toute perspective socio-environnementale pourtant nécessaire à son appréhension. Nous voudrions ici proposer une critique pour ainsi dire « interne » au projet neurocriminologique, en laissant de côté la question de la validité de la thèse ontologique. En se plaçant sur le terrain des neurocriminologues, c'est-à-dire en admettant que le cerveau joue un rôle privilégié dans la manifestation du comportement criminel, et en faisant abstraction du biais de confirmation résultant du cérébro-centrisme, nous allons chercher à comprendre s'il est légitime de déléguer toutes les capacités psychologiques d'un individu à son cerveau ou à ses régions cérébrales. Cette analyse nous amènera en définitive à montrer en quoi le projet d'identifier les conditions d'émergence du comportement criminel à des états cérébraux spécifiques n'a véritablement aucun sens. L'idée selon laquelle le criminel est « d'essence cérébrale » est largement contestable, tant du point de vue théorique que sur le plan empirique. Elle repose sur une mésinterprétation des données neurobiologiques, et sur une conception erronée de la relation entre les états cérébraux et les états psychologiques. Comme nous le montrerons dans la section suivante, le problème provient d'une erreur méréologique inhérente au projet explicatif de la neurocriminologie, et de la persistance d'une confusion héritée de Franz-Joseph Gall, entre les facultés psychologiques et les *conditions d'exercice* de ces facultés.

Dans la suite de cette section, nous développerons l'argument de l'erreur méréologique tel qu'il a été proposé par Max Bennett et Peter Hacker (2003) dans le cadre de leur analyse critique des neurosciences, avant de questionner sa pertinence en ce qui concerne le projet de la neurocriminologie et du neurodroit. Cet argument a en effet été

repris par les philosophes Michael S. Pardo et Dennis Patterson (2010, 2013) pour rejeter le projet d'intégration des neurosciences dans le système pénal américain. Bien que cette analyse ne fasse pas référence au champ de la neurocriminologie, elle mérite une attention particulière car elle rend bien compte des enjeux épistémologiques liés à l'exportation des neurosciences hors de leur cadre d'élaboration. À partir d'un exemple concret issu du champ du neurodroit, nous questionnerons à la fois les apports et les limites de l'argument méréologique pour réfuter la thèse internaliste telle qu'elle est défendue par les neurocriminologues. Cette analyse nous donnera l'occasion de mener une réflexion plus générale sur l'ambition de fonder des catégories normatives (relatives, par exemple, à la question de la responsabilité pénale) sur des critères neurobiologiques.

5.4.2. Les cerveaux ne commettent pas de crimes

Comme nous l'avons mentionné plus haut, les théories de l'énaction et de l'esprit étendu ne suffisent pas à invalider la thèse épistémologique de l'internalisme. Nous retiendrons néanmoins de ces théories l'idée que les phénomènes cognitifs désignent des évènements *que nous faisons*, plutôt que des états internes *qui se produisent en nous* (Noë 2009). Dans la mesure où la neurocriminologie s'intéresse à des capacités psychologiques qui incluent une large série de volitions (la capacité à former une intention d'action, à former des raisons d'agir, à contrôler son action, etc.), on peut se demander dans quelle mesure les états cérébraux répondent aux critères d'attribution de ces processus psychologiques. Sans verser dans un behaviorisme naïf qui refuserait d'admettre que ces derniers désignent autre chose que des comportements, on peut soutenir qu'ils ne se traduisent pas seulement par la manifestation d'un état cérébral donné : considérer que les conditions de l'activité mentale sont internes au cerveau ne nous dit rien sur la manière dont les processus psychologiques sont réalisés, car ces derniers doivent pouvoir s'exprimer au travers d'un comportement.

Dans cette perspective, il peut être intéressant de s'attarder sur l'argument de l'erreur méréologique défendu par Bennett et Hacker dans *Philosophical foundations of neuroscience* (2003). Cet argument s'inscrit dans la tradition néo-wittgensteinienne du langage ordinaire. Il prend pour cible tout type de discours consistant à attribuer au cerveau des prédicats psychologiques, une erreur que Bennett et Hacker jugent

constitutive du discours neuroscientifique : selon eux, les résultats neuroscientifiques s'accompagnent d'un « neurodiscours » qui consiste à faire du cerveau un sujet, c'est-à-dire à lui allouer des propriétés ou des capacités que l'on attribue généralement aux individus.[196] Un des exemples cités par Bennett et Hacker est donné par Antonio Damasio dans *L'erreur de Descartes* (1994), lorsqu'il proclame que « nos cerveaux peuvent prendre de bonnes décisions en quelques secondes ». On peut également citer la neurophilosophe Patricia Churchland − l'une des grandes figures de la philosophie éliminativiste − qui affirme que « c'est le cerveau [...] qui ressent, qui pense, qui décide » (Churchland 2002).[197] S'il est vrai que les neuroscientifiques cèdent parfois à ce type d'erreur grammaticale, il convient d'insister sur le fait que ce neurodiscours est surtout prégnant chez les réductionnistes éliminativistes. Il est d'ailleurs largement repris dans la littérature consacrée aux neurodisciplines, lesquelles ont tout intérêt à montrer que c'est le cerveau (et le cerveau seul) qui « apprend », « résout des problèmes », « répond aux stimuli visuels », « prend la décision d'acheter », etc.

Comme le soulignent Bennett et Hacker, ce type d'énoncé n'est pas faux sur le plan logique ; il n'a simplement aucune *signification cognitive*. Il est dénué de sens car le cerveau ne satisfait pas les critères d'application des prédicats psychologiques : le sujet du comportement n'est pas le cerveau mais l'individu lui-même, lequel doit être envisagé

[196] Cet argument rappelle bien entendu celui de Gilbert Ryle, mais aussi celui de Wittgenstein qui affirmait que c'est seulement d'un être humain que l'on peut dire « qu'il éprouve des sensations, qu'il voit, qu'il est aveugle, qu'il est sourd, malentendant, conscient ou inconscient. » (Wittgenstein 1953, I, § 281).

[197] On pourrait également citer les nombreux neuroscientifiques ayant contribué à fonder le champ des « neurosciences du libre arbitre » (Libet et al. 1983, 1985 ; Wegner and Wheatley 1999 ; Wegner 2002 ; Soon et al. 2008, 2013). Ces études proposent de démontrer l'illusion du libre arbitre en reconsidérant le rôle de la conscience dans la réalisation de nos actions. Le neurodiscours est régulièrement mobilisé dans l'interprétation des données car il sous-tend l'argument de l'épiphénoménisme de la conscience : la conscience ne jouerait aucun rôle dans la réalisation de nos actions car « le cerveau "décide" d'initier ou, au moins, de se préparer à initier l'acte avant qu'il n'y ait une conscience subjective qu'une telle décision a eu lieu. » (Libet 1985, p. 536). Précisons que ces recherches ont fait l'objet d'un long et vif débat, à la fois sur le plan méthodologique et empirique (Banks and Pockett 2007, Bayne 2011), mais aussi sur le plan conceptuel (Mele 2005, 2006, 2007 ; Pacherie and Haggard 2010 ; Pacherie 2013 ; Nahmias 2010). L'argument de l'épiphénoménisme de la conscience et la négation du libre arbitre reposent sur des erreurs conceptuelles et inférentielles car la construction du modèle expérimental ne tient pas compte de la dynamique *en amont* de l'initiation de l'action : l'action étudiée se rapporte à la dernière étape de la formation d'une intention en action, et n'apporte donc aucune information sur la notion de libre arbitre (voir en particulier Pacherie 2013).

dans toutes ses dimensions et sa complexité. Autrement dit, l'erreur méréologique est inhérente à toute tentative visant à allouer au cerveau des capacités psychologiques. Le cerveau ne réalise pas lui-même nos pensées ou nos comportements car il n'est qu'une partie de nous-mêmes. Il ne renferme pas en lui des intentions ou des raisons d'agir, mais seulement des activités électriques (Bennett et Hacker 2003 ; Noë 2009). En outre, comme le relève le juriste américain Stephen J. Morse, il n'y a aucun sens à faire du cerveau du criminel (ou des anomalies qui le caractérise) le véritable *sujet* de ses comportements : « *[b]rains do not commit crimes ; people commit crimes.* » (Morse 2006, p. 397). Les états cérébraux du criminel ne produisent pas à eux seuls les phénomènes psychologiques associés à leurs comportements car le cerveau ne satisfait pas les critères d'attribution des états psychologiques : ces derniers se rapportent à des *comportements*, et non à des états cérébraux.

L'argument de l'erreur de catégorie a été repris par les philosophes Michael S. Pardo et Dennis Patterson (2013), dans un ouvrage consacré à l'analyse des fondements conceptuels et empiriques du projet du neurodroit (voir aussi Pardo et Patterson 2010).[198] Selon eux, cet argument constitue le principal rempart contre tous les projets d'intégration des neurosciences dans le champ pénal : les recherches empiriques sont dénuées de sens car elles se déploient dans la confusion méréologique la plus totale. Sans partager entièrement leurs conclusions – qui s'apparentent à notre sens à une hostilité de principe − certaines méritent d'être mentionnées ici car elles permettent de mieux cerner le problème posé par l'ambition de fonder les catégories normatives du droit pénal sur des catégories neurobiologiques.

Leur analyse offre une perspective générale sur les erreurs conceptuelles et inférentielles qui sous-tendent la relation établie entre les trois niveaux de description mobilisés dans le champ du neurodroit : le niveau de description neurobiologique, les concepts psychologiques relatifs aux comportements individuels, et les concepts normatifs relevant du droit pénal. Les recherches en neurodroit sont en effet dictées par l'ambition d'utiliser les résultats neuroscientifiques qui portent sur des concepts mentaux ayant une pertinence vis-à-vis de la question de la responsabilité pénale (l'intention, le

[198] Cet ouvrage s'inscrit dans la continuité des thèses développées par Bennett et Hacker (2003, 2010). On peut d'ailleurs y voir une application directe de leurs thèses à la question de l'usage judiciaire des neurosciences, cet usage étant principalement analysé par le prisme de ses erreurs conceptuelles et catégorielles.

contrôle de soi, la connaissance morale, etc.). En clair, le neurodroit assoit sa légitimité sur le fait que le droit pénal américain mobilise un large éventail de concepts issus de la psychologie populaire, qui sont justement accessibles à l'investigation neuroscientifique. Pour Pardo et Patterson (2013), l'ambition d'intégrer les neurosciences dans le système pénal apparaît alors essentiellement comme un problème conceptuel ou catégoriel : celui de la traduction d'un concept empirique (relatif par exemple à un dysfonctionnement dans les régions impliquées dans le contrôle de soi) en un concept éminemment normatif (en l'occurrence une preuve de l'irresponsabilité pour cause psychiatrique).[199] Autrement dit, selon eux, le problème inhérent au projet du neurodroit est celui de l'inadéquation ontologique entre le langage neuroscientifique et le langage juridique issu de la psychologie populaire.

Pour reprendre un exemple régulièrement cité par Pardo et Patterson (2013), l'ambition d'intégrer les neurosciences dans le système pénal repose notamment sur l'idée qu'il serait possible d'inférer la possession d'une *connaissance* en enregistrant l'activité cérébrale d'un individu. Cette idée est largement partagée par les chercheurs en neurodroit car le problème de la responsabilité pénale pose un certain nombre de questions relatives à la notion de connaissance : de leur point de vue, les neurosciences apparaissent comme un outil précieux pour déterminer si le prévenu a une connaissance des règles morales ou juridiques (et donc évaluer sa capacité à apprécier la valeur morale de son acte), dans l'objectif de préciser le degré de *mens rea* (c'est-à-dire estimer si le prévenu a agi sciemment et en connaissance de cause),[200] ou encore pour déterminer s'il a connaissance de certaines informations relatives au crime pour lequel il est poursuivi.

---

[199] Nous faisons ici référence à l'un des projets du neurodroit, qui vise à utiliser les données issues de la neuroimagerie pour informer la condition volitionnelle de l'*Insanity Defense* (ou défense d'aliénation mentale), c'est-à-dire apporter la preuve que le prévenu est incapable de contrôler son comportement.

[200] La notion de connaissance est particulièrement importante lorsqu'il s'agit d'examiner l'élément moral de la responsabilité pénale (la *mens rea*) et se prononcer ainsi sur la culpabilité ou non d'un prévenu. Dans les juridictions de *common law*, il existe en effet quatre degrés d'intention coupable selon le niveau de connaissance (ou d'incertitude) face aux risques associés à l'acte : *purpose* / *knowledge* / *recklessness* / *negligence*. Des études publiées dans le domaine du neurodroit ont montré que ces degrés d'intention étaient associés à des patterns spécifiques d'activité cérébrale (Vilares et al. 2017 ; Ginther et al. 2018). Les auteurs de ces études ont alors évoqué la possibilité d'utiliser l'IRMf pour déterminer si le prévenu avait ou non connaissance des risques encourus, et évaluer avec précision son degré de culpabilité. En enregistrant son activité cérébrale pendant la reconstitution du scénario criminel, il serait selon eux possible d'inférer l'état

Cette dernière ambition s'est formée au début des années 1990, à partir des recherches menées par le neuroscientifique Lawrence Farwell sur ce qu'il a appelé lui-même le *Brain Fingerprinting Test* (Farwell and Donchin 1991). Ce test propose de reconnaître dans l'activité EEG d'un individu la « trace mémorielle » d'une information donnée, en enregistrant le profil d'activité de l'onde p300 : cette onde, qui survient 300 millisecondes après la présentation d'un stimulus, manifeste une réponse caractéristique lorsque ce dernier est familier (voir à ce sujet Polich and Kok 1995).[201] Aux Etats-Unis, ce test a été utilisé dans certaines procédures pénales, pour déterminer si le prévenu avait ou non connaissance des informations relatives au crime pour lequel il était poursuivi (l'arme du crime, la scène du crime, etc.) (Sharma, Verma, and Nagpal 2018). Pour ne citer qu'un seul exemple, ce test a été reconnu comme une preuve scientifique pour disculper un individu qui avait été condamné pour homicide volontaire, le rapport d'expertise scientifique réalisé par Lawrence Farwell (dont nous présentons ci-dessous un extrait) ayant largement contribué à sa libération en 2000, après 22 ans de détention :

> « *It is clear that Harrington's brain does not contain critical details about the crime […] his brain does, however, contain critical details about the events that actually took place that night. We can conclude scientifically that the record of the night of the crime stored in Harrington's brain does not match the crime scene, and does match the alibi.* » (Farwell 2012).[202]

S'il n'est pas de notre propos ici de discuter de la fiabilité de cette technique (ni de l'acceptabilité de la preuve elle-même), retenons qu'elle illustre parfaitement la confusion méréologique qui règne dans l'ambition d'intégrer les outils neuroscientifiques

---

mental dans lequel il se trouvait au moment de l'acte, afin de déterminer s'il a agi sciemment et en pleine connaissance de cause, ou par simple négligence ou imprudence.

[201] Cette technique repose sur ce que Lawrence Farwell a appelé la « P300-MERMER » (*Memory and Encoding Related Multifaceted Electroencephalographic Response*). Elle reprend globalement le même paradigme expérimental que le *Guilty Knowledge Test*, qui est utilisé dans le cadre de la neurodétection du mensonge. Il s'agit de présenter à l'individu trois types d'informations : une information dont il a connaissance, une information qu'il ignore, et une information que seul l'auteur du crime peut connaître. Si la réponse cérébrale associée à la première information recoupe celle associée à la troisième, on en déduit que le prévenu a connaissance des éléments du crime pour lequel il est poursuivi et qu'il est donc potentiellement coupable.

[202] Précisons que l'usage de l'EEG s'inscrit ici dans un contexte très particulier, favorisé par la pression de l'opinion publique pour libérer Terry Harrington en raison d'une insuffisance de preuves.

dans le système pénal. De toute evidence, le principe de fonctionnement du *Brain Fingerprinting Test* (BFT) repose sur l'idée que la connaissance d'une information est internalisée (emmagasinée) dans le cerveau sous la forme d'un engramme mémoriel. Ce présupposé rend légitime le neurodiscours consistant à faire du cerveau un sujet – « le cerveau de Harrington ne contient pas » les informations relatives au crime pour lequel il était poursuivi – et il justifie par la même occasion l'idée que le cerveau *seul* nous renseigne sur la question de savoir si l'individu a connaissance ou non des informations qui lui sont transmises.

Comme le montrent Pardo et Patterson (2013), le raisonnement consistant à inférer la possession d'une connaissance à partir des états cérébraux d'un individu relève d'une erreur méréologique, dans la mesure où les états neurologiques ne remplissent pas les critères d'attribution de la connaissance : la proposition selon laquelle « j'ai connaissance d'une certaine information » n'est pas satisfaite si l'on tient compte uniquement de mes états cérébraux. Selon eux, le raisonnement qui sous-tend le principe de fonctionnement du BFT n'a pas de sens car « la connaissance n'est pas localisée dans le cerveau ». Il repose sur une conception erronée de la connaissance, laquelle mobilise un ensemble de capacités cognitives (notamment la capacité à reconnaître une proposition comme vraie, la capacité à former des croyances, etc.), qui doivent pouvoir s'exprimer au travers d'un comportement. Les critères d'attribution de la connaissance se rapportent en effet à des comportements (ce que l'on fait ou ce que l'on dit) et non à des états neurologiques : les états internes de l'individu ne suffisent pas à lui attribuer une quelconque connaissance car ils ne rendent pas compte de la manière dont elle s'exprime sur le plan comportemental. Précisons d'ailleurs que les états neurologiques n'apportent aucune information sur les relations que la notion de connaissance entretient avec d'autres concepts tels que la croyance, le doute, la certitude, etc. Comme le soulignent Pardo et Patterson (2013), quand bien même les données constitueraient un « indicateur neurophysiologique » de la possession d'une connaissance (dans sa définition la plus générale), il resterait encore à déterminer si elles nous disent quoi que ce soit sur la connaissance au sens où elle est mobilisée dans la *mens rea*.

On peut ajouter que le raisonnement consistant à inférer l'exercice d'un état psychologique à partir de l'activité cérébrale correspond à une inférence inverse, qui est très largement discutée dans le champ des neurosciences et de la philosophie des neurosciences (e.g. Poldrack 2006, 2011 ; Machery 2014). Rappelons que l'inférence inverse est un raisonnement qui consiste à inférer l'exercice d'un comportement ou d'un

processus psychologique à partir d'un état cérébral donné. Un certain nombre de techniques neuroscientifiques développées dans le champ du neurodroit sont fondées sur ce raisonnement, notamment les neuro-détecteurs de mensonge qui reposent globalement sur le même paradigme que le BFT : si une région R est activée pendant l'interrogatoire du sujet, et si l'on sait que cette même région est impliquée dans l'élaboration d'un mensonge, alors cela signifie que le sujet est en train de mentir (e.g. Spence et al. 2005). Or, nous avons déjà eu l'occasion de montrer (section 3.6) que l'inférence inverse était limitée par la spécificité fonctionnelle des régions considérées : une telle inférence est valide si et seulement si la région R est activée uniquement lorsque la fonction F est engagée. On peut donc douter de la validité de ce raisonnement en ce qui concerne l'attribution d'une connaissance ou d'un mensonge, dans la mesure où l'on s'intéresse à des fonctions cognitives complexes et fortement intégrées, qui engagent de multiples réseaux d'interactions.

Ce dernier point nous donne l'occasion de questionner la pertinence d'une analyse centrée sur les problèmes conceptuels ou catégoriels posés par le projet du neurodroit, et qui fait fi des conditions expérimentales dans lesquelles les recherches sont conduites. Sans mettre en doute le bien-fondé de l'analyse proposée par Pardo et Patterson (2013), on peut s'interroger sur la force de l'argument de l'erreur méréologique pour rejeter définitivement l'ambition de redéfinir les catégories normatives du droit pénal à la lumière des données neuroscientifiques. Il semble en effet que la critique du projet normatif du neurodroit et de la neurocriminologie ne peut faire l'impasse sur les questions proprement empiriques concernant les modalités d'opérationnalisation des comportements ou des processus mentaux : la question de savoir si les données neuroscientifiques peuvent résoudre des problèmes normatifs n'est pas seulement conceptuelle ou catégorielle ; elle est aussi et surtout empirique.

5.4.3. Les raisons de dépasser le rejet de principe

L'argument de l'erreur méréologique nous a permis d'apporter quelques éléments de réponse pour invalider le modèle internaliste défendu par la neurocriminologie et contester l'idée que les états internes produisent à eux seuls les phénomènes psychologiques. On peut pourtant se demander si cet argument suffit à rejeter le projet

d'intégration des neurosciences dans le champ criminologique et judiciaire. Depuis la publication des *Philosophical foundations of neuroscience* par Bennett et Hacker (2003), les neuroscientifiques sont nombreux à contester la pertinence de l'argument de l'erreur méréologique, le neurodiscours auquel ils s'adonnent parfois n'étant selon eux qu'une simple « analogie métaphorique et inoffensive » (Levy 2014). On peut supposer que les chercheurs en neurodroit ou en neurocriminologie partagent cette position, et que l'erreur de catégorie n'est selon eux qu'un faux problème issu d'une philosophie « en fauteuil » qui refuse de s'engager sur le terrain empirique.

Plus encore, on peut regretter que la radicalité de cette thèse masque les enjeux réels posés par le neurodroit, précisément parce qu'elle n'est pas suffisamment informée sur le plan empirique. La question de la connaissance (et des divers usages qui sont envisagés dans le champ) est largement traitée dans l'ouvrage de Pardo et Patterson (2013). Elle éclipse pourtant la question plus fondamentale encore de l'apport des neurosciences pour informer les conditions volitionnelles et cognitives de l'*Insanity Defense*, c'est-à-dire déterminer si le prévenu possède les capacités requises pour être tenu pénalement responsable, à savoir : la capacité à contrôler son comportement et la capacité à apprécier la valeur morale de son acte. Dans les deux cas, nous concédons qu'il est essentiel de déterminer s'il est possible d'inférer la possession d'une connaissance à partir des états cérébraux de l'individu. Cette question est pertinente du point de vue de la recherche en neurodroit, dans la mesure où la connaissance des règles morales ou juridiques est un prérequis à la définition juridique de la responsabilité : l'évaluation de la responsabilité pénale nécessite de déterminer si le prévenu est capable de former un jugement moral (et donc s'il a une connaissance morale), et s'il est capable de contrôler son comportement *conformément* à cette connaissance morale.

Pardo et Patterson estiment que les neurosciences n'apportent aucune information vis-à-vis des concepts psychologiques mobilisés dans le champ pénal, car le principe méréologique des neurosciences interdit toute identification du sujet (de ses comportements ou de ses états psychologiques) à son cerveau ou à ses régions cérébrales : l'erreur méréologique (et les confusions conceptuelles qui en résultent) suffisent selon eux à invalider le projet du neurodroit. S'il est vrai que ces confusions soulèvent de sérieux doutes quant à sa validité épistémologique, nous voudrions ici nuancer la thèse selon laquelle les neurosciences n'apportent aucune information pertinente vis-à-vis du droit pénal. Pour le dire simplement, nous souscrivons à l'argument, mais nous estimons

qu'il est insuffisant pour tirer une conclusion définitive sur le projet d'intégrer les neurosciences dans le champ pénal. Deux raisons peuvent être invoquées.

