



**HAL**  
open science

## Neurobiologie de l'éthique

Thomas Boraud

► **To cite this version:**

Thomas Boraud. Neurobiologie de l'éthique. Revue de Bioéthique de Nouvelle-Aquitaine, 2020, 5, p.19-25. hal-03272911

**HAL Id: hal-03272911**

**<https://hal.science/hal-03272911>**

Submitted on 28 Jun 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Titre : Neurobiologie de l'éthique

Thomas Boraud,

Université de Bordeaux, UMR 5293, IMN, 33000, Bordeaux, France

CNRS, UMR 5293, IMN, 33000 Bordeaux, France

CHU de Bordeaux, IMN Clinique, 33000 Bordeaux, France

**Mots clés :** Éthologie – Sciences Cognitives – Neurosciences – Épistémologie – Utilitarisme – Déontologie

### Résumé : Neurobiologie de l'éthique

Depuis que le dogme des animaux-machines a volé en éclat au milieu du XXe siècle, éthologues et psychologues expérimentaux se sont peu à peu attachés à mettre en évidence différents degrés de conscience animale pouvant aller jusqu'à la conscience de soi chez certaines espèces. Ce constat a naturellement amené certains d'entre eux à se poser la question de préexistence de comportements éthiques dans d'autres espèces que chez l'homme et notamment chez les primates. Parallèlement, l'émergence de l'étude de la neurobiologie des processus de décisions a poussé des neurobiologistes à s'interroger sur le substrat cérébral des comportements éthiques et des jugements moraux. Le but de cet article est de faire une synthèse de ce double corpus et de mettre en évidence les lacunes ontologiques qui peuvent limiter la portée des conclusions avancées.

**Keywords :** Ethology – Cognitive Sciences – Neurosciences – Epistemology – Utilitarianism – Deontology

### Abstract : Neurobiology of Ethics

Since the refutation of the animal machines dogma in the middle of the 20th century, ethologists and experimental psychologists have gradually endeavored to highlight different degrees of animal consciousness that can go as far as self-awareness in certain species. This observation naturally led some of them to ask the question of the preexistence of ethical behavior in other species than in humans and in particular in primates. At the same time, the emergence of studies of the neurobiology of decision-making processes has led neurobiologists to question the cerebral substrate of ethical behavior and moral judgments. The aim of this article is to synthesize this double corpus and to highlight the ontological gaps which may limit the scope of the conclusions put forward.

## I- INTRODUCTION ET RAPPEL HISTORIQUE

Depuis une dizaine d'année, la question des bases neurobiologiques de l'éthique est devenue une discipline à part entière aux frontières de l'éthologie, des sciences cognitives et de la philosophie des sciences. Longtemps, le dogme des animaux-machines formulé par Descartes a prévenu l'attribution de comportement intentionnel chez les animaux. Mais sous l'influence du positivisme, il est apparu de plus en plus difficile de tenir une ligne aussi radicale et progressivement au cours du XXe siècle, s'est imposée l'idée qu'il existait un continuum entre la cognition animale et humaine, plutôt qu'une différence de nature irréductible. Éthologistes et psychologues expérimentaux se sont peu à peu attachés à mettre en évidence différents degrés de conscience animale culminant avec la reconnaissance de soi identifié chez une demi-douzaine d'espèces (mammifères et oiseaux) et des ébauches de théorie de l'esprit<sup>1</sup>. L'ensemble a culminé par une reconnaissance dans le code civil de l'animal comme "être vivant doué de sensibilité" dont on peut légitimement se demander s'il s'agit d'une bonne chose tant la définition vague de ce statut gomme les différences qu'il peut exister entre un mollusque et un grand singe et de ce fait en rend difficiles les applications concrètes. Quoi qu'il en soit, notre culture anthropocentrique pousse souvent les auteurs à évaluer ce continuum à l'aune du développement cognitif de l'enfant. Il n'est pas rare de lire que tel ou tel espèce a les mêmes capacités cognitives qu'un enfant de 18, 24 ou 36 mois...