La première est que la thèse de Pardo et Patterson se trompe de cible. Elle s'inscrit dans la continuité des critiques formulées à l'encontre des neurosciences et relève en ce sens davantage d'un rejet de principe de l'apport des neurosciences vis-à-vis des concepts mentaux, que d'une analyse approfondie des arguments empiriques avancés par les chercheurs en neurodroit. En outre, si le *projet* du neurodroit est bien marqué par des confusions conceptuelles, celles-ci ne sont pas suffisantes pour conclure que *les neurosciences* n'apportent aucune information pertinente vis-à-vis du droit pénal. Le problème ne concerne pas les neurosciences en tant que telles, mais la solidité du cadre théorique et méthodologique dans lequel les recherches en neurodroit sont conduites. D'ailleurs, et sans nécessairement chercher à sauver le projet du neurodroit, on peut noter que leur thèse sous-estime la complexité des études menées dans le champ, lesquelles ne peuvent se réduire à la recherche d'une localisation cérébrale des concepts juridiques issus de la psychologie populaire (l'intention, le contrôle de soi, la connaissance morale, etc.).[203] Globalement, le principal enjeu posé par le neurodroit est de déterminer dans quelle mesure les données neuroscientifiques peuvent informer les capacités requises pour être tenu pénalement responsable de son acte. Le problème est donc bel et bien épistémologique : il ne s'agit pas tant de savoir si les propriétés cérébrales peuvent être identifiées à des propriétés psychologiques, mais de déterminer si les données neurobiologiques peuvent nous dire quoi que ce soit sur les fonctions psychologiques des individus.

Cette remarque nous amène à la seconde raison pour laquelle l'argument de l'erreur méréologique ne suffit pas à invalider le projet du neurodroit : le problème est davantage empirique que conceptuel ou catégoriel. Il y a bien une erreur de catégorie dans le raisonnement consistant à inférer la possession d'un état psychologique à partir des états cérébraux de l'individu, puisque l'on confond les propriétés de la partie (les propriétés cérébrales) avec les propriétés du tout (les propriétés psychologiques ou comportementales). Mais le problème n'est pas tant l'inadéquation entre le langage

[203] À bien des égards, le projet du neurodroit qui se dessine depuis le début des années 2000 est très différent du projet neurocriminologique : non seulement du point de vue des finalités recherchées, mais également sur le plan des choix méthodologiques opérés par les chercheurs. Le neurodroit s'inscrit dans un cadre méthodologique nettement plus rigoureux et certaines recherches livrent une réflexion très fine sur la nécessité de proposer une clarification des concepts mobilisés dans le champ.

neuroscientifique et le langage issu de la psychologie populaire, que la confusion entretenue dans l'interprétation des données neurobiologiques, entre les états psychologiques et leurs conditions d'exercice. En cherchant à inférer la possession d'une connaissance (ou d'une intention) à partir des états cérébraux de l'individu, on assimile les capacités psychologiques des individus avec les conditions de possession de ces capacités. Cette inférence interprétative prend donc le risque de réactualiser le principe organologique de Gall qui, rappelons-le, reposait sur une instanciation directe des facultés dans le cerveau des individus.

Il semble donc que l'ambition d'utiliser les données neuroscientifiques pour informer les questions fondamentales du droit pénal recoupe certains enjeux que nous avons rencontrés dans le champ de la neurocriminologie. Les obstacles auxquels sont confrontés ces deux projets s'articulent finalement autour de la confiance que l'on peut raisonnablement placer à l'égard des outils et des données neuroscientifiques. Cette question ne peut se soustraire à une analyse approfondie des arguments empiriques qui sous-tendent le projet normatif. En clair, si l'on veut isoler les enjeux épistémologiques de ces usages (et éventuellement les résoudre), il convient de dépasser toute hostilité de principe pour interroger le cadre théorique et méthodologique qui régit les études publiées dans les deux champs. Certaines questions conceptuelles et empiriques doivent être discutées, notamment en ce qui concerne les critères de définition des comportements que l'on cherche à opérationnaliser et les différentes manières d'interpréter les données. Ainsi, quelle que soit la nature des finalités recherchées − utiliser les neurosciences pour informer la question de la responsabilité (dans le cas du neurodroit) ou la question de la dangerosité (dans le cas de la neurocriminologie) − l'ambition de fonder des catégories normatives sur des critères neurobiologiques peut être analysée par le prisme de trois questions : 1/ la question de la validité conceptuelle du modèle (le concept que l'on cherche à mesurer est-il suffisamment clarifié ?) ; 2/ la question de la validité de la mesure (le modèle expérimental représente-t-il le concept sous tous ses aspects ?) ; et 3/ la question de la portée normative des résultats (les données empiriques permettent-elles de tirer des conclusions normatives ?).

Ces trois niveaux d'analyse posent d'une part, la question de la conceptualisation et de l'opérationnalisation du comportement que l'on cherche à mesurer, et d'autre part, la question de la portée explicative et normative des résultats. Si l'on reprend l'exemple que nous avons mentionné concernant la possibilité d'inférer la possession d'un état

psychologique à partir des états cérébraux de l'individu, on peut dès lors multiplier les arguments pour invalider la portée normative des résultats. Le problème est d'abord conceptuel, car les chercheurs font l'impasse sur les questions relatives à la définition de leur objet expérimental : ils échouent à identifier les critères de définition du concept qu'ils cherchent à opérationnaliser (le contrôle de soi, la connaissance, l'intention, etc.). Les conclusions empiriques s'en trouvent affaiblies car les confusions conceptuelles peuvent introduire un *hiatus* entre l'objet qui est opérationnalisé et celui qui est réellement mesuré. En conséquence, et sans compter les erreurs inférentielles qui peuvent émerger de l'interprétation des données, la possibilité de tirer des conclusions normatives sur la seule base de ces données paraît doublement injustifiée : non seulement en raison des confusions conceptuelles et empiriques qui sous-tendent la construction du modèle expérimental, mais aussi parce que les modalités d'opérationnalisation excluent les aspects normatifs du comportement que l'on cherche à mesurer. Il ne s'agit pas seulement de dire que le concept n'est pas suffisamment clarifié, ou que le modèle expérimental ne permet pas de représenter le concept tel qu'il est mobilisé dans le droit pénal. Il s'agit aussi de s'interroger sur la valeur de preuve des outils et des données neurobiologiques, compte tenu du fait que les comportements opérationnalisés s'inscrivent dans des contextes sociaux et normatifs qui ne peuvent être réduits à une seule description neurobiologique.

Le champ de la neurocriminologie paraît exemplaire des critiques que l'on pourrait formuler à l'encontre des diverses neurodisciplines, non seulement parce qu'il s'inscrit dans la même tradition éliminativiste, mais aussi parce qu'il souligne l'impérieuse nécessité de résoudre les problèmes conceptuels et empiriques inhérents à un usage prescriptif des données neurobiologiques. Face au développement massif des domaines d'application des neurosciences, la réflexion critique a sans doute plus de chances d'être entendue si elle cesse d'aborder la question de leur pouvoir explicatif comme un problème essentiellement normatif, qui oppose aux investigations empiriques un anti-naturalisme de principe. Le problème posé par la neurocriminologie et le neurodroit ne tient pas seulement à la naturalisation du crime ou à l'inaccessibilité empirique des catégories juridiques, mais aux présupposés théoriques et méthodologiques qui structurent l'ensemble des études menées depuis l'émergence de ce programme de recherche, au début des années 1940. Il ne s'agit pas ici de présumer que ces enjeux pourraient être résolus sur le plan empirique par une plus grande rigueur méthodologique

ou par un examen plus attentif des données disponibles ; il s'agit plutôt de montrer que la question de l'apport épistémologique de ces neurodisciplines ne peut se soustraire à une analyse approfondie des arguments qui sous-tendent leur projet normatif, mais aussi des conditions d'opérationnalisation des processus psychologiques et des inférences interprétatives qui en découlent.

Nous avons eu l'occasion de montrer que la thèse neuroessentialiste concentrait une grande partie des difficultés posées par l'usage des techniques de neuroimagerie comme un instrument de connaissance des comportements individuels. La question de savoir s'il s'agit d'un présupposé inhérent aux neurosciences cognitives reste ouverte ; cependant, nous voudrions insister sur le piège du biais de confirmation auxquelles sont confrontées les neurodisciplines comme la neurocriminologie ou le neurodroit : le neuroessentialisme n'est pas problématique s'il s'en tient à l'affirmation selon laquelle le cerveau *rend possible* l'exercice des processus psychologiques, mais il peut introduire une grande confusion lorsqu'il s'impose comme la seule grille de lecture des données empiriques. Précisons d'ailleurs que ce biais constitue à la fois la source de l'impasse épistémologique dans laquelle est engagée la neurocriminologie, mais également la justification principale de son projet explicatif, prédictif et curatif : si le criminel est *d'essence cérébrale*, alors il parait légitime de déléguer entièrement la gestion de la criminalité aux neurosciences cognitives. On peut donc craindre que la tentation hégémonique qui anime le projet neurocriminologique dispense les chercheurs de toute réflexion épistémologique sur leurs inférences interprétatives ou leurs conclusions empiriques.

Nous avons exposé les raisons pour lesquelles il est épistémologiquement erroné de considérer que le criminel est d'essence cérébrale. Si le cerveau joue un rôle incontestable dans l'exercice des processus mentaux, cela ne signifie pas pour autant que les états cérébraux renferment toutes les manifestations comportementales des individus, leurs raisons d'agir, leur subjectivité ou leur identité personnelle. Non seulement il paraît peu judicieux de parler d'un « cerveau criminel » comme d'une catégorie naturelle objectivant l'existence d'une dangerosité immanente au sujet, mais cette représentation neuroessentialiste se révèle extrêmement préoccupante vis-à-vis du discours neuroscientifique dont les neurocriminologues se réclament. Nous avons d'ailleurs eu l'occasion de montrer que ce « cerveau criminel » (défaillant, déficient, irrégulier...) correspondait moins à une réalité empirique qu'à une construction normative visant à promouvoir les finalités poursuivies par les neurocriminologues : l'essentiel du projet

neurocriminologique est bien de réduire la figure de l'individu dangereux à un « cerveau à corriger », la recherche de l'explication étant supplantée par une logique neuro-interventionniste qui n'apporte elle-même aucune information sur le statut épistémologique de la relation entre les anomalies cérébrales et la manifestation du comportement criminel.

Conclusion

En renonçant à une hostilité de principe à l'égard de l'apport potentiel des neurosciences dans la compréhension du phénomène criminel, nous avons abordé la neurocriminologie comme un objet épistémologique, c'est-à-dire comme un domaine de recherche *scientifique* susceptible d'informer l'évaluation de la responsabilité pénale ou de la dangerosité d'un individu. Ce parti pris méthodologique nous a permis d'analyser la rationalité interne de ce projet, en examinant chacun des arguments qui pourraient justifier la prétention d'utiliser ces données à des fins prédictives, explicatives ou curatives. Car si le discours neurocriminologique cultive une certaine ambiguïté quant à la nature réelle des finalités poursuivies, l'analyse de ses présupposés théoriques et méthodologiques révèle bien l'ambition de redéfinir en termes neurobiologiques l'ensemble des questions relatives à l'explication, la prévention et la prise en charge de la criminalité.

On peut ainsi retracer les principales étapes de ce projet normatif, tel qu'il est promu par la neurocriminologie depuis son émergence dans les années 1940 : 1/ une définition neurophysiologique du crime dans lequel le critère de dangerosité est indexé sur le degré d'anomalies EEG ; 2/ l'élaboration d'un modèle éliminativiste qui exclut toute interprétation alternative à l'explication neurobiologique du comportement criminel ; et 3/ la construction neuroessentialiste d'un « cerveau à corriger », ultime ressort du projet pour justifier l'apport des outils neuroscientifiques dans la lutte contre la problématique du crime violent et de la récidive. Si ces trois projets n'affichent pas explicitement leur continuité, nous avons montré qu'ils contribuaient activement à l'agenda de ce programme de recherche : faire du crime un problème *exclusivement* neurobiologique dont la résolution ne peut s'opérer qu'à l'intérieur du champ de la neurocriminologie. On ne peut en effet pas manquer de relever la grande cohérence du projet au regard des finalités recherchées, la « vision neurocriminologique » ayant orienté chacun des choix méthodologiques opérés dans la conceptualisation et l'opérationnalisation du comportement criminel.

La confusion épistémologique entretenue par la neurocriminologie découle de la revendication d'une démarche scientifique inspirée de la méthodologie classique des neurosciences cognitives. La posture neuroessentialiste trahit pourtant tous les efforts

déployés pour justifier la légitimité scientifique de ce projet, anticiper les accusations de cérébro-centrisme et masquer les stigmates de son long héritage idéologique. Loin de marquer une rupture épistémologique dans l'histoire de la biologisation du phénomène criminel née au XIX^{e} siècle, la neurocriminologie ne fait que réactualiser des thèses dont la scientificité a été invalidée depuis plus d'un siècle : l'analyse de la littérature publiée depuis les années 1940 montre que cette continuité historique n'est pas seulement de l'ordre d'une filiation conceptuelle, méthodologique ou anthropologique, puisqu'elle prend la forme d'une *identité* épistémologique avec les thèses de Franz-Joseph Gall et Cesare Lombroso.

Il est d'ailleurs frappant de constater que les affinités conceptuelles et méthodologiques se partagent dans les deux moments-clés de l'histoire de la neurocriminologie : la neurocriminologie des années 1940-1980 (basée sur l'EEG) est particulièrement emblématique des présupposés qui sous-tendent la théorie du criminel-né ; tandis que la neurocriminologie contemporaine réactualise certains principes inhérents à la stratégie localisatrice de Gall, notamment le principe d'identité psychophysiologique et l'hypothèse du « déséquilibre cérébral ». La prégnance de la pensée lombrosienne se révèle dans des présupposés idéologiques tenaces en ce qui concerne l'« infériorité biologique » des criminels, la hiérarchisation des « types criminels » ou encore le problème de la monstruosité criminelle, dont la résolution ne semble être recherchée qu'en tant qu'elle permet de doubler le crime dans le cerveau du criminel.

Le présupposé selon lequel le criminel ressemble à son cerveau et à son crime est remarquable à plus d'un titre, car il introduit un biais systématique dans l'interprétation des données, conduisant par la même occasion à neutraliser tous les efforts déployés par les neurocriminologues pour s'implanter dans le champ des neurosciences cognitives. En cherchant à montrer que le criminel est « d'essence cérébrale », la neurocriminologie a construit son projet explicatif autour d'une série de pétitions de principe qui réhabilitent la dynamique de l' « instinct criminel » ou de la « force irrésistible » telle qu'elle a été formulée dans l'anthropologie criminelle du XIX^{e} siècle : les neurocriminologues laissent suggérer l'existence d'une mécanique pathologique à l'origine d'une pulsion criminelle irrépressible, d'une potentialité cérébrale largement indéfinie sur le plan fonctionnel, mais suffisamment saillante pour faire l'objet d'une intervention directe, neurobiologique et spécifique.

Finalement, la principale difficulté posée par la neurocriminologie n'est plus seulement celle d'un héritage inavouable (car essentiellement idéologique), mais le problème plus fondamental encore d'un programme de recherche dicté par une logique interventionniste, et qui ne fournit aucune explication des phénomènes psychologiques étudiés. La neurocriminologie est emblématique de cette « neuropolitique » qui contamine la recherche en neurosciences cognitives en mobilisant ses outils pour fonder son autorité scientifique, tout en déployant sa stratégie au-delà de ses moyens théoriques et méthodologiques. Face à l'impératif sécuritaire, et compte tenu de l'engouement porté aux neurosciences cognitives dans nos sociétés, cette confusion épistémologique pourrait donner une impulsion favorable à la réorientation des politiques pénales en matière de lutte contre la récidive, en encourageant une reconfiguration des problématiques liées à la prise en charge de la criminalité. Si l'on est encore loin de généraliser l'usage des techniques de neurostimulation comme un outil de réhabilitation des criminels, on peut en revanche craindre un *hiatus* entre d'un côté, les attentes sociales grandissantes à l'égard des neurosciences et de ses potentielles applications judiciaires, et de l'autre, l'état des savoirs neurocriminologiques et leurs faiblesses épistémologiques. Le postulat d'objectivité neuroscientifique et la tentation d'apporter une caution scientifique à un projet de politique pénale pourraient outrepasser ces limites, précisément parce qu'elles semblent apporter une solution concrète à des problèmes extrêmement complexes. Sans compter le risque de marginaliser les explications sociologiques de la criminalité, le déploiement de ce programme de recherche pourrait entretenir une grande confusion à l'égard du pouvoir explicatif *réel* des neurosciences cognitives.

Annexes

Annexe n° 1. Tableau récapitulatif de l'ensemble des anomalies physiques, motrices, sensorielles et psychiques recensées chez le criminel-né.

Tableau I

LISTE DES CRITÈRES PERMETTANT DE DISTINGUER LES CRIMINELS-NÉS (D'après [4], p. 450-1)

	Criminels	Épileptiques	Phénomènes ataviques	Phénomènes d'arrêt de développement	Phénomènes morbides	Phénomènes atypiques		Criminels	Épileptiques	Phénomènes ataviques	Phénomènes d'arrêt de développement	Phénomènes morbides	Phénomènes atypiques
Crâne							Absence de barbe	+	+	+		+	
Exagération de volume	+	+		+		+	Teint olivâtre	+	+	+			
Diminution de volume, microcéphalie	+	+	+	+	+	+	Tatouage	+		+			
Sclérose	+	+	+	+	+		Canitie et calvitie tardives	+	+	+			
Exostose	+	+			+		Cheveux noirs et crépus	+	+	+			
Asymétrie	+	+			+		***Anomalies motrices***						
Fossette occipitale médiane	+	+		+			Gaucherie et ambidextrisme	+	+	+			
Indice crânien exagéré	+	+			+		Altérations des réflexes	+	+			+	
Arcs sourciliers exagérés	+	+		+	+		Pupilles inégales	+	+			+	
Front bas, fuyant	+	+		+			Agilité exagérée	+	+	+			
Front hydrocéphalique	+	+					***Anomalies sensorielles***						
Ostéophites crâniens	+	+					Obtusité tactile	+	+	+		+	
Wormiens nombreux	+	+	+	+			Obtusité dolorifique et générale	+	+	+		+	
Suture métopique	+	+	+	+			Grande acuité visuelle	+	+	+			
Sinostoses précoces	+	+					Obtusité de l'ouïe, du goût et de l'odorat	+	+	+		+	
Suture crânienne simple	+	+	+				Gaucherie sensorielle	+	+	+		+	
Orbites obliques	+	+	+				Scotomes périphériques du ch. vis.	+	+			+	
Face							***Anomalies organiques***						
Appendice lémurien	+	+	+				Disvulnérabilité	+	+	+			
Mâchoires hypertrophiques	+	+	+				***Anomalies psychiques***						
Zygomes saillants volumineux	+	+	+				Intelligence limitée	+	+	+	+		
Diamètre biangulaire de la mâchoire	+	+	+				Superstition	+	+	+			
Oreilles à anses volumineuses	+	+	+				Obtusité affective	+	+	+		+	
Asymétrie faciale	+	+					Amour des animaux	+	+			+	
Strabisme	+	+					Obtusité du sens moral	+	+	+		+	
Physionomie virile chez la femme	+	+	+				Absence de remords	+	+	+			
Diastème dentaire	+	+	+				Impulsivité, cannibalisme, férocité	+	+	+	+	+	
Anomalie des os du nez	+	+				+	Pédérastie, onanisme, obscénité	+	+	+		+	
Anomalies dentaires	+	+	+				Croyances religieuses exagérées	+	+	+			
Développ. exagéré des os de la face	+	+	+				Vagabondage	+	+	+			
Cerveau							Précocité sexuelle exagérée, etc.	+	+	+		+	
Anomalie des circonvol. des scissures	+	+	+	+			Vanité	+	+	+			
Poids moindre	+	+	+	+			Esprit d'association	+	+				
Hypertrophie du cervelet	+	+	+				Simulation	+	+			+	
Altérations histologiques de l'écorce	+	+	+	+			Paresse, inertie	+	+	+		+	
Traces de méningite	+	+					Imprévoyance	+	+	+			
Corps							Lâcheté	+	+	+			
Asymétrie thoracique	+	+					Passion du jeu	+	+	+			
Pied préhensile	+	+	+				Manie, paranoïe, délire	+	+			+	
Gaucherie (Mancinismo) anatomique	+	+	+				Vertiges	+	+			+	
Fréquence des lésions viscérales	+	+					Mémoire exagérée	+	+			+	
Hernies	+	+	+	+			***Causes***						
Simplicité des plis palmaires	+	+	+			+	Héréditaires (dans les parents,						
Thorax à entonnoir	+	+				+	alcoolisme, folie, épilepsie, parents âgés)	+	+				
Peau							Alcoolisme	+	+				
Rides anormales	+	+	+				Égale distribution géographique	+	+				

Lombroso, Cesare. 1887. *L'Homme criminel : criminel-né, fou moral, épileptique : étude anthropologique et médico-légale*. Paris, Félix Alcan.

Annexe n° 2. Tableau issu de l'étude menée par Denis Hill et Donald Watterson (1942) recensant la proportion des anomalies EEG observées chez les différentes catégories de psychopathes (n = 151) et chez le groupe contrôle (n = 52).

TABLE 3

	TOTAL	E.E.G.					
		POSITIVE			DOUBTFUL	TOTAL ABNORMAL	
		REST ONLY	HYP. ONLY	REST + HYP.			PERCENT.
Aggression	66	10	14	19	4	43	65
Inadequacy	38	3	3	6	6	12	32
Delinquency	5	0	0	0	1	0	—
Sexual perversion ..	8	1	0	0	0	1	—
Mixed	34	6	5	6	4	17	50
Controls	52	6	0	2	3	8	15

Hill, D., and D. Watterson. 1942. "Electro-Encephalographic Studies of Psychopathic Personalities." *Journal of Neurology and Psychiatry* 5 (1–2): 47–65.

Annexe n° 3. Tableau récapitulatif de l'étude menée par Daniel Silverman (1944), classant les criminels selon leurs données cliniques et les données relatives à leur dossier judiciaire.

TABLE 1.—*Offenses Committed by Four Hundred and Eleven Prisoners*

	Homosexual Psychopaths	All Psychopaths	Prisoners with "Organic" Disease of Brain: Psychotic Prisoners	Prisoners with "Organic" Disease of Brain: Normal Prisoners	All Prisoners: Number	All Prisoners: Per Cent
1. Crimes of acquisition, without violence	..	..	..	..	285	69.3
A. Thefts	..	..	..	..	179	43.5
Automobile	0	87	11	15	124	
Interstate shipment	0	6	3	1	12	
Mail and postoffice	0	16	5	5	27	
Other types	0	9	4	2	16	
B. Swindles	..	..	..	..	56	13.6
Embezzlement and mail fraud	0	2	3	5	13	
Counterfeiting	0	2	4	2	15	
Forgery	0	6	1	3	13	
Impersonation	0	7	0	5	15	
C. Illicit business	..	..	..	..	50	12.2
Liquor	0	2	12	7	34	
Narcotics	0	0	4	1	10	
White slave traffic	0	3	0	0	6	
2. Crimes involving violence, actual or implied	..	..	..	..	97	23.6
A. For money	..	..	..	..	30	7.3
Bank robbery	0	1	5	0	7	
Other robbery	0	3	2	1	6	
Kidnapping and extortion	0	6	10	0	17	
B. Not for money	..	..	..	..	17	4.1
Assault	0	2	2	1	7	
Homicide	0	0	6	0	7	
Threatening letters	0	1	0	1	3	
C. Sexual	..	..	..	..	50	12.2
Sodomy	29	4	1	0	34	
Rape	0	4	0	0	4	
Violation of Mann Act	0	1	1	7	9	
Obscene letters	0	1	2	0	3	
3. Miscellaneous	..	..	..	..	29	7.1
Selective service violation	4	10	1	5	20	
Other offenses	1	1	4	1	9	

Silverman, Daniel. 1944. "The Electroencephalogram of Criminals: Analysis of Four Hundred and Eleven Cases." *Archives of Neurology & Psychiatry* 52 (1): 38–42.