A partir du moment où il a été mis en évidence que certaines espèces possèdent une conscience de soi et des autres, même embryonnaire, la question de préexistence de comportements éthiques s'impose d'elle-même. Frans de Waal en a popularisé l'idée à partir de ses travaux de la décennie 2000 chez les primates, puis ensuite chez d'autres espèces comme certains oiseaux ou l'éléphant.

Parallèlement, l'émergence de l'étude de la neurobiologie des processus de décisions qui s'est coagulé au début des années 2000 sous le néologisme de neuro-économie a poussé certains à s'interroger sur le substrat neurobiologique des comportements éthiques. En utilisant des paradigmes empruntés à la psychologie expérimental Joshua Greene et ses comparses ont cherché à identifier les structures cérébrales associées aux jugements moraux et en ont spéculé une évolution du jugement éthique à partir de ce que l'on sait du développement des structures cérébrales chez les mammifères et les primates.

L'ensemble et la multiplicité de ces travaux apportent un certain nombre de données qui d'un premier abord semble signaler un vaste progrès. Cependant, une analyse plus critique laisse apparaître des lacunes ontologiques qui limitent la portée des conclusions avancées de façon parfois péremptives. Le but de cet

<sup>1</sup> Rappelons que la théorie de l'esprit désigne la capacité à identifier des intentions et des besoins chez autrui.

article est d'en recenser les principales étapes et d'explorer leurs limites.

## II- L'ÉTHIQUE DU PRIMATE

Les premiers animaux, dont les capacités à avoir des comportements éthiques ont été testées, ont tout naturellement été d'autres primates. A notre connaissance, l'expérience la plus ancienne est celle qui a été proposée par Masserman, Whechkin & Terris (1964) dans les remous de l'expérience de Milgram (1965)<sup>2</sup> intitulée "comportements altruistes chez les singes rhésus". Les auteurs avaient conditionné des animaux à tirer sur une chaînette pour obtenir de la nourriture. Une fois les animaux entraînés, Les expérimentateurs avaient couplé la chaînette à un générateur de chocs électrique qui envoyait une décharge à un congénère situé dans une autre cage de la même pièce. Dans ces conditions, 75% des animaux préféraient se priver de nourriture en s'abstenant de tirer sur la chaînette afin d'éviter d'électrocuter leurs congénères. Les conclusions des auteurs étaient donc que les macaques rhésus étaient capables de comportement secourable, sous-entendu : contrairement aux êtres humains qui, sous couvert d'autorité, sont capable des pires exactions conformément à la thèse d'Hannah Arendt (2002) et aux travaux de Milgram. Mais Masserman et ses collaborateurs s'intéressaient plus au rôle des conflits internes dans la genèse des troubles psychiatriques et ces expériences furent oubliées.

Il faut attendre le tournant du XXIe siècle pour que la notion de bases biologiques de l'éthique émerge véritablement. C'est Frans de Waal qui en a popularisé l'idée à travers de nombreux articles scientifiques (Preston and de Waal, 2002; Brosnan and de Waal, 2003, 2004; Brosnan et al., 2005; Plotnik et al., 2006; de Waal, 2007) et d'ouvrages plus ou moins grand public (de Waal, 1997, 2006, 2011). Sa théorie dérive de ses travaux initiaux d'éthologie sur l'entraide menés sur les colonies de grands singes des zoos d'Arnhem et de San Diego (de Waal, 1990a, 1990b). Ils lui avaient permis d'observer que des coalitions d'animaux pouvaient se former afin de modifier les équilibres hiérarchiques au sein de la colonie. A partir de ces observations, il a décidé de mener des travaux expérimentaux sur les ressorts psychologiques de l'entraide chez les mammifères. Ses expériences sont destinées à démontrer que les notions de reconnaissance de soi (Plotnik et al., 2006), d'empathie (Preston and de Waal, 2002; Wynne, 2004; de Waal, 2007) et d'iniquité (Brosnan and de Waal, 2003, 2004; Brosnan et al., 2005) préexistent chez de nombreuses espèces de mammifères et sont à l'origine d'une éthique naturaliste.

<sup>2</sup> Rappelons que cette expérience faisait suite au procès d'Eichmann et visait à tester la théorie selon laquelle l'homme ordinaire pouvait infliger des souffrances à ses congénères si une autorité lui en donnait l'ordre. Les conditions dans lesquelles l'expérience a été réalisée et les conclusions qu'en a tiré Milgram sont toujours débattues, mais il a largement contribué à valider la théorie de la banalité du mal chère à d'Hannah Arendt.

Frans de Waal a récemment formalisé sa théorie sous la forme d'un processus gigogne (de Waal and Preston, 2017). Il existerait tout d'abord une capacité d'apprentissage inverse (qu'il appelle modèle perception-action) qui permet d'imiter un autre individu d'une même espèce. Les neurones miroirs<sup>3</sup> seraient au cœur de ce processus. Cette aptitude aurait ensuite permis de partager des ressentis émotionnels dont l'expression la plus basique serait le bâillement contagieux. En activant les circuits internes de récompense qui sont à la base des processus de renforcement par apprentissage, ce partage émotionnel aurait permis le développement des comportements empathiques appris pouvant aller chez certaines espèces jusqu'à l'anticipation des besoins d'autrui avec le développement d'un embryon de théorie de l'esprit. Cette approche, bien qu'élégante a cependant été critiquée pour son anthropomorphisme. Notamment l'interprétation des comportements décrits comme un refus de l'iniquité qui pourrait simplement être des manifestations de frustration (Henrich, 2004; Wynne, 2004).

### III- NEUROBIOLOGIE DU JUGEMENT MORAL<sup>4</sup>

Dans les années 1990, Antonio Damasio et ses collaborateurs ont mis en évidence l'importance des émotions dans la prise de décision (Bechara et al., 1994, 2000). Ces travaux ont contribué à l'étude neurobiologique des jugements moraux en établissant un lien entre les lésions cérébrales préfrontales et l'altération du raisonnement moral et de la cognition sociale (Anderson et al., 1999 ; Eslinger et al., 2004, 1992). A la suite de ces travaux, de nombreuses équipes se sont attachées à étudier les réseaux impliqués dans le sens de l'équité (Sanfey et al., 2003), la confiance (Andari et al., 2010), ou la théorie de l'esprit (Coricelli and Nagel, 2009), mais ce sont les équipes de Joshua Greene (2001) et Jorge Moll (2001, 2002, 2005) qui ont surtout cherché à préciser les structures impliquées dans la cognition morale grâce à l'utilisation de l'imagerie fonctionnelle, et ont conforté l'idée que les émotions jouent un rôle central dans la production des jugements moraux. Le paradigme expérimental consiste généralement à enregistrer l'activité cérébrale de sujets formulant un jugement à propos d'un dilemme moral ou non-moral (Moll et al., 2001)<sup>5</sup> ou à différentes

3 Les neurones miroir sont des neurones qui ont été identifiés pour la première fois dans le cortex pré-moteur de macaques rhésus par Giacomo Rizzolatti et son équipe (Gallese et al., 1996; Rizzolatti et al., 1996; Caggiano et al., 2009). Ils ont été appelés ainsi car ils ont la propriété de s'activer de façon identique pour un geste spécifique qu'il soit effectué par l'individu chez lequel les enregistrements sont effectués ou par un congénère qu'il observe. Depuis, Rizzolatti insiste sur le fait que ces neurones se retrouvent aussi dans d'autres aires cérébrales et joueraient un rôle vraiment central dans les mécanismes d'empathie (Rizzolatti and Caruana, 2017).