Annexe n°4. Graphe issu de l'étude menée par Hill et Pond (1952), qui illustre les résultats de Stafford-Clark et Taylor (1949) en y ajoutant trente nouveaux cas étudiés par les psychiatres F.H. Taylor et J.C. Matheson. L'ensemble des résultats montre une proportion d'anomalies EEG similaire entre les criminels ayant commis un crime sans motif apparent (mais reconnus pénalement responsables) et les criminels ayant bénéficié d'une irresponsabilité pour cause psychiatrique.

Fig. 1 shows the number of cases in each of these groups and the incidence of abnormal E.E.G.s among them.

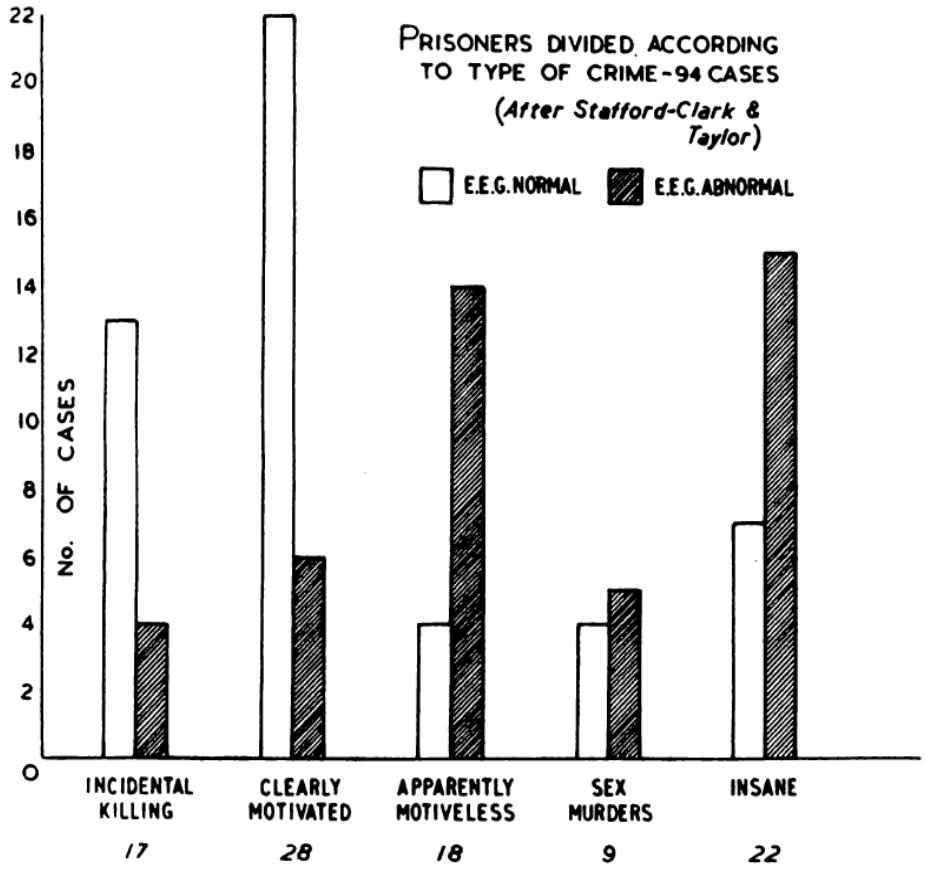


FIG. 1.

Hill, D., and D. A. Pond. 1952. "Reflections on 100 Capital Cases Submitted to Electroencephalography." *The Journal of Mental Science* 98 (410): 23–43.

Annexes n°5 et 6. Graphe représentant la relation entre la proportion d'anomalies EEG et l'âge de la première condamnation d'une part, et la sévérité du crime d'autre part (Levy and Kennard 1953). Les résultats montrent que les anomalies sont nettement plus fréquentes chez les sujets ayant commis leur premier crime avant l'âge de seize ans (Annexe n°5) et chez les criminels récidivistes et présentant une personnalité instable (Annexe n°6).

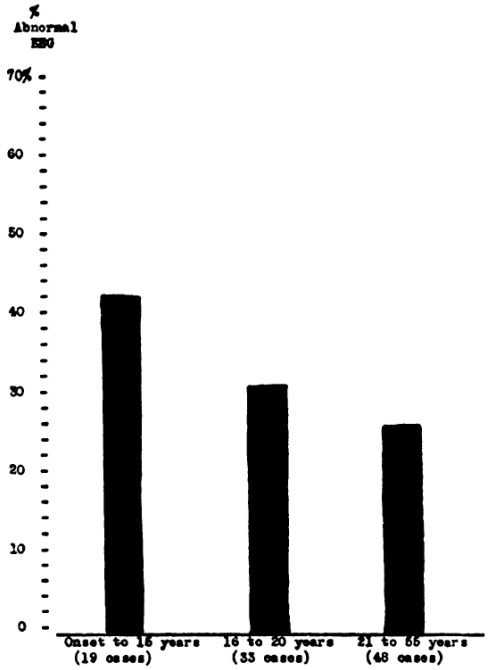


FIG. 3.—Relation of age at first crime to EEG.

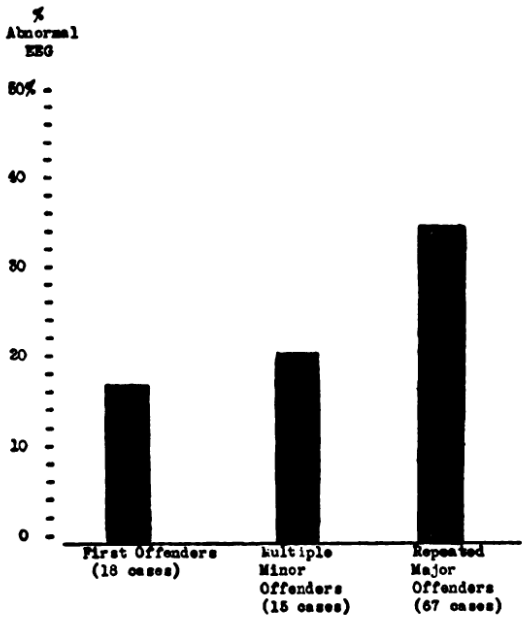


FIG. 4.—Relation of severity of crime to EEG

Levy, Sol, and Margaret Kennard. 1953. "A Study of the Electroencephalogram as Related to Personality Structure in a Group of Inmates of a State Penitentiary." *American Journal of Psychiatry* 109 (11): 832–39.

Annexe n°7, table I. Tableau comparant la proportion d'anomalies EEG recensées chez les criminels épileptiques et les criminels non-épileptiques, selon le type de condamnation criminelle (Gunn and Bonn 1971).

TABLE I

Current offences

	Epileptics		Controls	
	No.	%	No.	%
Property	105	*66*	108	*60*
Violent*	23	*15*	40	*22*
Sex	14	*9*	12	*7*
Other	16	*11*	20	*12*
Total	158	*100*	180	*100*

$\chi^2 = 3{\cdot}81$ (3d.f.) N.S.

*This includes 6 epileptics and 4 controls convicted of homicide (murder or manslaughter).

Gunn, J., and J. Bonn. 1971. "Criminality and Violence in Epileptic Prisoners." *The British Journal of Psychiatry: The Journal of Mental Science* 118 (544): 337–43.

Annexe n°7, tables IV et VI. Tableau comparant la proportion d'anomalies EEG recensées chez les criminels épileptiques et les criminels non-épileptiques, selon le degré de violence et d'impulsivité associé au crime (Gunn and Bonn 1971).

TABLE IV
Degrees of violence compared (epileptics v. controls)

	Epileptics		Controls	
Violence 0	123	(82%)	124	(77%)
Violence 1	3		5	
Violence 2	10		19	
Violence 3	15		14	
Total	151		162	

$\chi^2 = 2{\cdot}93$ (2d.f.) N.S.
(amalgamating violence 1 and violence 2)
(7 eps and 18 controls were 'violence unrateable')

TABLE VI
Epileptics and controls compared for impulsiveness

	Epileptics		Controls	
	N	%	N	%
Impulsiveness 0	97	70	104	73
Impulsiveness 1	14	10	31	22
Impulsiveness 2	28	20	7	5
Total	139		142	

$\chi^2 = 19{\cdot}24$ (2d.f.) $p < 0{\cdot}0005$
(19 epileptics and 38 controls were 'impulsiveness unrateable'.)

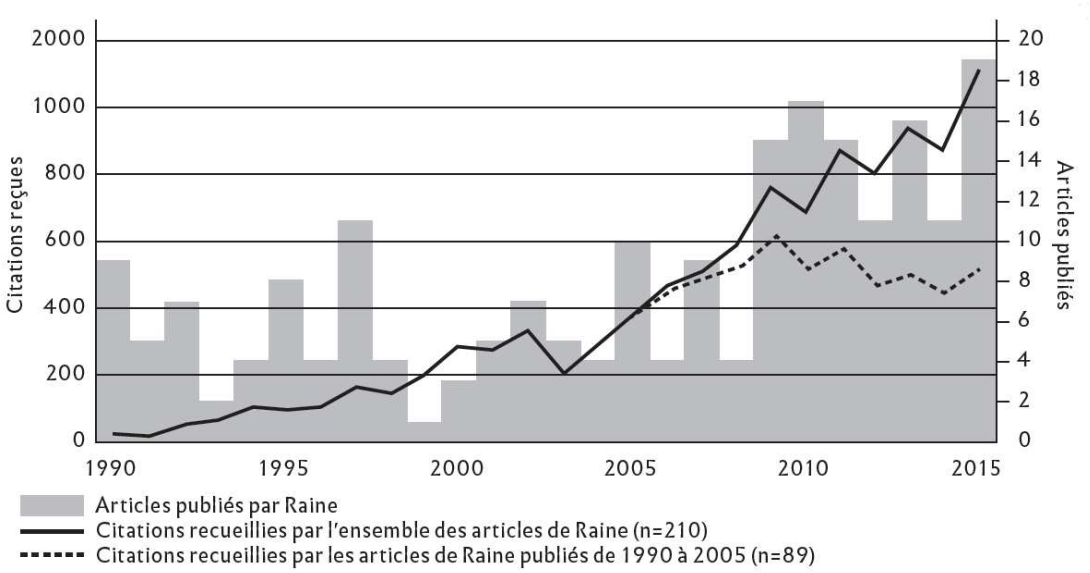
Annexe n°7, table II. Tableau associant les différentes formes d'épilepsie des criminels (idiopathique, temporale, etc.) avec le type d'infraction (violente ou non) commise (Gunn and Bonn 1971).

TABLE II
Property and violent offenders by epileptic diagnosis

	Property		Violence		Total
	N	%	N	%	N
Idiopathic ..	15	79	4	21	19
Temporal ..	34	83	7	17	41
Other focal ..	22	82	5	18	27
Unratable ..	34	83	7	17	41
Total	105		23		128

$\chi^2 = 0{\cdot}17$ (3d.f.) N.S.

Annexe n°8. Graphe représentant l'évolution du nombre de citations des articles publiés par Adrian Raine entre 1990 et 2015.



Larregue, Julien. 2019. "La nouvelle orange mécanique : la contribution des bio-criminologues à la médicalisation du contrôle social." *Zilsel* N° 5 (1): 56–87.

Annexe n°9. Tableau récapitulatif d'une série d'études utilisant la tomodensitométrie (CT-scan) sur des individus condamnés pour diverses infractions à caractère sexuel.

TABLE 1 Overview of Key Findings from CT Studies of Offender Groups[a]

Reference	Subjects	Technique	Key finding
Hucker *et al.* (1986)	39 pedophiles 14 property offenders	CT	Left and bilateral temporal horn dilation; dilation of anterior horn of lateral ventricles in pedophiles
Langevin *et al.* (1987)	18 murderers 21 assaulters 16 property offenders	CT	No significant differences
Herzbert & Fenwick (1988)	14 aggressive TLE[a] 7 nonaggressive TLE	CT	No significant differences
Hucker *et al.* (1988)	22 sadistic sex assaulters 21 nonsadistic sex assaulters 36 property offenders	CT	Right temporal horn abnormalities in sadists
Langevin *et al.* (1988)	91 incest offenders 36 property offenders	CT	No significant differences for main group comparisons, but violence within incest group associated with temporal lobe abnormalities
Langevin *et al.* (1989a)	84 pedophiles 32 property offenders	CT	No significant differences
Langevin *et al.* (1989b)	15 exhibitionists 36 property offenders	CT	No significant differences
Wright *et al.* (1990)	18 pedophiles 12 incest offenders 34 rapists 12 controls	CT	Smaller left frontal and temporal areas in sex offenders

[a]TLE, temporal lobe epilepsy.

Raine, A., M. S. Buchsbaum, J. Stanley, S. Lottenberg, L. Abel, and J. Stoddard. 1994. "Selective Reductions in Prefrontal Glucose Metabolism in Murderers." *Biological Psychiatry* 36 (6): 365–73.

Bibliographie

Adams, Fred, and Ken Aizawa. 2001. “The Bounds of Cognition.” *Philosophical Psychology* 14 (1): 43–64.

Aharoni, Eyal, Chadd Funk, Walter Sinnott-Armstrong, and Michael Gazzaniga. 2008. “Can Neurological Evidence Help Courts Assess Criminal Responsibility? Lessons from Law and Neuroscience.” *Annals of the New York Academy of Sciences* 1124 (March): 145–60.

Aharoni, Eyal, Gina M. Vincent, Carla L. Harenski, Vince D. Calhoun, Walter Sinnott-Armstrong, Michael S. Gazzaniga, and Kent A. Kiehl. 2013. “Neuroprediction of Future Rearrest.” *Proceedings of the National Academy of Sciences* 110 (15): 6223–28.

Aigner, Martin, Eher Reinhard, S. Fruehwald, Patrick Frottier, K. Gutierrez-Lobos, and S. Margretta Dwyer. 2000. “Brain Abnormalities and Violent Behavior.” *Journal of Psychology & Human Sexuality* 11 (3): 57–64.

Alschuler, Albert. 2003. “The Changing Purposes of Criminal Punishment: A Retrospective on the Past Century and Some Thoughts about the Next.” *University of Chicago Law Review* 70 (1).

Angus, Douglas J., Dennis J. L. G. Schutter, David Terburg, Jack van Honk, and Eddie Harmon-Jones. 2016. “A Review of Social Neuroscience Research on Anger and Aggression.” In *Social Neuroscience: Biological Approaches to Social Psychology*, 223–46. Frontiers of Social Psychology. New York, NY, US: Routledge/Taylor & Francis Group.

Arbuckle, Nathan L., and Matthew S. Shane. 2017. “Up-Regulation of Neural Indicators of Empathic Concern in an Offender Population.” *Social Neuroscience* 12 (4): 386–90.

Aspinwall, Lisa G., Teneille R. Brown, and James Tabery. 2012. “The Double-Edged Sword: Does Biomechanism Increase or Decrease Judges’ Sentencing of Psychopaths?” *Science* 337 (6096): 846–49.

Baker, D. A., Jillian M. Ware, N. J. Schweitzer, and Evan F. Risko. 2017. “Making Sense of Research on the Neuroimage Bias.” *Public Understanding of Science* 26 (2): 251–58.

Banks, William P., and Susan Pockett. 2007. “Benjamin Libet’s Work on the Neuroscience of Free Will.” In *The Blackwell Companion to Consciousness*, edited by Max Velmans and Susan Schneider, 657–670. Blackwell.

Barn, Gulzaar. 2019. “Can Medical Interventions Serve as ‘Criminal Rehabilitation’?” *Neuroethics* 12 (1): 85–96.

Bartholomew, Allen A., Kerry L. Milte, and Frank Galbally. 1978. “Epileptic Homicide.” *The British Journal of Psychiatry* 133 (6): 564–65.

Baskin-Sommers, Arielle R., Craig S. Neumann, Lora M. Cope, and Kent A. Kiehl. 2016. “Latent-Variable Modeling of Brain Gray-Matter Volume and Psychopathy in Incarcerated Offenders.” *Journal of Abnormal Psychology* 125 (6): 811–17.

Bassarath, L. 2001. “Neuroimaging Studies of Antisocial Behaviour.” *Canadian Journal of Psychiatry.* 46 (8): 728–32.

Battro, Antonio M., Kurt W. Fischer, and Pierre J. Léna. 2008. *The Educated Brain: Essays in Neuroeducation*. Cambridge University Press.

Batts, Shelley. 2009. “Brain Lesions and Their Implications in Criminal Responsibility.” *Behavioral Sciences & the Law* 27 (2): 261–72.

Bayne, Tim. 2011. “Libet and the Case for Free Will Scepticism.” In *Free Will and Modern Science*, edited by Richard Swinburne. Oxford: Oxford University Press.

Bear, D.M. 1989. “Hierarchical Neural Regulation of Aggression: Some Predictable Patterns of Violence.” In *D.A. Britzer & M. Crowner (Eds.), Current Approaches to the Prediction of Violence*, 85–100. Washington, D.C.: American Psychiatric Press.

Bechtel, William. 2002. “Decomposing the Mind-Brain: A Long-Term Pursuit.” *Brain and Mind* 3 (2): 229–42.

———. 2008. *Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience*. Taylor & Francis.

Bechtel, William, and Jennifer Mundale. 1999. “Multiple Realizability Revisited: Linking Cognitive and Neural States.” *Philosophy of Science* 66 (2): 175–207.

Bechtel, William P. 2002. “Decomposing the Brain: A Long Term Pursuit.” *Brain and Mind* 3 (1): 229–242.

Bechtel, William, and Robert C. Richardson. 1993. *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. Princeton. Princeton University Press.

Bell, Emily, Ghislaine Mathieu, and Eric Racine. 2009. “Preparing the Ethical Future of Deep Brain Stimulation.” *Surgical Neurology* 72 (6): 577–86.

Bellucci, Gabriele, Sergey Chernyak, Morris Hoffman, Gopikrishna Deshpande, Olga Dal Monte, Kristine M. Knutson, Jordan Grafman, and Frank Krueger. 2017. “Effective Connectivity of Brain Regions Underlying Third-Party Punishment: Functional MRI and Granger Causality Evidence.” *Social Neuroscience* 12 (2): 124–34.

Bennett, Christopher. 2018. “Intrusive Intervention and Opacity Respect.” In *Treatment for Crime: Philosophical Essays on Neurointerventions in Criminal Justice*. Oxford University Press.

Bennett, M. R., and P. M. S. Hacker. 2003. *Philosophical Foundations of Neuroscience*. Wiley.

Berlin, Leonard. 2014. “Neuroimaging, Expert Witnesses, and Ethics: Convergence and Conflict in the Courtroom.” *AJOB Neuroscience* 5 (2): 3–8.

Bertsch, Katja, Michel Grothe, Kristin Prehn, Knut Vohs, Christoph Berger, Karlheinz Hauenstein, Peter Keiper, Gregor Domes, Stefan Teipel, and Sabine C. Herpertz. 2013. “Brain Volumes Differ between Diagnostic Groups of Violent Criminal Offenders.” *European Archives of Psychiatry and Clinical Neuroscience* 263 (7): 593–606.

Beshai, James A. 1971. “Behavioral Correlates of the Eeg in Delinquents.” *The Journal of Psychology* 79 (1): 141–46.

Bigenwald, Ariane, and Valerian Chambon. 2019. “Criminal Responsibility and Neuroscience: No Revolution Yet.” *Frontiers in Psychology* 10 (June).

Birbaumer, Niels, Ralf Veit, Martin Lotze, Michael Erb, Christiane Hermann, Wolfgang Grodd, and Herta Flor. 2005. “Deficient Fear Conditioning in Psychopathy: A Functional Magnetic Resonance Imaging Study.” *Archives of General Psychiatry* 62 (7): 799–805.

Birks, David, and Alena Buyx. 2018. “Punishing Intentions and Neurointerventions.” *AJOB Neuroscience* 9 (3): 133–43.

Birks, David, and Thomas Douglas. 2018. *Treatment for Crime: Philosophical Essays on Neurointerventions in Criminal Justice*. Oxford University Press.

Blackburn, R. 1971. “Personality Types among Abnormal Homicides.” *The British Journal of Criminology* 11 (1): 14–31.

Blair, James, Derek Mitchell, and Karina Blair. 2005. *The Psychopath: Emotion and the Brain*. Malden: Blackwell Publishing.

Blair, R. J. 2001. “Neurocognitive Models of Aggression, the Antisocial Personality Disorders, and Psychopathy.” *Journal of Neurology, Neurosurgery, and Psychiatry* 71 (6): 727–31.

Blair, R. James R. 2003. “Neurobiological Basis of Psychopathy.” *The British Journal of Psychiatry* 182 (1): 5–7.