4 Une version plus développée de ce paragraphe peut être trouvée dans Penavayre, Brun & Boraud (Penavayre et al., 2019).

5 Dans l'étude publiée par Moll et coll. (Moll et al., 2001), les participants étaient chargés d'évaluer le caractère moral et la valence émotionnelle d'une série d'énoncés présentant un contenu moral explicite (ex : « Je transgresse la loi si nécessaire »), ou des énoncés dépourvus de connotation morale (ex : « Les pierres sont constituées d'eau »). Dans l'étude de Greene et coll. (Greene et al., 2001) les sujets devaient formuler des jugements à propos de dilemmes

versions d'un dilemme moral (Greene et al., 2001). L'activité cérébrale peut être recueillie par de l'imagerie par résonance magnétique fonctionnelle (IRMf Greene et al., 2001) ou l'électroencéphalographie (EEG ; e.g. Yoder and Decety, 2014). Il peuvent aussi interférer temporairement avec celle-ci à l'aide de la stimulation transcrânienne magnétique (TMS ; e.g. Young et al., 2010) ou à courant continu (tDSC ; e.g. Kuehne et al., 2015). Les régions impliquées dans la formation des jugements moraux identifiées grâce à la convergence des observations anatomo-cliniques et des données expérimentales obtenues sont les suivantes : le cortex préfrontal ventro-médian, le cortex orbito-frontal, le cortex temporal antérieur droit, l'insula et l'amygdale.

Greene introduit une distinction entre deux types de dilemmes moraux : les dilemmes "impersonnels" et les dilemmes "personnels" qu'il formalise à l'aide du paradigme du trolley. Sous sa forme impersonnelle, il se présente généralement ainsi : "Vous êtes au volant d'un trolley hors de contrôle qui approche un aiguillage. Sur la voie de gauche se trouve un groupe de cinq ouvriers de maintenance. Sur la voie de droite se trouve un seul ouvrier de maintenance. Si vous ne faites rien, le trolley poursuivra sa route vers la gauche, causant la mort des cinq ouvriers. La seule manière d'éviter la mort de ces ouvriers est d'actionner un levier sur le tableau de bord qui déviara le trolley sur la voie de droite, causant la mort de l'ouvrier isolé. Est-il selon vous approprié d'actionner le levier afin d'éviter la mort des cinq ouvriers ?"

La forme personnelle est la suivante : "Un trolley hors de contrôle se dirige vers cinq ouvriers qui seront tués s'il poursuit sa route. Vous vous trouvez sur une passerelle au-dessus de la voie entre le trolley et les cinq ouvriers. À côté de vous se trouve un homme de forte corpulence que vous ne connaissez pas. La seule manière de sauver la vie des cinq ouvriers est de pousser cet inconnu sur les voies où son corps arrêtera le trolley. L'inconnu mourra si vous faites cela mais les cinq ouvriers seront sauvés. Est-il selon vous approprié de pousser l'inconnu sur la voie afin de sauver les cinq ouvriers ?"

Bien que la situation puisse être décrite dans les deux cas comme le sacrifice d'une personne pour en sauver cinq, le changement de formulation entraîne des réponses différentes. Dans son étude publiée en 2013, Joshua Greene rapporte que 87% des sujets interrogés approuvent le fait d'actionner le levier et donc de sacrifier un ouvrier pour en sauver cinq dans le dilemme du trolley impersonnel, alors que seulement 31% des

moraux et non-moraux tandis que leur activité cérébrale était enregistrée par IRMf. Les dilemmes non-moraux portaient sur des choix de la vie ordinaire contraints par différents facteurs comme par exemple réaliser un seul trajet pour transporter des plantes au prix de salir les sièges arrière en cuir de votre véhicule, ou de faire deux trajets en n'utilisant que le coffre de votre véhicule que vous avez préalablement protégé par des bâches. Nous estimons que la plupart des dilemmes non-moraux utilisés par Greene et coll. (2001) ne sont pas des dilemmes mais des problèmes pragmatiques à l'image de celui que nous venons de mentionner (Voir McGuire et al., 2009 p.578).