Blake, P. Y., J. H. Pincus, and C. Buckner. 1995. “Neurologic Abnormalities in Murderers.” *Neurology* 45 (9): 1641–47.
Boccardi, Marina, Rossana Ganzola, Roberta Rossi, Francesca Sabattoli, Mikko P. Laakso, Eila Repo-Tiihonen, Olli Vaurio, et al. 2010. “Abnormal Hippocampal Shape in Offenders with Psychopathy.” *Human Brain Mapping* 31 (3): 438–47.
Bomann-Larsen, Lene. 2013. “Voluntary Rehabilitation? On Neurotechnological Behavioural Treatment, Valid Consent and (In)appropriate Offers.” *Neuroethics* 6 (1): 65–77.
Bonkalo, A. 1967. “Electroencephalography in Criminology.” *Canadian Psychiatric Association Journal* 12 (3): 281–86.
Bonnie, Richard J., ed. 1997. *Criminal Law*. University Casebook Series. Westbury, N.Y: Foundation Press.
———. 2002. “Responsibility for Addiction.” *The Journal of the American Academy of Psychiatry and the Law* 30 (3): 405–13.
Boraud, Thomas, Arthur Leblois, and Nicolas P. Rougier. 2018. “A Natural History of Skills.” *Progress in Neurobiology* 171 (December): 114–24.
Borum, R., and K. L. Appelbaum. 1996. “Epilepsy, Aggression, and Criminal Responsibility.” *Psychiatric Services (Washington, D.C.)* 47 (7): 762–63.
Bourgeois, M. L., and M. Bénézech. 2001. “Dangerosité criminologique, psychopathologie et co-morbidité psychiatrique.” *Annales Médico-Psychologiques, Revue Psychiatrique* 159 (7): 475–86.
Braunstein, Jean-François. 2010. “Une vision médicale du monde.” *Archives de Philosophie* Tome 73 (4): 631–54.
Brennan, Patricia A., Sarnoff A. Mednick, and Adrian Raine. 1997. “Biosocial Interactions and Violence.” In *Biosocial Bases of Violence*, edited by Adrian Raine, Patricia A. Brennan, David P. Farrington, and Sarnoff A. Mednick, 163–74. Nato ASI Series. Boston, MA: Springer US.
Brennan, Patricia A., and Adrian Raine. 1997. “Biosocial Bases of Antisocial Behavior: Psychophysiological, Neurological, and Cognitive Factors.” *Clinical Psychology Review*, Biophysical conceptualizations of human aggression, 17 (6): 589–604.
Brigo, Francesco. 2013. “Epilepsy and Homicide: Beware of the EEG (a Misdiagnosis Can Occur!).” *Epilepsy & Behavior* 28 (3): 521.
Brill, Norman Q., and Herta Seidemann. 1941. “The Electroencephalogram of Normal Children.” *American Journal of Psychiatry* 98 (2): 250–56.
Brill, Norman Q., Herta Seidemann, Helen Montague, and Ben H. Balser. 1942. “Electroencephalographic Studies in Delinquent Behavior Problem Children.” *American Journal of Psychiatry* 98 (4): 494–98.
Brill, Norman Q., and Edmund F. Walker. 1945. “Psychopathic Behavior with Latent Epilepsy.” *The Journal of Nervous and Mental Disease* 101 (6): 545.
Brower, M. C., and B. H. Price. 2001. “Neuropsychiatry of Frontal Lobe Dysfunction in Violent and Criminal Behaviour: A Critical Review.” *Journal of Neurology, Neurosurgery, and Psychiatry* 71 (6): 720–26.
Bublitz, Christoph. 2018. “‘The Soul Is the Prison of the Body’ – Mandatory Moral Enhancement, Punishment & Rights Against Neuro-Rehabilitation.” In *Treatment for Crime: Philosophical Essays on Neurointerventions in Criminal Justice*, Oxford University Press.
Bublitz, Christoph, and Reinhard Merkel. 2013. *Guilty Minds in Washed Brains?: Manipulation Cases and the Limits of Neuroscientific Excuses in Liberal Legal Orders*. Oxford University Press.

Buckholtz, Joshua, Valerie F. Reyna, and Christopher Slobogin. 2016. “A Neuro-Legal Lingua Franca: Bridging Law and Neuroscience on the Issue of Self-Control.” Rochester, NY: Social Science Research Network.

Buckholtz, Joshua W., Christopher L. Asplund, Paul E. Dux, David H. Zald, John C. Gore, Owen D. Jones, and René Marois. 2008. “The Neural Correlates of Third-Party Punishment.” *Neuron* 60 (5): 930–40.

Buckholtz, Joshua W., and René Marois. 2012. “The Roots of Modern Justice: Cognitive and Neural Foundations of Social Norms and Their Enforcement.” *Nature Neuroscience* 15 (5): 655–61.

Buckholtz, Joshua W., Justin W. Martin, Michael T. Treadway, Katherine Jan, David H. Zald, Owen Jones, and René Marois. 2015. “From Blame to Punishment: Disrupting Prefrontal Cortex Activity Reveals Norm Enforcement Mechanisms.” *Neuron* 87 (6): 1369–80.

Bueso-Izquierdo, Natalia, Juan Verdejo-Román, Oren Contreras-Rodríguez, Martina Carmona-Perera, Miguel Pérez-García, and Natalia Hidalgo-Ruzzante. 2016. “Are Batterers Different from Other Criminals? An fMRI Study.” *Social Cognitive and Affective Neuroscience* 11 (5): 852–62.

Bufkin, Jana L., and Vickie R. Luttrell. 2005. “Neuroimaging Studies of Aggressive and Violent Behavior: Current Findings and Implications for Criminology and Criminal Justice.” *Trauma, Violence & Abuse* 6 (2): 176–91.

Caldwell, J.M. Jr. 1941. “The Constitutional Psychopathic State.” *J. Crim. Psychopath*, 171–79.

Canavero, Sergio. 2014. “Criminal Minds: Neuromodulation of the Psychopathic Brain.” *Frontiers in Human Neuroscience* 8: 124.

Cantor, James M., Noor Kabani, Bruce K. Christensen, Robert B. Zipursky, Howard E. Barbaree, Robert Dickey, Philip E. Klassen, et al. 2008. “Cerebral White Matter Deficiencies in Pedophilic Men.” *Journal of Psychiatric Research* 42 (3): 167–83.

Carswell, Kenneth, Barbara Maughan, Hilton Davis, Franscesca Davenport, and Nick Goddard. 2004. “The Psychosocial Needs of Young Offenders and Adolescents from an Inner City Area.” *Journal of Adolescence* 27 (September): 415–28.

Castel, Pierre-Henri. 2010. *L’esprit malade. Cerveaux, folies, individus*. Ithaque.

Castel, Robert. 1976. *L’ordre psychiatrique: l’âge d’or de l’aliénisme*. Le sens commun. Paris: Les éditions de minuit.

———. 1983. “De la dangerosité au risque.” *Actes de la Recherche en Sciences Sociales* 47 (1): 119–27.

Châles-Courtine, Sylvie. 2005. “Le corps criminel dans les Archives d’anthropologie criminelle.” *Criminocorpus. Revue d’Histoire de la justice, des crimes et des peines*, January.

Chalmers, David J., and Frank Jackson. 2001. “Conceptual Analysis and Reductive Explanation.” *Philosophical Review* 110 (3): 315–61.

Chamak, Brigitte, et Baptiste Moutaud. 2014. *Neurosciences et société: Enjeux des savoirs et pratiques sur le cerveau*. Armand Colin.

Chemero, Tony, and Michael Silberstein. 2008. “After the Philosophy of Mind: Replacing Scholasticism with Science.” *Philosophy of Science* 75 (1): 1–27.

Chen, Chiao-Yun, Adrian Raine, Kun-Hsien Chou, I.-Yun Chen, Daisy Hung, and Ching-Po Lin. 2016. “Abnormal White Matter Integrity in Rapists as Indicated by Diffusion Tensor Imaging.” *BMC Neuroscience* 17 (1): 45.

Chen, Frances R., Yu Gao, Andrea L. Glenn, Sharon Niv, Jill Portnoy, Robert Schug, Yaling Yang, and Adrian Raine. 2015. “Biosocial Bases of Antisocial and Criminal Behavior.” In *The Handbook of Criminological Theory*, 355–79. Wiley.

Cherici, Céline. 2003. “L’innéisme organique des facultés intellectuelles chez Vincenzo Malacarne (1744-1816).” *Bulletin d’histoire et d’épistémologie des sciences de la vie* 10 (2): 245–63.

———. 2010. “La définition d’une entité clinique entre développements techniques et spécialisation médicale : épilepsie et épileptologie au XX^e^ siècle.” *Revue d’histoire des sciences* 63 (2): 409–437.

Cherici, Céline, et Jean-Claude Dupont. 2008. *Les querelles du cerveau : comment furent inventées les neurosciences*. Paris: Vuibert.

Chew, Christopher, Thomas Douglas, and Nadira S. Faber. 2018. “Biological Interventions for Crime Prevention.” In *Treatment for Crime: Philosophical Essays on Neurointerventions in Criminal Justice*. Oxford University Press.

Choudhury, Suparna, Saskia Kathi Nagel, and Jan Slaby. 2009. “Critical Neuroscience: Linking Neuroscience and Society through Critical Practice.” *BioSocieties* 4 (1): 61–77.

Choy, Olivia, Farah Focquaert, and Adrian Raine. 2018a. “Benign Biological Interventions to Reduce Offending.” *Neuroethics*.

Choy, Olivia, Adrian Raine, and Roy H. Hamilton. 2018b. “Stimulation of the Prefrontal Cortex Reduces Intentions to Commit Aggression: A Randomized, Double-Blind, Placebo-Controlled, Stratified, Parallel-Group Trial.” *Journal of Neuroscience*, 3317–17.

Churchland, Patricia S. 1986. *Neurophilosophy: Toward a Unified Science of the Mind-Brain*. MIT Press.

———. 2002. *Brain-Wise: Studies in Neurophilosophy*. MIT Press.

Clark, Andy, and David J. Chalmers. 1998. “The Extended Mind.” *Analysis* 58 (1): 7–19.

Clausen, Jens. 2010. “Ethical Brain Stimulation - Neuroethics of Deep Brain Stimulation in Research and Clinical Practice.” *The European Journal of Neuroscience* 32 (7): 1152–62.

Clauzade, Laurent. 2008a. “La « physiologie phrénologique » : une approche différentielle des fonctions mentales.” In *Les fonctions en psychologie*, 336. PSY-Théories, débats, synthèses. Mardaga.

———. 2008b. “Le cerveau chez Cabanis et Gall : la philosophie biologique du XVIII^e^ siècle en débat.” In *Les Querelles du cerveau : comment furent inventées les neurosciences*, 235–55. Vuibert.

Cleckley, Hervey M. 1941. *The Mask of Sanity: an Attempt to Reinterpret the so-Called Psychopathic Personality*. C.V. Mosby Publishing Company.

Cloninger, C. Robert, Dragan M. Svrakic, and Richard D. Wetzel. 1994. *The Temperament and Character Inventory (TCI): A Guide to Its Development and Use*. Center for Psychobiology of Personality, Washington University.

Coan, James A., and John J. B. Allen. 2004. “Frontal EEG Asymmetry as a Moderator and Mediator of Emotion.” *Biological Psychology* 67 (1–2): 7–49.

Connolly, William E. 2002. *Neuropolitics: Thinking, Culture, Speed*. University of Minnesota Press.

Conrad, Edwin C. 1959. “Electroencephalograph (EEG) as Evidence of Criminal Responsitility Police Science.” *Journal of Criminal Law, Criminology and Police Science*, no. 4(1960): 405–14.

Convit, A., P. Czobor, and J. Volavka. 1991. “Lateralized Abnormality in the EEG of Persistently Violent Psychiatric Inpatients.” *Biological Psychiatry* 30 (4): 363–70.

Cope, L. M., E. Ermer, L. M. Gaudet, V. R. Steele, A. L. Eckhardt, M. R. Arbabshirani, M. F. Caldwell, V. D. Calhoun, and K. A. Kiehl. 2014. “Abnormal Brain Structure in Youth Who Commit Homicide.” *NeuroImage. Clinical* 4: 800–807.

Cope, Rosemarie V., and W. Martin Donovan. 1979. “A Case of Insane Automatism?” *The British Journal of Psychiatry* 135 (6): 574–75.

Cova, Florian. 2011. “Neurosciences et droit pénal : le déterminisme peut-il sauver la conception utilitariste de la peine ?” *Klesis – Revue Philosophique*.
Craig, Jared N. 2016. “Incarceration, Direct Brain Intervention, and the Right to Mental Integrity – a Reply to Thomas Douglas.” *Neuroethics* 9 (2): 107–18.
Craver, Carl F. 2007. *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*. Clarendon Press.
———. 2009. “Levels of Mechanisms: A Field Guide to the Hierarchical Structure of the World.” In *The Routledge Companion to Philosophy of Psychology*, Routledge, 427–39.
Craver, Carl F., and Lindley Darden. 2013. *In Search of Mechanisms: Discoveries across the Life Sciences*. University of Chicago Press.
Cunha-Bang, Sofi da, Patrick M. Fisher, Liv Vadskjær Hjordt, Erik Perfalk, Anine Persson Skibsted, Camilla Bock, Anders Ohlhues Baandrup, et al. 2017. “Violent Offenders Respond to Provocations with High Amygdala and Striatal Reactivity.” *Social Cognitive and Affective Neuroscience* 12 (5): 802–10.
Curran, Desmond, and Paul Mallinson. 1944. “Psychopathic Personality.” *Journal of Mental Science* 90 (378): 266–86.
Damasio, Antonio R. [1994] 2006. *L’erreur de Descartes: la raison des émotions*. Paris: Odile Jacob.
Damasio, H., T. Grabowski, R. Frank, A. M. Galaburda, and A. R. Damasio. 1994. “The Return of Phineas Gage: Clues about the Brain from the Skull of a Famous Patient.” *Science* 264 (5162): 1102–5.
Dambacher, Franziska, Teresa Schuhmann, Jill Lobbestael, Arnoud Arntz, Suzanne Brugman, and Alexander T. Sack. 2015. “Reducing Proactive Aggression through Non-Invasive Brain Stimulation.” *Social Cognitive and Affective Neuroscience* 10 (10): 1303–9.
Danet, Jean. 2008. “La dangerosité, une notion criminologique, séculaire et mutante.” *Champ pénal* Vol. V. (Juin).
Darby, R. Ryan, and Alvaro Pascual-Leone. 2017. “Moral Enhancement Using Non-Invasive Brain Stimulation.” *Frontiers in Human Neuroscience* 11.
Dartigues, Laurent. 2017. *Le neurodroit est-il une nouvelle phrénologie ?* Presses universitaires de Paris Nanterre.
———. 2018. “Une irrésistible ascension ? Le neurodroit face à ses critiques.” *Zilsel* N° 3 (1): 63–103.
Davis, Hallowell. 1936. “Some Aspects of the Electrical Activity of the Cerebral Cortex.” *Cold Spring Harbor Symposia on Quantitative Biology* 4 (January): 285–91.
Davis, Hallowell, and Pauline A. Davis. 1936. “Action Potentials of the Brain: In Normal Persons and in Normal States of Cerebral Activity.” *Archives of Neurology & Psychiatry* 36 (6): 1214–24.
Dawkins, Richard. 1982. *The Extended Phenotype: The Long Reach of the Gene*. Oxford University Press.
De Ridder, Dirk, Berthold Langguth, Mark Plazier, and Tomas Menovsky. 2009. “Moral Dysfunction: Theoretical Model and Potential Neurosurgical Treatments.” In *The Moral Brain: Essays on the Evolutionary and Neuroscientific Aspects of Morality*, edited by Jan Verplaetse, Jelle Schrijver, Sven Vanneste, and Johan Braeckman, 155–83. Dordrecht: Springer Netherlands.
Decety, Jean, Laurie R. Skelly, and Kent A. Kiehl. 2013. “Brain Response to Empathy-Eliciting Scenarios Involving Pain in Incarcerated Individuals with Psychopathy.” *JAMA Psychiatry* 70 (6): 638–45.
Deckel, A. W., V. Hesselbrock, and L. Bauer. 1996. “Antisocial Personality Disorder, Childhood Delinquency, and Frontal Brain Functioning: EEG and Neuropsychological Findings.” *Journal of Clinical Psychology* 52 (6): 639–50.

DeGrazia, David. 2014. “Moral Enhancement, Freedom, and What We (Should) Value in Moral Behaviour.” *Journal of Medical Ethics* 40 (6): 361–68.

Delgado-Escueta, A. V., R. H. Mattson, L. King, E. S. Goldensohn, H. Spiegel, J. Madsen, P. Crandall, F. Dreifuss, and R. J. Porter. 1981. “Special Report. The Nature of Aggression during Epileptic Seizures.” *The New England Journal of Medicine* 305 (12): 711–16.

Delgado-Escueta, Antonio V., Richard H. Mattson, Lambert King, Eli S. Goldensohn, Herbert Spiegel, Jack Madsen, Paul Crandall, Fritz Dreifuss, and Roger J. Porter. 2002. “The Nature of Aggression during Epileptic Seizures.” *Epilepsy & Behavior: E&B* 3 (6): 550–56.

Demazeux, Steeves, et Vincent Pidoux. 2015. “Le projet RDoC - La classification psychiatrique de demain ?” *Médecine/Sciences* 31 (8–9): 792–96.

Deming, Philip, Carissa L. Philippi, Richard C. Wolf, Monika Dargis, Kent A. Kiehl, and Michael Koenigs. 2018. “Psychopathic Traits Linked to Alterations in Neural Activity during Personality Judgments of Self and Others.” *NeuroImage: Clinical* 18 (January): 575–81.

Denno, Deborah. 2002. “Crime and Consciousness: Science and Involuntary Acts.” *Minnesota Law Review* 87 (December): 269–399.

———. 2015. “The Myth of the Double-Edged Sword: An Empirical Study of Neuroscience Evidence in Criminal Cases.” *Boston College Law Review* 56 (2): 493.

Desmoulin-Canselier, Sonia. 2018. “Usages et interprétations judiciaires des images cérébrales.” *Revue de science criminelle et de droit penal comparé* N° 2 (2): 343–57.

———. 2019. “La France à « l’ère du neurodroit » ? La neuro-imagerie dans le contentieux civil français.” *Droit et société* N° 101 (1): 115–35.

———. 2020. “Another Perspective on ‘Neurolaw’: The Use of Brain Imaging in Civil Litigation Regarding Mental Competence.” In *Neuroscience and Law: Complicated Crossings and New Perspectives*, edited by Antonio D’Aloia and Maria Chiara Errigo, 529–47. Springer.

Devinsky, Orrin, and David M. Bear. 1984. “Varieties of Aggressive Behavior in Temporal Lobe Epilepsy.” *The American Journal of Psychiatry* 141 (5): 651–56.

Dodge, Kenneth A. 1991. “The Structure and Function of Reactive and Proactive Aggression.” In *The Development and Treatment of Childhood Aggression*, 201–18. Hillsdale, NJ, US: Lawrence Erlbaum Associates.

Douglas, Thomas. 2008. “Moral Enhancement.” *Journal of Applied Philosophy* 25 (3): 228–45.

———. 2013. “Enhancement, Biomedical.” In *International Encyclopedia of Ethics*. American Cancer Society.

———. 2014. “Criminal Rehabilitation Through Medical Intervention: Moral Liability and the Right to Bodily Integrity.” *The Journal of Ethics* 18 (2): 101–22.

———. 2015. “The Morality of Moral Neuroenhancement,” September, 1227–50.

———. 2018. “Neural and Environmental Modulation of Motivation: What’s the Moral Difference?” In *Treatment for Crime: Philosophical Essays on Neurointerventions in Criminal Justice*. Oxford University Press.

Douglas, Thomas, and Jonathan Pugh. 2016. “Neurointerventions as Criminal Rehabilitation: An Ethical Review.” *J Jacobs and J Jackson (Eds.), The Routledge Handbook of Criminal Justice Ethics*.

Driver, M. V., L. R. West, and M. Faulk. 1974. “Clinical and EEG Studies of Prisoners Charged with Murder.” *The British Journal of Psychiatry: The Journal of Mental Science* 125 (December): 583–87.

Dumit, Joseph. 1999. “Objective Brains, Prejudicial Images.” *Science in Context* 12 (1): 173–201.

Dupont, Jean-Claude. 2015. “Que faire de l’imagerie cérébrale ? Territoires anciens et nouveaux d’une technologie.” *Comptes Rendus Biologies*, Biologie et devenir technologique de l’homme, 338 (8): 607–12.
Dupont, Jean-Claude, et Céline Cherici. 2015. *L’exploration cérébrale: Histoire récente et nouveaux outils*. Paris: Hermann.
Durand, E., L. Watier, M. Fix, J. J. Weiss, M. Chevignard, and P. Pradat-Diehl. 2016. “Prevalence of Traumatic Brain Injury and Epilepsy among Prisoners in France: Results of the Fleury TBI Study.” *Brain Injury* 30 (4): 363–72.
Earp, Brian D., Thomas Douglas, and Julian Savulescu. 2017. “Moral Neuroenhancement.” In *The Routledge Handbook of Neuroethics*, edited by L. Syd M. Johnson and Karen S. Rommelfanger. New York: Routledge.
East, W. Norwood. 1945. “Psychopathic Personality and Crime.” *Journal of Mental Science* 91 (385): 426–46.
Eastman, Nigel, and Colin Campbell. 2006. “Neuroscience and Legal Determination of Criminal Responsibility.” *Nature Reviews Neuroscience* 7 (4): 311–18.
Eddy, Clare M., Andrea E. Cavanna, Hugh E. Rickards, and Peter C. Hansen. 2016. “Temporo-Parietal Dysfunction in Tourette Syndrome: Insights from an fMRI Study of Theory of Mind.” *Journal of Psychiatric Research* 81 (October): 102–11.
Ehrenberg, Alain. 2004. “Le sujet cérébral.” *Esprit*, no. 309 (11): 130–55.
———. 2018. *La Mécanique des passions: Cerveau, comportement, société*. Odile Jacob.
Ehrlich, S. K., and R. P. Keogh. 1956. “The Psychopath in a Mental Institution.” *A.M.A. Archives of Neurology and Psychiatry* 76 (3): 286–95.
Eichelberger, Rebecca, and J. C. Barnes. 2015. “Biosocial Criminology.” In *The Encyclopedia of Crime & Punishment*, 1–8. American Cancer Society.
Eisenschenk, Stephan, Harry Krop, and Orrin Devinsky. 2014. “Homicide during Postictal Psychosis.” *Epilepsy & Behavior Case Reports* 2 (January): 118–20.
Ermer, Elsa, Lora M. Cope, Prashanth K. Nyalakanti, Vince D. Calhoun, and Kent A. Kiehl. 2012. “Aberrant Paralimbic Gray Matter in Criminal Psychopathy.” *Journal of Abnormal Psychology* 121 (3): 649–58.
———. 2013. “Aberrant Paralimbic Gray Matter in Incarcerated Male Adolescents with Psychopathic Traits.” *Journal of the American Academy of Child and Adolescent Psychiatry* 52 (1): 94–103.e3.
Ervin, F. R., V. H. Mark, and W. H. Sweet. 1969. “Focal Cerebral Disease, Temporal Lobe Epilepsy and Violent Behavior.” *Transactions of the American Neurological Association* 94: 253–56.
Faigman, David L., John Monahan, and Christopher Slobogin. 2013. “Group to Individual (G2i) Inference in Scientific Expert Testimony.” Rochester, NY: Social Science Research Network.
Fallin, Mallory, Owen Whooley, and Kristin Kay Barker. 2018. “Criminalizing the Brain: Neurocriminology and the Production of Strategic Ignorance.” *BioSocieties*, September.
Fangerau, Heiner, Jörg M. Fegert, and Thorsten Trapp. 2011. *Implanted Minds: The Neuroethics of Intracerebral Stem Cell Transplantation and Deep Brain Stimulation*. Walter de Gruyter & Co.
Farwell, L. A., and E. Donchin. 1991. “The Truth Will out: Interrogative Polygraphy (‘lie Detection’) with Event-Related Brain Potentials.” *Psychophysiology* 28 (5): 531–47.
Farwell, Lawrence A. 2012. “Brain Fingerprinting: A Comprehensive Tutorial Review of Detection of Concealed Information with Event-Related Brain Potentials.” *Cognitive Neurodynamics* 6 (2): 115–54.
Faucher, Luc. 2018. “Stimulation cérébrale profonde et augmentation morale.” *Bulletin d’histoire et d’epistemologie des sciences de la vie.* Volume 25 (2): 221–38.

Faucher, Luc, et Simon Goyer. 2016. “Le Research Domain Criteria (Rdoc), le réductionnisme et la psychiatrie clinique.” *Revue de Synthèse* 137 (1): 117–49.
Faucher, Luc, et Pierre Poirier. 2001. “Psychologie évolutionniste et théories interdomaines.” *Dialogue: Revue Canadienne de Philosophie* 40 (3): 453–86.
Faugeras, Patrick. 2009. “Le criminel né.” *Sud/Nord* n° 24 (1): 55–67.
Fazel, Seena, Paul Lichtenstein, Martin Grann, and Niklas Långström. 2011. “Risk of Violent Crime in Individuals with Epilepsy and Traumatic Brain Injury: A 35-Year Swedish Population Study.” *PLoS Medicine* 8 (12).
Fazel, Seena, Johanna Philipson, Lisa Gardiner, Rowena Merritt, and Martin Grann. 2009. “Neurological Disorders and Violence: A Systematic Review and Meta-Analysis with a Focus on Epilepsy and Traumatic Brain Injury.” *Journal of Neurology* 256 (10): 1591–1602.
Fazel, Seena, Evangelos Vassos, and John Danesh. 2002. “Prevalence of Epilepsy in Prisoners: Systematic Review.” *BMJ (Clinical Research Ed.)* 324 (7352): 1495.
Feeley, Malcolm M., and Jonathan Simon. 1992. “The New Penology: Notes on the Emerging Strategy of Corrections and Its Implications.” *Criminology* 30 (4): 449–74.
Fenton, G. W., and E. L. Udwin. 1965. “Homicide, Temporal Lobe Epilepsy and Depression : A Case Report.” *The British Journal of Psychiatry: The Journal of Mental Science* 111 (April): 304–6.
Féré, Charles. 1888. *Dégénérescence et criminalité : essai physiologique*. Paris, Félix Alcan.
Ferri, Enrico. 1893. *La sociologie criminelle*. Paris, Félix Alcan.
Fessler, A. J., and D. M. Treiman. 2009. “Epilepsy and Aggression: Proceed with Caution.” *Neurology* 73 (21): 1720–21.
Fischer, John, and Mark Ravizza. 2000. “Responsibility and Control: A Theory of Moral Responsibility.” *Philosophy and Phenomenological Research* 61 (September).
Fletcher, P. C., F. Happé, U. Frith, S. C. Baker, R. J. Dolan, R. S. J. Frackowiak, and C. D. Frith. 1995. “Other Minds in the Brain: A Functional Imaging Study of ‘theory of Mind’ in Story Comprehension.” *Cognition* 57 (2): 109–28.
Flor-Henry, P., R.A. Lang, Z.J. Koles, and R.R. Frenzel. 1991. “Quantitative EEG Studies of Pedophilia.” *International Journal of Psychophysiology* 10 (3): 253–58.
Foot, Philippa. 1967. “The problem of abortion and the doctrine of double effect”. Oxford Review, 5.
Forest, Denis. 2008. “Localisation fonctionnelle et stratégies de recherche : le cerveau hier et aujourd’hui.” In *Les fonctions en psychologie*, 19–48. Mardaga.
———. 2012. “Le cerveau, une réputation bien surfaite ? La conception standard et ses ennemis.” *Revue philosophique de la France et de l’etranger* Tome 137 (4): 535–49.
———. 2014. *Neuroscepticisme*. Ithaque.
———. 2019. “Ni élimination, ni réduction, ni intégration alors quoi? Sciences de l’esprit, neurosciences et modèle instrumental.” *Intellectica*, no. 69: pp.167-185.
Forsberg, Lisa. 2018. “Crime-Preventing Neurointerventions and the Law : Learning from Anti-Libidinal Interventions.” In *Treatment for Crime: Philosophical Essays on Neurointerventions in Criminal Justice*. Oxford University Press.
Forssman, H., and T. S. Frey. 1953. “Electroencephalograms of Boys with Behavior Disorders.” *Acta Psychiatrica Et Neurologica Scandinavica* 28 (1): 61–73.
Foucault, Michel. 1972. *Histoire de la folie à l’âge classique*. Paris: Gallimard.
———. 1975. *Surveiller et punir: Naissance de la prison*. Paris: Gallimard.
———. 1981. “L’évolution de la notion d’‘individu dangereux’ dans la psychiatrie légale.” *Déviance et société* 5 (4): 403–22.

Franzini, Angelo, Giovanni Broggi, Roberto Cordella, Ivano Dones, and Giuseppe Messina. 2013. "Deep-Brain Stimulation for Aggressive and Disruptive Behavior." *World Neurosurgery* 80 (3–4): S29.e11-14.

Franzini, Angelo, Carlo Marras, Paolo Ferroli, Orso Bugiani, and Giovanni Broggi. 2005. "Stimulation of the Posterior Hypothalamus for Medically Intractable Impulsive and Violent Behavior." *Stereotactic and Functional Neurosurgery* 83 (2–3): 63–66.

Frazzetto, Giovanni, and Suzanne Anker. 2009. "Neuroculture." *Nature Reviews Neuroscience* 10 (11): 815–21.

Fuss, Johannes, Matthias K. Auer, Sarah V. Biedermann, Peer Briken, and Werner Hacke. 2015. "Deep Brain Stimulation to Reduce Sexual Drive." *Journal of Psychiatry & Neuroscience: JPN* 40 (6): 429–31.

Gall, Franz Josef, 1825. *Sur les fonctions du cerveau et sur celles de chacune de ses parties : avec des observations sur la possibilité de reconnaître les instincts, les penchans, les talens, ou les dispositions morales et intellectuelles des hommes et des Animaux, par la configuration de leur cerveau et de leur tête*. Paris: Jean-Baptiste Baillière.

Gall, Franz Josef, et Johann Gaspar Spurzheim. 1810. *Anatomie et physiologie du système nerveux en général et du cerveau en particulier, avec des observations sur la possibilité de reconnaître plusieurs dispositions intellectuelles et morales de l'homme et des animaux par la configuration de leurs têtes*. Paris: Frédéric Schoell.

Gallagher, Helen L., and Christopher D. Frith. 2003. "Functional Imaging of 'theory of Mind.'" *Trends in Cognitive Sciences* 7 (2): 77–83.

Gallagher, Roswell J., Erna L. Gibbs, and Frederic A. Gibbs. 1942. "Relation between the Electrical Activity of the Cortex and the Personality in Adolescent Boys." *Psychosomatic Medicine* 4 (2): 134.

Galski, Thomas, Kirtley Thornton, and David Shumsky. 1990. "Brain Dysfunction in Sex Offenders." *Journal of Offender Rehabilitation* 16 (December): 65–80.

Gao, Yu, Adrian Raine, Peter H. Venables, Michael E. Dawson, and Sarnoff A. Mednick. 2010. "Association of Poor Childhood Fear Conditioning and Adult Crime." *The American Journal of Psychiatry* 167 (1): 56–60.

Gao, Yu, Catherine Tuvblad, Anne Schell, Laura Baker, and Adrian Raine. 2015. "Skin Conductance Fear Conditioning Impairments and Aggression: A Longitudinal Study." *Psychophysiology* 52 (2): 288–95.

Garnett, E. S., C. Nahmias, G. Wortzman, R. Langevin, and R. Dickey. 1988. "Positron Emission Tomography and Sexual Arousal in a Sadist and Two Controls." *Annals of Sex Research* 1 (3): 387–99.

Garofalo, Raffaele. 1890. *La criminologie : étude sur la nature du crime et la théorie de la pénalité*. Deuxième édition. Paris, Félix Alcan.

Gershon, Michael. 1999. *The Second Brain: A Groundbreaking New Understanding of Nervous Disorders of the Stomach and Intestine*. Harper Perennial.

Gibbens, T. C. N. 1951. "Recent Trends in the Management of Psychopathic Offenders." *The British Journal of Delinquency* 2 (2): 103–16.

Gibbens, T. C., D. A. Pond, and D. Staffordclark. 1959. "A Follow-up Study of Criminal Psychopaths." *The Journal of Mental Science* 105 (438): 108–15.

Gibbs, Erna L., and Frederic A. Gibbs. 1951. "Electroencephalographic Evidence of Thalamic and Hypothalamic Epilepsy." *Neurology* 1: 136–44.

Gibbs, Frederic A., B. K. Bagchi, and Wilfred Bloomberg. 1945. "Electroencephalographic Study of Criminals." *American Journal of Psychiatry* 102 (3): 294–98.

Gibbs, Frederic A., and Erna L. Gibbs. 1951. *Atlas of Electroencephalography, Volume 1 : Methodology and Controls. 2nd Ed.* Second Edition edition. Addison-Wesley Press.

———. 1964. *Atlas of Electroencephalography, Volume 3 : Neurological and Psychiatric Disorders*. Addison-Wesley.
Ginther, Matthew R., Richard J. Bonnie, Morris B. Hoffman, Francis X. Shen, Kenneth W. Simons, Owen D. Jones, and René Marois. 2016. "Parsing the Behavioral and Brain Mechanisms of Third-Party Punishment." *Journal of Neuroscience* 36 (36): 9420–34.
Ginther, Matthew, Francis Shen, Richard Bonnie, Morris Hoffman, Owen Jones, and Kenneth Simons. 2018. "Decoding Guilty Minds: How Jurors Attribute Knowledge and Guilt." *Vanderbilt Law Review* 71 (1): 179.
Glannon, Walter. 2014a. "Intervening in the Psychopath's Brain." *Theoretical Medicine and Bioethics* 35 (1): 43–57.
———. 2014b. "Philosophical Reflections on Therapeutic Brain Stimulation." *Frontiers in Computational Neuroscience* 8 (May).
Glenn, A. L., A. Raine, and R. A. Schug. 2009. "The Neural Correlates of Moral Decision-Making in Psychopathy." *Molecular Psychiatry* 14 (1): 5–6.
Glenn, Andrea L., and Adrian Raine. 2014. "Neurocriminology: Implications for the Punishment, Prediction and Prevention of Criminal Behaviour." *Nature Reviews. Neuroscience* 15 (1): 54–63.
Glennan, Stuart. 2002. "Rethinking Mechanistic Explanation." *Philosophy of Science* 69 (S3): 342–53.
———. 2005. "Modeling Mechanisms." *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, Mechanisms in biology, 36 (2): 443–64.
———. 2008. "Mechanisms." In *The Routledge Companion to the Philosophy of Science*, Routledge, 375–84.
Glennan, Stuart S. 1996. "Mechanisms and the Nature of Causation." *Erkenntnis* 44 (1): 49–71.
Glimcher, Paul W. 2003. *Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics*. 1. MIT Press.
Gonon, François, Jan Pieter Konsman, et Thomas Boraud. 2014. "Neurosciences et médiatisation : entre argumentation de la preuve et rhétorique de la promesse." In *Neurosciences et Société : Enjeux des savoirs et pratiques sur le cerveau, Chamak B. and Moutaud B. (Dir.)*, Paris, Armand Colin.
Goodenough, Oliver R. 2001. "Mapping Cortical Areas Associated with Legal Reasoning and Moral Intuition." *Jurimetrics* 41 (4): 429–42.
Goodwin, Charles. 1994. "Professional Vision." *American Anthropologist* 96 (3): 606–33.
Gottfredson, Michael R., and Travis Hirschi. 1990. *A General Theory of Crime*. Stanford University Press.
Gottlober, A. B. 1938. "The Relationship between Brain Potentials and Personality." *Journal of Experimental Psychology* 22 (1): 67–74.
Gould, Stephen Jay. 1983. *La mal-mesure de l'homme*. Paris: Ramsay.
Graber, B., K. Hartmann, J. A. Coffman, C. J. Huey, and C. J. Golden. 1982. "Brain Damage Among Mentally Disordered Sex Offenders." *Journal of Forensic Science* 27 (1): 125–34.
Grafman, J., K. Schwab, D. Warden, A. Pridgen, H. R. Brown, and A. M. Salazar. 1996. "Frontal Lobe Injuries, Violence, and Aggression: A Report of the Vietnam Head Injury Study." *Neurology* 46 (5): 1231–38.
Granieri, Enrico, and Patrik Fazio. 2012. "The Lombrosian Prejudice in Medicine. The Case of Epilepsy. Epileptic Psychosis. Epilepsy and Aggressiveness." *Neurological Sciences: Official Journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology* 33 (1): 173–92.

Grant, Ryan A., Casey H. Halpern, Gordon H. Baltuch, John P. O'Reardon, and Arthur Caplan. 2014. "Ethical Considerations in Deep Brain Stimulation for Psychiatric Illness." *Journal of Clinical Neuroscience* 21 (1): 1–5.

Greely, H. T. 2012. "Direct Brain Interventions to 'Treat' Disfavored Human Behaviors: Ethical and Social Issues." *Clinical Pharmacology & Therapeutics* 91 (2): 163–65.

Greely, Henry T. 2008. "Neuroscience and Criminal Justice: Not Responsibility but Treatment." *Kansas Law Review*, June.

———. 2009. "Law and the Revolution in Neuroscience: An Early Look at the Field." *Akron Law Review* 42: 687.

Greene, Edith, and Brian S. Cahill. 2012. "Effects of Neuroimaging Evidence on Mock Juror Decision Making." *Behavioral Sciences & the Law* 30 (3): 280–96.

Greene, Joshua D., Leigh E. Nystrom, Andrew D. Engell, John M. Darley, and Jonathan D. Cohen. 2004. "The Neural Bases of Cognitive Conflict and Control in Moral Judgment." *Neuron* 44 (2): 389–400.

Greene, Joshua D., R. Brian Sommerville, Leigh E. Nystrom, John M. Darley, and Jonathan D. Cohen. 2001. "An fMRI Investigation of Emotional Engagement in Moral Judgment." *Science* 293 (5537): 2105–8.

Gregory, Sarah, Dominic ffytche, Andrew Simmons, Veena Kumari, Matthew Howard, Sheilagh Hodgins, and Nigel Blackwood. 2012. "The Antisocial Brain: Psychopathy Matters." *Archives of General Psychiatry* 69 (9): 962–72.

Gunn, J. 1978. "Epileptic Homocide: A Case Report." *The British Journal of Psychiatry: The Journal of Mental Science* 132 (May): 510–13.

Gunn, J., and J. Bonn. 1971. "Criminality and Violence in Epileptic Prisoners." *The British Journal of Psychiatry: The Journal of Mental Science* 118 (544): 337–43.

Gunn, J. C. 1969. "The Prevalence of Epilepsy among Prisoners." *Proceedings of the Royal Society of Medicine* 62 (1): 60–63.

Gunn, J., and G. Fenton. 1971. "Epilepsy, Automatism, and Crime." *Lancet (London, England)* 1 (7710): 1173–76.

Gunn, John, and George Fenton. 1969. "Epilepsy in Prisons: A Diagnostic Survey." *Br Med J* 4 (5679): 326–28.

Gurley, Jessica R., and David K. Marcus. 2008. "The Effects of Neuroimaging and Brain Injury on Insanity Defenses." *Behavioral Sciences & the Law* 26 (1): 85–97.

Gutmann, L. 2007. "Jack Ruby." *Neurology* 68 (9): 707–8.

Habermeyer, Benedikt, Fabrizio Esposito, Nadja Händel, Patrick Lemoine, Hans Christian Kuhl, Markus Klarhöfer, Ralph Mager, et al. 2013. "Response Inhibition in Pedophilia: An fMRI Pilot Study." *Neuropsychobiology* 68 (4): 228–37.

Hacking, Ian. 2001. "Degeneracy, Criminal Behavior, and Looping." In *Genetics and Criminal Behavior*, edited by David Wasserman and Robert Wachbroit, 141–68. Cambridge University Press.

Hampel, R., and H. Selg. 1957. "Fragebogen Zur Erfassung von Aggressivitatsfaktoren [Questionnaire for Factors of Aggressiveness]." *Hogrefe, Göttingen*.

Hardcastle, Valerie Gray. 2017. *My Brain Made Me Do It?* Routledge Handbooks.

———. 2020. "Group-to-Individual (G2i) Inferences: Challenges in Modeling How the U.S. Court System Uses Brain Data." *Artificial Intelligence and Law* 28 (1): 51–68.

Hare, Robert D. 1980. "A Research Scale for the Assessment of Psychopathy in Criminal Populations." *Personality and Individual Differences* 1 (2): 111–19.

———. 1991. *The Hare Psychopathy Checklist-Revised : Manual*. Multi-Health Systems.

———. 2003. *The Hare Psychopathy Checklist Revised (2nd Ed.)*. Multi-Health Systems.

Hare, Robert D., Timothy J. Harpur, A. R. Hakstian, Adelle E. Forth, Stephen D. Hart, and Joseph P. Newman. 1990. "The Revised Psychopathy Checklist: Reliability and Factor

Structure." *Psychological Assessment: A Journal of Consulting and Clinical Psychology* 2 (3): 338–41.
Hare, Robert D., Stephen D. Hart, and Timothy J. Harpur. 1991. "Psychopathy and the DSM-IV Criteria for Antisocial Personality Disorder." *Journal of Abnormal Psychology* 100 (3): 391–98.
Hare, Robert D., and Jeffrey W. Jutai. 1983. "Criminal History of the Male Psychopath: Some Preliminary Data." In *Prospective Studies of Crime and Delinquency*, edited by Katherine Teilmann Van Dusen and Sarnoff A. Mednick, 225–36. Longitudinal Research in the Behavioral, Social and Medical Sciences. Dordrecht: Springer Netherlands.
Harenski, Carla L., Keith A. Harenski, Matthew S. Shane, and Kent A. Kiehl. 2010. "Aberrant Neural Processing of Moral Violations in Criminal Psychopaths." *Journal of Abnormal Psychology* 119 (4): 863–74.
Harmon-Jones, E., and J. Sigelman. 2001. "State Anger and Prefrontal Brain Activity: Evidence That Insult-Related Relative Left-Prefrontal Activation Is Associated with Experienced Anger and Aggression." *Journal of Personality and Social Psychology* 80 (5): 797–803.
Harmon-Jones, Eddie. 2003. "Clarifying the Emotive Functions of Asymmetrical Frontal Cortical Activity." *Psychophysiology* 40 (6): 838–48.
Harmon-Jones, Eddie, and Philip A. Gable. n.d. "On the Role of Asymmetric Frontal Cortical Activity in Approach and Withdrawal Motivation: An Updated Review of the Evidence." *Psychophysiology* 55 (1): e12879.
Harmon-Jones, Eddie, Philip A. Gable, and Carly K. Peterson. 2010. "The Role of Asymmetric Frontal Cortical Activity in Emotion-Related Phenomena: A Review and Update." *Biological Psychology*, The biopsychology of emotion: Current theoretical and empirical perspectives, 84 (3): 451–62.
Harris, John. 2011. "Moral Enhancement and Freedom." *Bioethics* 25 (2): 102–11.
Hart, H.L.A. 1968. "IX. Postscript: Responsibility and Retribution." In *Punishment and Responsibility*, 210–37. Oxford: Clarendon Press.
Hart, Stephen D., Robert D. Hare, and Adelle E. Forth. 1994. "Psychopathy as a Risk Marker for Violence: Development and Validation of a Screening Version of the Revised Psychopathy Checklist." In *Violence and Mental Disorder: Developments in Risk Assessment*, 81–98. The John D. and Catherine T. MacArthur Foundation Series on Mental Health and Development. University of Chicago Press.
Hart, Stephen D., Philip R. Kropp, and Robert D. Hare. 1988. "Performance of Male Psychopaths Following Conditional Release from Prison." *Journal of Consulting and Clinical Psychology* 56 (2): 227–32.
Hecht, David. 2011. "An Inter-Hemispheric Imbalance in the Psychopath's Brain." *Personality and Individual Differences* 51 (1): 3–10.
Hempel, Carl G., and Paul Oppenheim. 1948. "Studies in the Logic of Explanation." *Philosophy of Science* 15 (2): 135–75.
Henderson, D. K. 1939. *Psychopathic States*. Psychopathic States. New York, NY, US: W W Norton & Co.
Henderson, Monika. 1982. "An Empirical Classification of Convicted Violent Offenders." *British Journal of Criminology*, no. 1: 1–20.
Hendricks, S. E., D. F. Fitzpatrick, K. Hartmann, M. A. Quaife, R. A. Stratbucker, and B. Graber. 1988. "Brain Structure and Function in Sexual Molesters of Children and Adolescents." *The Journal of Clinical Psychiatry* 49 (3): 108–12.
Herrnstein, Richard J., and Charles A. Murray. 1994. *The Bell Curve: Intelligence and Class Structure in American Life*. Free Press.

Herzberg, J. L., and P. B. Fenwick. 1988. “The Aetiology of Aggression in Temporal-Lobe Epilepsy.” *The British Journal of Psychiatry: The Journal of Mental Science* 153 (July): 50–55.

Hill, Denis. 1944. “Cerebral Dysrhythmia: Its Significance in Aggressive Behaviour.” *Proceedings of the Royal Society of Medicine* 37 (7): 317–30.

———. 1952. “EEG in Episodic Psychotic and Psychopathic Behaviour; a Classification of Data.” *Electroencephalography and Clinical Neurophysiology* 4 (4): 419–42.

Hill, Denis, and Desmond Pond. 1952. “Reflections on 100 Capital Cases Submitted to Electroencephalography.” *The Journal of Mental Science* 98 (410): 23–43.

Hill, Denis, and William Sargant. 1943. “A Case of Matricide.” *The Lancet* 241 (6243): 526–27.

Hill, Denis, and Donald Watterson. 1942. “Electro-Encephalographic Studies of Psychopathic Personalities.” *Journal of Neurology and Psychiatry* 5 (1–2): 47–65.

Hindler, C. G. 1989. “Epilepsy and Violence.” *The British Journal of Psychiatry: The Journal of Mental Science* 155 (August): 246–49.

Holtzheimer, Paul E., and Helen S. Mayberg. 2011. “Deep Brain Stimulation for Psychiatric Disorders.” *Annual Review of Neuroscience* 34: 289–307.

Hortensius, Ruud, Dennis J. L. G. Schutter, and Eddie Harmon-Jones. 2012. “When Anger Leads to Aggression: Induction of Relative Left Frontal Cortical Activity with Transcranial Direct Current Stimulation Increases the Anger-Aggression Relationship.” *Social Cognitive and Affective Neuroscience* 7 (3): 342–47.

Hosking, Jay G., Erik K. Kastman, Hayley M. Dorfman, Gregory R. Samanez-Larkin, Arielle Baskin-Sommers, Kent A. Kiehl, Joseph P. Newman, and Joshua W. Buckholtz. 2017. “Disrupted Prefrontal Regulation of Striatal Subjective Value Signals in Psychopathy.” *Neuron* 95 (1): 221–231.e4.

Howard, R. C. 1984. “The Clinical EEG and Personality in Mentally Abnormal Offenders.” *Psychological Medicine* 14 (3): 569–80.

Hübner, Dietmar, and Lucie White. 2016. “Neurosurgery for Psychopaths? An Ethical Analysis.” *AJOB Neuroscience* 7 (3): 140–49.

Hucker, S., R. Langevin, G. Wortzman, J. Bain, L. Handy, J. Chambers, and S. Wright. 1986. “Neuropsychological Impairment in Pedophiles.” *Canadian Journal of Behavioural Science* 18 (4): 440–48.

Hucker, S., Ron Langevin, and J. Bain. 1988. “A Double Blind Trial of Sex Drive Reducing Medication in Pedophiles.” *Annals of Sex Research* 1 (2): 227–42.