sujets approuvent le fait de pousser l'inconnu sur les rails dans le dilemme de la passerelle (Greene, 2013). Son interprétation est que la distinction entre les deux formules tient à ce que le dilemme du trolley impersonnel n'implique pas d'action de proximité ou en face à face avec la personne seule sur la voie (d'où sa qualification d'impersonnel, Greene et al., 2001) ; tandis que le dilemme personnel implique une action à proximité de ou en face à face avec l'homme de forte corpulence. Il en déduit que les dilemmes moraux personnels impliquent un traitement émotionnel plus important que les impersonnels et que ces différences d'engagement émotionnel affectent leur jugement. Selon lui, l'allongement des temps de réaction des participants lorsque l'action de sacrifice est jugée "moralement appropriée" dans la condition du dilemme personnel apportent un argument supplémentaire (Greene et al., 2001) : il suggère que ce contraste est lié à une interférence émotionnelle liée au contact physique que la décision impose.

L'IRMf associé aux expériences a mis en évidence une activité plus importante au niveau du cortex préfrontal dorso latéral (CPDL, une structure impliquée notamment dans la délibération rationnelle) dans la condition morale impersonnelle ; tandis que dans la condition morale-personnelle on observe une activité plus importante du cortex préfrontal ventro-médian (CPFVM, associé à la régulation émotionnelle).

L'ensemble de ces résultats les a amenés à conclure que les dilemmes moraux impliquaient une compétition de processus cognitifs et émotionnels ; ces processus s'inhibant l'un l'autre selon le type de dilemme présenté et la nature du jugement exprimé. Ces résultats posent les fondations d'un modèle duel du traitement des jugements moraux selon lequel la résolution des dilemmes résulte d'une opposition ou d'un conflit entre deux systèmes. D'un côté, la présentation d'un dilemme moral-impersonnel met en jeu des processus cognitifs qui donnent lieu à des jugements qualifiés d'"utilitaristes" (l'action de sacrifier une personne pour en sauver cinq est jugée "moralement appropriée") ; de l'autre, la présentation d'un dilemme moral-personnel produit une réponse émotionnelle immédiate qui, lorsqu'elle n'est pas inhibée par les processus cognitifs, favorise une option considérée comme "déontologiste" (l'action de sacrifier une personne pour en sauver cinq est jugée "moralement inappropriée").

Le modèle dualiste de Greene et al. (Greene et al., 2001, 2004; Greene, 2003, 2013, 2014) fait désormais figure de paradigme dans les domaines de la psychologie morale et de la neuropsychologie. Cette influence relève d'une ambition affichée par Joshua Greene (2015), : "La science peut faire progresser l'éthique en révélant les fonctionnements internes cachés de nos jugements moraux, particulièrement ceux que nous faisons intuitivement. Une fois ces fonctionnements internes révélés, nous pourrions être moins confiants dans certains de nos jugements et vis-à-vis des théories éthiques qui sont (explicitement ou implicitement) fondées sur ceux-ci."

#### IV- CAVEAT

Malgré leurs intérêts indéniables, ces approches sont néanmoins critiquées et critiquables. La quête des bases neurobiologiques de l'éthique est souvent entachée d'anthropomorphisme qui biaise l'interprétation des résultats, mais aussi la conception des expériences. Le premier problème touche au substrat neurobiologique des émotions. Pour opérationnaliser ce concept, les neurobiologistes ont choisi de ne pas tenir compte du compte-rendu subjectif des sujets de leur expérience se plaçant ainsi dans la continuité des principes établis par la psychologie expérimentale depuis Watson. Cette stratégie, qui vise à neutraliser les phénomènes de rationalisation post-hoc, conduit à retirer du champ d'étude l'expérience consciente des émotions qui en est pourtant une composante essentielle. L' "émotion" devient un état identifié fonctionnellement à partir de ses relations causales avec les autres états internes du système cognitif, les stimuli environnementaux et les comportements observables. Néanmoins, en définissant le concept de partage émotionnel, Frans de Waal réintroduit un ressenti subjectif indémontrable qui devient ainsi un maillon faible de sa théorie.

Greene, quant à lui, utilise le terme émotion pour désigner des structures cérébrales corrélées avec des états émotionnels dont l'activité est détectée lors de l'expérimentation. Si la corrélation entre les signaux recueillis et les processus émotionnels est avérée chez l'homme, leur lien causal est beaucoup plus problématique. Un exemple éclairant est celui de "l'hormone du plaisir" qui a un temps désigné la dopamine, alors qu'il est maintenant avéré qu'elle joue un rôle beaucoup plus complexe impliquant la motivation, l'évaluation de l'erreur, etc. (Schultz et al., 1997; Schultz, 2006). Par ailleurs, ce n'est pas parce qu'on observe un signal dans une structure que celle-ci est causalement impliquée. A fortiori cela n'implique pas que l'on puisse identifier fonctionnellement un état émotionnel dans le mécanisme de jugement moral et, plus encore, que l'on puisse attribuer à l'émotion un rôle prépondérant dans la formation d'un jugement "déontologiste".

Il y a donc de bonnes raisons de penser que le modèle du traitement duel du jugement moral est insuffisant pour trancher le débat entre utilitarisme et déontologisme sur des bases empiriques. Le fait même que les procédures expérimentales (et le cadre théorique dans lequel celles-ci prennent place) partent de préconceptions « utilitaristes » ou au moins favorables à l'utilitarisme en constitue un indice supplémentaire. Il est clair que nos jugements moraux (comme toutes nos décisions) reposent sur une dynamique complexe entre des systèmes distincts et qu'il est vain de chercher dans le cerveau (ou ailleurs) une moralité unifiée, un « nerf moral » sur la base duquel nous les formons. Les résultats issus d'expérimentations sur des dilemmes sacrificiels sont importants parce qu'ils mettent à l'épreuve les intuitions philosophiques discutées à leurs propos. Mais

ils échouent à clarifier, voire à clore, le débat entre déontologisme et utilitarisme parce que les dilemmes moraux constituent des situations limites qui ne peuvent être pleinement traitées par le déontologisme. Tout dilemme sacrificiel implique dans sa structure même une forme plus ou moins explicite d'utilitarisme (au moins un calcul de maximisation de l'utilité espérée et, souvent, une prise en compte de l'intérêt du plus grand nombre).

## V- CONCLUSIONS

L'ensemble de ces critiques ne doit pas nous amener à conclure à l'inutilité des recherches en neurobiologie des jugements moraux. Ces travaux permettent de mieux identifier ce qui, dans le processus même de leur opérationnalisation, demande à être plus précisément défini.

Une éthique naturalisée ne peut se réduire à l'élaboration d'une morale normative sur la base de données empiriques. La moralité est propre à chaque culture et a évolué en accommodant les intérêts personnels et collectifs des humains vivant en groupes interdépendants, permettant de réguler les conflits, la répartition des tâches et les organisations hiérarchiques. Les systèmes de normes morales ont permis de rendre possible la coopération au bénéfice de l'espèce. Afin de saisir cette généalogie dans toutes ses dimensions, nous ne pouvons, en tant que naturalistes, nous en remettre qu'aux données de la biologie évolutionniste, de l'éthologie, de la psychologie du développement des neurosciences cognitives et comportementales et de l'anthropologie culturelle. Ces éléments sont pertinents parce qu'ils contraignent la manière dont nos morales normatives se déploient dans nos pratiques concrètes. Le point aveugle principal de l'entreprise de fondation d'une morale normative sur la base des seules neurosciences cognitives et comportementales conduit à ignorer la dimension historique et sociale de notre moralité. La légitimité de l'intervention des neurosciences comportementales et cognitives dans l'entreprise de naturalisation de la morale n'est garantie qu'à la condition qu'elle limite ses prétentions à éclairer partiellement notre moralité. Prétendre qu'elles peuvent établir empiriquement la supériorité d'une morale normative sur une autre, revient à outrepasser les limites de sa pertinence épistémique.