Hughes, John R., Eugene D. Means, and Bernard S. Stell. 1965. “A Controlled Study on the Behavior Disorders Associated with the Positive Spike Phenomenon.” *Electroencephalography and Clinical Neurophysiology* 18 (4): 349–53.

Huneman, Philippe. 2010. “Topological Explanations and Robustness in Biological Sciences.” *Synthese* 177 (2): 213–45.

Hux, K., V. Bond, S. Skinner, D. Belau, and D. Sanger. 1998. “Parental Report of Occurrences and Consequences of Traumatic Brain Injury among Delinquent and Non-Delinquent Youth.” *Brain Injury* 12 (8): 667–81.

Jasper, Herbert H., Philip Solomon, and Charles Bradley. 1938. “Electroencephalographic Analyses of Behavior Problem Children.” *American Journal of Psychiatry* 95 (3): 641–58.

Jenkins, R. L., and B. L. Pacella. 1943. “Electroencephalographic Studies of Delinquent Boys.” *American Journal of Orthopsychiatry* 13 (1): 107–20.

Jeurissen, Danique, Alexander T. Sack, Alard Roebroeck, Brian E. Russ, and Alvaro Pascual-Leone. 2014. “TMS Affects Moral Judgment, Showing the Role of DLPFC and TPJ in Cognitive and Emotional Processing.” *Frontiers in Neuroscience* 8: 18.

Jones, Owen D., René Marois, Martha J. Farah, and Henry T. Greely. 2013. "Law and Neuroscience." *Journal of Neuroscience* 33 (45): 17624–30.
Jones, Owen D., and Francis X. Shen. 2012. "Law and Neuroscience in the United States." In *International Neurolaw: A Comparative Analysis*, edited by Tade Matthias Spranger, 349–80. T.M. Spranger, ed., Springer-Verlag.
Jones, Sara, Christian C. Joyal, Josh M. Cisler, and Shasha Bai. 2017. "Exploring Emotion Regulation in Juveniles Who Have Sexually Offended: An fMRI Study." *Journal of Child Sexual Abuse* 26 (1): 40–57.
Joyal, Christian C., Deborah N. Black, and Benoit Dassylva. 2007. "The Neuropsychology and Neurology of Sexual Deviance: A Review and Pilot Study." *Sexual Abuse: A Journal of Research and Treatment* 19 (2): 155–73.
Kaluszynski, Martine. 2008. "Le retour de l'homme dangereux. Réflexions sur la notion de dangerosité et ses usages." *Champ pénal* Vol. V. (Juin).
Kennard, M. A. 1956. "The Electroencephalogram and Disorders of Behavior; a Review." *The Journal of Nervous and Mental Disease* 124 (2): 103–24.
Kennard, Margaret A. 1949. "Inheritance of Electroencephalogram Patterns in Children with Behavior Disorders." *Psychosomatic Medicine* 11 (3): 151.
Keune, Philipp M., Linda van der Heiden, Bálint Várkuti, Lilian Konicar, Ralf Veit, and Niels Birbaumer. 2012. "Prefrontal Brain Asymmetry and Aggression in Imprisoned Violent Offenders." *Neuroscience Letters* 515 (2): 191–95.
Keune, Philipp M., Sarah V. Mayer, Aiste Jusyte, and Michael Schönenberg. 2018. "Frontal Alpha Asymmetry and Callous-Unemotional Traits in Imprisoned Violent Offenders: A Pilot Study." *Psychophysiology* 55 (1).
Kiehl, K. A., A. M. Smith, R. D. Hare, A. Mendrek, B. B. Forster, J. Brink, and P. F. Liddle. 2001. "Limbic Abnormalities in Affective Processing by Criminal Psychopaths as Revealed by Functional Magnetic Resonance Imaging." *Biological Psychiatry* 50 (9): 677–84.
Kiehl, Kent A., Andra M. Smith, Adrianna Mendrek, Bruce B. Forster, Robert D. Hare, and Peter F. Liddle. 2004. "Temporal Lobe Abnormalities in Semantic Processing by Criminal Psychopaths as Revealed by Functional Magnetic Resonance Imaging." *Psychiatry Research* 130 (3): 297–312.
Kim, Jeong-Min, Kon Chu, Keun-Hwa Jung, Soon-Tae Lee, Sang-Sub Choi, and Sang Kun Lee. 2011. "Characteristics of Epilepsy Patients Who Committed Violent Crimes: Report from the National Forensic Hospital." *Journal of Epilepsy Research* 1 (1): 13–18.
King, L. N., and Q. D. Young. 1978. "Increased Prevalence of Seizure Disorders among Prisoners." *JAMA* 239 (25): 2674–75.
Klaming, Laura, and Pim Haselager. 2013. "Did My Brain Implant Make Me Do It? Questions Raised by DBS Regarding Psychological Continuity, Responsibility for Action and Mental Competence." *Neuroethics* 6 (3): 527–39.
Kneer, Jonas, Viola Borchardt, Christian Kärgel, Christopher Sinke, Claudia Massau, Gilian Tenbergen, Jorge Ponseti, et al. 2019. "Diminished Fronto-Limbic Functional Connectivity in Child Sexual Offenders." *Journal of Psychiatric Research* 108 (January): 48–56.
Knott, John R., and Jacques S. Gottlieb. 1943. "The Electroencephalogram in Psychopathic Personality." *Psychosomatic Medicine* 5 (2): 139.
Knox, S. J. 1968. "Epileptic Automatism and Violence." *Medicine, Science and the Law* 8 (2): 96–104.
Koi, Polaris, Susanne Uusitalo, and Jarno Tuominen. 2018. "Self-Control in Responsibility Enhancement and Criminal Rehabilitation." *Criminal Law and Philosophy* 12 (2): 227–44.

Konicar, Lilian, Ralf Veit, Hedwig Eisenbarth, Beatrix Barth, Paolo Tonin, Ute Strehl, and Niels Birbaumer. 2015. “Brain Self-Regulation in Criminal Psychopaths.” *Scientific Reports* 5 (March): 9426.

Korponay, Cole, Maia Pujara, Philip Deming, Carissa Philippi, Jean Decety, David S. Kosson, Kent A. Kiehl, and Michael Koenigs. 2017. “Impulsive-Antisocial Dimension of Psychopathy Linked to Enlargement and Abnormal Functional Connectivity of the Striatum.” *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging* 2 (2): 149–57.

Kostić, Daniel. 2018. “The Topological Realization.” *Synthese* 195 (1): 79–98.

Krakauer, John W., Asif A. Ghazanfar, Alex Gomez-Marin, Malcolm A. MacIver, and David Poeppel. 2017. “Neuroscience Needs Behavior: Correcting a Reductionist Bias.” *Neuron* 93 (3): 480–90.

Langevin, R., G. Wortzman, R. Dickey, P. Wright, and L. Handy. 1988. “Neuropsychological Impairment in Incest Offenders.” *Annals of Sex Research* 1 (3): 401–15.

Langevin, R., G. Wortzman, P. Wright, and L. Handy. 1989a. “Studies of Brain Damage and Dysfunction in Sex Offenders.” *Annals of Sex Research* 2 (2): 163–79.

Langevin, Ron. 1990. “Sexual Anomalies and the Brain.” In *Handbook of Sexual Assault: Issues, Theories, and Treatment of the Offender*, edited by W. L. Marshall, D. R. Laws, and H. E. Barbaree, 103–13. Applied Clinical Psychology. Boston, MA: Springer US.

Langevin, Ron, Reuben A. Lang, George Wortzman, Roy R. Frenzel, and Percy Wright. 1989b. “An Examination of Brain Damage and Dysfunction in Genital Exhibitionists.” *Annals of Sex Research* 2 (1): 77–87.

Lanteri-Laura, Georges. 1970. *Histoire de la phrénologie: l’homme et son cerveau selon F. J. Gall*. Presses Universitaires de France.

Larregue, Julien. 2017. “La criminologie biosociale à l’aune de la théorie du champ. Ressources et stratégies d’un courant dominé de la criminologie états-unienne.” *Déviance et Societe* Vol. 41 (2): 167–201.

———. 2019. “La nouvelle orange mécanique : la contribution des bio-criminologues à la médicalisation du contrôle social.” *Zilsel* N° 5 (1): 56–87.

———. 2020. *Héréditaire: L’éternel retour des théories biologiques du crime*. Seuil.

Larrieu, Peggy. 2016. “Recours aux techniques biomédicales en vue de « neuro-amélioration » chez la personne non malade : enjeux éthiques.” *Droit, Sante et Societe* N° 2-3 (2): 13–24.

Le Bras, Anatole. 2018. “Entre contrôle social et assistance : la gestion du danger dans la pratique asilaire (seconde moitié du XIX^{e} siècle).” *Histoire@Politique*, no. 36(September).

Leutgeb, V., M. Leitner, A. Wabnegger, D. Klug, W. Scharmüller, T. Zussner, and A. Schienle. 2015. “Brain Abnormalities in High-Risk Violent Offenders and Their Association with Psychopathic Traits and Criminal Recidivism.” *Neuroscience* 308 (November): 194–201.

Leutgeb, Verena, Albert Wabnegger, Mario Leitner, Thomas Zussner, Wilfried Scharmüller, Doris Klug, and Anne Schienle. 2016. “Altered Cerebellar-Amygdala Connectivity in Violent Offenders: A Resting-State fMRI Study.” *Neuroscience Letters* 610 (January): 160–64.

Levy, Neil. 2014. “Is Neurolaw Conceptually Confused?” *The Journal of Ethics* 18 (2): 171–85.

Levy, S., R. H. Southcombe, J. R. Cranor, and R. A. Freeman. 1952. “The Outstanding Personality Factors among the Population of a State Penitentiary: A Preliminary Report.” *Journal of Clinical and Experimental Psychopathology* 13 (2): 117–30.

Levy, Sol, and Margaret Kennard. 1953. “A Study of the Electroencephalogram as Related to Personality Structure in a Group of Inmates of a State Penitentiary.” *American Journal of Psychiatry* 109 (11): 832–39.

Lewis, Dorothy O, Jonathan H. Pincus, Marilyn Feldman, Lori Jackson, and Barbara Bard. 1986. “Psychiatric, Neurological, and Psychoeducational Characteristics of 15 Death Row Inmates in the United States.” *The American Journal of Psychiatry* 143 (7): 838–45.

Lewis, Dorothy O., Jonathan H. Pincus, Shelley S. Shanok, and Gilbert H. Glaser. 1982. “Psychomotor Epilepsy and Violence in a Group of Incarcerated Adolescent Boys.” *The American Journal of Psychiatry* 139 (7): 882–87.

Lewontin, Richard C. 1994. *Inside and Outside: Gene, Environment, and Organism*. Clark University Press.

Libet, B., C. A. Gleason, E. W. Wright, and D. K. Pearl. 1983. “Time of Conscious Intention to Act in Relation to Onset of Cerebral Activity (Readiness-Potential). The Unconscious Initiation of a Freely Voluntary Act.” *Brain: A Journal of Neurology* 106 (September): 623–42.

Libet, Benjamin. 1985. “Unconscious Cerebral Initiative and the Role of Conscious Will in Voluntary Action.” *Behavioral and Brain Sciences* 8 (4): 529–66.

Lindberg, Nina, Pekka Tani, Matti Virkkunen, Tarja Porkka-Heiskanen, Björn Appelberg, Hannu Naukkarinen, and Tapani Salmi. 2005. “Quantitative Electroencephalographic Measures in Homicidal Men with Antisocial Personality Disorder.” *Psychiatry Research* 136 (1): 7–15.

Lindsley, D. B., and C. Bradley. 1939. “Electroencephalography as an Aid to Understanding Certain Behavior Disorders of Childhood.” *Zeitschrift Für Kinderpsychiatrie* 6: 33–37.

Lindsley, D.B., and K. Cutts. 1940. “Electroencephalograms of ‘Constitutionally Inferior’ and Behavior Problem Children.” *Arch. Neurol. & Psychiat.* 44 (1199).

Lippert-Rasmussen, Kasper. 2018. “The Self-Ownership Trilemma, Extended Minds, and Neurointerventions.” In *Treatment for Crime: Philosophical Essays on Neurointerventions in Criminal Justice*. Oxford University Press.

Lipsman, Nir, and Walter Glannon. 2013. “Brain, Mind and Machine: What Are the Implications of Deep Brain Stimulation for Perceptions of Personal Identity, Agency and Free Will?” *Bioethics* 27 (9): 465–70.

Littlefield, Melissa M., and Jenell Johnson. 2012. *The Neuroscientific Turn: Transdisciplinarity in the Age of the Brain*. Ann Arbor: The University of Michigan Press.

Livingston, Samuel. 1964. “Epilepsy and Murder.” *JAMA* 188 (2): 172–172.

Llorca, Aurore. 2005. “La criminologie, héritière paradoxale de l’école d’anthropologie criminelle.” *Raisons politiques* no 17 (1): 47–64.

Loeber, Rolf, and Dustin Pardini. 2008. “Neurobiology and the Development of Violence: Common Assumptions and Controversies.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 363 (1503): 2491–2503.

Lombroso, Cesare. 1878. *Delitto genio follia. Scritti scelti*. Torino: Bollati Boringhieri.

———. 1887. *L’homme criminel : criminel-né, fou moral, épileptique : étude anthropologique et médico-légale*. Paris, Félix Alcan.

———. 1894. *Les Palimpsestes des prisons*. Lyon-Paris, Storck-Masson.

Lombroso, Cesare, et Guglielmo Ferrero. 1896. *La femme criminelle et la prostituée*. Paris, Félix Alcan.

Lombroso, Cesare, et Rodolfo Laschi. 1892. *Le crime politique et les révolutions : par rapport au droit, à l’anthropologie criminelle et à la science du gouvernement*. Paris, Félix Alcan.

Loomis, A. L., E. N. Harvey, and G. A. Hobart. 1937. “Cerebral States during Sleep, as Studied by Human Brain Potentials.” *Journal of Experimental Psychology* 21 (2): 127–44.

Ly, Martina, Julian C. Motzkin, Carissa L. Philippi, Gregory R. Kirk, Joseph P. Newman, Kent A. Kiehl, and Michael Koenigs. 2012. "Cortical Thinning in Psychopathy." *The American Journal of Psychiatry* 169 (7): 743–49.
Machamer, Peter, Lindley Darden, and Carl F. Craver. 2000. "Thinking about Mechanisms." *Philosophy of Science* 67 (1): 1–25.
Machery, Edouard. 2014. "In Defense of Reverse Inference." *The British Journal for the Philosophy of Science* 65 (2): 251–67.
Mackay, Ronnie D. 1990. "Fact and Fiction about the Insanity Defence." *Crim.L.R.*, 247–55.
Malatesti, Luca, and John McMillan, eds. 2010. *Responsibility and Psychopathy: Interfacing Law, Psychiatry and Philosophy*. International Perspectives in Philosophy and Psychiatry. Oxford University Press.
Maoz, Uri, and Gideon Yaffe. 2015. "What Does Recent Neuroscience Tell Us about Criminal Responsibility?" *Journal of Law and the Biosciences* 3 (1): 120–39.
Marcus, Steven. 2002. *Neuroethics: Mapping the Field*. Edited by Charles A. Dana Foundation. New York: Dana Press.
Markowitsch, Hans J. 2008. "Neuroscience and Crime." *Neurocase* 14 (1): 1–6.
Marsh, Laura, and Gregory L. Krauss. 2000. "Aggression and Violence in Patients with Epilepsy." *Epilepsy & Behavior* 1 (3): 160–68.
Massau, Claudia, Christian Kärgel, Simone Weiß, Martin Walter, Jorge Ponseti, Tillmann Hc Krueger, Henrik Walter, and Boris Schiffer. 2017. "Neural Correlates of Moral Judgment in Pedophilia." *Social Cognitive and Affective Neuroscience* 12 (9): 1490–99.
Matravers, Matt. 2018. "The Importance of Context in Thinking About Crime-Preventing Neurointerventions." In *Treatment for Crime: Philosophical Essays on Neurointerventions in Criminal Justice*. Oxford University Press.
Maudsley, Henry. 1874. *Responsibility in Mental Disease*. King.
Maughs, S. 1941. "A Concept of Psychopathy and Psychopathic Personality: Its Evolution and Historical Development." *Journal of Criminal Psychopathology* 2: 329-356-499.
Mbanzoulou, Paul, Josefina Alvarez, Helene Basex, et Olivier Razac. 2008. *Les nouvelles figures de la dangerosité*. Paris: Editions L'Harmattan.
McCabe, David, and Alan Castel. 2008. "Seeing Is Believing: The Effect of Brain Images on Judgments of Scientific Reasoning." *Cognition* 107 (May): 343–52.
McCord, William Maxwell, and Joan McCord. 1964. *The Psychopath: An Essay on the Criminal Mind*. Van Nostrand.
McCord, William, and Joan McCord. 1956. *Psychopathy and Delinquency*. Psychopathy and Delinquency. Oxford, England: Grune & Stratton.
McGeer, Victoria. 2008. *Moral Psychology, The Neuroscience of Morality: Emotion, Disease and Development (Ed., Walter Sinnott-Armstrong)*. Vol. 3. MIT Press.
McKinley, J. C., and S. R. Hathaway. 1944. "The Minnesota Multiphasic Personality Inventory. V. Hysteria, Hypomania and Psychopathic Deviate." *Journal of Applied Psychology* 28 (2): 153–74.
Mednick, Sarnoff A., Jan Vola Vka, William F. Gabrielli Jr, and Turan M. Itil. 1981. "Eeg as a Predictor of Antisocial Behavior." *Criminology* 19 (2): 219–30.
Mednick, Sarnoff A., and Jan Volavka. 1980. "Biology and Crime." *Crime and Justice* 2: 85–158.
Mele, Alfred R. 2005. "Decisions, Intentions, and Free Will." *Midwest Studies In Philosophy* 29 (1): 146–62.
———. 2006. *Free Will and Luck*. Oxford University Press, USA.
———. 2007. "Free Will: Action Theory Meets Neuroscience." In *Intentionality, Deliberation, and Autonomy: The Action-Theoretic Basis of Practical Philosophy*, edited by C. Lumer. Ashgate.

Meloy, J. Reid. 1988. *The Psychopathic Mind: Origins, Dynamics, and Treatment*. The Psychopathic Mind: Origins, Dynamics, and Treatment. Lanham, MD, US: Jason Aronson.

———. 1997. "The Clinical Risk Management of stalking: Someone is watching over me…" *American Journal of Psychotherapy* 51 (2): 174–84.

Mier, Daniela, Leila Haddad, Kersten Diers, Harald Dressing, Andreas Meyer-Lindenberg, and Peter Kirsch. 2014. "Reduced Embodied Simulation in Psychopathy." *The World Journal of Biological Psychiatry: The Official Journal of the World Federation of Societies of Biological Psychiatry* 15 (6): 479–87.

Miller, C. Arden, and Margaret A. Lennox. 1948. "Electroencephalography in Behavior Problem Children." *The Journal of Pediatrics* 33 (6): 753–60.

Milne, H. B. 1979. "Epileptic Homicide: Drug-Induced." *The British Journal of Psychiatry* 134 (5): 547–48.

Milstein, Victor. 1988. "EEG Topography in Patients with Aggressive Violent Behavior." In *Biological Contributions to Crime Causation*, edited by Terrie E. Moffitt and Sarnoff A. Mednick, 40–52. NATO ASI Series. Dordrecht: Springer Netherlands.

Mischel, Walter, and Ebbe B. Ebbesen. 1970. "Attention in Delay of Gratification." *Journal of Personality and Social Psychology* 16 (2): 329–37.

Missa, Jean-Noël. 2006. *Naissance de la psychiatrie biologique : Histoire des traitements des maladies mentales au XX^e^ siècle*. Science, Histoire, Société. Presses universitaires de France.

———. 2013. "De la leucotomie à la stimulation électrique profonde : aspects historiques et éthiques de la psychochirurgie." *Bulletin d'histoire et d'épistemologie des sciences de la vie* Volume 20 (1): 103–19.

Moffitt, T. E., and Sarnoff A. Mednick, eds. 1988. *Biological Contributions to Crime Causation*. Springer Netherlands.

Moffitt, Terrie E. 1990. "The Neuropsychology of Juvenile Delinquency: A Critical Review." *Crime and Justice* 12 (January): 99–169.

Morange, Michel. 2005. "Quelle place pour l'épigénétique ?" *Médecine/Sciences, 2005, Vol. 21, N° 4; p. 367-369*.

Morse, Stephen J. 2003. "Diminished Rationality, Diminished Responsibility." *Faculty Scholarship at Penn Law* 525 (January).

———. 2006. "Brain Overclaim Syndrome and Criminal Responsibility: A Diagnostic Note." *State Journal of Criminal Law* Vol. 3 (April): 397.

———. 2007. "The Non-Problem of Free Will in Forensic Psychiatry and Psychology." *Behavioral Sciences & the Law* 25 (2): 203–20.

———. 2010. "Lost in Translation?: An Essay on Law and Neuroscience." *Faculty Scholarship at Penn Law* 368 (January).

———. 2011. "Avoiding Irrational NeuroLaw Exuberance: A Plea for Neuromodesty." *Law, Innovation and Technology* 3 (2): 209–28.

———. 2015. "Neuroscience, Free Will, and Criminal Responsibility." *Faculty Scholarship at Penn Law* 1604 (January).

Motzkin, Julian C., Joseph P. Newman, Kent A. Kiehl, and Michael Koenigs. 2011. "Reduced Prefrontal Connectivity in Psychopathy." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 31 (48): 17348–57.

Mucchielli, Laurent. 1994. *Histoire de la criminologie française*. Collection "Histoire des sciences humaines." Paris: Editions L'Harmattan.

———. 2000. "Criminologie, hygiénisme et eugénisme en France (1870-1914) : débats médicaux sur l'élimination des criminels réputés « incorrigibles », *Revue d'histoire des sciences humaines* no 3 (2): 57–88.

Müller, Jürgen L. 2010. “Psychopathy - an Approach to Neuroscientific Research in Forensic Psychiatry.” *Behavioral Sciences & the Law* 28 (2): 129–47.

Müller, Jürgen L., Monika Sommer, Verena Wagner, Kirsten Lange, Heidrun Taschler, Christian H. Röder, Gerhardt Schuierer, Helmfried E. Klein, and Göran Hajak. 2003. “Abnormalities in Emotion Processing within Cortical and Subcortical Regions in Criminal Psychopaths: Evidence from a Functional Magnetic Resonance Imaging Study Using Pictures with Emotional Content.” *Biological Psychiatry* 54 (2): 152–62.

Mullins-Sweatt, Stephanie N., Natalie G. Glover, Karen J. Derefinko, Joshua D. Miller, and Thomas A. Widiger. 2010. “The Search for the Successful Psychopath.” *Journal of Research in Personality* 44 (4): 554–58.

Mundale, Jennifer, and William Bechtel. 1996. “Integrating Neuroscience, Psychology and Evolutionary Biology through a Teleological Conception of Function.” *Minds and Machines: Journal for Artificial Intelligence, Philosophy and Cognitive Science* 6 (4): 481–505.

Mundy-Castle, A. C. 1953. “Electrical Responses of the Brain in Relation to Behaviour.” *British Journal of Psychology (London, England: 1953)* 44 (4): 318–29.

Nachshon, Israel. 1988. “Hemisphere Function in Violent Offenders.” In *Biological Contributions to Crime Causation*, edited by Terrie E. Moffitt and Sarnoff A. Mednick, 55–67. NATO ASI Series. Dordrecht: Springer Netherlands.

Nahmias, Eddy. 2010. “Scientific Challenges to Free Will.” In *A Companion to the Philosophy of Action*, edited by C. Sandis and T. O’Connor, 345–356. Wiley-Blackwell.

Niv, Sharon, Syed Ashrafulla, Anand Joshi, Adrian Raine, Richard Leahy, Laura A. Baker, and Catherine Tuvblad. 2018. “Relationships of Alpha, Beta, and Theta EEG Spectra Properties with Aggressive and Nonaggressive Antisocial Behavior in Children and Adolescents.” *The American Journal of Psychology* 131 (4): 429–37.

Niv, Sharon, Syed Ashrafulla, Catherine Tuvblad, Anand Joshi, Adrian Raine, Richard Leahy, and Laura A. Baker. 2015. “Childhood EEG Frontal Alpha Power as a Predictor of Adolescent Antisocial Behavior: A Twin Heritability Study.” *Biological Psychology* 105 (February): 72–76.

Noë, Alva. 2009. *Out of Our Heads: Why You Are Not Your Brain, and Other Lessons from the Biology of Consciousness*. Hill & Wang.

Noë, Alva, and Evan Thompson. 2004. “Are There Neural Correlates of Consciousness?” *Journal of Consciousness Studies* 11 (January): 3–28.

Nordstrom, Benjamin R., Yu Gao, Andrea L. Glenn, Melissa Peskin, Anna S. Rudo-Hutt, Robert A. Schug, Yaling Yang, and Adrian Raine. 2011. “Neurocriminology.” In *Advances in Genetics*, edited by Danika L. Bannasch and Patricia Brennan Robert Huber, 75:255–83. Aggression. Academic Press.

Okasha, Ahmed, A. Sadek, and S. Abdel Moneim. 1975. “Psychosocial and Electroencephalographic Studies of Egyptian Murderers.” *The British Journal of Psychiatry* 126 (1): 34–40.

Olson, William H., Frederic A. Gibbs, and Charles L. Adams. 1970. “Electroencephalographic Study of Criminals.” *Clinical Electroencephalography* 1 (3): 92–100.

Ostow, M. 1946. “Bilaterally Synchronous Paroxysmal Slow Activity in the Electroencephalograms of Non-Epileptics.” *The Journal of Nervous and Mental Disease* 103 (April): 346–58.

Oxley, Douglas R., Kevin B. Smith, John R. Alford, Matthew V. Hibbing, Jennifer L. Miller, Mario Scalora, Peter K. Hatemi, and John R. Hibbing. 2008. “Political Attitudes Vary with Physiological Traits.” *Science* 321 (5896): 1667–70.

Oyama, Susan. 2000. *The Ontogeny of Information: Developmental Systems and Evolution*. Duke University Press.

Pacherie, Elisabeth. 2013. “Can Conscious Agency Be Saved?” *Topoi* 33 (1): 33–45.
Pacherie, Elisabeth, and Patrick Haggard. 2010. “What Are Intentions?” In *Conscious Will and Responsibility. A Tribute to Benjamin Libet*, edited by L. Nadel and W. Sinnott-Armstrong, 70–84. Oxford University Press.
Pacholczyk, Anna. 2011. “Moral Enhancement: What Is It and Do We Want It?” *Law, Innovation and Technology* 3 (2): 251–77.
Pallone, Nathaniel J., and James J. Hennessy. 1998. “Brain Dysfunction and Criminal Violence.” *Society* 35 (6): 21–27.
Pandya, Neil S, Mirna Vrbancic, Lady Diana Ladino, and José F Téllez-Zenteno. 2013. “Epilepsy and Homicide.” *Neuropsychiatric Disease and Treatment* 9: 667–73.
Pardo, Michael S., and Dennis Patterson. 2011. “Minds, Brains, and Norms.” *Neuroethics* 4 (3): 179–90.
———. 2015. *Minds, Brains, and Law: The Conceptual Foundations of Law and Neuroscience*. Oxford University Press.
Patrick, Christopher J. 2008. “Psychophysiological Correlates of Aggression and Violence: An Integrative Review.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 363 (1503): 2543–55.
———. 2014. “Physiological Correlates of Psychopathy, Antisocial Personality Disorder, Habitual Aggression, and Violence.” In *Electrophysiology and Psychophysiology in Psychiatry and Psychopharmacology*, edited by Veena Kumari, Petr Bob, and Nash N. Boutros, 197–227. Current Topics in Behavioral Neurosciences. Springer.
Paul, G. M., and K. W. Lange. 1992. “Epilepsy and Criminal Law.” *Medicine, Science, and the Law* 32 (2): 160–66..
Penavayre, Marie, Cédric Brun, et Thomas Boraud. 2019. “Neurobiologie des jugements moraux, avancée épistémique ou voie sans issue ?” *Intellectica*, no. 70 (Janvier): 63–82.
Perron, Brian E., and Matthew O. Howard. 2008. “Prevalence and Correlates of Traumatic Brain Injury among Delinquent Youths.” *Criminal Behaviour and Mental Health: CBMH* 18 (4): 243–55.
Persson, Ingmar, and Julian Savulescu. 2008. “The Perils of Cognitive Enhancement and the Urgent Imperative to Enhance the Moral Character of Humanity.” *Journal of Applied Philosophy* 25 (3): 162–77.
———. 2013. “Getting Moral Enhancement Right: The Desirability of Moral Bioenhancement.” *Bioethics* 27 (3): 124–31.
Petersen, Thomas Søbirk. 2017. “Should Neurotechnological Treatments Offered to Offenders Be in Their Best Interests?” *Jounal of Medical Ethics*.
Peterson, Carly K., Alexander J. Shackman, and Eddie Harmon-Jones. 2008. “The Role of Asymmetrical Frontal Cortical Activity in Aggression.” *Psychophysiology* 45 (1): 86–92.
Philippi, Carissa L., Maia S. Pujara, Julian C. Motzkin, Joseph Newman, Kent A. Kiehl, and Michael Koenigs. 2015. “Altered Resting-State Functional Connectivity in Cortical Networks in Psychopathy.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 35 (15): 6068–78.
Piaget, Jean. 1967. *Biologie et connaissance: essai sur les relations entre les régulations organiques et les processus cognitifs*. Gallimard.
Pillmann, F., A. Rohde, S. Ullrich, S. Draba, U. Sannemüller, and A. Marneros. 1999. “Violence, Criminal Behavior, and the EEG: Significance of Left Hemispheric Focal Abnormalities.” *The Journal of Neuropsychiatry and Clinical Neurosciences* 11 (4): 454–57.
Pincus, J. H. 1980. “Can Violence Be a Manifestation of Epilepsy?” *Neurology* 30 (3): 304–7.

Poeppl, Timm B., Simon B. Eickhoff, Peter T. Fox, Angela R. Laird, Rainer Rupprecht, Berthold Langguth, and Danilo Bzdok. 2015. “Connectivity and Functional Profiling of Abnormal Brain Structures in Pedophilia.” *Human Brain Mapping* 36 (6): 2374–86.

Poeppl, Timm B., Joachim Nitschke, Beate Dombert, Pekka Santtila, Mark W. Greenlee, Michael Osterheider, and Andreas Mokros. 2011. “Functional Cortical and Subcortical Abnormalities in Pedophilia: A Combined Study Using a Choice Reaction Time Task and fMRI.” *The Journal of Sexual Medicine* 8 (6): 1660–74.

Poeppl, Timm B., Joachim Nitschke, Pekka Santtila, Martin Schecklmann, Berthold Langguth, Mark W. Greenlee, Michael Osterheider, and Andreas Mokros. 2013. “Association between Brain Structure and Phenotypic Characteristics in Pedophilia.” *Journal of Psychiatric Research* 47 (5): 678–85.

Poirier, Pierre, et Luc Faucher. 2011. *Des sciences cognitives évolutionnaires doublement externalistes*. Editions Matériologiques.

Poirier, Pierre, et Nicolas Payette. 2007. “Les gardiens du bon usage : étude critique de « Philosophical Foundations of Neuroscience », de P. M. R. Hacker et M. R. Bennett.” *Philosophiques* 34 (1): 183–200.

Poldrack, Russell A. 2011. “Inferring Mental States from Neuroimaging Data: From Reverse Inference to Large-Scale Decoding.” *Neuron* 72 (5): 692–97.

Poldrack, Russell A., John Monahan, Peter B. Imrey, Valerie Reyna, Marcus E. Raichle, David Faigman, and Joshua W. Buckholtz. 2018. “Predicting Violent Behavior: What Can Neuroscience Add?” *Trends in Cognitive Sciences* 22 (2): 111–23.

Polich, John, and Albert Kok. 1995. “Cognitive and Biological Determinants of P300: An Integrative Review.” *Biological Psychology* 41 (2): 103–46.

Ponseti, Jorge, Daniel Bruhn, Julia Nolting, Hannah Gerwinn, Alexander Pohl, Aglaja Stirn, Oliver Granert, et al. 2018. “Decoding Pedophilia: Increased Anterior Insula Response to Infant Animal Pictures.” *Frontiers in Human Neuroscience* 11.

Ponseti, Jorge, Oliver Granert, Olav Jansen, Stephan Wolff, Klaus Beier, Janina Neutze, Günther Deuschl, Hubertus Mehdorn, Hartwig Siebner, and Hartmut Bosinski. 2012. “Assessment of Pedophilia Using Hemodynamic Brain Response to Sexual Stimuli.” *Archives of General Psychiatry* 69 (2): 187–94.

Potochnik, Angela. 2017. *Idealization and the Aims of Science*. University of Chicago Press.

Prehn, Kristin, Florian Schlagenhauf, Lars Schulze, Christoph Berger, Knut Vohs, Monika Fleischer, Karlheinz Hauenstein, Peter Keiper, Gregor Domes, and Sabine C. Herpertz. 2013. “Neural Correlates of Risk Taking in Violent Criminal Offenders Characterized by Emotional Hypo- and Hyper-Reactivity.” *Social Neuroscience* 8 (2): 136–47.

Preu, P.W. 1944. “The Concept of Psychopathic Personality.” *Personality and the Behavior Disorders, Ed Hunt J. M., Editor. (New York, NY: The Ronald Press Company)*, 922–937.

Proust, Joëlle, et Elisabeth Pacherie. 2008. “Neurosciences et compréhension d’autrui.” In *Philosophie et Neurosciences,* Pierre Poirier et Luc Faucher. Bellarmin.

Pugh, Jonathan. 2018. “Coercion and the Neurocorrective Offer.” In *Treatment for Crime: Philosophical Essays on Neurointerventions in Criminal Justice*, David Rhys Birks and Thomas Douglas.

Pustilnik, Amanda C. 2009. “Violence on the Brain: A Critique of Neuroscience in Criminal Law.” *Wake Forest Law Review* 44 (1): 183–238.

Quirion, Bastien. 2006. “Traiter les délinquants ou contrôler les conduites : le dispositif thérapeutique à l’ère de la nouvelle pénologie.” *Criminologie* 39 (2): 137–64.

Racine, Eric. 2008. “Tendances récentes en neuroéthique.” In *Des neurosciences à la philosophie : neurophilosophie et philosophie des neurosciences*, Syllepse, 271–92. Pierre Poirier et Luc Faucher.

Racine, Eric, Ofek Bar-Ilan, and Judy Illes. 2005. “fMRI in the Public Eye.” *Nature Reviews. Neuroscience* 6 (2): 159–64.
Rafter, Nicole Hahn. 2008. *The Criminal Brain: Understanding Biological Theories of Crime*. New York: New York University Press.
Raine, Adrian. 1993. *The Psychopathology of Crime: Criminal Behavior as a Clinical Disorder*. Academic Press.
———. 2002. “Biosocial Studies of Antisocial and Violent Behavior in Children and Adults: A Review.” *Journal of Abnormal Child Psychology* 30 (4): 311–26.
———. 2013. *The Anatomy of Violence: The Biological Roots of Crime*. London: Penguin.
Raine, Adrian, Peter H. Venables, and Mark Williams. 1990. “Relationships between Central and Autonomic Measures of Arousal at Age 15 Years and Criminality at Age 24 Years.” *Archives of General Psychiatry* 47 (11): 1003–7.
Raine, Adrian, Monte Buchsbaum, Jill Stanley, Steven Lottenberg, Leonard Abel, and Jacqueline Stoddard. 1994. “Selective Reductions in Prefrontal Glucose Metabolism in Murderers.” *Biological Psychiatry* 36 (6): 365–73.
Raine, Adrian, Monte Buchsbaum, and Lori LaCasse. 1997a. “Brain Abnormalities in Murderers Indicated by Positron Emission Tomography.” *Biological Psychiatry* 42 (6): 495–508.
Raine, Adrian, Chandra Reynolds, Peter H. Venables, and Sarnoff A. Mednick. 1997b. “Biosocial Bases of Aggressive Behavior in Childhood.” In *Biosocial Bases of Violence*, edited by Adrian Raine, Patricia A. Brennan, David P. Farrington, and Sarnoff A. Mednick, 107–26. Nato ASI Series. Boston, MA: Springer US.
Raine, Adrian, D. Phil, Jacqueline Stoddard, Susan Bihrle, and Monte Buchsbaum. 1998a. “Prefrontal Glucose Deficits in Murderers Lacking Psychosocial Deprivation.” *Neuropsychiatry, Neuropsychology, and Behavioral Neurology* 11 (1): 1–7.
Raine, Adrian, J. Reid Meloy, Susan Bihrle, Jacqueline Stoddard, Lori LaCasse, and Monte Buchsbaum. 1998b. “Reduced Prefrontal and Increased Subcortical Brain Functioning Assessed Using Positron Emission Tomography in Predatory and Affective Murderers.” *Behavioral Sciences & the Law* 16 (3): 319–32.
Raine, Adrian, Sohee Park, Todd Lencz, Susan Bihrle, Lori LaCasse, Cathy Spatz Widom, Louai Al-Dayeh, and Manbir Singh. 2001. “Reduced Right Hemisphere Activation in Severely Abused Violent Offenders during a Working Memory Task: An fMRI Study.” *Aggressive Behavior* 27 (2): 111–29.
Raine, Adrian, Sharon S. Ishikawa, Estibaliz Arce, Todd Lencz, Kevin H. Knuth, Susan Bihrle, Lori LaCasse, and Patrick Colletti. 2004. “Hippocampal Structural Asymmetry in Unsuccessful Psychopaths.” *Biological Psychiatry* 55 (2): 185–91.
Raine, Adrian, Kenneth Dodge, Rolf Loeber, Lisa Gatzke-Kopp, Don Lynam, Chandra Reynolds, Magda Stouthamer-Loeber, and Jianghong Liu. 2006. “The Reactive–Proactive Aggression Questionnaire: Differential Correlates of Reactive and Proactive Aggression in Adolescent Boys.” *Aggressive Behavior* 32 (2): 159–71.
Raine, Adrian, and Yaling Yang. 2006. “Neural Foundations to Moral Reasoning and Antisocial Behavior.” *Social Cognitive and Affective Neuroscience* 1 (3): 203–13.
Redding, Richard E. 2006. “The Brain-Disordered Defendant: Neuroscience and Legal Insanity in the Twenty-First Century.” *The American University Law Review* 56 (1): 51–127.
Reimer, Marga. 2008. “Psychopathy Without (the Language of) Disorder.” *Neuroethics* 1 (3): 185–98.
Reinach, Joseph. 1882. *Les récidivistes*. Paris, Georges Charpentier.
Renneville, Marc. 1994. “L’anthropologie du criminel en France.” *Criminologie* 27 (2): 185–209.

———. 1995. “La réception de Lombroso en France (1880-1900).” In *Histoire de la criminologie française*, edited by Laurent Mucchielli, 107–35. Histoire des sciences de l’homme. Paris: Editions L’Harmattan.
———. 1997. “Rationalité contextuelle et présupposé cognitif. Réflexion épistémologique sur le cas Lombroso.” *Revue de Synthèse*, no. 4: 495–529.
———. 2000. *Le Langage des crânes : Une histoire de la phrénologie*. Paris, La Découverte, Empêcheurs de penser en rond.
———. 2003. *Crime et folie : Deux siècles d’enquêtes médicales et judiciaires*. Paris: Fayard.
Reuber, Markus, and Ronnie D. Mackay. 2008. “Epileptic Automatisms in the Criminal Courts: 13 Cases Tried in England and Wales between 1975 and 2001.” *Epilepsia* 49 (1): 138–45.
Ridder, Dirk, Berthold Langguth, Mark Plazier, and Tomas Menovsky. 2009. “Moral Dysfunction: Theoretical Model and Potential Neurosurgical Treatments.” In *The Moral Brain: Essays on the Evolutionary and Neuroscientific Aspects of Morality*, 155–83.
Riva, Paolo, Leonor J. Romero Lauro, C. Nathan DeWall, David S. Chester, and Brad J. Bushman. 2015. “Reducing Aggressive Responses to Social Exclusion Using Transcranial Direct Current Stimulation.” *Social Cognitive and Affective Neuroscience* 10 (3): 352–56.
Rizzolatti, G., R. Camarda, L. Fogassi, M. Gentilucci, G. Luppino, and M. Matelli. 1988. “Functional Organization of Inferior Area 6 in the Macaque Monkey.” *Experimental Brain Research* 71 (3): 491–507.
Rockwell, W. Teed. 2005. *Neither Brain Nor Ghost: A Nondualist Alternative to the Mind-Brain Identity Theory*. MIT Press.
Rocque, Michael, Brandon C. Welsh, and Adrian Raine. 2012. “Biosocial Criminology and Modern Crime Prevention.” *Journal of Criminal Justice* 40 (4): 306–12.
Rodin, Ernst A. 1973. “Psychomotor Epilepsy and Aggressive Behavior.” *Archives of General Psychiatry* 28 (2): 210–13.
Rodman, Alexandra M., Erik Kastman, Hayley M. Dorfman, Arielle Baskin-Sommers, Kent A. Kiehl, Joseph P. Newman, and Joshua W. Buckholtz. 2016. “Selective Mapping of Psychopathy and Externalizing to Dissociable Circuits for Inhibitory Self-Control.” *Clinical Psychological Science: A Journal of the Association for Psychological Science* 4 (3): 559–71.
Rose, Nikolas, and Joelle M. Abi-Rached. 2013. *Neuro: The New Brain Sciences and the Management of the Mind*. Princeton University Press.
Roskies, Adina L. 2002. “Neuroethics for the New Millenium.” *Neuron* 35 (1): 21–23.
———. 2007. “Are Neuroimages Like Photographs of the Brain?” *Philosophy of Science* 74 (5): 860–872.
———. 2015. “Agency and Intervention.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 370 (1677): 20140215.
Rowlands, Mark. 1999. *The Body in Mind: Understanding Cognitive Processes*. Cambridge University Press.
Rudo-Hutt, Anna, Yu Gao, Andrea Glenn, Melissa Peskin, Yaling Yang, and Adrian Raine. 2011. “Biosocial Interactions and Correlates of Crime.” *The Ashgate Research Companion to Biosocial Theories of Crime*, January, 17–44.
Ryberg, Jesper. 2012. “Punishment, Pharmacological Treatment, and Early Release.” *International Journal of Applied Philosophy* 26 (2): 231–244.
Ryle, Gilbert. 1949. *The Concept of Mind*. University of Chicago Press.
Saks, Michael J., N. J. Schweitzer, Eyal Aharoni, and Kent A. Kiehl. 2014. “The Impact of Neuroimages in the Sentencing Phase of Capital Trials.” *Journal of Empirical Legal Studies* 11 (1): 105–31.

Salmon, Wesley C. 1984. *Scientific Explanation and the Causal Structure of the World*. Princeton University Press.
Sartorius, Alexander, Matthias Ruf, Christine Kief, Traute Demirakca, Josef Bailer, Gabriele Ende, Fritz A. Henn, Andreas Meyer-Lindenberg, and Harald Dressing. 2008. "Abnormal Amygdala Activation Profile in Pedophilia." *European Archives of Psychiatry and Clinical Neuroscience* 258 (5): 271–77.
Saul, L. J., H. Davis, and P. A. Davis. 1949. "Psychologic Correlations with the Electroencephalogram." *Psychosomatic Medicine* 11: 361–76.
Savulescu, Julian, and Ingmar Persson. 2012. "Moral Enhancement, Freedom and the God Machine." *The Monist* 95 (3): 399–421.
Sayed, Z. A., S. A. Lewis, and R. P. Brittain. 1969. "An Electroencephalographic and Psychiatric Study of Thirty-Two Insane Murderers." *The British Journal of Psychiatry* 115 (527): 1115–24.
Scarpa, Angela, and Adrian Raine. 2007. "Biosocial Bases of Violence." In *The Cambridge Handbook of Violent Behavior and Aggression*, 151–69. New York: Cambridge University Press.
Schechtman, Marya. 2010. "Philosophical Reflections on Narrative and Deep Brain Stimulation." *The Journal of Clinical Ethics* 21 (2): 133–39.
Schermer, Maartje. 2011. "Ethical Issues in Deep Brain Stimulation." *Frontiers in Integrative Neuroscience* 5.
Schiele, B. C., A. B. Baker, and S. R. Hathaway. 1943. "The Minnesota Multiphasic Personality Inventory." *Journal-Lancet* 63: 292–97.
Schienle, Anne, Albert Wabnegger, Mario Leitner, and Verena Leutgeb. 2017. "Neuronal Correlates of Personal Space Intrusion in Violent Offenders." *Brain Imaging and Behavior* 11 (2): 454–60.
Schiffer, Boris, Tillmann Krueger, Thomas Paul, Armin de Greiff, Michael Forsting, Norbert Leygraf, Manfred Schedlowski, and Elke Gizewski. 2008. "Brain Response to Visual Sexual Stimuli in Homosexual Pedophiles." *Journal of Psychiatry & Neuroscience : JPN* 33 (1): 23–33.
Schiltz, Kolja, Joachim G. Witzel, Josef Bausch-Hölterhoff, and Bernhard Bogerts. 2013. "High Prevalence of Brain Pathology in Violent Prisoners: A Qualitative CT and MRI Scan Study." *European Archives of Psychiatry and Clinical Neuroscience* 263 (7): 607–16.
Schiltz, Kolja, Joachim Witzel, Georg Northoff, Kathrin Zierhut, Udo Gubka, Hermann Fellmann, Jörn Kaufmann, Claus Tempelmann, Christine Wiebking, and Bernhard Bogerts. 2007. "Brain Pathology in Pedophilic Offenders: Evidence of Volume Reduction in the Right Amygdala and Related Diencephalic Structures." *Archives of General Psychiatry* 64 (6): 737–46.
Schirmann, Felix. 2014. "'The Wondrous Eyes of a New Technology'—a History of the Early Electroencephalography (EEG) of Psychopathy, Delinquency, and Immorality." *Frontiers in Human Neuroscience* 8.
Schmalleger, Frank. 1996. *Criminology Today*. Prentice Hall.
Schweitzer, N. J., and Michael J. Saks. 2011. "Neuroimage Evidence and the Insanity Defense." *Behavioral Sciences & the Law* 29 (4): 592–607.
Schweitzer, N. J., Michael J. Saks, Emily R. Murphy, Adina L. Roskies, Walter Sinnott-Armstrong, and Lyn M. Gaudet. 2011. "Neuroimages as Evidence in a Mens Rea Defense: No Impact." *Psychology, Public Policy, and Law* 17 (3): 357–93.
Secunda, Lazarus, and Knox H. Finley. 1942. "Electroencephalographic Studies in Children Presenting Behavior Disorders." *New England Journal of Medicine* 226 (21): 850–54.