## Bibliographie

- Andari E, Duhamel J-R, Zalla T, Herbrecht E, Leboyer M, Sirigu A (2010) Promoting social behavior with oxytocin in high-functioning autism spectrum disorders. *Proc Natl Acad Sci* 107:4389–4394 Available at: <http://www.pnas.org/content/107/9/4389.abstract>.
- Anderson SW, Bechara A, Damasio H, Tranel D, Damasio AR (1999) Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nat Neurosci* 2:1032.
- Arendt H (2002) *Les origines du totalitarisme ; Eichmann à Jérusalem*.
- Bechara A, Damasio AR, Damasio H, Anderson SW (1994) Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition* 50:7–15 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8039375>.
- Bechara A, Damasio H, Damasio AR (2000) Emotion, decision making and the orbitofrontal cortex. *Cereb Cortex* 10:295–307 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10731224>.
- Brosnan SF, de Waal FBM (2003) Monkeys reject unequal pay. *Nature* 425:297–299.
- Brosnan SF, de Waal FBM (2004) Socially learned preferences for differentially rewarded tokens in the brown capuchin monkey (*Cebus apella*). *J Comp Psychol* 118:133–139 Available at: [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=15250800](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15250800).
- Brosnan SF, Schiff HC, de Waal FBM (2005) Tolerance for inequity may increase with social closeness in chimpanzees. *Proc R Soc B Biol Sci* 272:253–258 Available at: <http://rspb.royalsocietypublishing.org/cgi/doi/10.1098/rspb.2004.2947>.
- Caggiano V, Fogassi L, Rizzolatti G, Thier P, Casile A (2009) Mirror neurons differentially encode the peripersonal and extrapersonal space of monkeys. *Science* (80- ) 324:403–406.
- Coricelli G, Nagel R (2009) Neural correlates of depth of strategic reasoning in medial prefrontal cortex. *Proc Natl Acad Sci U S A* 106:9163–9168 Available at: [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=19470476](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19470476).
- de Waal FBM (1990a) *De la réconciliation chez les primates*. Flammarion.
- de Waal FBM (1990b) *La politique du Chimpanzé*. Editions du Rocher.
- de Waal FBM (1997) *Le bon singe : Les bases naturelles de la morale*. Bayard Editions.
- de Waal FBM (2006) *Le Singe en nous*. Paris, France: fayard.
- de Waal FBM (2007) With a Little Help from a Friend. *PLoS Biol* 5:e190.
- de Waal FBM (2011) *L'âge de l'empathie*. Babel.
- de Waal FBM, Preston SD (2017) Mammalian empathy: Behavioural manifestations and neural basis. *Nat Rev Neurosci* 18:498–509 Available at: <http://dx.doi.org/10.1038/nrn.2017.72>.
- Eslinger PJ, Flaherty-Craig C V, Benton AL (2004) Developmental outcomes after early prefrontal cortex damage. *Brain Cogn* 55:84–103.
- Eslinger PJ, Grattan L, Damasio H, Damasio AR (1992) Developmental consequences of childhood frontal lobe damage. *Arch Neurol* 49:764–769.
- Gallese V, Fadiga L, Fogassi L, Rizzolatti G (1996) Action recognition in the premotor cortex. *Brain* 119 ( Pt 2:593–609.
- Greene J (2003) From neural “is” to moral “ought”: what are the moral implications of neuroscientific moral psychology? *Nat Rev Neurosci* 4:846–850.
- Greene J (2014) *Cognitive Neurosciences of Moral Judgement and Decision Making*. In: *Cognitive Neurosciences* (Gazzaniga MS, Mangun GR, eds). Cambridge, Massachusetts, London, England: A Bradford Book / The MIT Press.