Sellaro, Roberta, Berna Güroğlu, Michael A. Nitsche, Wery P. M. van den Wildenberg, Valentina Massaro, Jeffrey Durieux, Bernhard Hommel, and Lorenza S. Colzato. 2015. "Increasing the Role of Belief Information in Moral Judgments by Stimulating the Right Temporoparietal Junction." *Neuropsychologia* 77 (October): 400–408.

Senon, Jean-Louis. 2008. "Troubles de la personnalité et psychiatrie face au courant d'insécurité de la société : de la nécessité de penser le champ du soin face aux peurs collectives." *L'information psychiatrique* Volume 84 (3): 241–47.

Senon, Jean-Louis, Mélanie Voyer, Christelle Paillard, and Nemat Jaafari. 2009. "Dangerosité criminologique : données contextuelles, enjeux cliniques et expertaux." *L'information psychiatrique* Volume 85 (8): 719–25.

Serafetinides, E. A. 1965. "Aggressiveness in Temporal Lobe Epileptics and Its Relation to Cerebral Dysfunction and Environmental Factors." *Epilepsia* 6 (1): 33–42.

Sergent, Justine. 1994. "Brain-Imaging Studies of Cognitive Functions." *Trends in Neurosciences* 17 (6): 221–27.

Seruca, Tânia, and Carlos F. Silva. 2015. "Recidivist Criminal Behaviour and Executive Functions: A Comparative Study." *The Journal of Forensic Psychiatry & Psychology* 26 (5): 699–717.

Seward, James D. 1987. "Epilepsy, Violence and Aggression: Legal Implications." *Journal of Police and Criminal Psychology* 3 (1): 35–43.

Sharma, Nikita, Srashti Verma, and Pooja Batra Nagpal. 2018. "A Review of Brain Fingerprinting Process." *A Review of Brain Fingerprinting Process* 12 (1): 5–5.

Shaw, Elizabeth. 2014. "Direct Brain Interventions and Responsibility Enhancement." *Criminal Law and Philosophy* 8 (1): 1–20.

———. 2018. "Against the Mandatory Use of Neurointerventions in Criminal Sentencing." In *Treatment for Crime: Philosophical Essays on Neurointerventions in Criminal Justice*. Oxford University Press.

———. 2019. "The Right to Bodily Integrity and the Rehabilitation of Offenders Through Medical Interventions: A Reply to Thomas Douglas." *Neuroethics* 12 (1): 97–106.

Shen, Francis X. 2016. "The Overlooked History of Neurolaw." *Fordham Law Review* 85 (2): 667.

Shen, Francis X., and Owen D. Jones. 2011. "Brain Scans as Evidence: Truths, Proofs, Lies, and Lessons." SSRN Scholarly Paper ID 1736288. Rochester, NY: Social Science Research Network.

Shih, Jerry J., Thabele LeslieMazwi, Germano Falcao, and Jay Van Gerpen. 2009. "Directed Aggressive Behavior in Frontal Lobe Epilepsy: A Video-EEG and Ictal Spect Case Study." *Neurology* 73 (21): 1804–6.

Shoemaker, David. 2011. "Psychopathy, Responsibility, and the Moral/Conventional Distinction." *The Southern Journal of Philosophy* 49 (September): 99–124.

Shook, John R. 2012. "Neuroethics and the Possible Types of Moral Enhancement." *AJOB Neuroscience* 3 (4): 3–14.

Siep, Nicolette, Franca Tonnaer, Vincent van de Ven, Arnoud Arntz, Adrian Raine, and Maaike Cima. 2018. "Anger Provocation Increases Limbic and Decreases Medial Prefrontal Cortex Connectivity with the Left Amygdala in Reactive Aggressive Violent Offenders." *Brain Imaging and Behavior*, August.

Siever, Larry J. 2008. "Neurobiology of Aggression and Violence." *The American Journal of Psychiatry* 165 (4): 429–42.

Silverman, Daniel. 1943. "Clinical and Electroencephalographic Studies on Criminal Psychopaths." *Archives of Neurology & Psychiatry* 50 (1): 18–33.

———. 1944. "The Electroencephalogram of Criminals: Analysis of Four Hundred and Eleven Cases." *Archives of Neurology & Psychiatry* 52 (1): 38–42.

Sinnott-Armstrong, Walter, Adina Roskies, Teneille Brown, and Emily Murphy. 2008. “Brain Images as Legal Evidence.” *Episteme* 5 (3): 359–73.

Sirgiovanni, Elisabetta. 2017. “Criminal Heredity: The Influence of Cesare Lombroso’s Concept of the ‘born Criminal’ on Contemporary Neurogenetics and Its Forensic Applications.” *Medicina Nei Secoli* 29 (1): 165–88.

Sisson, Boyd D., and Robert J. Ellingson. 1955. “On the Relationship between ‘normal’ EEG Patterns and Personality Variables.” *Journal of Nervous and Mental Disease* 121: 353–58.

Smith, S. M., L. Honigsberger, and C. A. Smith. 1973. “E.E.G. and Personality Factors in Baby Batterers.” *British Medical Journal* 3 (5870): 20–22.

Snaith, R. P. 1994. “Psychosurgery: Controversy and Enquiry.” *The British Journal of Psychiatry* 165 (5): 582–84.

Soderstrom, H., M. Tullberg, C. Wikkelsö, S. Ekholm, and A. Forsman. 2000. “Reduced Regional Cerebral Blood Flow in Non-Psychotic Violent Offenders.” *Psychiatry Research* 98 (1): 29–41.

Soon, Chun, Marcel Brass, Hans-Jochen Heinze, and John-Dylan Haynes. 2008. “Unconscious Determinants of Free Decisions in the Human Brain.” *Nature Neuroscience* 11 (June): 543–45.

Soon, Chun Siong, Anna Hanxi He, Stefan Bode, and John-Dylan Haynes. 2013. “Predicting Free Choices for Abstract Intentions.” *Proceedings of the National Academy of Sciences* 110 (15): 6217–22.

Spence, Sean A, Mike D Hunter, Tom F D Farrow, Russell D Green, David H Leung, Catherine J Hughes, and Venkatasubramanian Ganesan. 2004. “A Cognitive Neurobiological Account of Deception: Evidence from Functional Neuroimaging.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 359 (1451): 1755–62.

Stafford-Clark, D., and F. H. Taylor. 1949. “Clinical and Electro-Encephalographic Studies of Prisoners Charged with Murder.” *Journal of Neurology, Neurosurgery, and Psychiatry* 12 (4): 325–30.

Stafford-Clark, D., Desmond Pond, and J.W. Lovett Doust. 1951. “The Psychopath in Prison: A Preliminary Report of a Co-Operative Research.” *British Journal of Delinquency* 2: 117–29.

Steinberg, Laurence, and Elizabeth S. Scott. 2003. “Less Guilty by Reason of Adolescence: Developmental Immaturity, Diminished Responsibility, and the Juvenile Death Penalty.” *American Psychologist* 58 (12): 1009–18.

Steinberg, Laurence. 2013. “The Influence of Neuroscience on US Supreme Court Decisions about Adolescents’ Criminal Culpability.” *Nature Reviews Neuroscience* 14 (7): 513–18.

Sternberg, Eliezer J. 2010. *My Brain Made Me Do It: The Rise of Neuroscience and the Threat to Moral Responsibility*. Prometheus Books.

Stevens, J. R., and B. P. Hermann. 1981. “Temporal Lobe Epilepsy, Psychopathology, and Violence: The State of the Evidence.” *Neurology* 31 (9): 127–127.

Strauss, Hans, W. E. Jr Rahm, and S. E. Barrera. 1940. “Studies on a Group of Children with Psychiatric Disorders. I. Electroencephalographic Studies.” *Psychosomatic Medicine* 2 (1): 34.

Strickland, Justin C., and Matthew W. Johnson. 2020. “Rejecting Impulsivity as a Psychological Construct: A Theoretical, Empirical, and Sociocultural Argument.” *Psychological Review*.

Sullivan, Harry Stack. 1924. “Schizophrenia: Its Conservative and Malignant Features.” *American Journal of Psychiatry* 81 (1): 77–91.

Syndulko, K. 1978. “Electrocortical Investigations of Sociopathy.” In *Psychopathic Behaviour. Approaches to Research*, John Wiley & Sons: New York, 145–56. R. D. Hare and D. Schalling.

Tarde, Gabriel. 1885. “Le type criminel.” *Revue philosophique de la France et de l’étranger* 19: 593–627.

Tassy, Sébastien, Olivier Oullier, Yann Duclos, Olivier Coulon, Julien Mancini, Christine Deruelle, Sharam Attarian, Olivier Felician, and Bruno Wicker. 2012. “Disrupting the Right Prefrontal Cortex Alters Moral Judgement.” *Social Cognitive and Affective Neuroscience* 7 (3): 282–88.

Taylor, J. Sherrod, J. Anderson Harp, and Tyron Elliott. 1991. “Neuropsychologists and Neurolawyers.” *Neuropsychology* 5 (4): 293–305.

Taylor, Stuart P. 1967. “Aggressive Behavior and Physiological Arousal as a Function of Provocation and the Tendency to Inhibit aggression.” *Journal of Personality* 35 (2): 297–310.

Terrel, Jean. 2003. *Foucault au Collège de France: un itinéraire*. Presses Universitaires de Bordeaux.

Thijssen, Sandra, and Kent A. Kiehl. 2017. “Functional Connectivity in Incarcerated Male Adolescents with Psychopathic Traits.” *Psychiatry Research* 265 (July): 35–44.

Tierney, Ann Jane. 2000. “Egas Moniz and the Origins of Psychosurgery: A Review Commemorating the 50th Anniversary of Moniz’s Nobel Prize.” *Journal of the History of the Neurosciences* 9 (1): 22–36.

Tiihonen, Jari, Roberta Rossi, Mikko P. Laakso, Sheilagh Hodgins, Cristina Testa, Jorge Perez, Eila Repo-Tiihonen, et al. 2008. “Brain Anatomy of Persistent Violent Offenders: More rather than Less.” *Psychiatry Research* 163 (3): 201–12.

Tonnaer, Franca, Nicolette Siep, Linda van Zutphen, Arnoud Arntz, and Maaike Cima. 2017. “Anger Provocation in Violent Offenders Leads to Emotion Dysregulation.” *Scientific Reports* 7 (1): 1–9.

Topinard, Paul. 1887. “L’anthropologie criminelle.” *Revue d’anthropologie*, 658–91.

Treiman, D. M. 1986. “Epilepsy and Violence: Medical and Legal Issues.” *Epilepsia* 27: 77-104.

———. 1999. “Violence and the Epilepsy Defense.” *Neurologic Clinics* 17 (2): 245–55.

Turner, Andrew. 2010. “Are Disorders Sufficient for Reduced Responsibility?” *Neuroethics* 3 (July): 151–60.

Umbach, Rebecca, Colleen M. Berryessa, and Adrian Raine. 2015. “Brain Imaging Research on Psychopathy: Implications for Punishment, Prediction, and Treatment in Youth and Adults.” *Journal of Criminal Justice* 43 (4): 295–306.

Uttal, William R. 2003. *The New Phrenology: The Limits of Localizing Cognitive Processes in the Brain*. MIT Press.

Vallentyne, Peter. 2018. *Neurointerventions, Self-Ownership, and Enforcement Rights*. Oxford University Press.

Varela, Francisco J. 1992. “Whence Perceptual Meaning? A Cartography of Current Ideas.” In *Understanding Origins: Contemporary Views on the Origin of Life, Mind and Society*, edited by Francisco J. Varela and Jean-Pierre Dupuy, 235–63. Boston Studies in the Philosophy and History of Science. Dordrecht: Springer Netherlands.

Vargas, Manuel. 2010. “Are Psychopathic Serial Killers Evil? Are They Blameworthy for What They Do?” In *Serial Killers and Philosophy*, edited by Sarah Waller. Blackwell.

Verdejo-Román, Juan, Natalia Bueso-Izquierdo, Julia C. Daugherty, Miguel Pérez-García, and Natalia Hidalgo-Ruzzante. 2019. “Structural Brain Differences in Emotional Processing and Regulation Areas between Male Batterers and Other Criminals: A Preliminary Study.” *Social Neuroscience* 14 (4): 390–97.

Vidal, Fernando. 2005. “Le sujet cérébral: une esquisse historique et conceptuelle.” *Psychiatrie Sciences Humaines Neurosciences* 3 (1): 37–48.
Vidal, Fernando, and Francisco Ortega. 2017. *Being Brains: Making the Cerebral Subject*. Fordham University Press.
Vilares, Iris, Michael J. Wesley, Woo-Young Ahn, Richard J. Bonnie, Morris Hoffman, Owen D. Jones, Stephen J. Morse, Gideon Yaffe, Terry Lohrenz, and P. Read Montague. 2017. “Predicting the Knowledge–recklessness Distinction in the Human Brain.” *Proceedings of the National Academy of Sciences* 114 (12): 3222–27.
Villa, Renzo. 1985. *Il deviante e i suoi segni. Lombroso e la nascita dell’antropologia criminale*. Franco Angeli.
Vincent, Nicole A. 2008. “Responsibility, Dysfunction and Capacity.” *Neuroethics* 1 (3): 199–204.
———. 2009. “Responsibility: Distinguishing Virtue from Capacity.” *Polish Journal of Philosophy* 3: 111–26.
———. 2010. “On the Relevance of Neuroscience to Criminal Responsibility.” *Criminal Law and Philosophy* 4 (1): 77–98.
———. 2011. “Capacitarianism, Responsibility and Restored Mental Capacities.” In *Technologies on the Stand: Legal and Ethical Questions in Neuroscience and Robotics.*, Nijmegen: Wolf Legal Publishers., 41–65. L. Klaming & B. Van den Berg.
———. 2013. “Enhancing Responsibility.” In *Neuroscience and Legal Responsibility*, 305–333. Oxford University Press.
———. 2014. “Neurolaw and Direct Brain Interventions.” *Criminal Law and Philosophy* 8 (1): 43–50.
Volavka, Jan. 1987. “Electroencephalogram among Criminals.” In *A. Mednick, T. E. Moffitt and S. Stack (Eds.), The Causes of Crime: New Biological Approaches*, Cambridge University Press, 137–45.
———. 1990. “Aggression, Electroencephalography, and Evoked Potentials: A Critical Review.” *Cognitive and Behavioral Neurology* 3 (4).
———. 1999. “The Neurobiology of Violence: An Update.” *The Journal of Neuropsychiatry and Clinical Neurosciences* 11 (3): 307–14.
———. 2002. *Neurobiology of Violence*. 2nd Revised edition. Washington, DC: American Psychiatric Association Publishing.
———. 2011. “Violent Crime, Epilepsy, and Traumatic Brain Injury.” *PLoS Medicine* 8 (12).
Volkow, N. D., and L. Tancredi. 1987. “Neural Substrates of Violent Behaviour. A Preliminary Study with Positron Emission Tomography.” *The British Journal of Psychiatry: The Journal of Mental Science* 151 (November): 668–73.
Waldbauer, Jacob R., and Michael S. Gazzaniga. 2001. “The Divergence of Neuroscience and Law.” *Jurimetrics* 41 (3): 357–64.
Walker, Charlotte F., and Barbara B. Kirkpatrick. 1947. “Dilantin Treatment for Behavior Problem Children with Abnormal Electroencephalograms.” *American Journal of Psychiatry* 103 (4): 484–92.
Walter, Martin, Joachim Witzel, Christine Wiebking, Udo Gubka, Michael Rotte, Kolja Schiltz, Felix Bermpohl, et al. 2007. “Pedophilia Is Linked to Reduced Activation in Hypothalamus and Lateral Prefrontal Cortex during Visual Erotic Stimulation.” *Biological Psychiatry*, Neural Mechanisms of Personality Disorders, Alcoholism, and Addiction, 62 (6): 698–701.
Walter, Sven, and Miriam Kyselo. 2009. “Fred Adams, Ken Aizawa: The Bounds of Cognition.” *Erkenntnis* 71 (September): 277–81.

Walters, Glenn D., and Kent A. Kiehl. 2015. “Limbic Correlates of Fearlessness and Disinhibition in Incarcerated Youth: Exploring the Brain-Behavior Relationship with the Psychopathy Checklist: Youth Version.” *Psychiatry Research* 230 (2): 205–10.

Wegner, D. M., and T. Wheatley. 1999. “Apparent Mental Causation. Sources of the Experience of Will.” *The American Psychologist* 54 (7): 480–92.

Weiger, W. A., and D. M. Bear. 1988. “An Approach to the Neurology of Aggression.” *Journal of Psychiatric Research* 22 (2): 85–98.

Weinberg, Carl D. 1976. “Epilepsy and the Alternatives for a Criminal Defense Note.” *Case Western Reserve Law Review*, no. 3(1977): 771–803.

Weisberg, Deena Skolnick, Frank C. Keil, Joshua Goodstein, Elizabeth Rawson, and Jeremy R. Gray. 2008. “The Seductive Allure of Neuroscience Explanations.” *Journal of Cognitive Neuroscience* 20 (3): 470–77.

Welsh, George Schlager. 1956. *Basic Readings on the MMPI in Psychology and Medicine*. Edited by W. Grant Dahlstrom. Basic Readings on the MMPI in Psychology and Medicine. Minneapolis, MN, US: University of Minnesota Press.

Whitman, S., T. E. Coleman, C. Patmon, B. T. Desai, R. Cohen, and L. N. King. 1984. “Epilepsy in Prison: Elevated Prevalence and No Relationship to Violence.” *Neurology* 34 (6): 775–775.

Widom, Cathy S. 1977. “A Methodology for Studying Noninstitutionalized Psychopaths.” *Journal of Consulting and Clinical Psychology* 45 (4): 674–83.

Williams, Denis. 1941. “The Significance of an Abnormal Electroencephalogram.” *Journal of Neurology and Psychiatry* 4 (3–4): 257–68.

———. 1969. “Neural Factors Related to Habitual Aggression. Consideration of Differences between Those Habitual Aggressives and Others Who Have Committed Crimes of Violence.” *Brain* 92 (3): 503–20.

Wittgenstein, Ludwig. 1953. *Philosophical Investigations*, New York: Macmillan.

Witzel, Joachim G., Bernhard Bogerts, and Kolja Schiltz. 2016. “Increased Frequency of Brain Pathology in Inmates of a High-Security Forensic Institution: A Qualitative CT and MRI Scan Study.” *European Archives of Psychiatry and Clinical Neuroscience* 266 (6): 533–41.

Wonderly, Monique. 2016. “Treating Psychopaths Fairly.” *AJOB Neuroscience* 7 (3): 158–60.

Wong, M. T. H., J. Lumsden, G. W. Fenton, and P. B. C. Fenwick. 1994. “Electroencephalography, Computed Tomography and Violence Ratings of Male Patients in a Maximum-Security Mental Hospital.” *Acta Psychiatrica Scandinavica* 90 (2): 97–101.

Woods, Sherwyn M. 1961. “Adolescent Violence and Homicide: Ego Disruption and the 6 and 14 Dysrhythmia.” *Archives of General Psychiatry* 5 (6): 528–34.

Woodward, James. 2003. *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press, USA.

———. 2008. “Causation and Manipulability.” In *Stanford Encyclopedia of Philosophy*.

Wright, Percy, Jose Nobrega, Ron Langevin, and George Wortzman. 1990. “Brain Density and Symmetry in Pedophilic and Sexually Aggressive Offenders.” *Annals of Sex Research* 3 (3): 319–28.

Yang, Yaling, and Adrian Raine. 2009. “Prefrontal Structural and Functional Brain Imaging Findings in Antisocial, Violent, and Psychopathic Individuals: A Meta-Analysis.” *Psychiatry Research* 174 (2): 81–88.

Yang, Yaling, Adrian Raine, Patrick Colletti, Arthur W. Toga, and Katherine L. Narr. 2010. “Morphological Alterations in the Prefrontal Cortex and the Amygdala in Unsuccessful Psychopaths.” *Journal of Abnormal Psychology* 119 (3): 546–54.

Yang, Yaling, Adrian Raine, Todd Lencz, Susan Bihrle, Lori LaCasse, and Patrick Colletti. 2005. "Volume Reduction in Prefrontal Gray Matter in Unsuccessful Criminal Psychopaths." *Biological Psychiatry* 57 (10): 1103–8.

Ye, Hang, Shu Chen, Daqiang Huang, Haoli Zheng, Yongmin Jia, and Jun Luo. 2015. "Modulation of Neural Activity in the Temporoparietal Junction with Transcranial Direct Current Stimulation Changes the Role of Beliefs in Moral Judgment." *Frontiers in Human Neuroscience* 9.

Yeudall, Lorne T., Orestes Fedora, and Delee Fromm. 1987. "A Neuropsychosocial Theory of Persistent Criminality: Implications for Assessment and Treatment." *Advances in Forensic Psychology & Psychiatry* 2: 119–91.

Yeudall, Lorne T. 1980. "A Neuropsychosocial Perspective of Persistent Juvenile Delinquency and Criminal Behavior: Discussion." *Annals of the New York Academy of Sciences* 347 (1): 349–55.

Young, Liane, Joan Albert Camprodon, Marc Hauser, Alvaro Pascual-Leone, and Rebecca Saxe. 2010. "Disruption of the Right Temporoparietal Junction with Transcranial Magnetic Stimulation Reduces the Role of Beliefs in Moral Judgments." *Proceedings of the National Academy of Sciences of the United States of America* 107 (15): 6753–58.

Young, Myla H., Jerald V. Justice, and Philip Erdberg. 2010. "Sexual Offenders in Prison Psychiatric Treatment: A Biopsychosocial Description." *International Journal of Offender Therapy and Comparative Criminology* 54 (1): 92–112.

———. 2012. "A Comparison of Rape and Molest Offenders in Prison Psychiatric Treatment." *International Journal of Offender Therapy and Comparative Criminology* 56 (7): 1103–23.

Zawidski, Tadeusz, and William P. Bechtel. 2004. "Gall's Legacy Revisited: Decomposition and Localization in Cognitive Neuroscience." In *Mind as a Scientific Object*, edited by Christina E. Erneling and David Martel Johnson. Oxford University Press.

Zeki, S., O. R. Goodenough, and Robert M. Sapolsky. 2004. "The Frontal Cortex and the Criminal Justice System." *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 359 (1451): 1787–96.

Zhang, Sheng, and Chiang-shan R. Li. 2012. "Functional Connectivity Mapping of the Human Precuneus by Resting State fMRI." *NeuroImage* 59 (4): 3548–62.

Zukov, Ilja, Radek Ptacek, and Slavomil Fischer. 2008. "EEG Abnormalities in Different Types of Criminal Behavior." *Activitas Nervosa Superior* 50 (4): 110–13.

Zurawicki, Leon. 2010. *Neuromarketing: Exploring the Brain of the Consumer*. Springer.