- Greene J, Sommerville RB, Nystrom LE, Darley JM, Cohen JD (2001) An fMRI investigation of emotional engagement in moral judgment. *Science* (80- ) 293:2105–2108.
- Greene JD (2013) *Moral tribes : emotion, reason, and the gap between us and them*. New York: The Penguin Press.
- Greene JD (2015) Beyond Point-and-Shoot Morality: Why Cognitive (Neuro)Science Matters for Ethics. *Law Ethics Hum Rights* 9:141–172.
- Greene JD, Nystrom LE, Engell AD, Darley JM, Cohen JD (2004) The neural bases of cognitive conflict and control in moral judgment. *Neuron* 44:389–400.
- Henrich J (2004) Animal behaviour: inequity aversion in capuchins? *Nature* 428:139; discussion 140.
- Kuehne M, Heimrath K, Heinze H-J, Zaehle T (2015) Transcranial Direct Current Stimulation of the Left Dorsolateral Prefrontal Cortex Shifts Preference of Moral Judgments Zalla T, ed. *PLoS One* 10:e0127061.
- Masserman JH, Wechkin S, Terris W (1964) Altruistic Behavior in Rhesus Monkeys. *Am J Psychiatry* 121:584–585.
- McGuire J, Langdon R, Coltheart M, Mackenzie C (2009) A reanalysis of the personal/impersonal distinction in moral psychology research. *J Exp Soc Psychol* 45:577–580.
- Milgram S (1965) Some Conditions of Obedience and Disobedience to Authority. *Hum Relations* 18:57–76.
- Moll J, de Oliveira-Souza R, Eslinger PJ, Bramati IE, Mourão-Miranda J, Andreiuolo PA, Pessoa L (2002) The Neural Correlates of Moral Sensitivity: A Functional Magnetic Resonance Imaging Investigation of Basic and Moral Emotions. *J Neurosci* 22:2730–2736.
- Moll J, Eslinger PJ, Oliveira-Souza R de (2001) Frontopolar and anterior temporal cortex activation in a moral judgment task: preliminary functional MRI results in normal subjects. *Arq Neuropsiquiatr* 59:657–664.
- Moll J, Zahn R, de Oliveira-Souza R, Krueger F, Grafman J (2005) The neural basis of human moral cognition. *Nat Rev Neurosci* 6:799–809.
- Penavayre M, Brun C, Boraud T (2019) Neurobiologie des jugements moraux, avancée épistémique ou voie sans issue ? In: *Éthique et sciences cognitives, Intellectica*, 70 (Khamassi M, Chatila R, Mille A, eds), pp 63–82.
- Plotnik JM, de Waal FBM, Reiss D (2006) Self-recognition in an Asian elephant. *Proc Natl Acad Sci U S A* 103:17053–17057.
- Preston SD, de Waal FBM (2002) Empathy: Its ultimate and proximate bases. *Behav Brain Sci* 25:1–71.
- Rizzolatti G, Caruana F (2017) Some considerations on de Waal and Preston review. *Nat Rev Neurosci* 18:769 Available at: <http://dx.doi.org/10.1038/nrn.2017.139>.
- Rizzolatti G, Fadiga L, Gallese V, Fogassi L (1996) Premotor cortex and the recognition of motor actions. *Brain Res Cogn Brain Res* 3:131–141.
- Sanfey AG, Rilling JK, Aronson JA, Nystrom LE, Cohen JD (2003) The neural basis of economic decision-making in the Ultimatum Game. *Science* (80- ) 300:1755–1758.
- Schultz W (2006) Behavioral theories and the neurophysiology of reward. *Annu Rev Psychol* 57:87–115 Available at: [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=16318590](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=16318590).
- Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science* (80- ) 275:1593–1599.
- Wynne CD (2004) Animal behaviour: fair refusal by capuchin monkeys. *Nature* 428:140; discussion 140.
- Yoder KJ, Decety J (2014) Spatiotemporal neural dynamics of moral judgment: A high-density ERP study. *Neuropsychologia* 60:39–45.
- Young L, Camprodon JA, Hauser M, Pascual-Leone A, Saxe R (2010) Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. *Proc Natl Acad Sci* 107:6753–6758.